Introduction

The economic forces that determined profitability change whenever

- technology
    - cassettes, the media changed. Also people like to book their own vacation instead of going to a travel agency.
- regulation
    - when Uber got to Belgium, it was forbidden!
- market information
    - customers have more market information, and that also changes to profitability of activities. For example, we now know all the prices of items in all countries. So countries cannot exploit customers like that anymore.
- consumer preferences
- relative costs

change. Consequently, companies that grow profitably in changing markets often need to break old rules and create new pricing models. A few examples:

- Netflix: they made use of new technology and completely changed a segment in the market. In the past you paid per video. With Netflix you have a subscription, and unlimited movies and series.
- Ryanair: tried to offer the cheapest flight possible. Before Ryanair, they sold tickets differently. They took all the services, and took them apart instead of selling a ticket with all the services included. They priced all the services separately. So customers only paid for these services that they valued, and not for the ones they didn't.
- Online advertising: in print media, you paid for the number of customers, the size of the readership. But of course they never knew if the reader actually read the add! Today in online advertising, they pay per click they get on their advertisements.

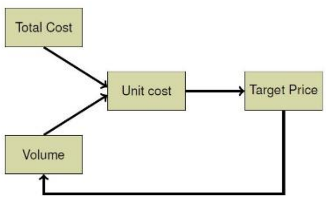
The 3 C's

Imagine you have an entrepreneur friend and he went on Alibaba and he has bought the whole lot of wooden trains. But at which price should he sell them? What strategies could you use to sell these trains?

- Look at the competitors. But you rely on the ideas of your competitor and this might not always be the same as your ideas.
- Ask consumers how much they are willing to pay for the product.
- You want to sell your trains for more than what they have cost you.

Cost-plus

You calculate the total cost of your product or service. You have the volume of your product or service. Based on that we can calculate what the unit cost is, and they we decide upon the mark-up (how much profit do you want). There is a problem with this reasoning though: there is a feedback loop from the target price to the volume.



EXAMPLE

The next table gives the projected costs and revenues at an expected sales of 1 000 000 units. So if we sell 1 000 000 trains, it will costs us €7,50 per train, so we will have to sell the train for € 9. Here the

	Total	Per Unit
Direct Variable Costs	€3,000,000	€3.00
Direct Fixed Costs	€3,000,000	€3.00
Administrative Overhead	€1,500,000	€1.50
Full Cost	€7,500,000	€7.50
Revenue	€9,000,000	€9.00
Profit	€1,500,000	€1.50

profit margin is €1,5 so we have a profit of €1 500 000. What is the danger here? We have to work with an expected number of sales. We have to make a prediction on how many sales we will have.

Now it turns out that you can only sell 750 000 trains! How does the table change then? We would have a company who is going bread even instead of profitable. Cost-plus tells us the unit cost before putting a price on the product. So the total fixed costs stay the same, but the cost per unit will become higher! We now have a full cost of €9 per unit, which means that we have a break even. We are not making any money anymore.

	Total	Per Unit
Direct Variable Costs	€2,250,000	€3.00
Direct Fixed Costs	€3,000,000	€4.00
Administrative Overhead	€1,500,000	€2.00
Full Cost	€6,750,000	€9.00
Revenue	€6,750,000	€9.00
Profit	€0	€0

	Current	5% ↓ Unit Sales	10% ↓ Unit Sales
Price	€9.00	€10.50	€10.50
Unit Sales	750,000	712,500	500,000
Variable Costs	€3.00	€3.00	€3.00
Fixed Costs	€4.00	€4.21	€6.00
Admin. Overhead	€2.00	€2.11	€3.00
Unit Cost	€9.00	€9.32	€12.00
Unit Profit	€0	+€1.18	-€1.50
Total Profit	€0	€843,750	-€750,000

In order to make a good expectation of the number of sales, you have to take the price into account. But also the other way around! In the second example, what would be the decision with the price? We have to increase again if we want to make profit, but you've already sold less! So why should you increase more? Well it depends on the drop of the sales. The picture on the left is the projected costs and revenues with the price increased to €10,5 per unit. If the demand drops only with 5%, we would make profit. If the demand drops with 10% then we are losing money! Here come price elasticities in handy: how will demand change, for a change in price. If the elasticity is high, that means that the slightest change will give a strong reaction in the demand. If the elasticity is low, there will not much reaction in the demand when the price changes. Same is true if we decrease the price. In the picture on the right, they show the financial implications of a 10% price cut. If the sales only go up by 5%, you are still in red. If they go up with 33%, then you have profit again! So the problem with cost-plus pricing is fundamental: in most industries it is impossible to determine a product's unit cost before determining its price. Why? Because unit costs change with volume.

	Current	5% ↑ Unit Sales	33% ↑ Unit Sales
Price	€9.00	€8.10	€8.10
Unit Sales	750,000	787,500	1,000,000
Variable Costs	€3.00	€3.00	€3.00
Fixed Costs	€4.00	€3.81	€3.00
Admin. Overhead	€2.00	€1.90	€1.50
Unit Cost	€9.00	€8.71	€7.50
Unit Profit	€0	-€0.61	+€0.60
Total Profit	€0	-€480,375	€600,000

Customer driven

There are many versions of customer-based pricing. When looking at B2B, the salesforce allows the purchasing agents to dictate the prices. And when we look at B2C, they can give a way a valuable product for free. But there are also problems with customer based pricing. People don't always do what they say they would do:

- Customers do not always reveal how much they value the product. They can say a price higher than they would actually pay, because they don't want to hurt the person. Or they can say a price lower than they would actually pay, to get a good deal.
- Customers need to be educated about the value of the product. Maybe they don't know the product. For example, Steve Jobs had to educate people about the iPad.
- Customers used to dictate prices will revolt to price changes.

Marketing approach: we should ask the customer what they want, because we want happy customers. But then customers know they can dictate the prices. Yves Rocher: they gave the customers lots of discounts. It was a mechanism that was set up in the past, so they couldn't stop it! As soon as they stopped the discounts, the customers stopped buying. They were too focussed on making profit, that the customer started to dictate the prices. If you offer a reduction, it should be very clear that it is rare and won't be coming back any time soon.

The purpose of pricing is NOT simply to create satisfied customers! Marketing should raise the willingness-to-pay of customers! They should not be focussed on raising the market share. It should also not trick customers into making a one time purchase.

<u>Competition based</u>

Finally, consider the policy of letting pricing be dictated by competitive conditions. This is a pricing strategy that does not reflect the product, but the pricing strategies of competitors. You copy the pricing of your competitors, and you focus on market share. Why should an organization want to achieve market-share goals? Because managers believe that more market share usually produces greater profit. Priorities are confused, however, when managers reduce the profitability of each sale, simply to achieve the market-share goal. Prices should be lowered only when they are no longer justified by the value offered in comparison to the value offered by the competition. If you copy the price of your competitor, the danger might be that your quality might not be the same quality as yours. If you have more to offer for example, then it is fair that this is reflected in the price. This pricing technique can also lead to price wars. If you competitor also decides to lower their price because of yours, then you get a price war. The one with the lowest cost structure will win the price ware. This is only a good idea as an initial strategy to attract some new customers. But after a while you really have to increase the price, so it matches the quality that you offer.

Colruyt has been optimizing their supply system, so as a new company it makes no sense to compete with them based on price. Walmart's diapers are offered at an extremely low price. They can buy them at huge volumes an then they can sell them at a low price. They do that to attract customers to the store. The goal is not to be profitable of these diapers, but to be profitable on other products customers buy when they go for diapers.

The problem with competitor based pricing is that it incourages firms to ignore their unique value proposition, it can lead to price wars (which end in a monopoly, which is bad for customers in the end), it makes companies focus to much on market share and it does not necessarily lead to maximum profits.

Strategic pricing

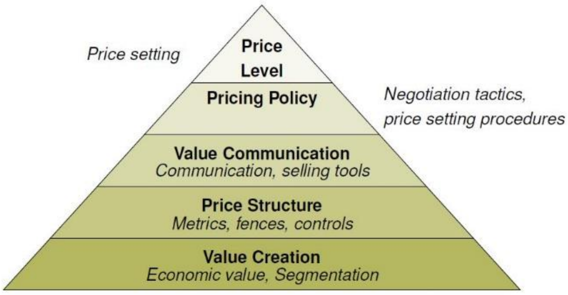
This is a combination of the 3 C's we have seen.

The objective of strategic pricing is profitability. Achieving exceptional profitability requires managing much more than just price levels. It requires ensuring that products and services include just those features that customers are willing to pay for, without those that unnecessarily drive up cost by more than they add value. Although more than one strategy can achieve profitable results, even within the same industry, nearly all successful pricing strategies embody three principles:

- **Value-based**. This means that differences in pricing across customers and changes over time reflect differences or changes in the value to customers. For example, many managers ask whether they should lower prices in response to reduced market demand during a recession. The answer: if customers receive less value from your product or service because of the recession, then prices should reflect that. But the fact that fewer customers are in the market for your product does not necessarily imply that they value it less than when they were more numerous. Unless a close competitor has cut its price, giving customers a better alternative, there may be no value-based reason for you to do so.
- **Proactive**. This means that companies anticipate disruptive events and develop strategies in advance to deal with them.

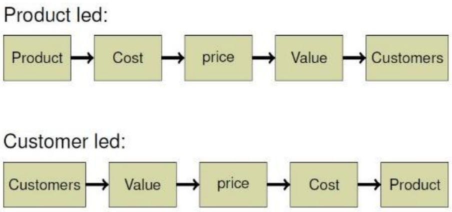
- **Profit-driven**. This means that the company evaluates its success at price management by what it earns relative to alternative investments rather than by the revenue it generates relative to its competitors.

A good pricing strategy involves five distinct but very different sets of choices that build upon one another. The choices are represented graphically as the five levels of the strategic pricing pyramid, with those lower in the pyramid providing the necessary support, or foundation, for those above.



1. Value creation

The next picture illustrates the flawed logic that leads many companies to product good quality, but poor value. Engineering and manufacturing departments design and make what they consider a "better" product. In the process, they make investments and incur costs to add features and services. Finance then totals these costs to determine "target" prices. Only at this stage does marketing enter the process, charged with the task of demonstrating enough value in these better products and services to justify premium prices to customers. But this is wrong! Solving the problems of cost-based pricing requires more than a quick fix in the mark-up. It requires a complete reversal of the process, starting with the customers! The target price is based on estimates of the value of features and services given the competitive alternatives and the portion of it that the firm can expect to capture in its price by segment. The job of financial management is not to insist that prices recover costs. It is to insist that costs are incurred only to make products that can be priced profitably given their value to the targeted customers. Most companies are product led. They have something that works, and they don't want to change it.



2. Price structure

Once you understand how value is created for different customer segments, the next step in building a pricing strategy is to create a pricing structure. The most simple price structure is a price per unit and is perfectly adequate for commodity products and services. The purpose of more complicated price structures is to reflect differences in the potential contribution that can be captured from different customer segments by capturing the best possible price from each segment, making the sale at the lowest possible cost, or both. Here are some structures explained short, they will be explained further throughout the cursus:

- Bundling: sell the PlayStation and the games together instead of apart.
- Two part tariffs: the price is split into different parts. For example Cambio car sharing. You have a fixed subscription for a year and then you pay per km that you drive with one of the cars.
- Customized pricing: different prices for different customers.

3. Price and value communication

Understanding the value your products create for customers and translating that understanding into a value based price structure can still result in poor sales unless customers recognize the value they are obtaining. A successful pricing strategy must justify the prices charged in terms of the value of the benefits provided. Developing price and value communications is one of the most challenging tasks for marketers because of the wide variety of product types and communication vehicles. In some instances, marketers might employ traditional advertising media to convey their differential value, as was the case with the now famous “I’m a Mac” ads created by Apple. The ads, featuring the actors Justin Long posing as a Mac and John Hodgman as a PC, highlighted common problems for PC owners not faced by Mac owners. In other instances, value messages will be communicated directly during the sales process with the aid of illustrations of value experienced by customers within a market segment or with the aid of a spreadsheet model to quantify the value of an offering to a particular customer.

4. Pricing policy

Ultimately, the success of a pricing strategy depends upon customers being willing to pay the price you charge. The rationale for value-based pricing is that a customer’s willingness-to-pay for one product versus another should track closely with differences in the relative value of those products. When customers become increasingly resistant to whatever price a firm asks, most managers would draw one of three conclusions:

- That the product is not offering as much value as expected,
- That customers do not understand the value
- Or that the price is too high relative to the value.

But there is another possible and very common cause of price resistance. Customers sometimes decline to pay prices that represent good value, simply because they have learned that they can obtain even better prices by exploiting the sellers’ pricing process.

Pricing policy refers to rules or habits, either explicit or cultural, that determine how a company varies its prices when faced with factors other than value and cost to serve that threaten its ability to achieve its objectives. Good policies enable a company to achieve its short-term objectives without causing customers, sales reps, and competitors to adapt their behavior in ways that undermine the volume or profitability or future sales. Poor pricing policies create incentives for customers, sales reps, or competitors to behave in ways that will undermine future sales or customers’ willingness-to-pay.

In the terminology of economics, good policies enable prices to change along the demand curve without changing expectations in ways that cause the demand curve to “shift” negatively for future purchases.

5. Price level

According to economic theory, setting prices is a straightforward exercise in which the marketer simply sets the price at the point of the demand curve where marginal revenues are equal to the marginal costs. But it is not so simple. On the one hand, it is impossible to predict how revenues will change following a price change because of the uncertainty about how customers and competitors will respond. On the other hand, the accounting systems in most companies are not equipped to identify the relevant costs for pricing strategy decisions, often causing marketers to make unprofitable pricing decisions. Price setting should be an iterative and cross-functional process led by marketing that includes several key actions.

- The first action is to set appropriate pricing objectives, whether that means to use price to drive volume or to maximize margins.
- The second action is to calculate price-volume trade-offs. Example: a 10% price cut for a product with a 20% contribution margin would have to result in a 100% increase in sales volume to be profitable.

Once the price-volume trade-offs are made explicit for a particular pricing move, the next activity is to estimate the likely customer response by assessing the drivers of price sensitivity that are unrelated to value.

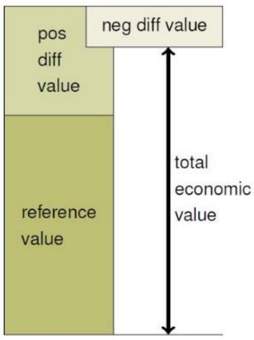
Value creation

At the foundation of the pyramid, in what should be the first task of any strategic marketing organization, is gaining a deep understanding of how products and services create value for customers, the essential initial input to pricing strategy. Many firms fail to understand and leverage their potential to create value through their products, services, and customer relationships. They erroneously assume that merely adding features or improving performance will lead to profitable gains in price, volume, or both. But more and better features will not lead to greater profitability unless those features translate into higher monetary and/or psychological value for the customer. In the marketing organization, understanding how value differs across segments provides the essential insight needed to make more profitable offer design and bundling choices.

The value at the heart of pricing strategy is the **total economic value**. This depends on the alternatives customers have available to satisfy the same need. The economic value accounts for the fact that the value one can capture for commodity attributes of an offer is limited to whatever competitors charge for them. Only the part of economic value associated with differentiation, which we call differentiation value, can potentially be captured in the price. **The differentiation value** comes in two forms:

- **Use value**: the overall satisfaction that a customer receives from using a product or service offering.
- **Psychological valu**e: the many ways that a product creates satisfaction for the customer that can not easily be quantified.

Example: you have manufactured a machine to cut wood and your machine can cut wood 2 times faster than the reference machine. This is more easily quantified.
Example: you work in a beach bar and you have to decide about the price of Coca-Cola.



- Positive differentiation value: how does your product differentiate positively in turn of the reference product (the supermarket Coca-Cola). What do we offer that the supermarket doesn't. The atmosphere, there's music at the beach, ice cubes,...
- Negative differentiation value: how does your product differentiate negatively in turn of the reference product. You have to wait for a waiter, you have to wait a long time.
- **Reference value**: the value of the best possible alternative that you have. A supermarket where you can also buy a Coca-Cola. In the supermarket the price is €1,5 for that drink.

If you then take the sum of these three, you get the total economic value. The difficulty is "how do you quantify these things?".

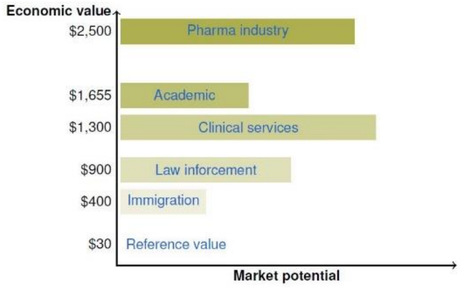
<u>EXAMPLE: DYNA-TEST versus GENETICORP</u>

The total positive differentiation value is $2498. The best competitor at the moment has a price of $30, which is very cheap! The product is actually worth $2498, so we are going to have great difficulties to sell our products at this price. What are the options?

- We can lower the price: we give more than what we ask. If we decrease our price, then we are giving more than the customers have to pay for.
- We can educate the people: for this you have to make sure that you have a good marketing system. So that potential customers know that our product is much better than EnSyn.

- We can change the price structure: $30 gives you 1 test at EnSyn. Now instead of asking so much $$$ per test, we can have a different structure. Maybe a price per year is better, for an unlimited number of tests a year.

You can do this for different segments to find the best segments to focus on. Whatever the service, there will be segments that will value your products more than other segments, and these segments can be of different sizes. In this case: Pharma should be willing to pay a large price for the product. Imagine you are a manager for Dyna tests, and you have to decide 'how much should we ask for our product?'. Because if we set $2500 then we will only sell to the pharma industry. If we set the price on $40, we can sell to all the segments, but a lot of the segments will get more than what they pay for, so we leave some money at the table. So the question is: "how can we sell the same product to different segments, for a different price".



<u>EXERCISE: SPRAYCO</u>
Situation: SprayCo, an $800M surface treatment and coating manufacturer, primarily serves the automotive industry. Recently, they unveiled a coating product that is superior to anything else on the market: it can be applied quickly, has better durability, and binds more effectively to curved surfaces. However, SprayCo's new product would be impossible to produce without patented materials they obtain from their chemical supplier. How could the chemical supplier define and quantify SprayCo's value drivers?

	Value Driver	Feature	Benefit	Value Formula	Value	
					Auto	Other
Costs	Decreased Materials Costs	Strong bonding coat	Only one coat necessary	(Percent reduction in materials used) X (Current cost of materials)	$12,500	$2,600
	Reduction in Product Failure Costs	Effectively adheres over other coatings on first attempt	Reduced likelihood that coating process will fail on initial application	(Total #Applications per year) X (Percent reduction in projects requiring reapplication) X (Total #of hours per reapplication) X (Avg $ cost of FTE per hour)	$6,300	$2,600
	Reduction in Inventory Costs	Non-Toxic	Storage takes up less space, as protective casing is not required	(Percentage reduction in standing inventory) X (Current Cost of Maintaining Inventory)	$1,000	$500
Revenue	Increased revenue due to expanded product use	Binds to sharply curved surfaces	Effectively coats headlight fixtures	(Margin of product) X (#of new units sold to coat headlight fixtures)	$4,000	$0
	Increased revenue due to better product durability	Durable, weather resistant	Longer time allowable between the application of new coats	(Percentage increase in price premium due to improved product durability) X (Current price of product) X (#of units sold)	$3,500	1,250
	Increased revenue due to rapid fulfillment times	Non-Toxic	Decreased processing time for non-toxic materials allows for same day fulfillment	(Number of new orders that require same day fulfillment) X (Premium paid for same day fulfillment)	$700	$600

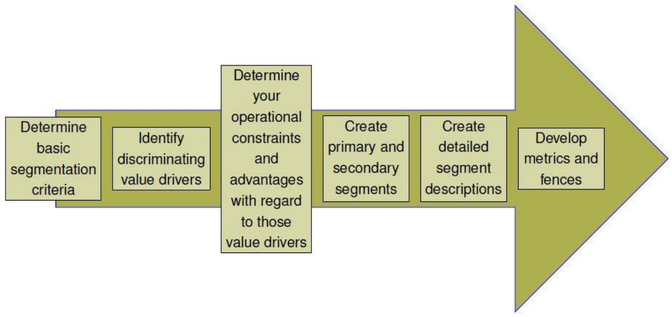
Telkens de twee values optellen bij elkaar en dan da took nog eens helemaal optellen, en dan kom je aan een total economic value van $800 035 550.

Psychological value: conjoint

Psychological value drivers such as satisfaction and security do not lend themselves to estimation via qualitative research techniques like in-depth interviewing. Instead, pricing researchers must rely on a variety of quantitative techniques to estimate the worth of a product's differentiated features. The most widely used of these techniques is conjoint analysis, a technique that can discern the hidden values that customers place on product features. The basic approach is to decompose a product into groups of features and then provide customers with a series of choices among various feature sets to understand which they prefer. Using conjoint analysis makes it possible to estimate the value of different feature sets in driving willingness-to-pay and, ultimately, the purchase decision.

Customer segmentation

Segmentation is one of the most important tasks in marketing. Identifying and describing market subgroups in a way that guides marketing and sales decision-making, makes the marketing and pricing process much more efficient and effective. To conduct a value-based segmentation, we recommend a six-step process.



1. ***Determine basic segmentation criteria***

The goal of any market segmentation is dividing a market into subgroups whose members have common criteria that differentiate their buying behaviours. Choosing appropriate segmentation criteria starts with a descriptive profile of the total market to identify obvious segments and differences among them. In consumer markets, basic demographics of age, gender and income provide obvious discriminators. Enterprise firmographics such as revenue, industry and number of employees clearly separate firms into nominally homogenous groups.

2. ***Identify discriminating value drivers***

Having preliminary segmentations in hand, you identify those value drivers - the purchase motivators - that vary the most among segments but which have more or less homogenous levels within segments. This allows you to zoom in on what's most important to each customer segment.
In-depth interviews probing how and why buyers choose among competitive suppliers provide the additional input required. The outputs of this step include a number of useful building blocks for value-based market segmentation, including a list of value drivers ranked by their ability to discriminate among customers, an explanation of why each driver adds value, and whether customers in each segment recognize that value.

3. ***Determine your operational constraints and advantages***

In this step, you examine where you have operational advantages. Which value drivers can you deliver more efficiently and at lower cost than others? Also, which drivers are constrained by your resources and operations? Use the discipline of activity-based costing to build a customer behaviour spectrum mapping your true costs serving different customers.
With this data, you can cross-reference and compare lists of customer needs served and unserved, the seller's advantages and resource limitations, and competitors' abilities.

4. ***Create primary and secondary segments***

This step combines what you've learned so far about how customer values differ and about your costs and constraints in serving different customers. Unless you're comfortable with multivariate statistical analyses accounting for several value drivers simultaneously, you'll find it most convenient to segment your marketplace in multiple stages, value driver by value driver.

The number of stages depends on the number of critical drivers that create substantial differences in value delivery among customer groups. Your primary segmentation is based on the most important criterion differentiating your customers. Your secondary segmentation divides primary segments into distinct subgroups according to your second most important criterion. Your tertiary segmentation divides second segments based on the third most important criterion, and so on.

5. ***Create detailed segment descriptions***

The segments should be described in everyday business terms so that salespeople and marketing communications planners know what kinds of customers each segment represents.

6. ***Develop metrics and fences***

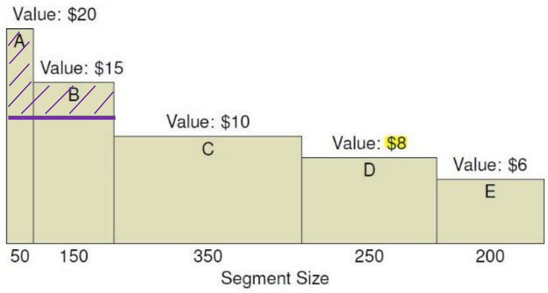
Segmentation isn't truly useful until you develop the metrics of value delivery to market segments and devise fences that encourage customers to accept price policies for their segments.
Metrics are the basis for tracking the value customers receive and how they pay for it.
Fences are those policies, rules, programs, and structures that customers must follow to qualiry for price discounts or rewards.
Choose fences and metrics that establish and enforce premium prices for high value segments, and allow feature repackaging and unbundling to appeal to low-value and low-cost-to-serve segments.

Price structure

After developing products or services that create value, a marketer must then determine how most profitably to capture that value in both volume and margin. The challenge in doing so is that customers value products differently because of different abilities to pay, different preferences, and different intended uses. Moreover, the timing of customers' needs, the speed of their payments, and the level of service and support they require can drive significant differences in the cost to serve them. A single price per unit is rarely the best way to generate revenues, because it does not take the above requirements into account.

<u>Motivation.</u> Assume the following picture. There is a supplier that is faced with five different segments, all willing to pay a different price to get the benefits they sought from a product. Now what price should the firm set? The right answer in principle is that price that maximizes profit contribution. Let's calculate the profit contribution for each segment:

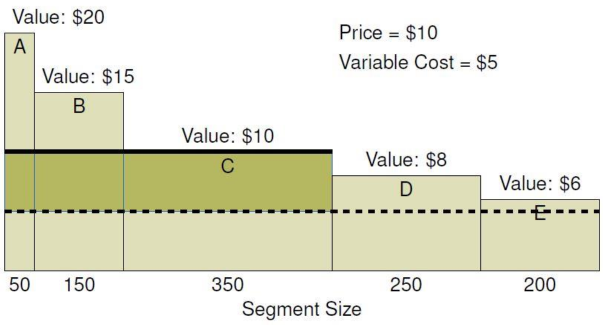
- A = 20 * 50 = 1000
- B = 15 * 150 = 3000
- C = 10 * 350 = 5500
- **D = 8 * 250 = 6400**
- E = 6 * 200 = 6000



The profit maximizing price would be $8. Now why is it impossible that a price between for example 15 and 10 will never be optimal? Because we lose customers if we do that! We lose the purple part in the picture. The optimal price will always be equal to one of the segments prices.

Now imagine we have an unit cost of $5, and we again want the profit maximizing price. Let's again calculate the profit contribution for each segment:

- A = 20 * 50 – 5 * 50 = 750
- B = 15 * 150 – 5 * 150 = 1500
- **C = 10 * 350 – 5 * 350 = 1750**
- D = 8 * 250 – 5 * 250 = 750
- E = 6 * 200 – 5 * 200 = 200



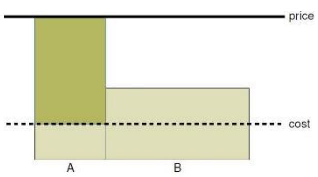
Now the optimizing price is $10. But since we only charge a single price, some things are suboptimal for us, we take the point of view of the seller, not the buyer. What is the opportunity we miss when we sell for $10? We lose the customers in segment D and E. These customers have nevertheless a valuation that is higher than our cost. We could sell profitable to these segments, but we don't do it because we can only set one price. So segment D and E are **unused market potential**. Everything that we are not making profit we lose because we don't sell to them. The danger of ignoring this potential, is that you allow the competitors to target the customers in these segments. They will grow and will gain trust in the segments you ignore! A second missed opportunity are the customers that are willing to pay more in segment A and B. The areas above the $10 line is profit that we don't take because our price is lower than the price that they are willing to pay. This is **the customer surplus**: 10*50+5*150=1250. The dark yellow part is the profit you make, $2750.

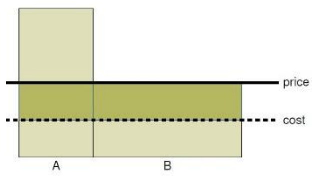
Now let’s see what happens if we sell at **2 different prices**. These prices would be $15 and $8. We again have some customer surplus, in segment A and in segment C: 350*2+50*5=950, but less then the previous example. The unused market potential is still above the cost line, but only in segment E. It has also decreased a lot compared to the previous example. The profit number is now $3800.

If you give each segment its price than that’s the most optimal. But don’t forget that not every customer cost the same. And also don’t forget: all customers will try to pay the cheapest price! So now **how can sellers charge different prices to different customers?** The answer is by creating a segmented price structure that varies not just the price, but also adjusts the offer or the criteria to qualify for it. A segmented price structure is one that causes revenues to vary with differences in the two key elements that drive potential profitability; the economic value that customers receive and the incremental cost to serve them. There are 3 mechanisms that one can use to maintain such a segmented structure: price-offer configuration, price metrics and price fences.

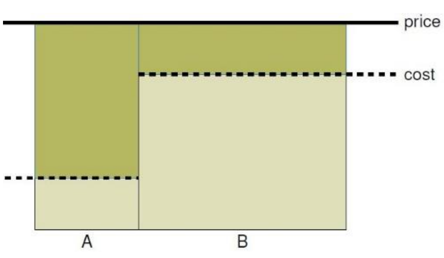
Price-offer configuration

When differences in the value of an offer across segments is caused by differences in the value associated with features, services, or both, a seller can segment the market by configuring different offers for different segments. Using offer design to implement segmented pricing requires minimal enforcement of the segments because customers self-select the offers that determine their prices. Office for example. They realised that many customers are interested in them, but not everybody wanted to pay the same. For example home owners and students don’t want to pay much for it. Software users will pay more for it. They value it much higher than then home and student users do. How can we charge a different amount of money for different segments? **Bundling**!
But why not just offer each feature separately? If you have to sell all these products apart, then you have to negotiate for every product apart. It is cheaper to do it all at once. Also reduced production costs when you bundle. If you sell them apart, you have to send them away separately. Why are people more price sensitive when we offer the products separately? When it’s in a bundle, people don’t always think about each component separately.



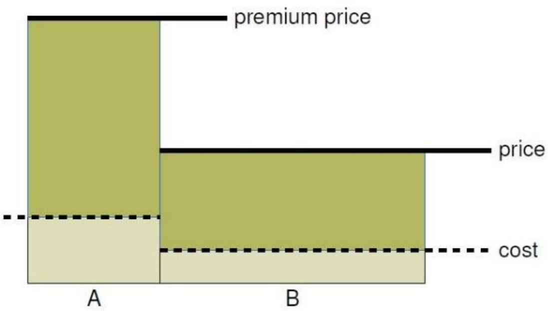


There are different ways to bundle. Imagine you have 2 segments who each have a different willingness-to-pay for your product. If you just charge the price from segment A, you will have unused market potential. But if you just charge the price from segment B, then you have customer surplus! How do you charge the price from segment A, but you make sure that segment B also buys at that price?



You can add something to your product in segment B, so that they want to pay the same price as segment A. So it adds value for segment B, and not for segment A. Example: when you buy an iPad, they offer you a couple of hours of iPad training in addition to the product. Segment A is most of us, they value such a product and we are willing to pay a bit more for it. Segment B are our parents or some older people. They feel not at ease with touchscreens etc.

They do not completely understand how this product add value, so they are initially not willing to pay as much for the iPad as segment A. So Apple adds something to that product so that the value increases. Now the people in segment B have an increased value because now they know how to use it! And for segment A it doesn't really matter if they get a training. Very often it is a combination of a product and a service. Another example is when you go on a holiday and you rend a car, then you need kids seats. There are some rental car agencies that offer these seats for free. By doing so, they increase the value for a specific segment and not for the other segments. For Ikea there is the kids corner.



Another option are **price premiums**, like Microsoft Office does. Here you create different versions of your product that are sold at different prices. For example, gym memberships. Or when you buy a car, you can get different options. Tesla had different versions of their cars with different battery capacity. When there was a hurricane they increased the capacity of the battery on the ones that had lower battery capacity. Which meant that they could have had better batteries all the time, Tesla just hadn't enabled it. When a new iPhone comes out you have 2 versions: 64GB or 32GB. Apple noticed that some people that are willing to pay a high price and some are not. So they came up with two versions on the phone where the only difference was the memory. But the price difference was way more than just the memory. So you create different versions of you product with different prices. If the customer wants more/a better product, the cost is higher and so is the price.

EXAMPLE

Concert segment	"Must see" performances	Innovative performances
Music connaisseurs	$40	$40
General public	$60	$25

Musical entertainment can provide an ideal opportunity for profitable bundling, where the "features" valued differently by different segments are the different types of performances. There are 2 types of performances: must see performances and innovative performances. The challenge is that there are two large customer segments to which these concerts appeal. Based upon past research and experimentation, assume that the concert promoters believe that the ticket prices in the next table represent roughly the acceptable price that would optimize price and attendance by each segment alone at each concert type. If the prices were set at $60 per ticket for must see performances, much of the music connaiseurs would be priced out, leaving a lot seats empty. If the innovative performances were priced at $40 per ticket, the large general segment would fail to show, and those performances would probably not be viable. Charging $40 per ticket for must see and $25 for innovations would fill the halls for both types of concerts but would leave a lot of potential revenue on the table.
Because of this reversal of preference, it is possible to price tickets more profitably as a bundle. After establishing single ticket prices of $60 for must see concerts and $40 for innovative concerts, the promoter can offer a series of headliner and innovative performances at a discount from those prices that fill the halls. Since the music connaisseurs would pay up to $80 for one headline plus one innovative performance, and the general segment would pay $85 for the pair. The venues can then be filled an generate more revenue per patron from each bundle of concerts than would be possible with single ticket pricing only. The magic behind this is that the different segments are paying the additional $15 per pair of performances for different reasons. Giving them both a reason to pay more within the same bundle facilitates the capture of that value without forgoing volume.

EXERCISE

As manager of “De Bijloke”, a concert hall in Ghent, you are responsible for arranging classical concerts. In general, performances are either more traditional or more avant garde. The public can be segmented in the conservative, market size of 700 individuals, versus the more progressive, market size of 300 individuals, concert goers. A concert typically has 3 performances and the vanue can host at maximum 1000 visitors. What is the ideal combination of 3 performances?

	Conservatives	Progressives	Revenue	
Width of bar	700	300		
TTT	9	6	700*9 = 6300	700*6+300*6 = 2700
TTA	**7**	9	**700*7+300*7 = 7000**	300*9 = 2700
TAA	5	12	700*5+300*5 = 5000	300*12 = 3600
AAA	3	15	700*3+300*3 = 3000	300*15 = 4500

The ideal combination of 3 events is TTA with a profit of 7000.

Price metrics

Not all differences in value across segments reflect differences in the features or services desired. Value received is sometimes not even related to differences in the quantity of the product consumed, necessitating a price metric unrelated to quantity or product or service provide.

Price metrics are the units to which the price is applied. They define the terms of exchange, what exactly will the buyer receive per unit of price paid. There are often a range of possible options? For example, a health club could charge per hour of use, per visit, per “annual membership” for unlimited access, or per some measure of benefit. The club might also vary those prices by time of day, or by season of the year to reflect differences in the opportunity cost of capacity. Finally, it might have a multi-part metric: an annual membership with an additional hourly charge for use of the tennis courts.

There are five criteria for determining the most profitable price metrics for an offering.

1. **It tracks with differences in value across segments**

While offer design facilitates different pricing differently based upon what people chose to buy, a price metric not based upon units of purchase can facilitate different pricing for the same offer. For example, it often makes more sense to price drugs per day of therapy rather than per milligram of the drug.

2. **It tracks with differences in cost-to-serve**

When customers’ behaviour influences the incremental cost to serve them and those costs are significant, a profit maximizing price metric needs to reflect that as well. The cost to deliver a service is significant if it exceeds the cost of measuring, monitoring, and charging for differences in its usage.

3. **It is easy to implement without any ambiguity about what charge the customer has incurred.**

Profit sharing or performance-based pricing are theoretically ideal ways to achieve the first two criteria for a good metric, tracking with value and cost. But in practice, these methods often end in rancorous debate about how profit or performance should be measured. At minimum, it is important to have absolute clarity in advance about what the metric is and who will measure it. That generally means that the metric must be objectively measured or verified.

4. **How the metric makes your pricing appear in comparison with competitors’ pricing, and the impact of that on the perceived attractiveness of your offer.**

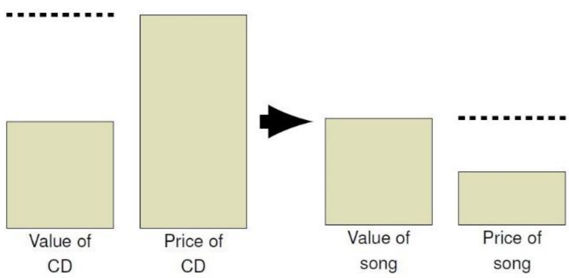
A new, hosted voice-recognition software that enabled a call center to process more callers without as much need for human intervention promised to create huge differential economic value for purchasers. Unfortunately, the traditional metric for pricing and evaluating hosted call center software was a price per minute of use. Since voice-recognition software processes callers faster, minutes using traditional call center software were not comparable to minutes using the voice-recognition software. A value-based pricing using that per-minute metric would need to be at least three times the price per minute for traditional software, inviting resistance from purchasers.

	Traditional caller response software	Natural voice recoginition software	Percent difference
Call length	7.2 min	4.4 min	-39%
Price/min	$0.90	$1.55	+72%
Price/call	$6.48	$6.82	+5%
% intervention	47%	12%	
Cost/intervention	$3.50	$3.50	
Cost/min	$1.13	$1.65	+46%
Cost/call	$8.14	$7.26	-11%

To overcome that, the company adopted a new metric; "cost per call processed". That metric naturally required conversion of the competitors' cost-per-minute metric into a cost-per-call metric. While the new software was still more expensive, its percent price premium was much smaller when framed in terms of cost per call than in terms of a cost per minute. Moreover, the differentiation value of the avoided operator intervention was much more dramatic when framed in terms of cost per call rather than cost per minute basis. The total cost per call was less with the new software, despite being higher on a per minute basis. While the favourable economics of the new software was exactly the same using either metric, the per call basis of comparison made the sales effort a lot easier. There are some special cases:

Performance based metrics. This metric would tie what the customer pays for a product or service directly to the economic value received and the incremental cost to serve. So you only let customers pay when some performance is achieved. For example when you take a 3 month prescription and you don't lose 10kg, then you get your money back.
Tie-in metrics. This is when a product is split up into two parts. One part is sold very cheap and the other part is not. For example, a printer itself is quite cheap, but the card riches are very expensive.
Two-part tariffs. When you have a fixed subscription and then you pay also per use.

<u>EXAMPLE: ITUNES</u>
Why had the new pricing model such an impact on the sales? Before iTunes, you could only buy a song as a part of a CD. Which meant that the price you paid for the CD was actually very much higher than the actual value you had for the CD, because you only bought it for that one song. But with iTunes, you could just buy one single song. Which was cheaper and also more valuable. But then came Spotify along with a subscription with which you could listen to as much songs as you wanted.



<u>EXERCISE: JOB POSTING SITE</u>
You are manager of a popular online job posting site. Potential employees can post their resumes and job aspirations anonymously, and employers can do the same for job listings.

- Define 2 employee and 2 employer segmentation criteria.

Employee: high profile, experienced work-seekers and recently graduated, looking for first job (always ask yourself; can this be measured easily? I can act as a less experienced employee. But in this case it is a good distinction because you get better job offers when you have a high profile). Or: white collar or blue collar workers, do you need higher education? Or: people looking for full time or part time job. Employer: searching for long term workers and searching for short-term/interim/temporary projects. Or: how urgent it is, but this is difficult to measure. Or: size of the company (bigger companies must pay more). Or: division in different industries.

- What is the value or cost driver that led you believe that the segment would be profitable?
- What design options or metrics could you employ?

Pay per job opening that you have as a company. Performance based metric: only pay when there is a match but then you have to do an offer to see if there really was a match. Subscription/membership.

Price fences

Sometimes value differs between customer segments even when all the features and measurable benefits are the same. Value can differ between customer segments and uses simply because they involve different "formulas" for converting features and benefits into economic values. Unless there is a good "proxy" metric that just happens to correlate with the resulting differences in value, the seller needs to find a price fence: a means to charge different customers different price levels for the same products and services using the same metrics.
Price fences are fixed criteria that customers must meet to qualify for a lower price. Sometimes they are based on age, on educational status or possession of a coupon from a local paper. All three types of customers have the same needs and cost to serve them, but perceive a different value from the purchase. There are different types.

<u>Buyer identification fences</u>. Occasionally pricing goods and services at different levels across segments is easy because customers have obvious characteristics that sellers can use to identify them. So you charge different prices to different buyers based on observable characteristics that signal buyers' price sensitivity. Buyers in different segments must have different characteristics that either are obvious, or that buyers can be induced to reveal. Induced to reveal: some hotels give discounts to those who will show an American associations of retired persons card. In the cinema you can only get a student discount if you show your student cart.
Often a buyer's relative price sensitivity does not depend on anything immediately observable or on factors a customer freely reveals. It depends instead on how well informed about alternatives a customer is and on the personal values the customer places on the differentiating attributes of the seller's offer. In such cases, the classification of buyers by segment usually requires an expert salesperson, trained in soliciting and evaluating the information necessary for segmented pricing.

<u>Purchase location fences</u>. When customers who perceive different values buy at different locations, they can be segmented by purchase location. Countries can be very different for prices. You pay more in city centers than in suburbs. A special case for bulky industrial products such as steel and coal is **freight absorption**. This is the agreement by the seller to bear part of the shipping costs of the product, the amount of which depends upon the buyer's location. The purpose is to segment buyers according to the attractiveness of their alternatives. Seller: "we don't want you to buy at our competitors, even though transportation costs to you are very high. But to make sure you will still buy from us, we will bear part of the shipping costs. The part of these shipping costs that are made by delivering further than the original segment".

<u>Time of purchase fences</u>. When customers in different market segments purchase at different times, one can segment them for pricing by time of purchase. Restaurants usually charge more in the evening than during lunch, because demand is more price sensitive for the midday meal. Why? There are more numerous inexpensive substitutes for lunches than there are for dinners. Example; Euro-tunnel: with your own car to England. When you only stay for a short period of time, having you car there is less valuable than when you stay a long time. Time is also a useful fence when demand varies significantly with the time of purchase but the product or service is not storable. Special cases: **peak-load pricing** is to reduce the peak-loads in your service. For example, everybody goes to the restaurant in the evening, so it gets way to busy. Make it cheaper to dine at lunch. Higher price when high demand, lower price for low demand. **Priority pricing/price skimming** is when new products in a retail store are offered at full price, or sometimes premium surcharges over full price in the case of extreme excess demand. Over time, as product appeal fades in comparison to newer competitive alternatives, buyers discount the product's value until they are willing to pay only a fraction of its original price for leftover models.

Purchase quantity fences. When customers in different segments buy different quantities, one can sometimes segment them for pricing with quantity discounts. There are four types of quantity discount tactics:

- **Volume discounts**. Are most common when selling products to business customers. They are based on the customer's total purchases over a month or year rather than on the amount of purchased at any one time. To attract large customers.
- **Order discounts**. Often sellers vary prices by the size of an order rather than by the size of a customer's total purchase volume. The logic for this is that many of the costs of processing an order are unrelated to the size of it. Consequently, the per unit cost of processing and shipping declines with the quantity ordered. For this reason, sellers generally prefer that buyers place large, infrequent orders, rather than small frequent ones. To encourage them to be do so, sellers give discounts based on the order quantity. Almost all office supplies are sold with order discounts.
- **Step discounts**. They differ from volume or order discounts in that they do not apply to the total quantity purchased, but only to the purchase beyond a specified amount. The rationale is to encourage individual buyers to purchase more of a product without having to cut the price on smaller quantities for which they would pay a higher price. So if you get to an specified amount, you get a discount. People who heat their whole house with electricity for example, they get a discount because they get over a specific amount, they need electricity at a low price.

Minicase: Genvet pharmaceuticals

Genvet has a new product: Tymacin. It's a growth hormone for cattle and hogs.

1. **What is the economic value of Tymacin in the cattle market?**

The advantages of Tymacin:

- Reduce time kept in feedlot before slaughter
- Increase feed efficiency, animals gain weight faster

The costs of Tymacin:

- There is a fixed cost of $32MM for developing and testing
- There is a variable cost of $7/lbs (so per pound) for producing and marketing

There are 2 potential markets: hog farmers and cattle ranchers. They both have a different total economical value.

The total economic value for the hog farmers is easy: $22/lbs = reference value (11) + positive differentiation value (11). You only need 3lbs/ton feed instead of 6lbs/ton feed, so it costs double. For 6lbs/ton feed of Tymacin, you would pay $22. So the positive differentiation value is $11.

1 Hog farmers

	Tymacin	Competitor
Quantity needed	3lbs/ton feed	6lbs/ton feed
Price	$22/lbs	$11/lbs

2 Cattle ranchers

	Tymacin	Competitor
Savings/ton feed	$76	$64
Quantity	2lbs/ton feed	2lbs/ton feed
Savings/lbs product	$38	$32
Price	??	$24

The reference value for the cattle ranchers is $24/lbs. The differentiation value is $76 - $64 = $12/ton, and you need 2 pounds per ton feed, so $6/lbs Tymacin. Which makes the total economic value for the cattle ranchers 24 + 6 = 30/lbs.

2. **How would you segment this market and price this product to maximize profitability?**

The estimated sales volumes for Tymacin:

- Hog market: 3,1 million lbs
- Cattle market: 1,8 million lbs

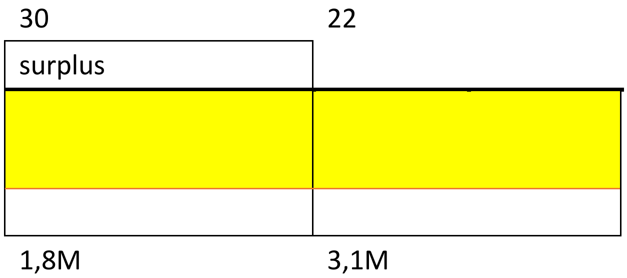
Market characteristics:

- Both markets buy with same distributors
- FDA approves by chemical composition

Now let's do profit calculations for 3 situations. First, what if we set the price from the hog market: $22. Also remember, we have a cost of $7.

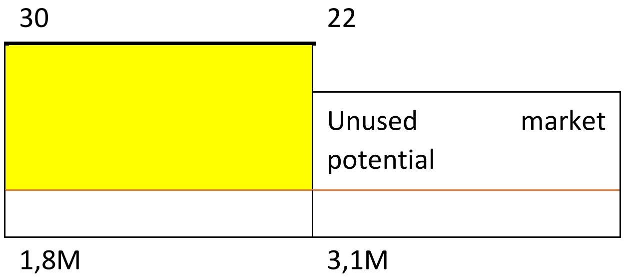
So we set the single price to $22.
- 3,1MM + 1,8MM = 4,9MM
- 4,9MM*(22 – 7) = $73,5MM/year

The yellow part is the profit part. The orange line is the cost at $7. As you can see we have a little bit customer surplus in the cattle segment.



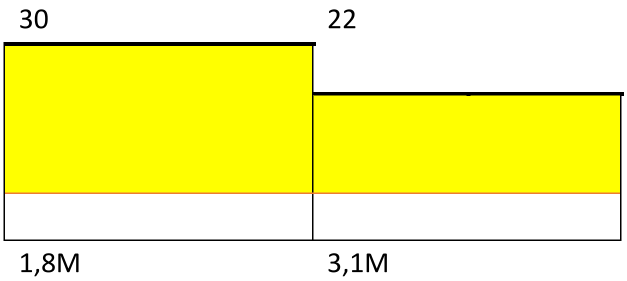
Next let’s set the single price to $30.
- 1,8MM * (30 – 7)
- = $41,4MM/year

The yellow part is again the profit part. The orange line is the cost at $7. We don’t have consumer surplus anymore, but we have unused market potential, the whole hog market is not served now.



Lastly, let’s set both prices for both segments.
- 3,1MM * (22 – 7) = $46,5MM/year
- 1,8MM * (30 – 7) = $41,4MM/year
- Total: $87,9MM/year

The yellow part is the profit part. As you can see, we have no longer customer surplus nor unused market potential, and the profit is the largest. But how do you keep the cattle segment from buying it for the price of the hog segment?



There are 2 possible solutions:
- **Value added bundling**: add a hog specific medicine (separate bundle). Or offer hog targeted service (same bundle).
- **Selective uglification**: maybe we can add something that is repulsive for the cows, but the hogs don’t seem to notice it. So when the cattle herders buy the cheaper version, it is not eaten by the cows!

Minicase Apple

PRICE METRIC: they pay the same amount eventually, but in a 21 month. Apple has his own pricing strategy: Apple’s iPhone Upgrade Purchase Plan combines the retail prices of both a selected iPhone 6S and its AppleCare+ enhanced warranty – then divides this total amount into 24 equal monthly payments, with no interest added on. After the 24 payment commitment is fulfilled, the purchaser owns the phone. However, after 12 months of payments, the buyer has the option to upgrade to a new phone. Doing so resets the agreement clock, meaning the buyer is on the hook for 24 more payments. Plus, the phone is unlocked, meaning purchasers can select whichever cellular carrier is the cheapest and switch at any time.
Unlock the phone: so you can switch providers all the time.
why are frequent upgrades so important? They have a history of increasing sales when they release it. Everybody expects that they will do better and better. But if they don’t, it will have an impact on the stock.
Pricing strategy: the actual price and how you pay → price metric.

Introduction to R

See ppt's.

Measurement of price sensitivity

Price sensitivity: the range at which your demand goes down as your price goes up. The most basic version of a pricing experiment is when you have a range of price points for your product, and then you figure out how much you would sell given it was sold at a certain price point. This is not always so convenient. Think for example about houses.

Procedures for estimating price sensitivity differ on two major dimensions; the conditions of measurement and the variable being measured. The conditions of measurement range from a completely uncontrolled to a highly controlled research environment. When making **uncontrolled** measurements, researchers are only observers. They measure what people actually do, or say they would do, in a situation not of the researcher's making. In contrast, when making **controlled** measurements, researchers manipulate the important variables that influence consumer behaviour to more precisely observe their effect. Generally, controlled research produces more accurate estimates of the effects of the controlled variables on price sensitivity, but depending on the level of realism, it is often costly to implement in a "real-world" setting. A laboratory setting is often used to better control other factors that may affect price sensitivity as well as to reduce costs, but these improvements come at the expense of realism.

	Uncontrolled	Experimentally controlled
Preferences and intentions	• Direct questioning • Buy-response survey • Depth interview	• Simulate purchase experiments • Conjoint
Actual purchases	• Historical sales data • Panel data • Scanner data	• In-store experiments • Laboratory purchase experiments

The dependent variable for estimating price sensitivity is either actual purchases or purchase preferences and intentions. **Actual-purchase** studies measure behaviour, whereas **preference-intention** studies measure the intended choices that people claim they would make in a hypothetical purchase situation. Since the ultimate goal of the research is to estimate how people respond to price changes in actual-purchase situations, research that measures actual behaviour is generally more desirable, but it is also more costly, time-consuming and sometimes impractical, given the need to move products to market quickly.

Uncontrolled – intentions

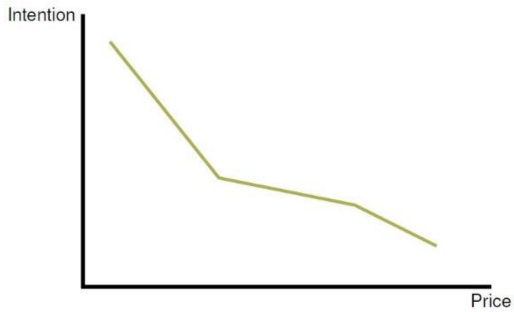
The most common research technique for directly estimating price sensitivity is the survey of brand preferences of purchase intentions. Companies prefer to measure preferences or intentions, rather than actual purchases, for a number of reasons:

- Survey data costs much less to collect than purchase data
- Survey data can be measured for large durable goods, such as automobiles or photocopiers, for which in-store or laboratory experiments at various prices are impractical
- Survey data can be collected even before a product is designed, when the information is most valuable in directing product development. So you want to check on the price sensitivity of a product that is not on the market yet so there is no actual purchase data
- The results can be collected quickly

The problem with survey research is that many consumers do not provide answers that are a reliable guide to their actual purchase behaviour. The reasons are varied, but one of the main issues is that surveys require a level of abstraction that the respondent my or may not be able to perform. This is especially true for new products that are wholly unfamiliar or whose application is not readily apparent.

Direct questioning. Just ask customers outright, “What is the most you would be willing to pay for this product?”. But this calls for bargaining behaviour, with consumers stating a lower price than they would actually pay. Or sometimes the respondent feels the desire to please the researcher, so they state a higher price that they would actually pay. They are reluctant to say that they don’t like the product. It is easy and cheap, but the results of such studies are at best useless and are potentially highly misleading.

Buy-response surveys. Showing consumers a product at a preselected price and asking if they would purchase it at that price. When the answers given by different consumers for different price levels are aggregated, they produce what looks like a demand curve for market share, sometimes called a purchase probability curve. But, monadic: you only evaluate the price of the product of interest, not at the competitors. The alternatives can have a big impact on the buying decision. And the respondent can still try to please the researcher.

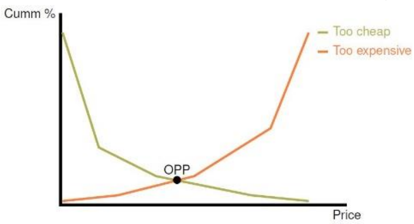


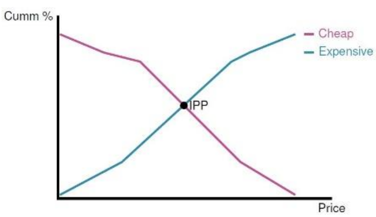
EXERCISE IN R.

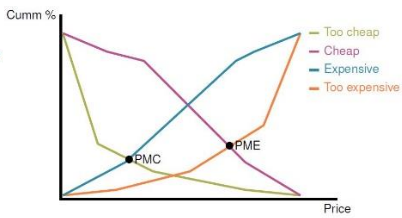
The price sensitivity meter of Van Westendorp. It is a market technique for determining consumer price preferences. This approach asks four price related questions, which are then evaluated as a series of four cumulative distributions, one distribution for each question. These are the questions:

1. At what price would you consider the product/service to be priced so low that you feel that the quality can’t be very good? (too cheap)
2. At what price would you consider this product/service to be a bargain – a great buy for little money? (cheap)
3. At what price would you say this product/service is starting to get expensive – it’s not out of the question, but you’d have to give some thought to buying it? (expensive)
4. At what price would you consider the product/service to be so expensive that you would not consider buying it?

The cumulative frequencies are then plotted, and this approach claims that you can interpret the intersects of the cumulative frequencies for each of the four price categories.







The OPP: the optimal price point. This is the point at which an equal number of respondents describe the price as exceeding either their upper or lower limits.
The IPP: the indifference price point. The IPP refers to the price at which an equal number of respondents rate the price point as either “cheap” or “expensive”.
The PMC: the point of marginal cheapness. This is the point where “too cheap” and “expensive” cross. It can be the lower bound of an acceptable price range.

The PME: the point of marginal expensiveness. This is the point where “too expensive” and “cheap” cross. This can be viewed as the upper bound of an acceptable price range.
This approach is quick, easy and cheap. It is the preliminary analysis to find the price-fork. It allows us to find changes in the price-response relationship, the curve gives us information about how demand would change with price. But you are only asking if the respondents find it cheap/expensive and not ‘would you buy it at that price?’. So you don’t have information about buying intentions. It’s not because they think it is cheap that they would actually buy it.
EXERCISE IN R.

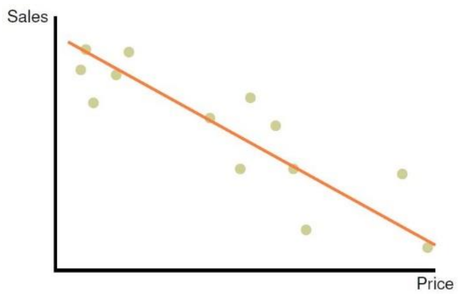
In-depth interviews. This is a “semi-structured” method that is used to get responses from customers on how they use products and services, from which the research infers value rather than asking bout value directly. The interview is often conducted one-on-one with a respondent and lasts for one to two hours. In a consumer environment, it is used to understand how individuals and families use products and how they might value different features or positioning approaches. In a business-to-business environment, the interviewers attempt to understand how businesses gain revenues or reduce costs by using a specific product or service. Example: Gabor-Granger. This is a face-to-face interview where you constantly raise or lower the price asked to find the optimal price. To find the value that is attracting the customer to your product.
In-depth interviews in pricing research are useful in:

- Understanding which product or service features and benefits are important to a customer
- Assessing the monetary or psychological value of these features and benefits that a customer receives as a result of using the product or service
- And assessing roughly what a customer might be willing to pay to obtain these features and benefits.

In-depth interviews do not ask customers directly how much they would be willing to pay. Instead, the interview focuses on the financial benefits to the customer that a product or service could influence. It is also possible to get a sense of perceived value by identifying other items that the customer buys to achieve the same benefit. For example, **evoked anchoring** asks respondents to identify items in their budget that they might consider a trade-off in order to obtain the value and benefits promised by a supplier’s proposed product or service solution. For example, when helping a software client to price relationship management software, we identified that one benefit was reduced customer turnover. By asking potential buyers of the software, to identify the costs to acquire new customers, we could infer the value to retain them.
The interview must be conducted outside the context of a selling opportunity, since customers are unlikely to reveal value at such times. However, the data gathered often form the basis of a value-based selling approach in which sales people, armed with an understanding of how their products differ from those of competitors and how those differences create value for customers, can justify their pricing to the customer and to themselves.

Uncontrolled – actual purchases

The company keeps track of all sales, but you don’t have information about what the competition is doing, that’s why it’s uncontrolled data. There are many factors that might have an impact that are outside your own company database. That is a problem because you just wanted to filter out these biases! Typically, we use **regression techniques** for that. We try to fit a straight line through the



data: $y = a + b * x$. X is price and y is amount. A and b relate to beta parameters, these relate to the slope and the intercept op the regression.

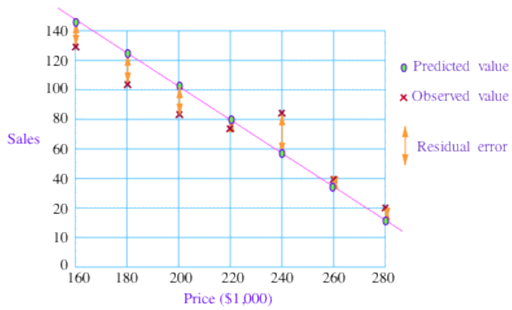
Formula: $d(q) = \beta_0 + \beta_1 * p$. Beta 0 is the intercept, beta 1 is the slope, the steepness of the line (the difference in Y when the price goes from 0 to 1). If we have a price of p_1 we aspect that we see a whole bunch of price values, but the function gives the mean, the expected value of y given specific value of price. So random variation around the line, $d(q) = \beta_0 + \beta_1 * p + \varepsilon$ (random error). We make some assumptions about this error. The expected value is equal to 0 and follows a specific distribution (normally distributed with mean 0 and some standard deviation that we will try to estimate from the data).
Whenever we have this 2D dataset and we want to fit a regression, we can draw multiple lines. How do we decide which one we take? When we are dealing with lots of data, it turns out normal. We use the normal distribution because we already made the assumption that the error is normally distributed. Formula:

$$\frac{1}{\sigma\sqrt{2\pi}}^{\left\{\frac{-1}{2\sigma}(x-\mu)^2\right\}}$$

Errors follow the normal distribution, this means that we can plug in sigma and epsilon in the formula:

$$\frac{1}{\sigma\sqrt{2\pi}}^{\left\{\frac{-1}{2\sigma}(\varepsilon-0)^2\right\}}$$

(epsilon-0)² = epsilon² so $d(q) = \beta_0 + \beta_1 * p + \varepsilon \rightarrow \varepsilon^2 = (d(q) - (\beta_0 + \beta_1 * p))^2$



Epsilon is the residual error in the picture, the difference between d(q) and the fitted or predicted demand $\beta_0 + \beta_1 * p$. The difference between the real data and the predicted data based on our model. If the difference is big that means that the observation is far from the regression line, which is not good! We are taking the square of the epsilons because we don't want errors to cancel out: if an observation is smaller than the prediction, it is negative. If it is bigger, it is positive: -1 +1 = 0!
Now we will calculate the value of this entire expression, this is the likelihood of that point. if it has a high value that means that the point is likely under the model. If it has a low value, than it means that it is unlikely. We take the likelihood for each point in the dataset and then we multiply them:

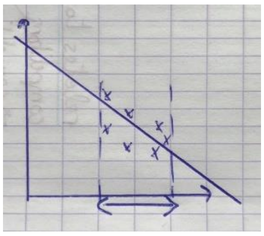
$$\prod_{i=1}^{N} \frac{1}{\sigma\sqrt{2\pi}}^{\left\{\frac{-1}{2\sigma}(\varepsilon-0)^2\right\}} \rightarrow likelihood \rightarrow we\ want\ to\ maximize\ this$$

Another approach is OLS: trying to minimize the sum of squared differences. It is identical to the previous approach: in previous we multiplied with -1 and therefore we maximized there and here we minimize. $\min \sum_{i=1}^{N}(d(q) - (\beta_0 + \beta_1 * p))^2$.

In reality this has a lot of problems. For example, football Ghent tickets.

- If there is almost no variation in price, then all the data points will just be plotted on each other. The demand only got bigger when they played against a top club. So there is a positive relationship between price and demand, but the error here is that they did not take into account that the value of the product was not constant. When we do price analysis, we need to make sure that the value of the product remains the same.

- We want to have a lot variation in the price. Most companies have only minor changes in the price. This is crucial in price analysis, this is what gives stability in our regression. If we only have a few points, it is much more difficult to make the regression line, it can almost be turned around. The further away the points are from each other, the better that we can create the regression line. The analysis of the error will indicate in a much clearer way which regression line is fitting the data best with points that are further away from each other.
- When independent variables vary/change together, you have collinearity. Football example: the price is perfectly correlated with the value of the game, the popularity of the visiting team. You don't want this. Whenever there is a variable correlated to the price, you want to capture that variable and bring it to the analysis. Another variable that is typically related to price is advertisement: whenever there is a reduction in price, that will go together with an increase in advertisements.



Imagine we have only data here (blue points). When we make a regression line, we can only make predictions for prices that are more or less in that price range. The further you go away, the more unrealistic it is that the fit is still linear, that the predictions will still be good.

One way to estimate price sensitivity is to analyse past sales data. We can do this with linear regression (described above). And it is a very powerful method to assess price sensitivity but we make the assumption that everything stays the same.... Still there are changes in:

- The number of brands on the market,
- How recently competitors offered price promotions,
- The amount and effectiveness of advertising by each brand,
- Increased price sensitivity of more-educated consumers,
- And general economic conditions

That can undermine the ability of historical data analysis to diagnose the true effects of a price change.

Historical sales data. Sales data collected as part of a company's regular operation are cheap and available for all products that have prior sales histories. But you need to be careful in recognizing that sales data only allow for the estimation of price elasticity of the next level in the channel. For example, in a retail environment, unless a manufacturer sells directly to the end-user, its sales data reflect shipments to retailers, not actual retail sales during the period. Retailers my stockpile products purchased at promotional prices with no intention of passing the savings on to the consumer, or in anticipation of increases in demand on the part of the consumer in a later period.
In the past, using historical data for any product not sold directly to the end consumer was problematic. Sales data was usually available only at an aggregated level for a long period of time, say a week. In any given week, some stores will charge higher prices than others. Over time, the same store will put the product on sale for a week and then return its price to the regular level. These price variations influenced sales but were masked by the aggregation!
Given the aggregation in the data, the researcher is forced to explain sales variations by looking at only the average retail price across stores and throughout the time period. Average prices have less variation and fewer observations than actual prices at individual stores in particular weeks, so that data have less statistical power than data on individual purchase prices. The data will be much more smoothed out than in real life.

Panel data. A number of marketing research companies collect individual purchase data from panels of a few thousand households. Each household keeps a daily record of all brands purchased and price paid or uses a special credit card that tracks purchases. Since products are purchased daily, the data for each household must be aggregated to produce a series on weekly or bi-weekly purchases. Such data have a number of advantages:

- More data and quicker compared to bimonthly/quarterly sales data
- Observe the real price paid versus the average of the retail prices over the different stores, more price variation, making the effects of price changes easier to detect
- Get data on sales and prices of competing products also
- You can correlate price sensitivity with various demographic classifications of consumers and possibly identify opportunities for segmentation

Disadvantages:

- The panel data may not be representative. Only 5% of the households contacted accept the offer, so it is a biased sample of the population.
- Panel members must record their purchases, this tends to make them more price aware, and thus more price sensitive
- Typically only one member of the household agrees to participate in the panel, yet in most households multiple people perform shopping duties. So you miss data from the ones that do not keep a record of their purchases.
- Expensive

Scanner data. This is very similar to historical sales data, but now based on individual sales. Retailers generate such data as part of their normal operations. And although scanner data lacks the balanced and complete demographics of consumer panel data, loyalty programs have made it possible to infer demographics and to track purchases over time. Scanner data also costs a lot less than panel data, but the link with the customer is not always possible. Also no competitor information.
Bid responses: when a company asks for 'offertes' of different suppliers. Those suppliers then place a bid to deliver the product. And the data that comes with those bids, is kept for the future. With that data you can calculate the probability that you would win the next bid.

Controlled – actual purchases

A researcher might attempt to estimate price sensitivity by generating experimental purchase data. Such data may come from pricing experiments conducted in a store without the buyers' knowledge or from pricing experiments conducted in a laboratory. Since the researcher controls the experiment, price variations can be created as desired to generate results while holding constant other marketing variables.

In-store experiments. This relies on actual purchase data collected when buyers are unaware that they are participating in an experiment. Customers don't like to realise that they are part of an experiment. Although it is called "in-store", it can also be done online, or through mail-catalogs for example.
The simples design for an in-store pricing experiment involves monitoring sales at the historical price to obtain a base level of sales and then initiating a price change to see how sales change from that base level. Problem: you need the ability to control for external factors. Fortunately, the addition of an experimental control store can reduce this problem! The researcher finds a second store in which sales tend to vary over the base period in the same way that they vary in the first store, indicating that factors other than price influence both stores' sales in the same way. The researcher then changes price only in the first store, but continues to monitor sales in both stores. Any change in sales in the control store indicates the researcher that some factor other than price is also causing a change in sales!

One of the greatest benefits of in-store experimentation is the ability to test for interactions between price and other marketing variables that, in historical data, tend to change together, but the cost of such experiment is very high. The greatest disadvantage is the high cost of monitoring sales, analysing the data, and securing the cooperation of the retailers. There is also the potential loss of consumer goodwill when some buyers are charged higher prices than others. Goodwill is the relationship between the customer and the company, builded through the products (see example bol.com). And lastly there is also the very real risk of being discovered by a competitor. They often take steps to influence (bias) the experiment.

<u>Laboratory purchase experiments</u>. Laboratory purchase experiments attempt to duplicate the realism of in-store experimentation without the high cost or the possible exposure of competitors. A typical laboratory experiment takes place in a research facility at a shopping mall. Interviewers intercept potential participants who are walking by and screen them to select only those who are users of the product category being researched. This is also a disadvantage: external validity, proportionate sampling. You want to be able to doortrekken the results to the bigger population. The laboratory researcher can control who participates and can quickly manipulate prices and other elements in the purchase environment, all at a single location. Disadvantages, it is very artificial, participants know that they're being watched and will pay extra attention to prices and to what they buy.

Controlled – intentions

To solve some of the problems of bias and extraneous factors when measuring preferences and intentions, researchers try to exercise some control over the purchase situation presented to respondents. The questions must be designed to make the survey respondents consider the questions in the same way they would consider an actual purchase decision.

<u>Simulate purchase experiment</u>. Example: imagine you are in your favorite grocery store and you want to buy some yogurt. You notice that there are 3 brands that come at different prices: €2,24, €1,99 and €2,49. Which offer to you prefer? The goal is to present an acceptable alternative for in-store or laboratory experiments. It is very quick and cheap, but it has still the same disadvantages as in store experiments.

<u>Conjoint analysis.</u> The particular strenght of trade-off analysis is its ability to disaggregate a product's price into the values consumer attach to each attribute. You split your product up into different attributes with different levels. And then you take a particular selection of each of the levels, and you ask the respondent to what extend they think the combination is attractive. Of course, you can't ask all the possible combinations (that would be a full profile method)

Location	Cabin	Price
Italy	Basic inside	1000/person
Scotland	Basic outside	1250/person
Norway	Luxury cabin	1500/person

Imagine we ask all the possible combinations: 3 x 3 x 3 trade-offs = 27 trade-offs! This is the full profile method. Instead decide at what the most important attributes and levels are and evaluate only 2 attributes at a time: (3x3) + (3x3) + (3x3) → again 27 trade-offs.

Fractional factorial design:

- Reduce the number of evaluations
- Maximize amount of information in non complete design

- Every level of a factor appears the same number of times at every level of each of the other factors

Adaptive conjoint:

- Respondents don't have to make trade offs for every (level of an) attribute
- Maximum number of attributes can go up (up to 30)
- The combinations that have to be evaluated are not fixed at the start of the experiment

Methods: multiple regression.

Experimental design

What we want to know is how much our customers like every configuration. We have a special interest in the price variable. We'll focus on the impact of price on how customers evaluate our product. We'll put every variable in a linear regression model and that allows us to access what the value is for the combination.

Example: A chain of movie theatres is planning to offer a service to buy your ticket online. The company wants to know what a feasible price for an online ticket would be. (what is the monetary value of the reduction in price). Together with management, you concluded that the following three attributes are relevant:

Payment	**Time**	**Price**
Online	Weekday (Mon – Thu)	€ 8
Offline	Weekend (Fri – Sun)	€ 10
		€ 12

<u>Conjoint analysis</u> is a multivariate technique developed specifically to understand how respondents develop preferences for any type of object (products, services, or ideas). It is based on the simple premise that consumers evaluate the value of an object by combining the separate amounts of value provided by each attribute. Moreover, consumers can best provide their estimates of preference by judging objects formed by combinations of attributes. **Utility**, a subjective judgement of preference unique to each individual, is the most fundamental concept in conjoint analysis and the conceptual basis for measuring value.

Experimental design

An experimental design is a plan for running an experiment, and it is often displayed as a matrix.
<u>Traditional conjoint</u>:

- Columns: variables or factors with 2 or more treatment levels
- Rows: treatment combinations (also runs)

This is when a small set of attributes (typically 4 to 5) are used to create profiles that are shown to respondents, often on individual cards. Respondents then rank or rate these profiles. Using relatively simple dummy variable regression analysis the implicit utilities for the levels can be calculated.
So there are different combinations that you show to respondents, and then you ask them 'how much do you like each combination from 0 to 10'. If you have this info of many different customers than we have a dependent variable (the score they give) but we also have independent variables: variables that you use to predict the pricing (the product configuration). We have to code them into quantitative variables: dummy coding.
The total number of product configurations would make **a full factorial design**. This would be ideal, but we'll never have that in practice! This design shows all possible combinations of the levels of factors to the respondents, let's you estimate main effects an interactions. Also all main effects and (higher-order) interactions are uncorrelated.

<u>Choice-based conjoint:</u>

- Columns: idem as traditional conjoint
- Rows: product alternatives
- Blocks: choice sets

We realized that the questions that we are asking our respondents are very boring. So we have to restrict the number of questions that we ask our respondents. We can only ask 10 to 12 questions. In the movie theatre case: 12 different product configurations (2 x 2 x 3) for extremely easy example. If we add one attribute with 2 levels we go from 12 to 24 (2 x 2 x 3 x 2). The number of possible combinations grows exponentially. We have to choose 12 questions that we can ask out of the total number of product configurations there are → **fractional-factorial**

For example:

Labeled	
Acme	$1.99
Acme	$2.99
Ajax	$1.99
Ajax	$2.99
Comet	$1.99
Comet	$2.99

Unlabeled	
1	1
1	2
2	1
2	2
3	1
3	2

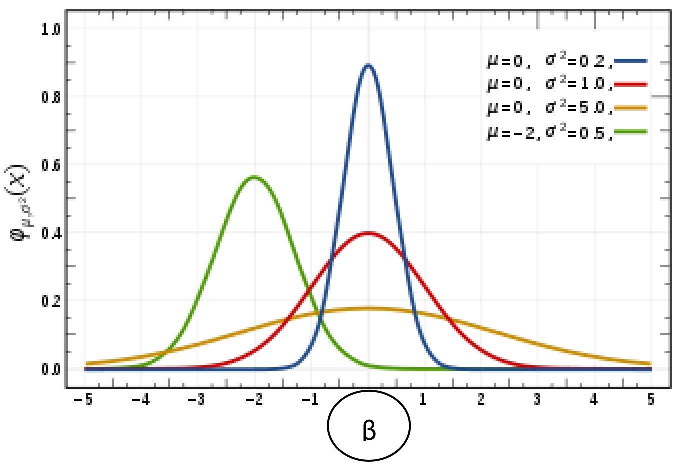
Choice-based		
1	Acme	$2.99
	Ajax	$1.99
	Comet	$1.99
2	Acme	$2.99
	Ajax	$2.99
	Comet	$2.99

1	1	1
2	1	2
1	2	2
2	2	1

So it is no longer a full factorial design, but just a factorial design. This design should be orthogonal (every pair of levels occurs equally often across all pairs of factors) and balanced (every level occurs equally often within each factor).

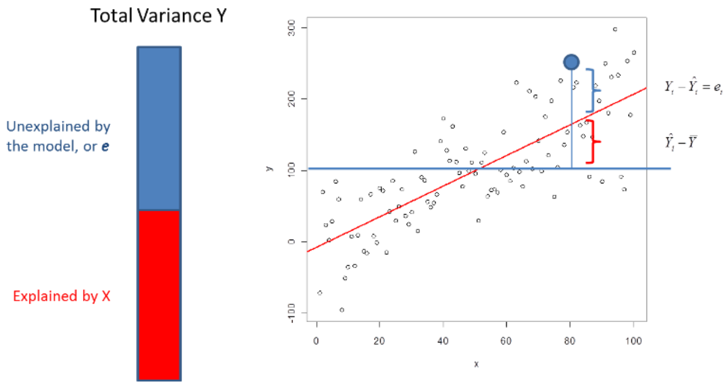
What does it mean to maximize the information? This is more tangible in statistical terms. What does it mean to have a lot of information in a regression analysis? When an analysis is more informative than another, that means that we have low standard errors.

The whole idea of statistics: our data is random, the uncertainty that we have is the result of the randomness of the data. If we take a set of customers we can calculate the beta coefficient. But if we take another set of customers and do the same, we get another beta. We choose the second set of customers based on random sampling because we try to avoid to take a set of customers that are completely different in the effects that we are estimating. But still we find different betas. But if you do this many times, you'll see a pattern in the estimates we make: most of the betas will be close to the true value, few will lay further away from it. This is called a normal distribution. Sampling distribution tells us to what extent beta would vary if we collect many sets. The more variation, the less certain we can be about the true beta coefficient. We would like to minimize tis variation, the variance of this distribution should be as small as possible.



A typical trick to reduce the variance of the distribution is to increase the number of observations because the variance is linked to this number: the number of observations occurs in the denominator of the variance formula. The square root of the variance is the standard deviation and is a measure of the amount of variation or dispersion of a set of values. The higher the variance, the more the individual values will differ from each other, and therefore also the more the values deviate from the "average".

In linear models, the covariance matrix of $\hat{\beta}$ equals $\sigma^2(X'X)^{-1}$. But this σ^2 is different from the sigma we just discussed. This is the variance of how the data points vary around our regression line, the irreducible error. (blue)



You would have slightly different regression lines if you take other sets of respondents, but if you do this a lot of times, we get close to the real value. This means that values closer to the true value are more likely to come up than values far from the true value. The spread we get there is a bit annoying? We prefer to be more certain about our estimates. Our sampling distribution should be as peaked as possible (see first graph, blue line), then we can focus on a smaller range of values for the true range.

Now instead of 1 beta, we use a vector of beta. We want to construct the covariance matrix of beta. The SE^2 of the first graph equals the $\sigma^2(X'X)^{-1}$. The width of the sampling distribution is the variance around the regression line times the X matrix. We want to minimize the width of the normal distribution, not by minimizing σ^2 because we can't influence it (irreducible error), but by choosing our X right. X is something we choose ourselves, it is the different product configurations that we show to our respondents.

*IN SHORT: Regression analysis: Y and X. regression line = summary of the data. Line is characterized by 2 parameters: intercept and slope. We have to extimate beta 0 and beta 1 from the points. If we repeat the data for different observations → slightly different data points. If you would do this many times, and you take the average of all the betas, you would be spot on to the true beta's! there is structure in how they are different! We will more often come up with an estimate that is close to the true value, than with estimates that is farther away from the true value. It allows us to say something about how likely some hypotheses are true → inference. The spread that you get in the normal distribution is kindof annoying. We do not observe the true beta coefficient, only estimates. You want to onderzoek how certain you are about an estimate, laying close to the true beta. We want to be more certain then less certain ofcourse. We want our sampling distribution as peaked as possible → variance as small as possible! We want to minimize the variance. What can we do for that. How do we calculate the variance of the sampling distributions? Beta hoedje = Sigma²(unchangeable) * (matrix multiplication X transpose X)inverse of that. →* ***COVARIANCE MATRIX***

Covariance is the same as correlation but covariance is not normalized + the correlation stays the same if the measurement changes, scale is independent. This is not the case with covariance. Covariance is an unnormalized correlation; how variables change together. In a covariance matrix you have columns and rows. For example a matrix with 3 variables. On the diagonal we have the variances of the variables. The other values are the covariances. The covariance matrix is a symmetric matrix.

	1	**2**	**3**
1	[O	X	Y]
2	[X	O	Z]
3	[Y	Z	O]

Now how to we calculate the covariance matrix? X'X divided by n for mean centred X'es. So the variance formula is:

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

We see the 1/n, but then we also see the next part: mean centring. Standardizing the X matrix, so that the mean of the variable is 0 → average and then every observation subtracted by the mean. So X'X should be large, because we want to have a lot of variation in our data. If we don't have that, we don't have a lot of information about our X! We want to choose our X matrix so that X'X is very large, so that $X'X^{-1}$ would be very small.

Design			X				$X'X$				$(X'X)^{-1}$			
1	1	1	1	1	1	1	4	0	0	0	.25	0	0	0
1	2	2	1	1	-1	-1	0	4	0	0	0	.25	0	0
2	1	2	1	-1	1	-1	0	0	4	0	0	0	.25	0
2	2	1	1	-1	-1	1	0	0	0	4	0	0	0	.25

Until now we talked about linear regression. We also discussed choice-based conjoint. Now which one of these 2 do you prefer? The biggest advantage of choice-based is that it is easier and a more realistic task to do. But now our dependent variable is no longer a scale from 0 to 10! Now we get if 1 or B was chosen. Unfortunately, that relationship becomes more difficult, it gets more difficult to define an X matrix. For ***choice designs*** there are 4 principles that need to be fulfilled:

- **Orthogonality**: satisfied when the joint occurrence on any two levels of different attributes appear in profiles with frequencies equal to the product of their marginal frequencies. Thus, if there is level balance, the joint occurrence of any combination of a three- and a four-level attribute must occur in exactly one-twelfth of the cases. (no correlation among levels of an attribute)
- **Level balance**: the levels of an attribute occur with equal frequency. For example, each level of a three-level attribute should occur in precisely one-third of the cases. (each level in a factor appears the same number of times)
- **Minimal overlap**: the contrasts between attribute levels are only meaningful as differences within a choice set. The probability that an attribute level repeats itself in each choice set, should be as small as possible.
- **Utility balance**: you should avoid showing 2 options where one option clearly dominates the other option. When there is an extreme unbalance in the utility. In order to make balanced choice options, you should already know the beta options a little bit!

There are 2 algorithms that can deal with this: catalog-based approach and mix and match approach.
- ***Catalog based approach:***
    - o Orthogonal main-effect array as first alternative in each choice.
    - o Additional alternatives by adding a constant to each attribute level of the first alternative.
    - o Essentially: orthogonal designs that can be constructed without the assistance of special software

- Catalog, fold-over, and other do-it-yourself approaches involve manually constructed experimental designs often based on OMEPs. Designs based on OMEPs support independent estimation of main-effect parameters for linear statistical models. OMEP designs do not allow independent estimation of interactions among attributes. Researchers have tended to favor the use of OMEPs because these designs are the most parsimonious in terms of the numbers of alternatives and choice questions required to obtain identification of the main effects. These designs also exhibit the two desirable design properties of orthogonality and level balance. The increased availability of software that facilitates the construction of more complicated designs has resulted in fewer studies that rely on catalog-based designs.
- It creates a choice experiment using an orthogonal array
- Tables of orthogonal main-effect designs are widely available. These catalog plans provide efficient flat designs from which to construct choice sets.

- ***Mix-and-match:***
  - Catalog-based approach
  - Calculate all possible alternatives for the orthogonal array
  - Randomly select 1 (or more) alternatives
  - Amai hier vind ik niks over

R code to create an experimental design:

```
# Load required library
library("support.CEs")

# Define attributes and levels
atts <- list(payment=c("online","offline"),
             time=c("weekend","weekday"),
             price=c("8","10","12"))

# Create choice design
des <- rotation.design(attribute.names=atts,
                       nalternatives=2,
                       nblocks=1,
                       randomize=TRUE,
                       seed=345)
```

```
# Print design
des

creating full factorial with 12 runs ...
Choice sets:
alternative 1 in each choice set
   BLOCK QES ALT payment    time price
1      1   1   1  online weekday   10
2      1   2   1  online weekend    8
3      1   3   1 offline weekday    8
4      1   4   1 offline weekday   10
5      1   5   1 offline weekend   10
6      1   6   1 offline weekday   12
...
```

"support.CEs" is a R package that provides seven basic functions that support an implementation of choice experiments. Next you define the attributes and the levels. Then you create a choice design.

```
alternative 2 in each choice set
   BLOCK QES ALT payment    time price
1      1   1   2  online weekend   10
2      1   2   2  online weekday   12
3      1   3   2  online weekday   10
4      1   4   2  online weekday    8
5      1   5   2  online weekend    8
6      1   6   2 offline weekday   10
7      1   7   2 offline weekend   10
8      1   8   2 offline weekend   12
9      1   9   2 offline weekday    8
10     1  10   2 offline weekday   12
11     1  11   2  online weekend   12
12     1  12   2 offline weekend    8
```

```
Candidate
  design:
   A B C
1  1 2 1
2  2 2 1
3  1 2 3
4  2 1 2
5  2 1 3
6  1 1 3
7  2 2 3
8  1 1 2
9  2 1 1
10 1 1 1
11 1 2 2
12 2 2 2
```

```
class=design, type= full factorial

Design information:
nbr of blocks = 1
nbr of questions per block = 12
nbr of alternatives per choice set = 2
nbr of attributes per alternative = 3
```

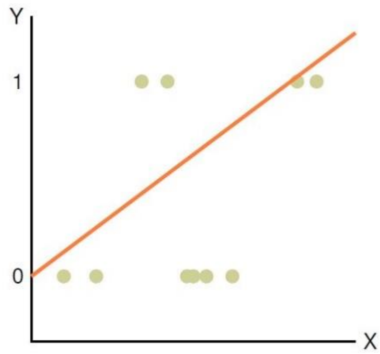
Choice models

There are two types of choice models: binary and multinomial choice. Our focus lays on the logistic regression. The dependent variable is binary and the independent variable can be categorical, interval, ratio.... For example: what are the determinants of lung cancer? Or why do customers buy a product? Or who is likely to churn?

Because the dependent variable is binary, we use dummy coding:

- 0 → non-event
- 1 → event

But what now if we put this variable in a simple linear regression?



$$E(Y) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

Look at the graph on the right. Fitting a straight line through that can be motivated as the expected value of choice given X. That expectation gives you something that can be interpreted as a probability. So, it could make sense to draw a line! But not completely. It doesn't work for many examples: for large values of p we get negative values for the choice. Then we would get negative probabilities and that is impossible. It is also not possible for small values of p, then we get probabilities larger than 1, what is also not possible. So, linear regression on binary data gives us some problems:

- There are only 2 values for the residuals, and this does not follow the assumption we make in linear regression about errors: they have to be **normally distributed**. Normally distributed implies having many different values for the error, and not just 2.
  - $1 - (\alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)$
  - $0 - (\alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)$
- The assumption of **IID residuals**. The standard regression model assumes that the residuals, or ϵ's, are independently, identically distributed (usually called 'IID' for short) as normal with $\mu = 0$ and variance σ^2. The size of the residuals depends on the value of p, epsilon will increase with p and epsilon 2 will decrease with p. The residuals depend on X. There will be smaller residuals for $\hat{Y}$ closer to 1 or 0.

So that's why we are going to use a logistic regression model. As explained somewhere above, conjoint analysis is all about **utility**.

Utility models:

$$U_1 = \beta_{1,0} + \beta_1 X_{1,1} + \beta_2 X_{1,2} + \cdots + \beta_k X_{1,k} + \epsilon_1$$
$$U_2 = \beta_{2,0} + \beta_1 X_{2,1} + \beta_2 X_{2,2} + \cdots + \beta_k X_{2,k} + \epsilon_2$$

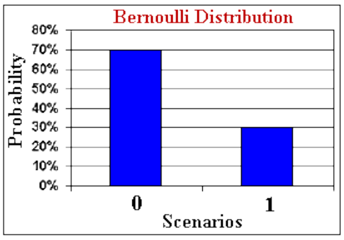
Differenced utility models:

$$(U_{i,1} - U_{i,2}) = (\beta_{1,0} - \beta_{2,0}) + \beta_1 (X_{1,1} - X_{2,1}) + \beta_2 (X_{1,2} - X_{2,2}) + \cdots + \beta_k (X_{1,k} - X_{2,k}) + \varepsilon$$

with $\varepsilon \sim Logistic(\mu = 0, \sigma = 1)$

It's not the absolute value of the product configurations that matters, it's the difference in the prices that counts. Therefore we take the difference of the x'es. We take the difference of the utility, if that is positive: product 1 was liked the most.

So we are trying to model a variable that only takes on 2 values: 1 and 0. The distribution (this is a function that tells us how likely every possible event is) for binary variables is ***Bernoulli***. For example when you toss a coin. This is the Bernoulli formula:



$$p(y|\pi) = \pi^y (1-\pi)^{(1-y)}$$

This is a likelihood function. Pi is the parameter of that distribution, y is the data. So pi is 'what is the probability of the event that we have coded by 1?'. What happens when we plug y = 1 in this formula? $p(y|\pi) = \pi^1(1-\pi)^{(1-1)} = \pi$. So if we observe a tail (= 1), than the probability will be pi. If we observe a head (= 0), than the probability will be $p(y|\pi) = \pi^0(1-\pi)^{(1-0)} = 1-\pi$. The Y and 1 – Y are switches. If y is 1 then the first term switched on, if y is 0 then the second part switches on.
So now we can calculate the probability of a single coin toss. But what if there are 2 coin tosses? Then we multiply the probabilities. If we observe many zero one random variables and we want to calculate the entire probability, then we have to multiply them all together:

$$\prod_{i=1}^{n} \pi^{y_i}(1-\pi)^{(1-y_i)}$$

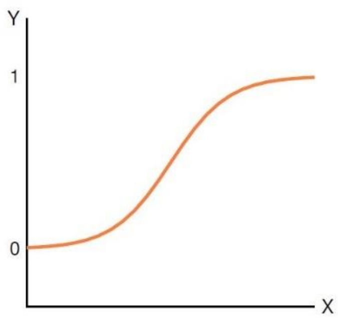
This is a **likelihood function**, the probability of data points given the parameters of the distribution, pi in this case. No back to the first formula, without the multiplier. We look for pi's that model the data as good as possible. If y is 1, pi should be as close to 1 as possible, if y is 0, pi should be as close to 0 as possible. If y is 1 and pi is 0,99, then the outcome of the equation is 0,99 which is very good. If y is 1 and we predict pi to be 0,01, then the outcome of the equation is 0,01 which is very bad. If y is 0 and we predict pi is 0,99 then the outcome is 0,01 which is also very bad. A bad prediction gives a low value for the equation, whilst a good prediction gives a high value for the equation. That is for a single prediction. But what now for all the data, all the predictions? Then we go back to the multiplied formula. We multiply all the values together. The larger that value for the likelihood, the better our predictions were. That's how we use the likelihood function to look for values of pi and to look for that value that maximizes the likelihood function. So we have again explained the maximum likelihood, but now for classification instead of linear regression.

There's one more step. The pi's can be different, when every data point has another probability (pi). Then the multiplying formula looks like this:

$$\prod_{i=1}^{n} \pi_i^{y_i}(1-\pi_i)^{(1-y_i)}$$

Imagine we have X's from the utility models from earlier, and we want to link them to the pi's. Pi is the probability of the event (between 0 and 1). The X's can range from minus infinity until plus infinity! We want a function that maps the X's to values between 0 and 1. Any function that maps real numbers to a number between 0 and 1 can be used here. The logit function that transforms any number X beta to a number between 0 and 1:

$$\pi = \frac{e^{X\beta}}{1+\ e^{X\beta}} = \frac{1}{1+e^{-X\beta}}$$



And then after we calculated all the pi's, we can plug that data into the last multiplying formula to get an S-shaped function.

Model performance: the ***Akaike Information Criterion***. The Akaike information criterion (AIC) is an estimator of out-of-sample prediction error and thereby relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection.
AIC is founded on information theory. When a statistical model is used to represent the process that generated the data, the representation will almost never be exact; so some information will be lost by using the model to represent the process. AIC estimates the relative amount of information lost by a given model: the less information a model loses, the higher the quality of that model.
In estimating the amount of information lost by a model, AIC deals with the trade-off between the goodness of fit of the model and the simplicity of the model. In other words, AIC deals with both the risk of overfitting and the risk of underfitting. Formula:

$$AIC = 2k - 2\log(L)$$

With K the number of estimated parameters in the model, and L the maximum value of the likelihood function for the model.

Data analysis

First we **create the design matrix**, using the support.CEs package:

```
ldmat <-
  make.design.matrix(choice.experiment.design=des,
  categorical.attributes=c("payment","time","price"),
  optout=FALSE,
  unlabeled=TRUE,
  binary=TRUE)
```

Don't forget the variable coding!
Binary variables: 0 – 1 coding
Categorical variables (for example 3 categories:

0	1
1	0

```
> ldmat
```

	BLOCK	QES	ALT	ASC	offline	weekday	X.10	X.12
1	1	1	1	1	0	1	0	0
2	1	2	1	1	0	−1	0	−1
3	1	3	1	1	1	0	−1	0
4	1	4	1	1	1	0	1	0
5	1	5	1	1	1	0	1	0
6	1	6	1	1	0	0	−1	1
7	1	7	1	1	−1	1	−1	0
8	1	8	1	1	−1	0	0	0
9	1	9	1	1	0	−1	0	0
10	1	10	1	1	−1	−1	1	−1
11	1	11	1	1	0	1	0	0
12	1	12	1	1	0	0	0	1

Then you **estimate the model parameters**, by loading your data into R:
Y <- read.csv("choices.csv")

Out <- glm(unlist(y)~0+as.matrix(ldat), family=binomial(logit))

Summary(out)

Next you **calculate importances**. Conjoint analysis can assess the relative importance of each factor. Because part-worth estimates are typically converted to a common scale, the greatest contribution to overall utility, and hence the most important factor, is the factor with the greatest range (low to high) of part-worths. The importance values of each factor can be converted to percentages summing to 100 percent by dividing each factor's range by the sum of all range values. So first you need to find the range (max value minus min value) for each attribute and then divide each range value by the total for the importance value.

```
# Calculate importances
## create function that calculates the ranges
calc_range <- function(x){
  x <- c(0,x)
  return(max(x)-min(x))
}
## use function for every attribute in our analysis
a <- calc_range(coef(out)[1])
b <- calc_range(coef(out)[2])
c <- calc_range(coef(out)[3:4])

## normalize
import1 <- a/(a+b+c)
```

Then you find the optimal profile. ←

```
## Create full-factorial
full <- cbind(rep(c(0,1),each=6),
              rep(rep(c(0,1),each=3),2),
              rep(c(1,0,0),4),
              rep(c(0,1,0),4))
full

## Calculate utility for every profile
utilities <- full%*%coef(out)
utilities

## find optimum
which.max(utilities)
```

Then we calculate the linear in parameter price effect? It is very similar as the previous example: see the R code. You calculate the compensatory price for online buying.

```
# Utility for offline + weekend + 10 (normalprice)
(u1 <- c(1,0,10)%*%coef(out))
# Utility for online + weekend + 10
(u2 <- c(0,0,10)%*%coef(out))
# Compensating reduction for online + weekend
discount <- (u2-u1)/coef(out)[3]
c(0,0,10-discount)%*%coef(out) # check
```

And lastly, calculate the individual level parameters.

```
# prepare data for individual level parameters
ldat_ind <- replicate(n, ldmat[5:7], simplify=FALSE)
y_ind <- split(unlist(y),f=rep(c(1:n),each=12))
ind <- mapply(list,y_ind,ldat_ind,SIMPLIFY=FALSE)
# create function so that we can use lapply
logreglist <- function(x){
  out <- glm(x[[1]]~0+as.matrix(x[[2]]),family = binomial(logit))
  return(coef(out))
}
# estimate parameters for every item in the list
ind_coef <- lapply(ind,FUN=logreglist)
```

Guest lecture Boobook

Using customer data to build strategic marketing plans. Research methods to support product development and pricing.
So what do they do: data driven customer strategy consulting. They develop a deep insight into the customers' behaviours, motivations, attitudes and demographics.

- They discover the wisdom through customer analytics
- And they spread the wisdom through visualisation and story telling

Using a toolkit of analytical approaches and methodologies, they get an holistic customer view (all around):

- ***<u>Customer experience and journey management</u>***
    - **Propensity modelling**
        - Propensity modelling attempts to predict the likelihood that visitors, leads, and customers will perform certain actions.
    - **CHAID**
        - Chi-square automatic interaction detection (CHAID) is a decision tree technique, based on adjusted significance testing (Bonferroni testing). In practice, CHAID is often used in the context of direct marketing to select groups of consumers and predict how their responses to some variables affect other variables, although other early applications were in the field of medical and psychiatric research. Like other decision trees, CHAID's advantages are that its output is highly visual and easy to interpret. Because it uses multiway splits by default, it needs rather large sample sizes to work effectively, since with small sample sizes the respondent groups can quickly become too small for reliable analysis.
    - **Text analytics**
        - Text analytics is the automated process of translating large volumes of unstructured text into quantitative data to uncover insights, trends, and patterns. Combined with data visualization tools, this technique enables companies to understand the story behind the numbers and make better decisions.
    - **Key drivers analysis**
        - A key driver analysis (KDA) allows you to identify what features or aspects have the biggest impact on an outcome variable such as likelihood to recommend, brand attitudes, and UX quality. It's one of the more powerful techniques we use to help prioritize findings in surveys.
- ***<u>Customer value proposition development</u>***
    - **Conjoint analysis**
        - 'Conjoint analysis' is a survey-based statistical technique used in market research that helps determine how people value different attributes (feature, function, benefits) that make up an individual product or service.
    - **Gabor Granger**
        - The Gabor–Granger method is a method to determine the price for a new product or service. To use the Gabor-Granger method in a survey, one must find the highest price that respondents are willing to pay. There are many ways to do this but the most common is usually done by choosing 5 price points for

the survey and then asking the respondent a 5-point purchase intent question for a random price from those 5 established price points. If the respondent answers in the top 2 choices - 'Definitely Buy' or 'Probably Buy' for this question, they are then asked the same question for a random price that is higher than was just asked. If it is not in the top 2 then the respondent is asked the same question for a random lower price. This is done until you find the highest price the respondent is in top 2 on Purchase Intent Scale. If they are not in top 2 for the lowest of the 5 prices, the respondent is usually coded as a zero or deleted from the analysis. Once you have this Gabor-Granger variable, the results can be used to produce a demand chart (where x-axis are the prices and y axis the percentage of people willing to pay that price) and a revenue curve (where y-axis is the percentage of optimal revenue and x-axis is still price).

- **Experimental design**
    - An experimental design is a plan for running an experiment, and it is often displayed as a matrix.
- **Max diff**
    - The MaxDiff is a long-established academic mathematical theory with very specific assumptions about how people make choices: it assumes that respondents evaluate all possible pairs of items within the displayed set and choose the pair that reflects the maximum difference in preference or importance.
- **Key drivers analysis**
    - A key driver analysis (KDA) allows you to identify what features or aspects have the biggest impact on an outcome variable such as likelihood to recommend, brand attitudes, and UX quality. It’s one of the more powerful techniques we use to help prioritize findings in surveys.

- ***Customer profiling and targeting***
    - **Segmentation, cluster analysis etc**
        - Market segmentation is the activity of dividing a broad consumer or business market, normally consisting of existing and potential customers, into sub-groups of consumers (known as segments) based on some type of shared characteristics.
        - Cluster analysis or clustering is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters).
    - **Propensity modelling**
        - Propensity modelling attempts to predict the likelihood that visitors, leads, and customers will perform certain actions.
    - **Decision trees (CHAID etc)**
        - A decision tree is a decision support tool that uses a tree-like model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements. Decision trees are commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal, but are also a popular tool in machine learning.

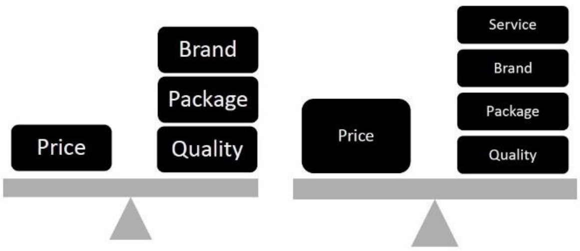
Leveraging multiple data sources for maximum insight:

- WHO (profiling, contextual data)
  - Primary research
  - 3rd party data
- WHAT (behavioural data)
  - Transactional data
  - Digital data
  - Interaction data
- WHY (attitudinal data)
  - Primary research
  - Social media
  - User generated content

Customer value proposition

First of all, what is conjoint? It is a range of methods in which we assess products or services by testing a series of potential concepts build up by combining levels from different attributes. An attribute is a potential feature of a product or service. Levels are different degrees of those attributes. It is based on a trade off principle. You can't have the best product for the lowest price.

- What are the features without even worrying about the price
- Different concepts are shown and evaluated
- You want to find the balance for every person (see picture)
- And while finding that balance, you can initiate segmentation as well.



Conjoint is typically used for product optimisation/NPD, portfolio optimisation, pricing strategy, competitive changes (how will my profit shift?) and cannibalisation (make sure that we don't lose the people that are willing to pay less as a result of bundling for example).

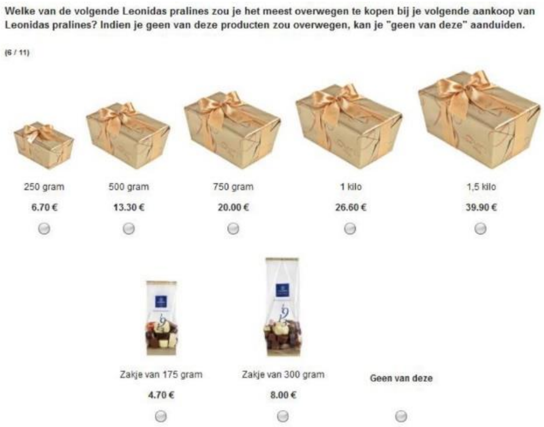
The key components for a successful conjoint exercise:

- Identifying the right business question
- Developing an appropriate conjoint grid
- Choosing the right methodology
- Making it easy for the respondent
- Correctly interpreting the results

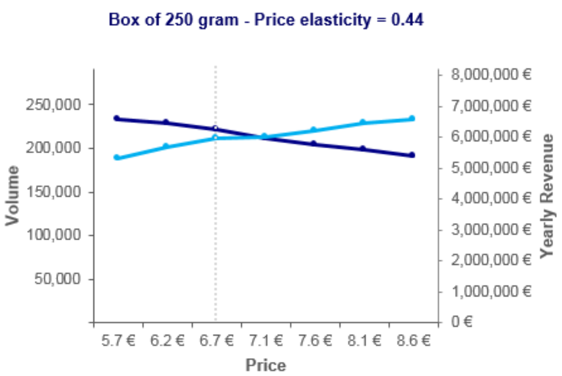
Case study Leonidas

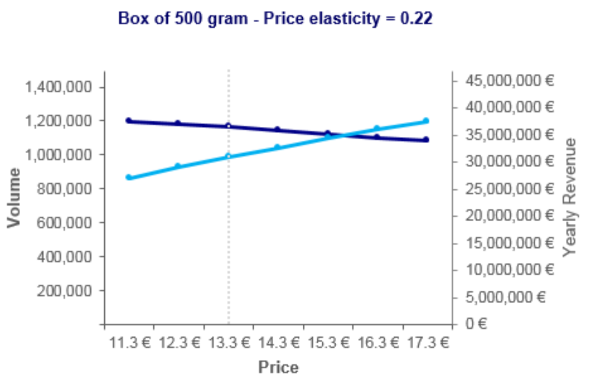
So Leonidas was questioning its pricing strategy. They are pretty cheap, cheaper than most competitors on the market. Why? Well they are **volume driven**, if you want to target a lot of customers, it generally means you can't set your price too high. They also have the **same price in all locations**, in city centers like Ghent for example, the price stays the same. And they have a **linear price setting**, depending on the quantity there is a linear increase in price.

But so they wanted to know more about the price elasticity and the buying behaviour of their customers so they did a primary market research, because there wasn't data that they could use so they had to get the data themselves. There were 14000 Belgian participants and 1600 French participants. The participants all bought Leonidas pralines in the last 12 months. The sample was spread across large cities, small cities, large villages and small villages. It was a questionnaire and the topics were about the purchase behaviour, the chocolate brands: brand & price perception and the brand-price trade-off. Example question: see image. You want to understand as of which price point, a product becomes too expensive (or too cheap) for a specific customer. There are two attributes: price and amount of chocolate.



It turned out that Leonidas has quite a good position in the market: good quality and low price. The key output are multiple what if scenarios.





The light blue line is the price, the dark blue line is the demand. Customers are more price sensitive for a box of 250g than for a box of 500g.

Case study home automation

Blended learning creates a mix of real world learning and virtual/online world learning.
Conjoint analysis will lead to understanding which impact each of the tested elements (attributes) has on the purchase consideration. And it can also show you how your market share is divided. 31% still has to be convinced of home automation, 42% of the respondents are energy-cost savers and 27% are domotica lovers.

Case study Disney

What should the price setting be for a one-day visitor pass?
In China, you have very fixed holidays and people often travel all at the same time during these periods. Basically, Disney Shanghai wanted to ask the highest price, but also wanted to offer some extra benefits in order to **<u>spread the volume of the visitors across the year</u>**.
→ *Conjoint analysis*:
Different price points (see image conjoint)
<u>Recommendations on pricing</u>:

1. There was not a lot of price sensitivity observed: there was some room to ask for a pretty high price, since customers are willing to pay a high price.
2. Rejected demand = number of tickets you would lose

What’s your revenue and lost revenue when people cannot visit the park.

Estimating basic price response functions

Intro: article about prices from hairdressers. We've seen that it is optimal to segment the market between male & female customers as both segments attribute a different value to your offered services. But is this ethical/legal based on gender jurisdiction?
Cost-plus pricing is not a price strategy that is optimal in terms of financial results, but it is the most digestible pricing strategy for customers. If companies change prices, they will often try to link this to a cost-plus pricing schema, what customers believe easily. For example, Leonidas increased prices for chocolate after they found in a research that demand would not drop significantly, publicly they announced that the price of raw cacao went up.

Introduction

Pricing and revenue optimizations deals with how companies should set and adjust their prices in order to maximize profitability.

- Pricing helps in decision making support system
- Sets analytical techniques in Dynamic environment

Before classical economists, people thought that price was an intrinsic property of a product. For example, the price of bread would be attached to the bread, like the colour, weight, taste.... another example: a bread is worth 10 apples. Also, everyone was forced to sell at prices set by ruler in Medieval stages, so there were non-fluctuating prices.

After classical economists was the price setting an interplay of supply and demand. The intrinsic value, cost or labour are secondary resources to set prices. Also, there is no "right" price, only actual prices based on willingness of sellers to sell and buyers to buy. In 1630 the Tulip became very popular in the Netherlands, people started buying more and more tulips and the prices of this product went up! Tulip mania was the first example of an economic bubble: the price grew extremely rapidly in this industry until the bubble burst at a certain time. (price at its highest was 10 times the wage of an average worker in a year for a tulip).

But there rose some new questions: why are prices stable? Answer of 20th century economists is that individuals maximize utility. There are 2 players.

- Workers get paid: trade labour for wage. With this wage they buy stuff and try to max utility
- Company buys labour and transfers it in products which they sell to the workers. They want to max their profitability.

These two create an equilibrium → stable prices. (if the price is too high, no one buys and no labour is needed.
Equations of the household (trading labour for wage) and equations of the companies (buying labour, producing products and services and selling those on the market) lead to an equilibrium.

But why are still variations in the prices then? Sometimes buyers behave irrational. Also long-term profit is not always maximized by sellers, they also want market share. Information asymmetries can exist where the seller knows more than the buyer (but the internet destroys this) and arbitrage is not always immediate! Arbitrage means that when people find a loophole to buy at a cheaper place and sell it at a more expensive place, price difference will go to zero.

Now what are the four things that helped pricing and revenue management to take off in companies?

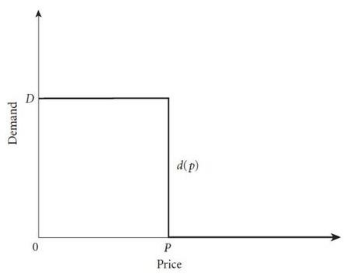
- Success of revenue management in airlines industry
    - The ticket prices were first regulated, then freed up. There was a difference between leisure and business travellers. Leisure travellers typically book way ahead of flying because of holiday planning: prices are considerably lower when you book very early. Also if there is a weekend is included, the software often concludes that the traveller is leisure instead of business.
    - That's why they do: the earlier you book, the lower the price. If they would do the later the lower the price, then everyone would postpone, which is a logistical hell. Even when it's to fill up spots, they won't drop the price so the customer wouldn't get used to that.
    - Why do they overbook? They always assume that 1% à 2% will not show up. And if everybody does show up, they pay someone to take a later flight with this extra money they made of all those that did not show up.
- Widespread adoption of enterprise resource planning (ERP) and customer relationship management (CRM) tools
    - Digital data is necessary to do revenue management! These IT tools helped companies to digitalize data and made revenue management possible.
    - These ERP and CRM tools gave good results, so they started thinking where they could also implement them → pricing!
- The rise of e-commerce
    - Suddenly so much data was available.
- Success of supply chain management analytics
    - Because for real management we need data, which ERP and CRM provided

First there was McKinsey in 1992. For most companies, better management of pricing is the fastest an most cost-effective way to increase profits. Next A.T. Kerney in 1999, price improvement has by far the largest impact on operating profit. Last Bain & company in 2013, price optimization models now allow companies to use pricing as a powerful profit lever, which often is underdeveloped.
Optimal pricing often means that you should not focus too much on costs, but on other perspectives!

Now how does pricing analytics work in practice? First you estimate the price-response function. Then you define an objective function (e.g. maximize revenue,...), the thing you want to analyse. Next you identify the constraints (e.g. production capacity,...). And last, you implement, analyse and reiterate.

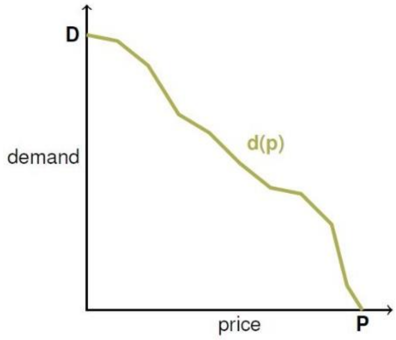
Basics of price optimization

<u>The price response function</u> specifies demand for the product of a single seller as a function of the price offered by that seller. It is not the same as the market response curve because we have control over the price with the price response function, in the market response curve (MRC), we don't. In a MRC, the price results from the demand.



The image on the left is a price response curve in a **perfectly competitive market.** If the price is higher than P, you sell nothing. And when the price is lower then P, there is demand in the entire market. The pricing is set by the market.

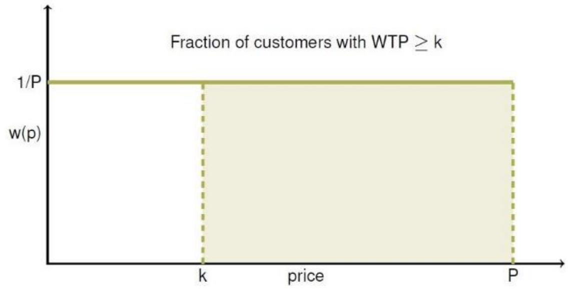
You only decide how much output you need, there is only one price. = true commodities, but these are rare.
Assumptions of this function. **The price response function** will only live in the quadrant where the X axis AND the y axis are positive. The prices are positive, demand goes from 0 to a positive number. Price response functions will also always be downwards sloping. **Luxury goods** deviate from this example. Sometimes the rich will stop buying a product if it becomes too cheap, because part of the reason that they buy the luxury good is to show off that they can afford the item in question. There is also something as a **Griffin-good**. For example, you have 8 euros each week and you have to eat each day (so 7 times in total). Steak = € 2 and burger = € 1. You want to eat as much steak as possible, so you eat 1 steak and 6 burgers. You can only eat steak one time each weak because you have to eat every day. When you eat steak 2 times, that means that you only have 4 euros left for 5 days. Even if you only eat burgers, you are still going to have a day that you cannot eat. Now imagine that the burger prices go up by 10%, so now a burger costs € 1,1. Now you have to eat 7 burgers and no steaks because of budget constraint. So the price of hamburgers goes up, but you still buy more = Griffin-good. BUT in the context of the course, we only assume negative relationships between price and demand!
A third assumption is that our price response functions are **differentiable**. What continuous functions are not differentiable? When you have multiple values for one particular point.



Willingness-to-pay is the maximum price a consumer is willing to pay for a given product. An alternative term is the reservation price. The price response function has a one-to-one relationship with the assumption that we make on the distribution of the willingness-to-pay. The price response function is based on a number of assumptions. Most importantly: the distribution of willingness-to-pay. So it is based on an assumption of consumer behaviour. A distribution is a function that has inputs and an output (number). Discrete distributions have a probability as output. Continuous distributions have a probability density as output. Basically, the distribution tells us how likely it is that an event will happen.
An example of a **discrete distribution** is a Bernoulli distribution. It has discrete probabilities on heads and tails. Probability can be put into the distribution (encoded, so not H and T but 0 and 1): $\pi^y(1-\pi)^{1-y}$. Assume Heads = 0, tails = 1 and pi = 0,5 because a coin toss: $0.5^0(1-0.5)^{(1-0)} = 0.5\ AND\ 0.5^1(1-0.5)^{(1-1)} = 0$.
For a **continuous distribution** you take an interval for which you want to calculate, because the probability that within a continuous spectrum an outcome is exactly equal to a single number is zero. For example, the height of a person: the probability that a person is exactly 183,12314cm tall is 0. You can however calculate the probability that a person is between 182,9cm and 183,2cm.

The question now is: is the willingness-to-pay (wtp) a discrete or continuous variable? P is the highest willingness to pay that any customer has in my population. No one is willing to pay a higher price for my product, or the demand would be 0. Where the demand is 0, that is the satiating price.
Now what is 1/P? Area of rectangle:

$$P * ? = 1 \rightarrow ? = 1/P$$



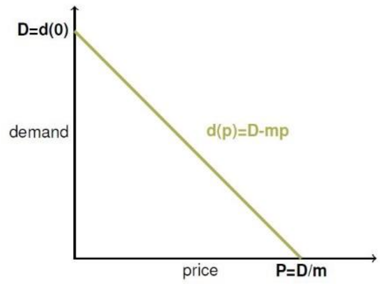
We add a price point k. The willingness-to-pay distribution should tell us something about k. it should make clear how many people have a WTP that is above or below k. The distribution would tell us what proportion of customers would buy our product at price k.
The area to the right of k is now of interest. Only people that have a WTP that is equal or higher than k will actually buy the product! Every customer has his own WTP and will only buy if the price is lower then his or her WTP.
Let D be the total market size. Then the estimated demand at price k can be written as:

$$d(k) = D * \int_{K}^{P} w(p)dp$$

D(k) is the demand expected to materialize at price p = k. It is also the number of people who have a WTP greater than price p = k. The integral after D * is the area of the WTP distribution to the right of price p = k.
In statistical terms, we are searching here for the cumulative distribution function. In mathematics, this is the primitive function w(p).

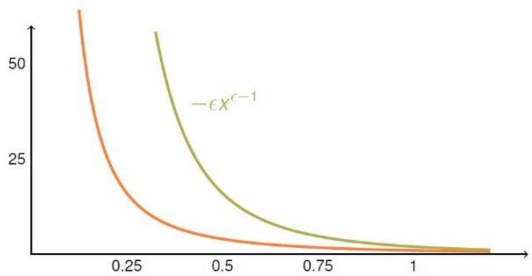
$$\left[W(x) = \frac{1}{p} * x + c\right]_{K}^{P} = W(P) - W(k) = 1 - (\frac{1}{p} * k)$$

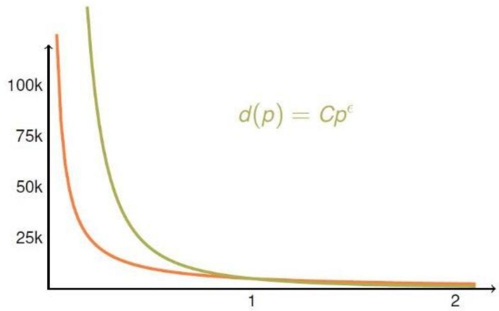
We know that 1 is the surface of the total rectangle [w(p)] because we are dealing with distributions so the total surface under the curve is 1, and 1/p * k is the surface of the rectangle left from k. so this essentially results in that area right from k up to P or the green marked area in the figure. If we then multiply this by D, we get the total number of customers that would buy instead of the proportion of customers that would buy. $D\left[1 - \frac{k}{P}\right] = D - \frac{Dk}{P}$. The picture on the right is a linear price-response function.



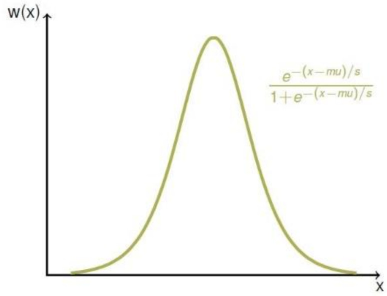
We now described an uniform WTP distribution, but this might not be very realistic in practice.

The first picture is the constant elasticity WTP distribution, the second is the constant elasticity price-response function.

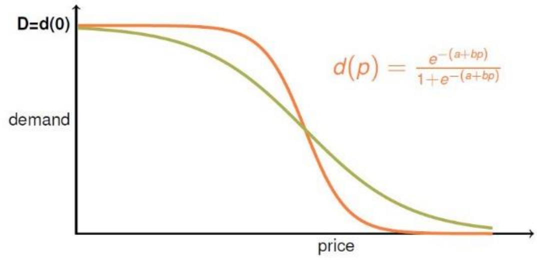




A logistic WTP distribution is easier to investigate. A normal distribution for WTP is the most logical. Normally the investigated price ranges are relatively small. For example, Leonidas will not investigate what would happen if they would sell their chocolate boxes for 3000 euros instead of 25 euros, or for 20 cents. When the investigated price ranges are relatively small, then the assumption of linearity is not that unrealistic (see 3 pictures up). But in general, the picture on the right is a more realistic price-response function:

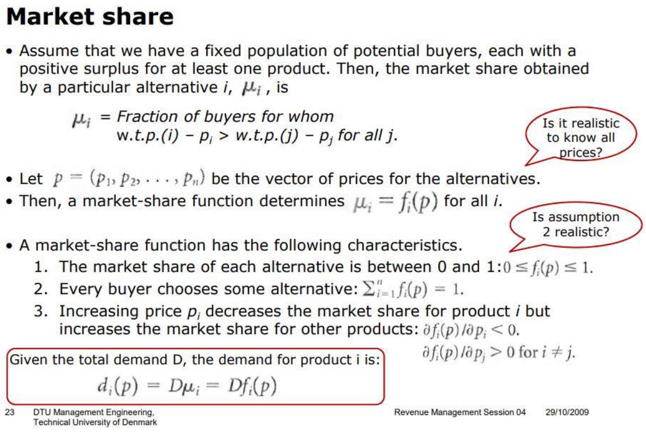


→



There are other price response functions, like the power price-response function. $d(p) = \frac{\alpha D}{p^{\beta}+\alpha}$. Where D>0, α>0 and β>0 are parameter values estimated by fitting this equation to the price/demand data. Higher values of beta represent more price-sensitive markets. As beta grows larger, the market approaches the perfectly competitive price-response function. Different values of alpha, alpha>0, shift the curves left and right along the horizontal price-axis. It investigates small price differences, small price ranges. And because of that we can assume linearity.

Another kind are the market share functions.



Measures of price sensitivity

Imagine we change our price by some number. How much change in demand does this cause? Price sensitivity is about the ***slope of your price-response curve***. By definition this slope for us will be negative. A 'large' slope, for us meaning a more negative slope, will indicate a higher price-sensitivity.

$$\delta(p_1, p_2) = difference\ between\ two\ exact\ price\ points, p_1\ and\ p_2$$

$$\delta(p_1, p_2) = \frac{d(p_2) - d(p_1)}{p_2 - p_1}$$

Function underneath: we let the distance between p1 and p2 become infinitely small, so we are basically calculating in one specific point. And this results in the equation of the derivative of a function in a certain point = the slope of a tangent line to that point. So if you want to calculate the slope but you have only given one point, this is the formula:

$$\delta(p_1) = lim_{h\to 0} \frac{d(p_1 + h) - d(p_1)}{h} = d'(p_1)$$

QUESTION

Assume that the demand for a chemical is 50 000 tons at 10 cent, but drops to 40 000 at 11 cents. 1. Calculate the slope for the price-response function (assume linear). 2. What would the slope be in kilogram per euro? 3. What is the expected sales at 15 cents?

1. $\frac{d(p_2)-d(p_1)}{p_2-p_1} = \frac{40\ 000-50\ 000}{11-10} = -10\ 000$
2. $\frac{40\ 000\ 000 - 50\ 000\ 000}{0{,}11-0{,}10} = -1\ 000\ 000\ 000 \frac{kg}{€}$
3. So we know the slope is -10 000. $-10\ 000 = \frac{50\ 000 - x}{10 - 15} \to x = 0$. The expected sales at 15 cents is 0.

Another measure of price sensitivity is ***elasticity***. It is similar to the slope, but independent of the measurement level. The slope is the absolute change in demand due to absolute change in price. While the elasticity is the relative change in demand due to relative change in price. For example, if I up my price with 10%, then this will lead to a decrease in demand of ...%. If the elasticity is large: then there will be a big difference in demand when there is a price change. When the elasticity is small: then the difference in demand when the price changes will not be that large. Elasticity is the ratio of the percentage in demand to a percentage change in price. Formula:

$$\epsilon(p_1, p_2) = \frac{100\frac{d(p_2) - d(p_1)}{d(p_1)}}{100\frac{p_2 - p_1}{p_1}} = \frac{p_1(d(p_2) - d(p_1))}{(p_2 - p_1)d(p_1)}$$

Again, if you have to calculate the elasticity and you have only one point given:

$$\epsilon(p_1) = lim_{h \to 0} \frac{p_1[d(p_1 + h) - d(p_1)]}{hd(p_1)} = \frac{p_1 d'(p_1)}{d(p_1)}$$

Price elasticity is a local estimate and is always smaller than 0. It is independent of units. Low elasticity (inelastic) → $|\epsilon| < 1$. High elasticity (elastic) → $|\epsilon| > 1$. Short-run versus long-run elasticity: elasticities are often lower in the short run than in the long run. In the short run, it is not easy for a person to make substantial changes in the energy consumption. Maybe you can carpool to work sometimes or adjust your home thermostat by a few degrees if the cost of energy rises, but that is about all. However, in the long-run you can purchase a car that gets more miles to the gallon, choose a job that is closer to where you live, buy more energy-efficient home appliances, or install more insulation in your home. As a result, the elasticity of demand for energy is somewhat inelastic in the short run, but much more elastic in the long run.

Market elasticity versus brand elasticity: if the price goes up for the whole market (for example all milk gets more expensive), it is way harder to switch for people and elasticity will be smaller. But when one particular brand raises its prices, people simply change to a different brand, so the elasticity will be higher.

Good	short-run	long-run
Salt	0.1	–
Airline travel	0.1	2.4
Tires	0.9	1.2
Restaurant meals	2.3	–
Automobiles	1.2	0.2
Chevrolet	4.0	–

Demand is less elastic on the short term than on the long term. In the long term they will find alternatives. In the short term, when the price increases, they will keep buying out of habit.

The impact of elasticity on revenue:

- $|\epsilon| < 1$: inelastic. Raising price will increase revenue
- $|\epsilon| = 1$: revenue is independent of price
- $|\epsilon| > 1$: elastic. Raising price will decrease revenue

<u>QUESTION</u>

Assume that the demand for a chemical is 50 000 tons at 10 cents, but drops to 40 000 tons at 11 cents. 1. Calculate the elasticity of the price response function (assume linear). 2. What would the elasticity be in kilogram per euro?

1. $\frac{p_1(d(p_2) - d(p_1))}{(p_2 - p_1)d(p_1)} = \frac{10*(40\,000 - 50\,000)}{(11-10)*50\,000} = -2$
2. The elasticity would be the same, because elasticity is independent of the measurement level.

To estimate the price response functions, you first have to find relevant data. This can be done through price experiments, from expert opinions, out customer surveys or sales data. Once you have the data, you have to take care of outliers, missing values and out of stock events. And then you have to choose the best price-response model. You do this by training lots of models and then testing the best performing model.

EXCERSISES IN R!

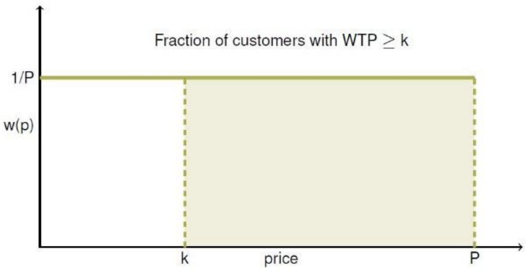
Basic price optimization

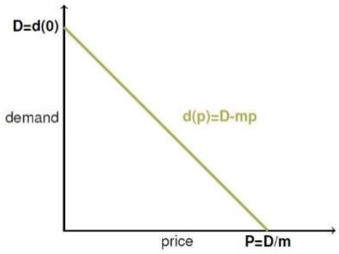
Little recap of last chapter

The price-response function is the core concept in pricing. We must understand the relationship between price and demand. The price-response function tells us what the expected demand would be for a given price. A linear price-response function goes down, is a downwards sloping function. Demand goes up if prices go down, and also the other way around. This is one of the most simple functions that we can use, but it might not be so realistic.

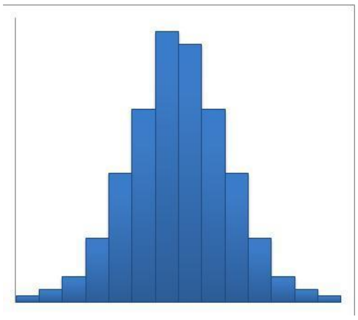
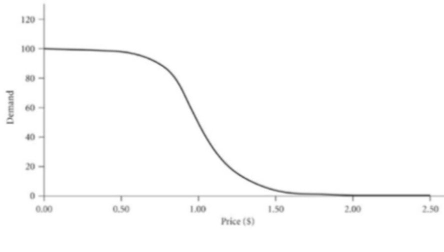
There is a one-to-one relationship between the price response function and the distribution of the willingness-to-pay. A distribution is a function and we can plug events in that function an it will say how likely that event will happen. We used a distribution over the random variable willingness-to-pay. The distribution tells us how likely every possible value for the wtp is in the population. For the linear price response function we get a very simple distribution.

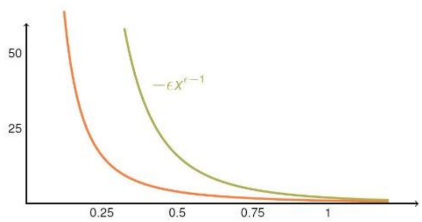
The willingness-to-pay: $\int_k^p wtp(p)dp$. This is a number between 0 and 1. Multiply it by D to get the whole population. So we are simply calculating the surface of a rectangle. Simplifying things further, and we get the formula of a straight line: $d(k) = D * \int_k^p wtp(p)dp$.





This might not be a very realistic wtp function. Wtp is a price point itself, it's not 'related' to price. We can assume that there are fewer people with a very high wtp. The higher the wtp, the fewer people we will find with this wtp. An uniform wtp distribution tells us the histogram we see on the right: a straight line. But this is not realistic!
A normal distribution is a bell shaped curve. A logistic price-response curve is an inverse S shape! This relates to the constant elasticity price function, which has an L shape.



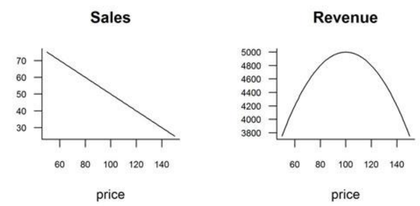
Knowing that our true price response function looks like an inverse S, how is it possible that we can estimate it with a straight line? Well you often only investigate small price ranges, so then the linear approximation gets better.
How price sensitive is our market now? We get a different measure of price sensitivity at different price points:

- Slope: -10. If the price goes up with 1, demand goes down with 10
- Price elasticity: -1. If I increase my price with 1%, the demand will go down with 1%

Basic price optimization

The end goals is to optimize the price. If we want to do this, we have to determine our variable that we want to optimize (the objective function). Do we for example want the maximize revenue, profit or revenue? Or do we want to minimize the prices?

Let's start with ***<u>the revenue function</u>***: $r(p) = d(p) * p$. It is impossible to get a negative value, because we are multiplying 2 positive values, demand and price. The revenue function is 0 when the price is 0 or when the demand is 0. The revenue function will cross the x-axis when the price is 0 or when the price is P (satiating price).

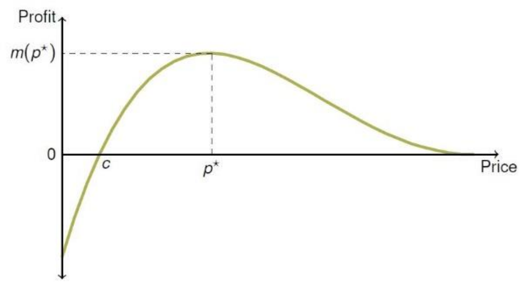


Now we want to find the maximum revenue (the objective function), what are the mathematics you use to arrive at this point? We want to find the top of the revenue curve. At the top the slope of that price point is 0. What is the formula for the calculation of the slope in 1 specific point? The marginal revenue!

$$\frac{dr(p)}{d(p)} = r'(p) = d'(p) * p + d(p) * 1 = marginal\ revenue\ (MR)$$

Now we want the marginal revenue to be 0, so we solve at p to get the optimal price to maximize revenue.

Next is ***<u>the profit function</u>***: $m(p) = d(p) * (p - c) - C$. Where c = unit cost price and C = total cost. But we don't care about C. If C increases, the only effect on the profit function is that it will shift up or down. The optimum price P^* that we are looking for, will remain unaltered.



If our price is smaller than the unit cost, the profit will be negative. If our price is larger than the satiating price, the profit will also be negative.

So now we want to maximize profit (the objective function), we want to find the optimum price P^*. In other words, we want to find that price point where the slope is 0. The formula that calculates the slope in 1 specific point is the marginal profit!

$$m'(p) = d'(p) * (p - c) + d(p) = marginal\ profit\ (MP)$$

So we want the p where the marginal profit is 0:

$$0 = d'(p) * p - d'(p) * c + d(p)$$

But here we recognize something! The cost formula: $c(p) = d(p) * c$. The marginal cost formula (MC): $c'(p) = d'(p) * c + d(p) * 0 = d'(p) * c$. If we then take the marginal profit = 0, and separate the marginal cost:

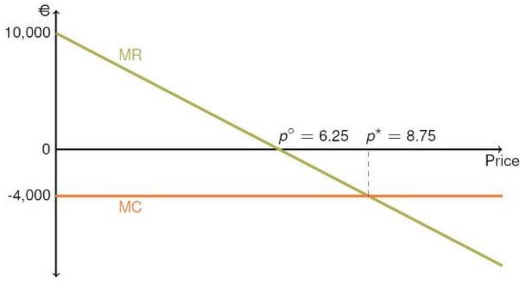
$$marginal\ cost\ (MC) = d'(p) * c = d'(p) * p + d(p) = marginal\ revenue\ (MR)$$

<u>EXAMPLE</u>

Consider a seller with price-response function given by: $d(p) = 10\,000 - 800p$ with incremental cost c = €5. 1. Calculate marginal revenue curve. 2. Calculate marginal cost curve. 3. Plot both curves. 4. Identify revenue maximizing price. 5. Identify profit maximizing price.

The linear price-response function is: d(p) = D – mp. So:

D = 10 000 (intercept)
M = -800 (slope)
C = 5 (unit cost)
The revenue function is d(p)*p so: r(p) = (10 000 – 800*p)*p
The marginal revenue curve is the deviate from the revenue function: r'(p) = -800*p + (10 000 – 800*p)
The cost function is c(p) = d(p)*c = (10 000 – 800*p)*5
The marginal cost function is de deviate from the cost function: c'(p) = -800*5 + (10 000 – 800*p)*0
The profit function is m(p) = (10 000 – 800*p) * (p – 5)
The marginal profit function is: m'(p) = -800*(p – 5) + 10 000 – 800*p
Revenue maximizing price:
0 = -800*p + 10 000 – 800*p → p = 6,25
Profit maximizing price:
0 = -800*p + 4 000 + 10 000 – 800*p → p = 8,75



Now let's go back to the marginal revenue formula: $r'(p) = d'(p) * p + d(p)$. The first term looks somewhat like price elasticity: $\varepsilon(p) = \frac{d'(p)*p}{d(p)}$! Let's see if we can insert the elasticity in the MR:

$$r'(p) = \frac{\boldsymbol{d(p)} * d'(p) * p}{\boldsymbol{d(p)}} + d(p) = d(p) * \varepsilon(p) + d(p) = d(p) * (\varepsilon(p) + 1)$$

The marginal revenue tells you something about the steepness of the revenue curve. If we calculate the marginal revenue in a certain point and we obtain a positive value: what does it mean for the price? Is the revenue maximizing value than larger or lower than the observed point?

What did we actually calculate with the marginal revenue in that point? the slope! So when this value is positive, that means that we are in a point where the price is lower than the revenue maximizing price. When the value is negative that means that the slope is now downwards, and that we are in a point that is larger than the revenue maximizing price.
If the value is positive, that means that if we increase our price we can increase our revenue. If the value is negative, that means that we can increase our revenue by decreasing our price.
Now we have rewritten our marginal revenue function to get the elasticity into the function. When can this function be negative? D(p) can't be negative. So for the marginal revenue to be negative, the term with the elasticity should be negative. When you have elastic prices, so when you have elasticities that are below -1. If you have inelastic prices, the function will become positive so you can increase the price to increase your revenue.

Now let's see if we can link the elasticity to the marginal profit function where the price is optimized.

$$0 = d'(p^*) * (p^* - c) + d(p^*)$$
$$d(p^*) = -d'(p^*) * (p^* - c)$$
$$\frac{-d'(p^*)}{d(p^*)} = \frac{1}{(p^* - c)}$$

Now let's multiply both sides with p^*.

$$\frac{-d'(p^*) * p^*}{d(p^*)} = \frac{p^*}{(p^* - c)}$$

Now we got the elasticity!

$$\varepsilon(p^*) = -\frac{p^*}{(p^* - c)}$$

Thus, at the profit maximizing price, the price elasticity is equal to the reciprocal of the contribution margin ratio. So to get the contribution margin ratio you do 1/elasticity. This also means that the profit maximizing price can be derived from price elasticity and incremental cost. Also:

$$p^* = \frac{\varepsilon(p^*)}{\varepsilon(p^*) + 1} * c$$

The profit maximizing price can be derived from price elasticity and incremental cost.

<u>EXERCISE 1</u>

An electronics goods retailer faces a constant-elasticity price-response function with ϵ = -2,5 for a popular model of television. It costs the manager € 180 apiece to purchase the television wholesale. What is the optimal contribution margin and price?

$$p^* = \frac{\varepsilon(p^*)}{\varepsilon(p^*) + 1} * c = \frac{-2{,}5 * 180}{-2{,}5 + 1} = 300$$

The profit maximizing price is €300.
The optimal contribution margin is 1/elasticity so 1/2,5 = 0,4 = 40%.

<u>EXERCISE 2</u>

A seller believes he is pricing optimally, and his contribution margin ratio is 20%. What should the price elasticity be?
Well the price elasticity is 1/contribution margin ratio, so 1/0,2 = 5.

Pricing with constrained supply

Most sellers face supply constraints at least some of the time. When there are fixed supply intervals for example. So in between replenishment times, there are limited to selling their current inventory. Or the seller may technically have a limited amount of inventory on hand, but he does not need to consider a supply constraint when setting the price (example, bottled water during a hurricane). In most cases, retailers stock enough inventory that a stock-out is unlikely. A drugstore will typically have enough toothpaste, shaving cream and aspirin (unless for extra ordinary circumstances).

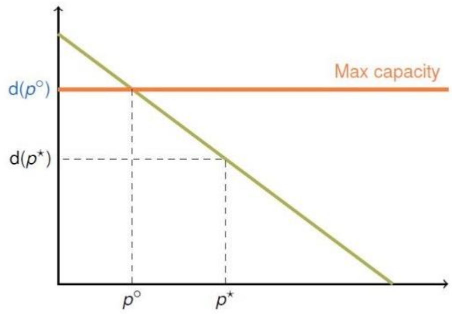
There are many cases in which a seller needs to consider the constrained nature of supply in order to calculate the optimal price.

- Service providers almost always face capacity constraints. A hotel is constrained by the number of rooms, a gas pipeline by the capacity of its pipes, a barbershop by the number of seats and barbers it has available, and the cinema also by the number of seats available.
- Manufacturers face physical constraints on the amount they can produce during a particular period. For example, Ford Motor Company in North America can produce about 475 000 vehicles per month. In any particular month, Ford can only sell 475 000 vehicles plus whatever inventory it had on hand at the beginning of the month.
- Retailers and wholesales often sell goods that are not replenishable. These include fashion goods that are typically ordered once, or electronic goods that are near the end of their life cycle.
- Intrinsically scarce or unique items, such as beachfront property, flawless blue diamonds, van Gogh paintings, and Stradivarius violins, command premium prices because of their scarcity. In these cases, marginal cost is not an important determinant of price either because it is meaningless (van Gogh paintings for example) or extremely low relative to the scarcity rent that sellers can command (diamonds).

Example: for a certain company, the optimal unconstrained price equals €19. At this price the company faces a demand of 3000 units. When the supplier faces supply constraints, there are three basic options:

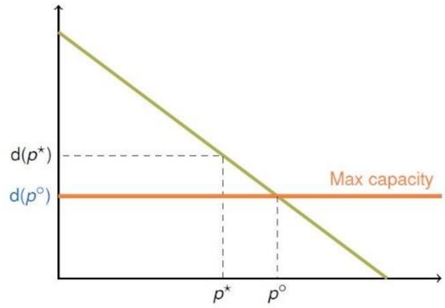
1. **Do nothing**. Keep the original price and let customers buy on a first-come-first-serve basis until supply is exhausted.
2. **Allocate the limited supply to favoured customer**. This is predictable from historical data: you know who and how many can enter the business.
3. **Raise the price until demand falls to meet supply**.

Alternatives two and three are not mutually exclusive: the seller could use a combination of allocations and price rises to manage the shortage. And, if he has segmented his market effectively, he could raise his average selling price by allocating most or all of the limited supply to higher-paying customers. Therefore, it is the basic idea behind revenue management.



Let b be the maximum supply available and p* the optimal unconstrained price. If d(p*) < b, the supplier doesn't need to do any further calculations: p* is also the optimal constrained price. In the first picture this is the case, the capacity limit is not constraining our optimal price, so we should not change our optimal price. When your capacity is not forcing you to change the optimal price, you can keep on selling at the optimal price.

If, on the other hand, d(p*) > b, then you need to charge a higher price in order to maximize contribution (=profit). This is the case in the second picture, the capacity constraint is important! You don't have enough capacity to sell at the optimal price! So you should increase the price to p^0. P^0 is the price at which you demand will be the same as you max capacity. It is the run-out price.



The profit maximising price under a supply constraint is equal to the maximum of the run-out price and the unconstrained profit-maximizing price. As a consequence, the profit-maximizing price under a supply constraint is always greater than or equal to the unconstrained profit-maximizing price.

Now let's calculate the formula of the runout price. It is based on the formula of the linear price-response curve: $d(p) = capacity = D - m * p \rightarrow capacity - D = -m * p \rightarrow p = \frac{capacity - D}{-m}$. Where p is the runout price, D is the intercept, and capacity(b) is the supply constraint/max capacity. And m is the slope.

Example of the supplier with: d(p) = 10 000 – 800*p, now with a max capacity of b = 2000:

$$runout\ price = p = \frac{(2000 - 10\ 000)}{-800} = 10$$

!! IN R GOED KIJKEN HOE JE DE M DEFINIEERD, DAAR HANGEN UW FORMULES VAN AF

<u>EXERCISE</u>

Consider a retailer selling SD-cards who faces a linear price-response function (8000 – 600*p) and an incremental cost of €2.

D = 8000

M = 600

C = 2

1. <u>Calculate the revenue-maximizing price, p*</u>

Revenue is: (8000 – 600*p)*p. Marginal revenue is: -600*p + 8000 – 600*p.

Revenue maximizing price is 0 = -600*p + 8000 – 600*p → **p = 6,66667 = pstar**

Or, another way: the revenue maximizing price is also $\frac{d}{2*m} = \frac{8000}{2*600} = 6{,}66667$

2. <u>Calculate demand at p*</u>

Demand = 8000 – 600*6,66667 → **demand = 3999,9998 = dpstar**

3. <u>Calculate revenue at p*</u>

Revenue = (8000 – 600*p)*p → **revenue = 26 666,67 = revpstar**

4. <u>Consider capacity constraints b = {100, 200, ... , 1000}</u>

In R: ccont <- seq(100, 1000, 100)

5. <u>Calculate optimal price, demand and revenue for every b</u>

Corresponding runout prices: runout <- (ccont – D)/-M

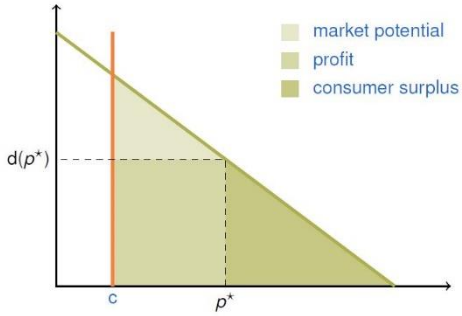
Revenue maximizing price under capacity constraint:

Pcirc <- ifelse(runout > pstar, runout, pstar) # if the runout price is larger than the revenue maximizing price than that means that the capacity constraint is interfering! That's why you have to put the larger price. If it is smaller than you can just keep the revenue maximizing price.
Corresponding demand: dpcirc <- D – M * pcirc
Corresponding revenue: revpcirc <- (pcirc – c) * dpcirc

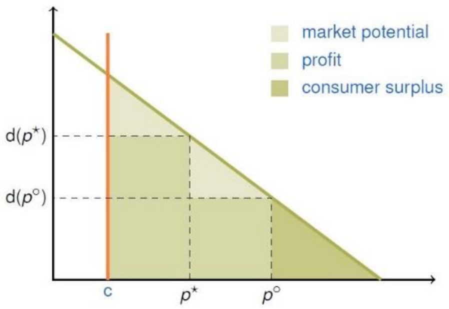
6. **Compare this with demand and revenue under p* and every b**

Price differentiation

Price differentiation refers to the practice of a seller charging different prices to different customers, either for exactly the same good or for slightly different versions of the same good. We use the term price differentiation to refer to the ways that additional profit can be extracted from a marketplace by charging different prices. Tactics for price differentiation include charging different prices to different customers, for exactly the same product, charging different prices for different versions of the same product, and combinations of the two.



The total potential opportunity for profit improvement from price differentiation is shown in the figure on the right. With a single price, the **profit** is the middle-light green square in the graph. The **consumer surplus** is the profit you lose because in that area are people who are willing to pay more than the price that is set, but they will buy it at the lower price. And the **lost market potential** are the people who are willing to pay more than the cost, but less than the price that is set. The sum of the three regions is the total contribution = profit that you would make if you were able to charge every potential customer exactly at his/her willingness to pay. While it is unrealistic to assume that this potential could ever be captured, the sheer magnitude of the potential gain means there is a powerful motivation for sellers to tailor different prices to different buyers according to their willingness to pay. If you would charge 2 prices, the surfaces of the areas change.



If price differentiation is such a powerful way for sellers to increase contribution, why don't all of them do it? The reason is that there are powerful real-world limits to price differentiation.

- **Imperfect segmentation**. The brain-scan technology required to determine the precise willingness to pay of each customer has not yet been developed. The best that can be done is to create market segments such that the average willingness to pay is different for each segment.
- **Cannibalization**. Under differential pricing, there is a powerful motivation for customers in high-price segments to find a way to pay the lower price.
- **Arbitrage**. Price differentials create a strong incentive for third-party arbitrageurs to find a way to buy the product at the low price, and resell to high-wtp customers below the market price, keeping the difference for themselves.

The presence of any one of these factors can eliminate the benefits of price differentiation.

Next we describe some of the most common and effective approaches to price differentiation used in different markets.

Group pricing

Group pricing is the tactic of offering different prices to different groups of customers for exactly the same product. The idea is to offer a lower price to customers with a low willingness to pay and a higher price to those with a high willingness to pay. In practice, "pure" group pricing requires determining

whether or not a prospective buyer belongs to a particular group and using that information to determine which price to charge. Examples are student discounts and senior citizen discounts.
Four criteria must hold for group pricing to be successful:

- **There must be an unambiguous indicator of group membership**. For example a student ID cart. Also it must be difficult or impossible for members of one group to masquerade as members of another. Otherwise, cannibalization could easily reduce or destroy the benefits of price differentiation.
- **Group membership must strongly correlate with price sensitivity**. Senior citizen discounts are predicated on the belief that senior citizens, as a whole, are more price sensitive than the public in general.
- **The product or service should not be easily traded or exchanged among purchasers**. This is necessary to avoid arbitrage, in which customers with access to low prices resell to customers who are quoted higher prices.
- **The segmentation must be both culturally and legally accepted**. While group pricing on the basis of age is broadly accepted, differentiating prices on the basis of other characteristics such as race and gender are controversial or illegal.

Pure group pricing is rare in direct consumer sales. It is most common in services.

Channel pricing

Channel pricing is the practice of selling the same product for different prices through different distribution channels. For example, Barnes and Noble sells books for different prices online than through its outlets.
There can be more than one reason why a seller might charge different prices through different channels. One is **cost**. For many companies, selling through the internet is cheaper then selling through traditional channels. However, it is also the case that customers arriving via different channels have **different price sensitivities**. For personal loans, it has been shown that customers inquiring through the internet are more price sensitive than those contacting a call center, who are in turn more price sensitive than those who apply for a loan at a retail branch. Furthermore, internet customers who access a consolidator website or use a shopping bot, tend to be more price sensitive than those who go directly to a bank's website. This is not surprising given the characteristics of the channels: it is generally easier and more convenient to shop and compare prices during a single internet session than by making many phone calls. Thus, **differential willingness to pay** is also a motivation for channel pricing.

Regional pricing

Regional pricing is an extremely common price differentiation technique. For example: in Latin America, McDonalds sells hamburgers for higher prices in wealthy neighbourhoods than in poorer ones. Or a glass of beer costs more at an airport bar than at the corner bar.
In each case, the price difference is based on **the supplier's desire to exploit differences in price sensitivity between locations**. After all, travellers at an airport are essentially a captive market and have few alternatives.

Couponing and self-selection

We have seen that group pricing is often both difficult and unpopular. Difficult because it requires the seller to categorize customers on the basis of price sensitivity before quoting them a price, and unpopular because it often seems 'unfair' to customers. It is often much more convenient to differentiate prices in ways that allow customers to self-select**. In a self-selection approach, both the**

list price and a discounted price are available to all customers, but it takes additional time, effort or flexibility to obtain the discounted price. The idea is that those willing to make the additional effort to get the discount are generally more price sensitive than those who are not. For example, retailers commonly offer discount coupons through newspapers, direct mail and magazines. Or movie theaters that charge lower prices for a weekday matinee than for a Saturday night show.
The common thread among these examples is that the seller has chosen a mechanism that allows customers to self-select, depending on the value they place on time or flexibility. Any customer can obtain an item at a discount if she is willing to take some additional effort.

Product versioning

When pure group pricing is not feasible, companies use other strategies to differentiate prices. The most notable of these is designing or developing products (either virtual or real) that may have only minor differences but enable the seller to exploit differences in price sensitivity among customer segments. This can involve developing an 'inferior' variant and/or a 'superior' variant of an existing product. Examples are:

- **Inferior goods.** Consider the following case. brand-name vegetable canners sell their products under their own brand but also sell to retailers who sell the product to consumers as a "house brand" or a generic brand. The motivation of the seller: a desire to sell a product cheaply to customers with lower willingnesses to pay without cannibalizing sales of the full-price product. This is achieved by creating an inferior version of the "standard product".
- **Damaged goods.** This is a term coined to refer to the situation in which a manufacturer or supplier creates an inferior good by damaging, degrading, or disabling a standard good. Since this process starts with the standard good, the supplier is actually paying more to create the inferior good it will sell at a lower price. Complex application software packages such as supply-chain software or enterprise resource planning software are often sold at different prices, depending on the number of features purchased by a customer: the more features purchased, the more expensive the license. In many cases, the software is configured for a particular customer by starting with the complete package and then disabling the features that the customer did not purchase.

  The concept of "damaging" a good in order to create an inferior good to be sold at a lower price may initially seem somewhat bizarre. However, it is really only a special case of the more general category of "inferior" goods.
- **Superior goods.** Spendrups is the largest brewery in Sweden. Traditionally Spendrups brewed medium or low priced lagers aimed at the mass market. In the 1980s, they created Spendrups Old Gold, which they advertised as a premium beer and sold in a special highly distinctive bottle. Although Old Gold did not generally fare better than Spendrups' other brands in comparative taste tests, Spendrups was able to establish Old Gold as a premium brand and maintain a price 25% to 50% higher than that of its other brands. This is the obvious complement to the inferior-good strategy: creating a superior good in order to extract a higher price from less price-sensitive customers.

  In some ways, a superior good strategy is safer than an inferior-good strategy because it does not threaten cannibalization of existing sales. Of course, it presumes an ability to create and establish a product that the market perceives as truly superior to the existing product and that there is a customer segment willing to pay a premium for the superior product.
- **Product lines.** Establishing a product line is the natural extension of creating inferior or superior products. A product line is a series of similar products serving the same general market but sold at different prices. For our purposes, we will consider vertical product lines,

where almost all customers would agree that a higher-priced product is superior to a lower-priced one. This applies, for example, to a hotel that charges more for an ocean-view room than a parking-lot-view room: almost all customers would prefer the ocean view to the parking lot view. It also applies to personal computers offered by Dell, where each product in the line as higher performance than the product just below it in the line.

This can be contrasted to horizontal product lines, where different customers would prefer different products within the line, even at the same price. Coca-Cola offers a horizontal profit line with "Classic" Coke, Diet Coke, Cherry Coke, Diet Cherry Coke, etc. This is a horizontal product line because no one of the products is unambiguously higher quality or more desirable than another.

An important advantage of pricing differentiation by establishing a product line is that consumers perceive it as fair. Consumers get to choose among the alternatives, and the concept of "paying more to get more" is widely accepted. This makes product-line pricing more acceptable than group pricing in most customers' minds.

Time-based differentiation

Time-based differentiation is a very common form of product versioning. For example, Amazon offers 5- to 9-day "Super-saver" shipping free while charging $3.97 for "standard shipping". Fashion goods cost more during the beginning of the season and are marked down toward the end of the season. Passenger airlines offer discount rates to customers who book a week or more prior to departure. In each of these cases, **companies have created differentiated products that allow customers to self-select**. In the case of Amazon, customers who are willing to wait for delivery can pay less. For passenger airlines, customers who have the flexibility to book earlier can pay less. Of course, it may be that the higher price charged by Amazon for early delivery exactly matches the incremental cost. But it is highly likely that Amazon is also using time of delivery as a segmentation variable, relying on the fact that some of their customers will willingly pay a premium in order to have the product in their hands sooner.

Time-based differentiation plays a very important role at passenger airlines, hotels, and rental car companies, in which time of booking and other factors, such as willingness to accept a Saturday night stayover, are used as indicators of whether or not a potential customer is traveling for leisure purposes or business purposes.

Product versioning or group pricing?

There is no clear line separating the two approaches and many price differentiation strategies contain elements of both. For example, consider the classic airline example of a roundtrip ticket from San Francisco to Chicago costing $250 if purchased a week in advance and including a Saturday stay-over versus $750 if purchased at the last minute without restrictions. Is this group pricing or product versioning? Disgruntled customers might argue that it is simply group pricing, since different customers are paying different amounts for exactly the same service, namely, a roundtrip coach seat San Francisco – New York. This is the "Why am I paying $500 more than the person sitting beside me for exactly the same flight?" objection. The airline would reply that the two types of tickets are distinct products and that the added cost of the full-fare ticket is fully justified by the flexibility of being able to purchase late and return without staying over a Saturday night.

The reality is, of course, that airline pricing, like many successful examples of price differentiation, includes elements of both group pricing and product versioning. The airlines consciously created restricted discount fares as an inferior product. They did so, however, as a way to enable them to offer different fares to different customer groups: lower fares to leisure travellers, who are more price

sensitive but more flexible, and higher fares to business travellers, who are less price sensitive but less flexible. Viewed one way, we could say that the airlines created an inferior product as the most efficient and leas controversial way to institute group pricing.
The airline example illustrates a very important point. Pure group pricing is very difficult to pull off in consumer markets. There are few cases in which consumers can unambiguously be identified as belonging to a particular group. Airlines have no reliable objective marker to tell them whether a particular customer is flying for business or for pleasure. In the absence of such a marker, they rely on the very imperfect criterion of whether or not a customer can book early. This works well enough, but it is imperfect. For example, there are plenty of highly price-sensitive leisure customers who would love to book late. Furthermore, the airlines have lots of empty seats they would like to fill with these customers even at a very low price. But, at least until recently, there has been no systematic way to sell to these customers without cannibalizing the full-fare business customers.

<u>EXAMPLE: THEME PARK</u>

Consider a hypothetical theme park that can serve up to 1000 customers per day. The theme park charges a single admission price, and all rides are free after admission. During the summer, demand follows a stable and predictable pattern, with higher demand on weekends than during weekdays. We assume that demands for different days of the week are independent, that demand curves are linear, with intercepts, slopes, and satiating prices as shown below, and that the theme park has a marginal cost of zero per customer.

1. What price should the theme park charge without variable pricing?
2. What price should the theme park charge with variable pricing?

Day of week	Intercept (D_i)	Slope (β_i)
Monday	1500	-50
Tuesday	1400	-40
Wednesday	1510	-42
Thursday	2000	-52.6
Friday	2500	-55.6
Saturday	3300	-60
Sunday	3100	-62

Day of week	Intercept (D_i)	Slope (β_i)	p	d(p)	m(p)
Monday	1500	-50	15	749	11250
Tuesday	1400	-40	18	700	12250
Wednesday	1510	-42	18	755	13572
Thursday	2000	-52.6	19	994	19011
Friday	2500	-55.6	26	1000	26978
Saturday	3300	-60	38	1000	38333
Sunday	3100	-62	34	1000	33871
				TOTAL	155265

Day of week	Intercept (D_i)	Slope (β_i)	p	d(p)	m(p)
Monday	1500	-50	25	226	5751
Tuesday	1400	-40	25	381	9698
Wednesday	1510	-42	25	440	11202
Thursday	2000	-52.6	25	659	16805
Friday	2500	-55.6	25	1000	25488
Saturday	3300	-60	25	1000	25488
Sunday	3100	-62	25	1000	25488
				TOTAL	119919

Picture on the right is without variable prices, picture on the left is with variable prices.
Intercept = (1500, 1400, 1510, 2000, 2500, 3300, 3100)
Slope = (-50, -40, -42, -52.6, -55.6, -60, -62) (we included the minus sign in the slope variable!!)
Now step by step!
First set your price of interest: price <- 15
Then calculate demand at that price: dp <- mean(intercept) + mean(slope)*price → dp = 1411
Now don’t forget the capacity constraint: dp <- ifelse(dp > 1000, 1000, dp) → dp = 1000
Now we calculate the revenue: dp * price → revenue = 15 000
Find the maximum of average price response curve
Mean(intercept) = 2187,143
Mean(slope) = -51,74286
Now we want the optimal price: $\frac{-mean(intercept)}{2*mean(slope)}$ → p* = 21,13473

Don't forget the capacity constraint: $\frac{1000-mean(intercept}{mean(slope)} \rightarrow\ p^0 = 22{,}94313$

Now we want the max of the two prices: $\max(\frac{-mean(intercept)}{2*mean(slope)}, \frac{1000-mean(intercept)}{mean(slope)})$ → which gives us $p^0 = 22{,}94313$.

But this analysis is simplified to much. A more realistic version would be the following.
First you create a function that calculates revenue:

```
Rev1 <- function(price, intercept, slope) {
        Dp <- intercept + slope*price
        Dp <- ifelse(dp > 1000, 1000, dp)
        Return(sum(dp*price))}
```

Now you test the function

```
Rev1(price = 15, intercept = intercept, slope = slope)
```

Next you use optimize function to find the maximum of average price-response curve

```
optimize(f=rev1,interval=c(0,1000),intercept=mean(intercept),slope=mean(slope),maximum=TRUE)
```

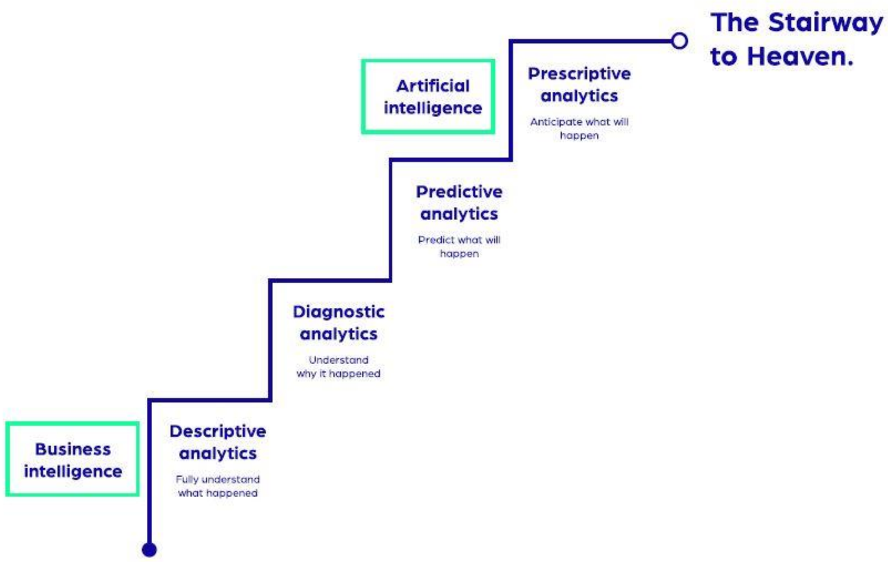
Then you find the optimal single price, given the different price-response curves

```
# Find optimal single price, given the different price-response curves
optimize(f=rev1,interval=c(0,1000),intercept=intercept,slope=slope,maximum=TRUE)

# Calculate optimal prices for each price response curve
# The optimize function does not allow for multi-parameter optimization (at least not in this case)
library("alabama")
# again start with creating the function to optimize
rev2 <- function(price,intercept,slope){
        dp <- intercept + slope*price
        return(-sum(dp*price))
}
# Set the inequality constraints
# I.e., demand should not exceed 1000
# (every element of the vector created by hin should be > 0)
hin <- function(price,intercept,slope){
        1000.1-(intercept + slope*price)
}
out <- auglag(par=rep(1,7),fn=rev2,hin=hin,intercept=intercept,slope=slope)
out$par
out$val

# Same but now analytically
(-intercept/slope)/2 # optimal price
(1000-intercept)/slope # runout price
apply(cbind((-intercept/slope)/2,(1000-intercept)/slope),FUN=max,MARGIN=1) # max of the two
```

Crunch guest lecture

The crunch team exist of technical experience (data scientists, data engineers), training experience (academy lead, teachers), strategy experience (retail lead, project managers, data strategists) and business (managers, business development). Crunch helps organisations with their pricing, based on data.

Crunch helps companies evolve through **the stairway to heaven**:



Business intelligence: they empower people to make better decisions.
Artificial intelligence: they build algorithms that automatically make optimal decisions.
Now why should organizations create impact with data? To make better business decisions and because monetizing data in essence means changing your product. (for example the automated trend spotter for ZEB).

Use cases

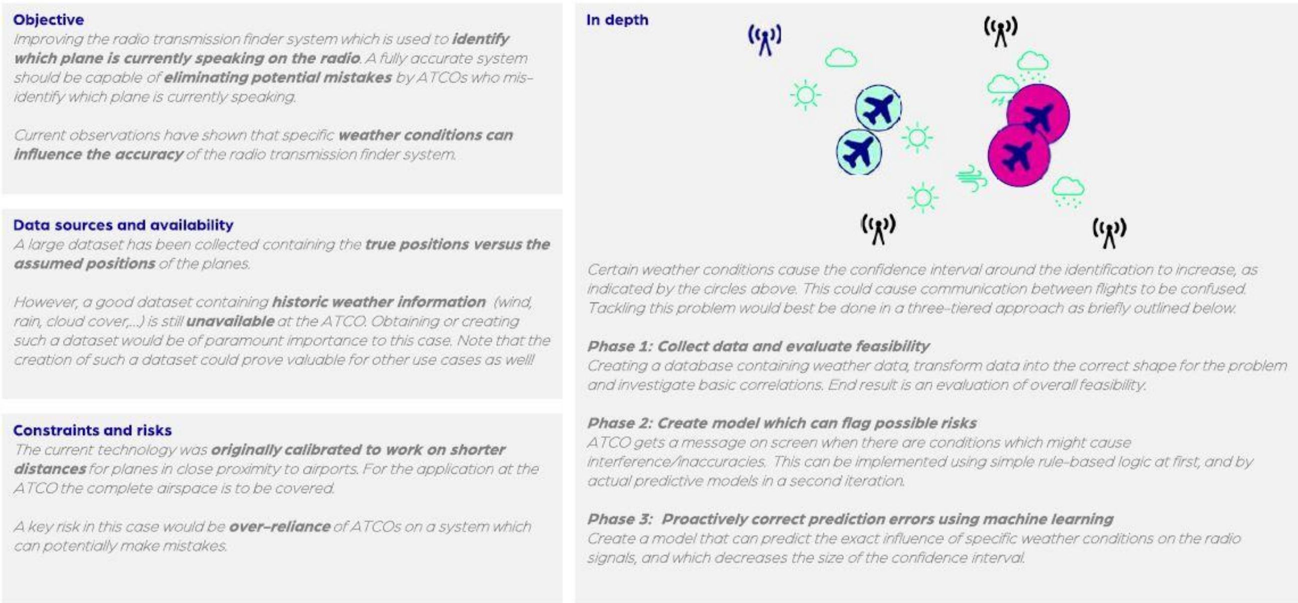
Forecasting and stock redistribution. Manage your inventory automatically and reduce out of stock events, while minimizing overstock. This can be achieved by:

- Continuously track inventory movements
- Monitor over-or underselling products
- Proactively shift stocks between shops
- Adjust algorithm for specific requirements
- Link to forecast, pricing and marketing system

A solution like this implemented at an omnichannel shoe retailer, reduced process time from days to minutes while driving both sales and decreasing leftover stock.

Applicance detection. Smart home applications are mainly centered around home automation and energy conservation. Many of these applications rely on knowing which appliances are connected to specific outlets. Crunch has created algorithms which are capable of classifying which device is plugged into a specific outlet based on its power usage statistics; in large part using only low frequency data and only using high frequency samples for “confusing” devices.

Detecting possible interference of Radio Direction Finders. Radio direction finders allow air traffic control to detect who (which plane) is actually talking on the radio. However, this device can be affected by weather conditions, which can of course lead to dangerous situations.
Phase 1: what types of weather are causing problems?
Phase 2: warn ATCO of possible inconsistencies.
Phase 3: proactively correct the signal.



An overview of pricing challenges

Their global viewpoint of pricing problems, specifically for retailers in today’s economic environment. First, a key distinction between pricing problems.

Level 1: strategic pricing. What is the right price for a given product at a given time. The objective of this first level is **to determine the standard price point which maximizes the total gross margin which can be earned from the product**. Two prerequisites to this analysis are a demand function which can correctly estimate product price elasticity, and a correct calculation of the GMDF for the product.
Predictive analytics models are not very useful in predicting price-response functions often due to a shortage in data. Solution: cluster data to create demand function. Pricing games are played between competitors that offer similar 1-to-1 products using web crawling.

Level 2: tactical pricing responses. Too much stock/running out of stock. Keeping the theoretically optimal price fixed is not always desirable. Specifically there are two scenarios in which short term deviations are desirable:

- **Suboptimal inventory positions**: when inventory is unexpectedly higher or lower due to variations in demand or supply side, it can be desirable to adjust prices to increase or decrease demand accordingly.
- **Competitor price changes**: maintaining a specific pricing position is of obvious strategic importance. Th speed ad magnitude of reaction however is a tactical decision. A typical tit-for-tat strategy might be optimal from a market share perspective, but it might result in an undesirable inventory position.

What is theoretically the best price I can set? Colruyt versus Delhaize: colruyt's value proposition is about at all times delivering the lowest price to customers by keeping costs very low. In contrast, Delhaize coming from a premium segment in the past, could have troubles in competing at low prices, therefore cutting their margins immensely.

Case 1: markdown management

<u>Decreasing the prices of articles to clear the inventory of seasonal and/or discontinued products.</u>
Why do we do this? Well the value of product has decreased over time, and because the physical space of the product is valuable. This aspect of retail has become more important as product life cycles have become shorter! Physical space is costing money for retailers. Life cycles have become shorter and shorter, which you see in practically every segment of retailing. Think for example about the phone market, each year there's a new phone coming out!
Business rules. Price ladders = always have for example reductions at 10%, 20% and 30%. Minimal price reduction, maximum price reduction, maximal sell-through at exit date (outlet shopping. This offers a way to access customers that do not care about the atmosphere while shopping, or where they shop, but are only interested in paying the lowest price for a certain product. → The objective of the retailer shifts from maximizing gross margin to maximizing turnover, the current inventory is to be regarded as a "sunk cost".
This problem is solved using two key building blocks. First of all the <u>demand model</u>, that predicts what the likely outcome of a specific markdown will be. There is no demand model that works for every company. Each company has specific drivers, specific constraints that delivers a unique demand model. Key factors influencing a demand model:

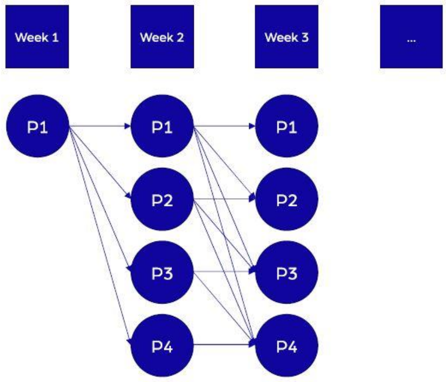
$$D_w = PLC_w \cdot PD_w \cdot INVEN_w \cdot SEAS_w \cdot PROMO_w$$

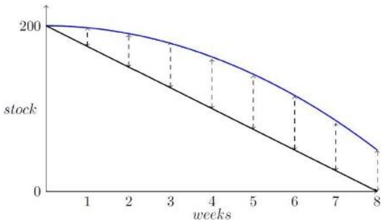
- PLC: Product life cycle effect for a product
- PD: Price discount
- INVEN: Remaining inventory levels
- SEAS: Specific weeks, special events (Holiday period,...)
- PROMO: Advertised promotions (Black Friday,...)
- ...

In reality:

1. **Focus on factors which are most important for the specific client**
2. **Pool items which are similar to get a decent number of observations to fit a function**

The second building block is <u>the markdown optimization model</u> that, given the response function, choses the best possible markdown to optimize the objective (total turnover), given business constraints.

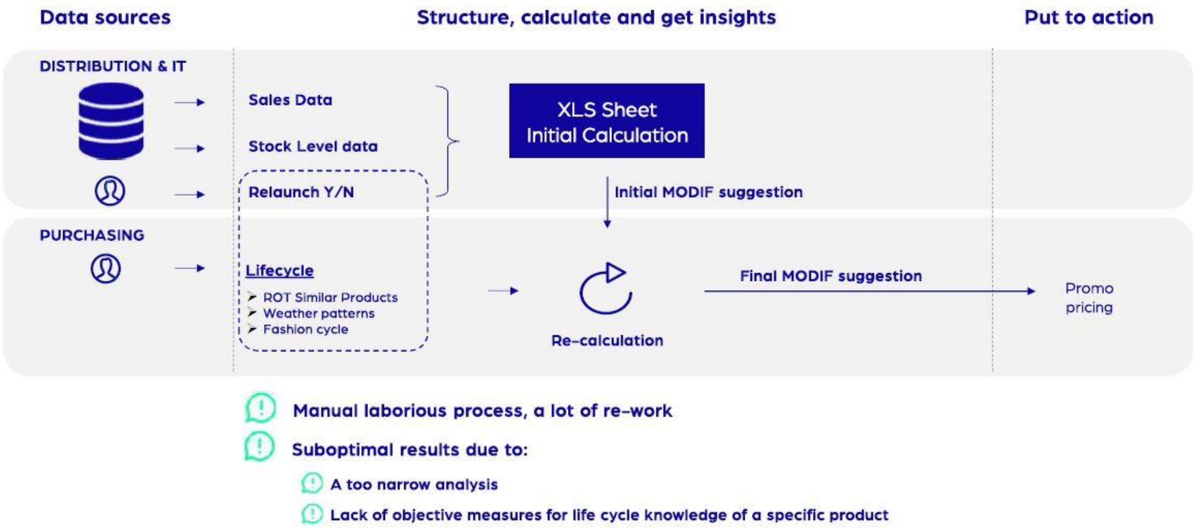




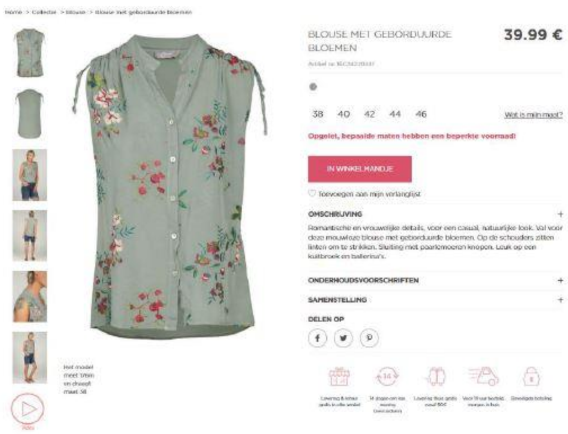
The picture on the right is the solution using dynamic programming, while the other picture (with the graph, is a solution using approximation of ideal promo path.

CASE 1: Casis Paprika

Example process currently in place at a client



Initial MODIF suggestion based on data (MODIF = product modification information)



REF_COL: 1TA91111115

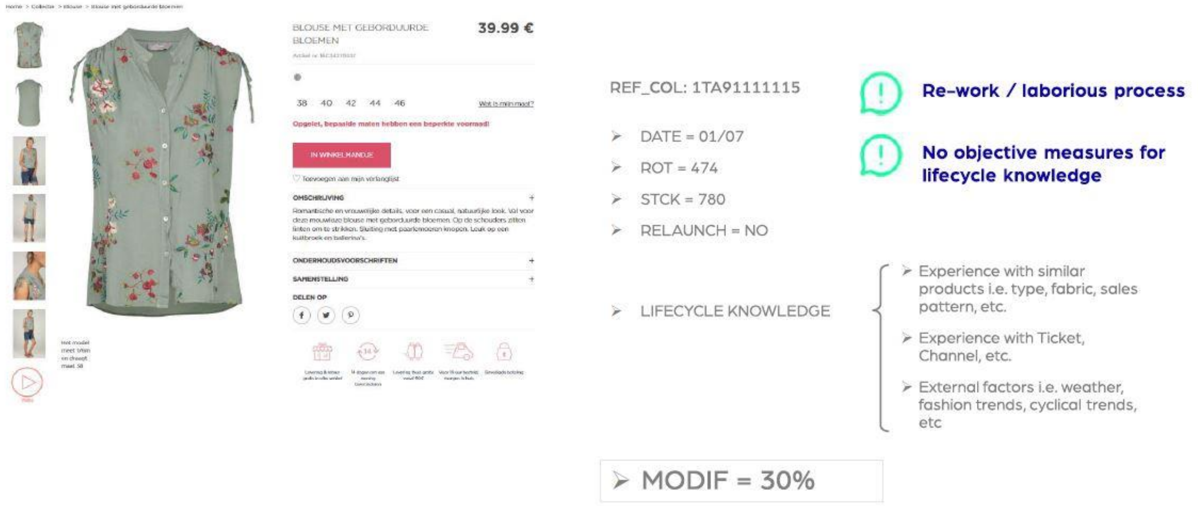
- DATE = 01/07
- ROT = 474
- STCK = 780
- RELAUNCH = NO

Too narrow analysis

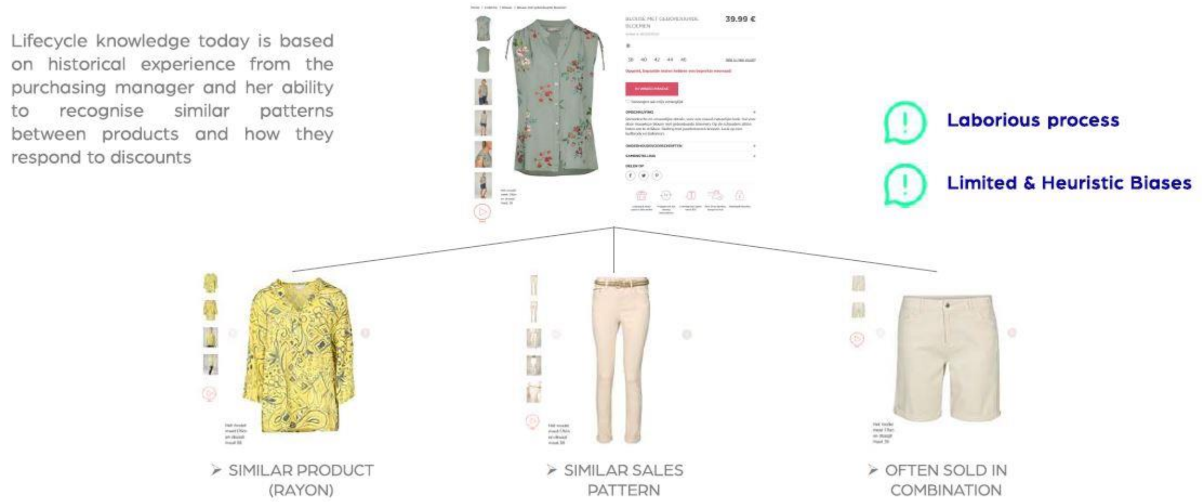
No product specific knowledge

- MODIF = 40%

Refining MODIF based on lifecycle knowledge



Manual linking is based on experience



There are lots of excel sheets for promo pricing, which also means lots of re-work and laborious process. Solution: algorithm that performs clustering based on data (grouping very similar clothing pieces). It is based on historical data, meaning that how customers have reacted to reductions on certain clothing pieces is incorporated.

The algorithm performs a deeper clustering based on data

All historical data is taken into account

Another algorithm identifies the actions that have been taken (discount prices) and the results those generated (sales) for the clustered products.

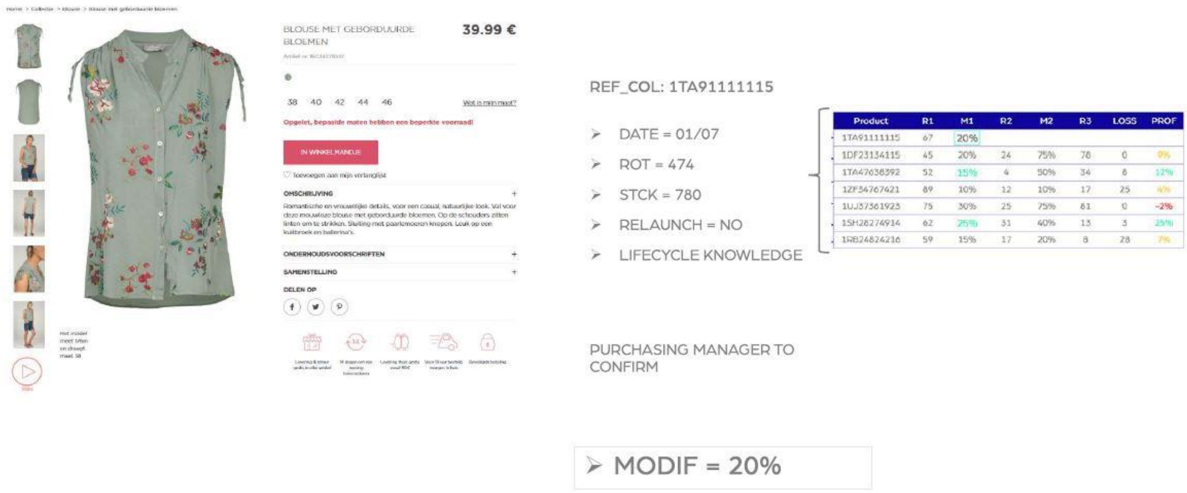
Product	R1	M1	R2	M2	R3	LOSS
T-shirt	67	?				
Blousse	45	20%	24	75%	78	0
Sunglasses	52	15%	4	50%	34	8
Dress	89	10%	12	10%	17	25
Blousse	75	30%	25	75%	81	0
Trousers	62	20%	31	40%	13	3
Blousse	59	15%	17	20%	8	28

Automatic MODIF suggested by the model

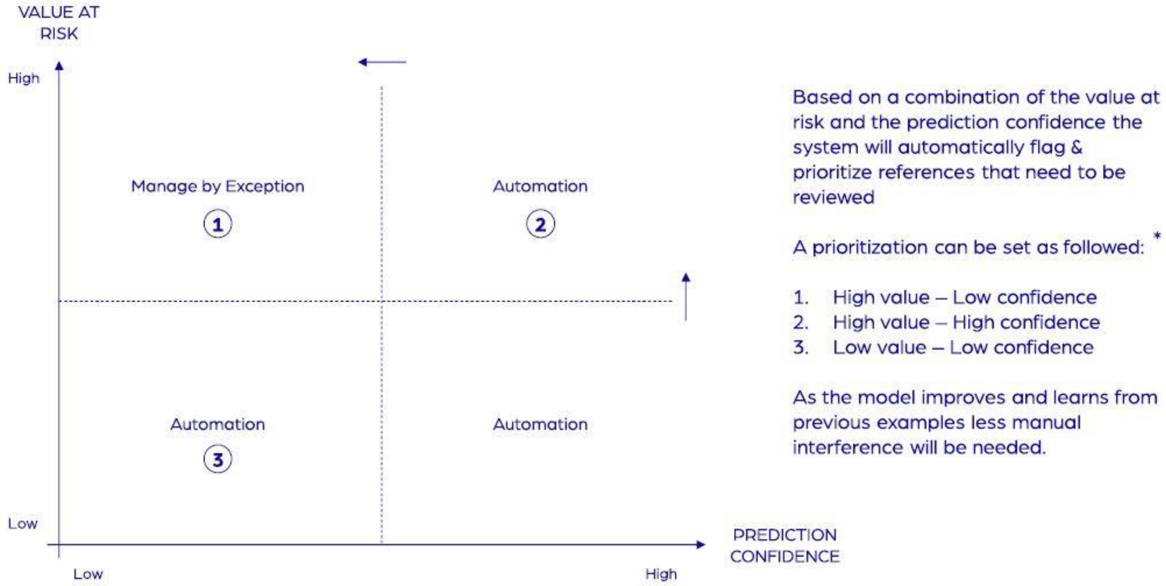
Another algorithm identifies the actions that have been taken (discount prices) and the results those generated (sales) for the clustered products.

Product	R1	M1	R2	M2	R3	LOSS	PROF
T-shirt	67	20%					
Blousse	45	20%	24	75%	78	0	9%
Sunglasses	52	15%	4	50%	34	8	12%
Dress	89	10%	12	10%	17	25	4%
Blousse	75	30%	25	75%	81	0	-2%
Trousers	62	25%	31	40%	13	3	25%
Blousse	59	15%	17	20%	8	28	7%

A MODIF suggestion based on a broad range of data

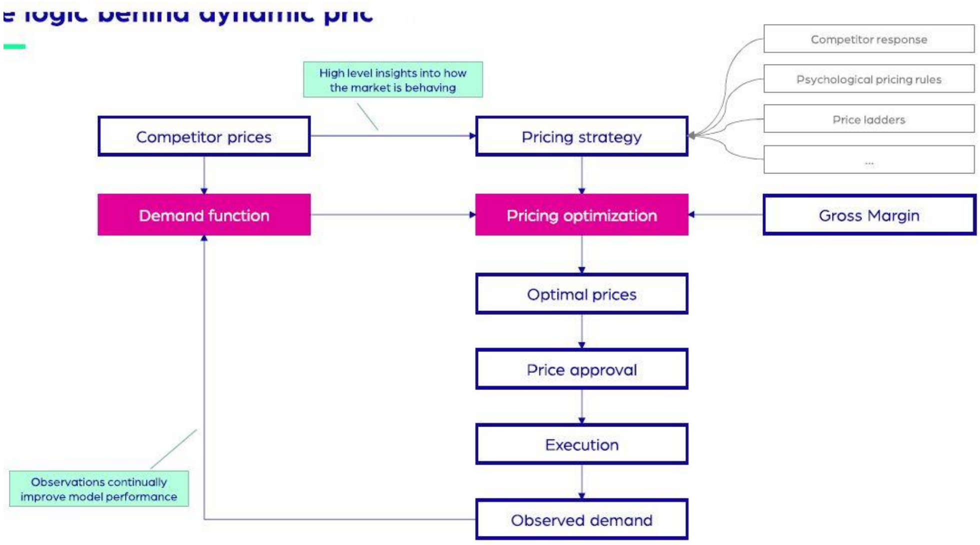


Manage by exception



The y-axis is the value as risk is high: products of which there are a lot left and are expensive. When you know the predicted demand function is not very accurate for products of which the value at risk is very high, the manage by exception quadrant allows for overruling the model. If there is an underlying reason for overruling the model, the reason why could be taken into account which therefore improves the model (and prevents future overruling).

Case 2: dynamic pricing



When you have demand, observe it. When you don't, use the nearest-neighbour method. For example, if you have no actual demand for a beer, use similar beers with for instance the same bitterness. Eventually you will end up with a reference price in the market. GMDF = Gross Margin after Delivery and Fulfillment (due to high delivery/shipping costs).

Defining pricing constraints

When you are defining the pricing constraints, get the minimal and maximal reference price in the market, so you end up with a range from which you can set the price. But the question now is, what range do we have left? Or how much of the range did we use so far?

$$\frac{€1{,}79 - €1{,}75}{€2{,}59 - €1{,}75} = 4{,}76\%$$

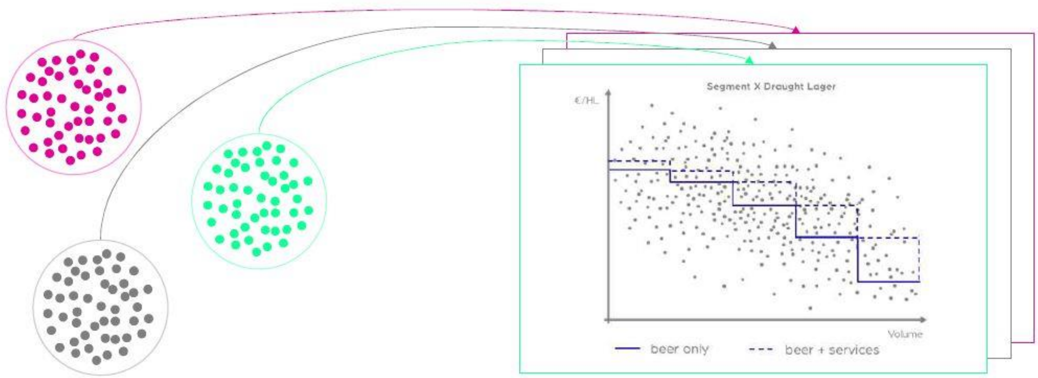
€ 2,09 was the suggested price level in the MPV solution.

Infeasibilities surface quite often, as there are often lots of constraints that need to be taken into account. There are hard and soft constraints, the hard ones can never ever be violated, and the soft ones are more flexible. Constraints at a lower level always overrule constraints on a higher level (and are generic). Deliver those results from the process above through dashboarding views (make dashboards interactive!). dashboards display what's happening in the market in terms of price changes of the company itself or its competitors. Prices change often overnight at for instance 3am when there is lowest customer traffic, hence we prevent the chance of customers noticing price changes in their basket.

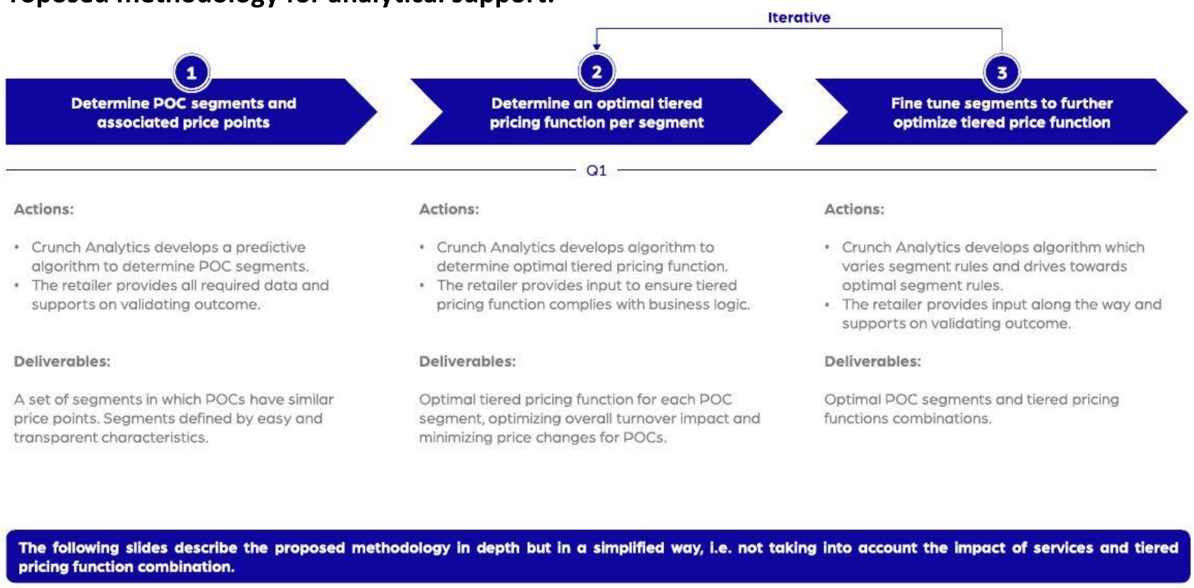
Case 3: transparent pricing

Project objective: determine consistent and transparent rules for setting prices for various point of consumption (POC) segments. Consistent: there are no big discrepancies between POCs which are similar on relevant dimensions. Transparent: the model is explainable and in line with intuitive understanding of the market. This allows to increase trust from POC owners and gain market share.

Consistent and transparent tiered pricing functions, determined for different identified POC segments. These tiered price functions are optimized for turnover and for a minimal required amount of price adjustments.



Proposed methodology for analytical support:



This is les conventional:

1. Determine POC segments and associated price points. Identify segments. K-means clustering techniques (cluster groups together based on the price customers are currently paying).

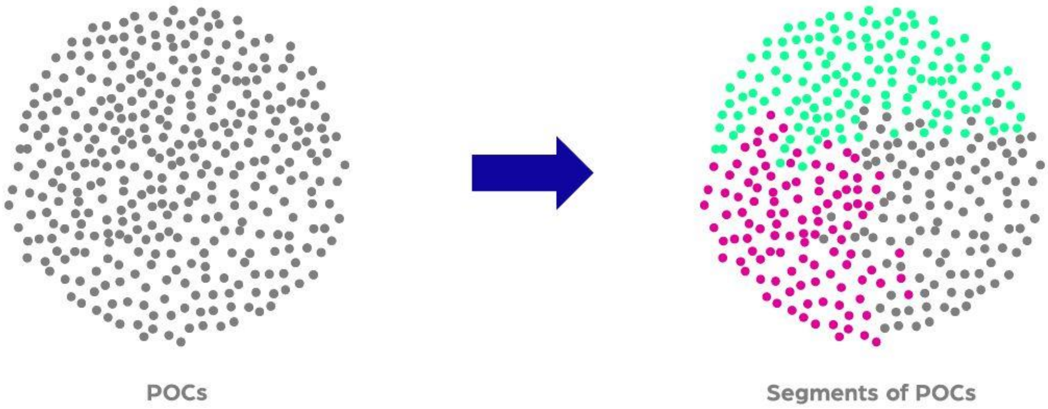
    Rule fit: Delivers rather decision tries

    Random forests: delivers variable importances

Objective of clustering:

Find the best understandable narrative for current price difference between different POCs.

The logic itself is constructed from simple variables (e.g. location, capacity,...) – the algorithm determines which combination of these variables results in the best possible trade-off between complexity and amount of price variance explained.



POCs can be segmented by using classic clustering algorithms or interpretable prediction models. The latter is undoubtly the preffered technique in this specific case.

Classic clustering algorithms

Description

We want to identify patterns and compare each POC to comparable counterparts to define the right price. The explanation can be limited to "similar POCs have a higher price, so your price should also be higher"

Advantages

- Mathematically simple

Disadvantages

- Clusters cannot always be explained
- Interactions between dimensions cannot be used
- Relative importance of dimensions must be quantified
- Does not work well with categorical dimensions (e.g. cities)

Best fit for this project

Interpretable prediction model

Description

A prediction model is created which tries to predict the current price level of current POCs, based on the properties which are to be used to motivate the differential prices to POCs. Because of the types of models which are used, a meaningful clustering of POCs can be created.

Advantages

- Explainable model
- Controllable dimensionality (e.g. maximal # rules)
- Interaction between dimensions are possible
- Importance and interpretation of dimensions is not required to be known beforehand

Disadvantages

- Assumes that the "average" price is correct for various POCs, and that systematic variations of prices are justified
- More dependent on amount of data available

The proposed approach starts with the creation of a predictive model, the dependent variable of which is the current price paid by a client. The independent variables (= information used to make the prediction) are descriptive properties of the client (denoted by the x variables), as well as the volume of product sold to the client (V).

We want to explain the current price

$$P = \hat{f}(x_1, x_2, ..., x_n, V)$$

Based on various properties of the POC

It is of great importance that the volume sold is also included as an explanatory variable, since the customers currently receive differential pricing based on the volume they purchase. As such, it will be important to remove any effects which are due to the volume sold from the interpretation of the predictive model.

The model used to make these predictions must be an interpretable model, the simplest example of which is a simple linear regression. However, in this case we believe that it would be worthwhile to experiment with other models as well, one example is the RuleFit model, which creates a model from a small set of decision trees - which are then aggregated to make a single prediction. This is of special interest for this case since in this situation this would allow for interpretability as well as interaction effects between terms. Moreover, the number of trees used and the complexity thereof can be controlled to result in a limited number of clusters at the end of this analysis.

How RuleFit works

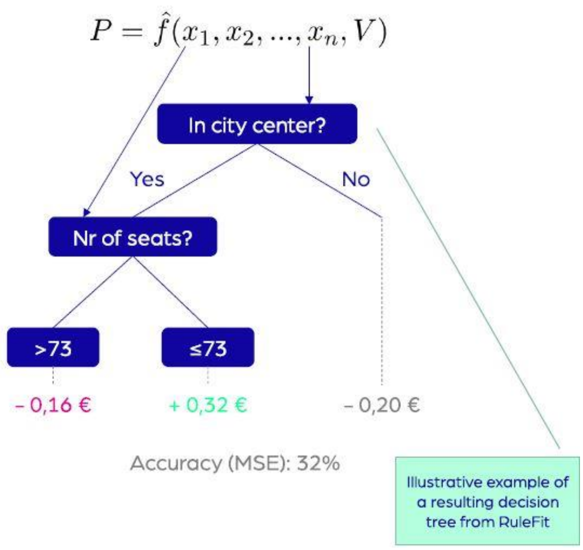
The outcome of RuleFit is an understandable decision tree that explains the variance in a variable

The algorithm is presented with:
- **The current price paid by a POC**
- **A wide range of possible explanatory variables**

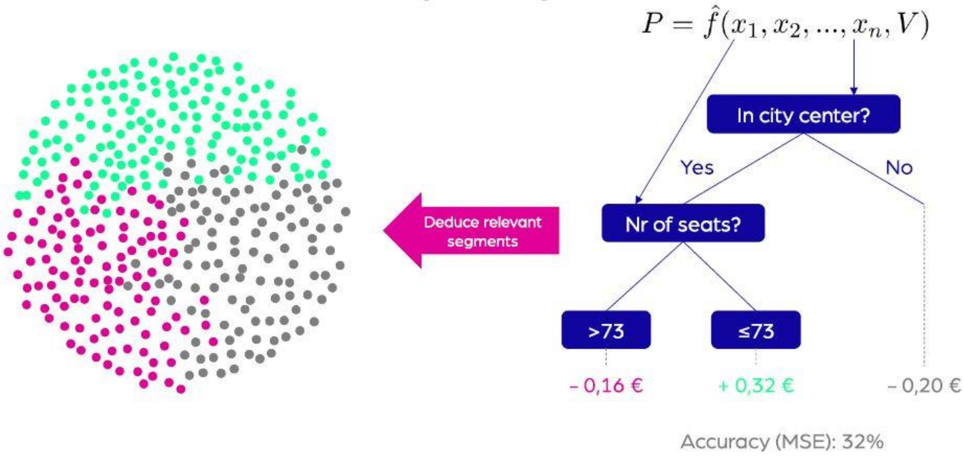
Parameters are set to determine the maximal complexity of the model:
- **Maximal number of variables used to explain the result**
- **Maximal number of decision trees used**
- **...**

Experimentation with these variables yields the right balance between model complexity and goodness of fit (i.e. how consistent the price is within a segment).

The algorithm calculates the best possible interrelationship and cutoff point between variables to explain variance.



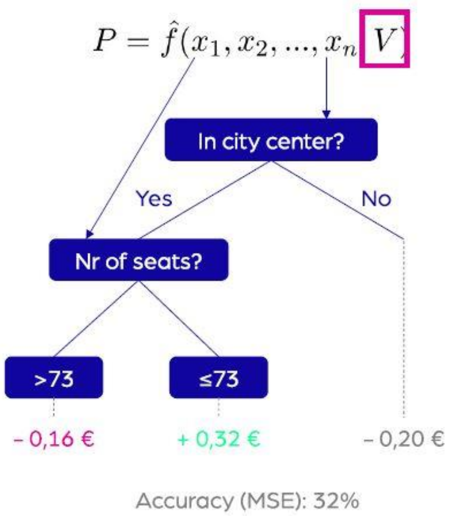
Decision trees are not used to predict price, but the logic of the decision trees is used to create segments, in this situation there are three leaf nodes in the decision tree – resulting in three segments.

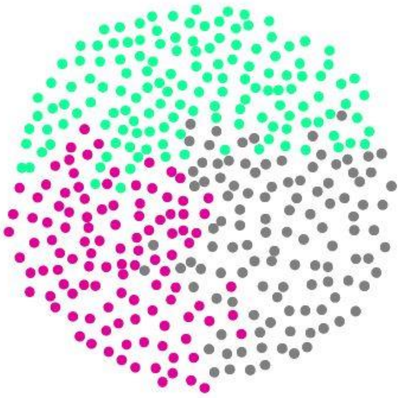


The volume purchased by a POC is included as an independent variable, since this is an important factor in the price a POC currently pays.

However, when this variable is used as a node in the decision tree, this node is ignored – since the price discrimination based on volume is handled by the step function which is introduced in the ***second*** step.

It is however highly important to include this as a variable in the predictive model since otherwise this would be a very prominent confounding variable; effectively this could result in a model which is in a large part predicting the volume sold to a POC, rather than the relative price point charged to a POC.



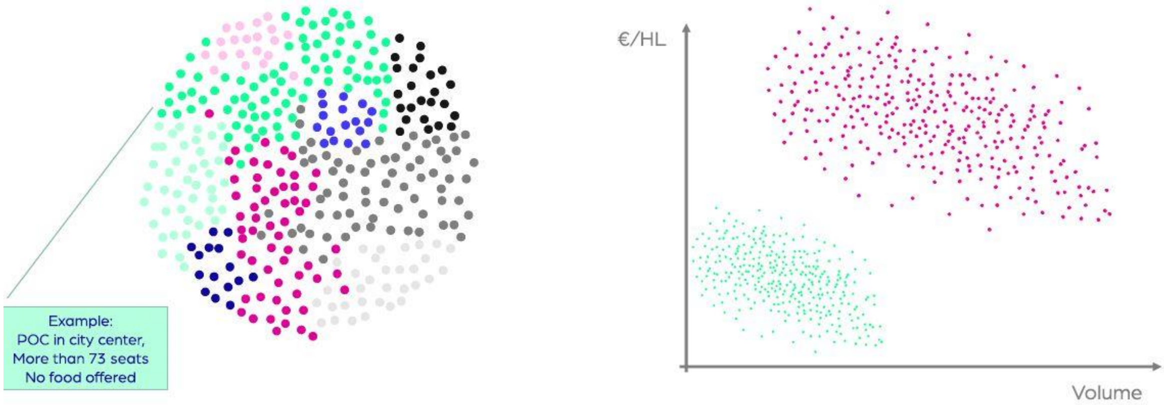


The use of White-Box algorithms such as RuleFit allows us to leave the selection of the right price discrimination criteria to an algorithmic procedure, removing the effect of historic bias towards specific variables and overall resulting in a more optimal pricing strategy.

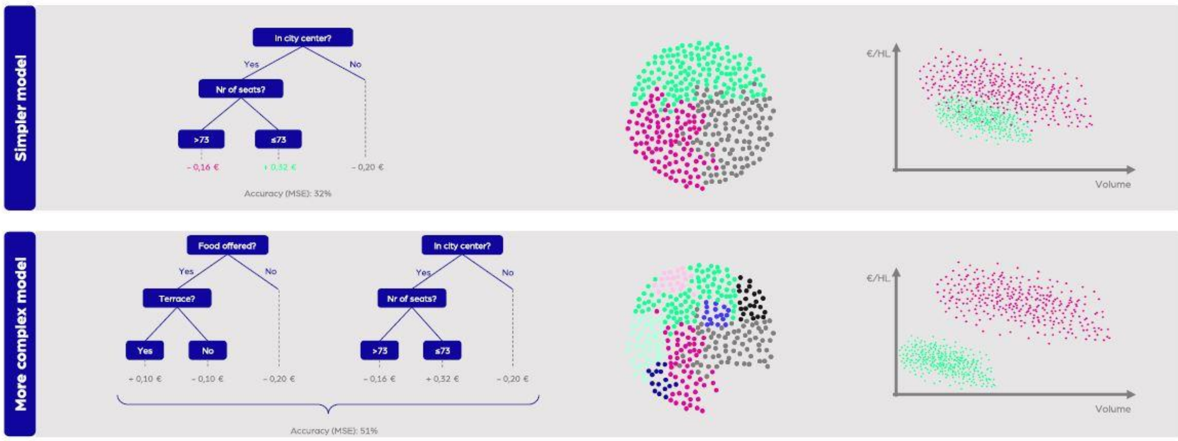
This is opposed to "traditional" segmentation where clusters contain similar POCs, but the similarities might be irrelevant for the pricing offered to these POCs. (e.g. the fact that two POCs have yellow facades might render them "similar" but does not mean anything for the price they pay).

The result of the first phase serves as a solid starting point for the further optimization of the pricing strategy, the various segments having:

1. A structurally different price relation between volume and price
2. An understandable motivation for being different segments

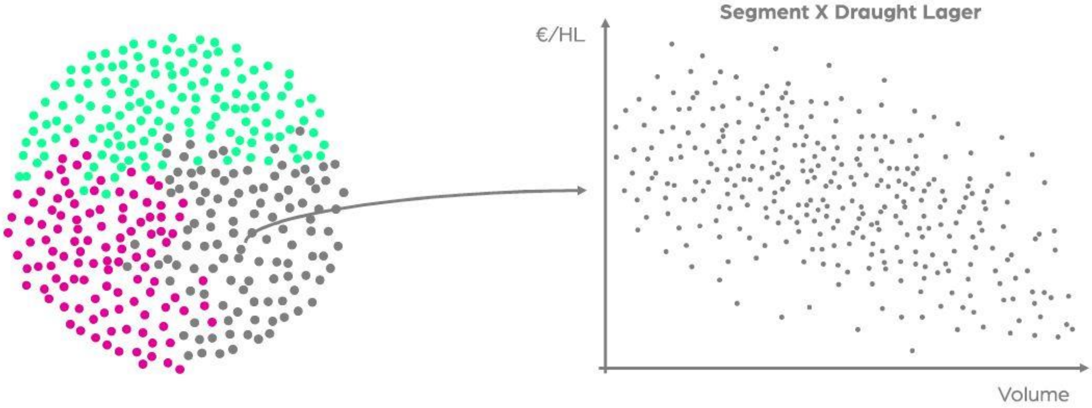


The core challenge in tuning this algorithm is finding the right balance between model complexity and explained price variance. Allowing the model to use more variables and more complex variables will result in more consistent groups – but these groups will inevitably be higher in number.



2. Determine an optimal tiered pricing function per segment

Now that we determined a set of POC segments, it is time to reintroduce the "Volume" variable. This provides us an understanding of the relationship between volume and pricing within each of the determined segments. In this step of the process we will build an algorithm that maps volume / price tiers in such a way that revenue per segment remains equal and the difference between current POC price points and new POC price points is minimized.



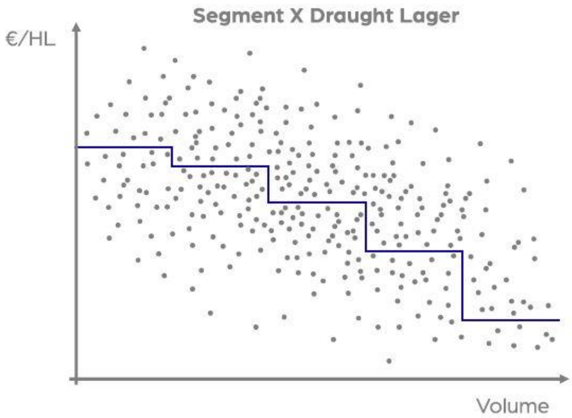
Mathematically finding the optimal volume / price tiers, means fitting a "piecewise linear function" taking into account following objectives:

1. Avoiding impact on overall turnover
2. Minimizing the required price changes per POC when adopting the new pricing model

In mathematical (simplified) formulas, this comes down to:

1. min $|\sum_c(P_c * V_c) - \sum_c(P'_c * V_c)|$
2. min $\sum_c(P'_c - P_c)^2 * V_c$

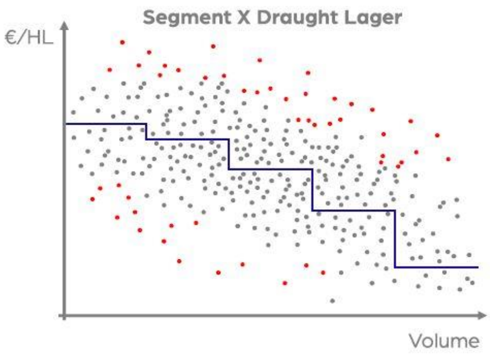
Where P_c means selling price per POC and V_c for volume per POC.



Note: In both equations we assume that the volume V_c is not impacted by the price, an assumption that could be relaxed to gain a more realistic estimate.

3. Fine tune those segments to further optimize tiered price function. Adjust the initial way in which you defined the x-variables in the initial function. Play around with variables.

The main objective of this third phase is to identify undesirable outliers, and apply adjustments to the segmentation rules to improve the performance in terms of overall turnover and absolute price change as defined in phase 2.



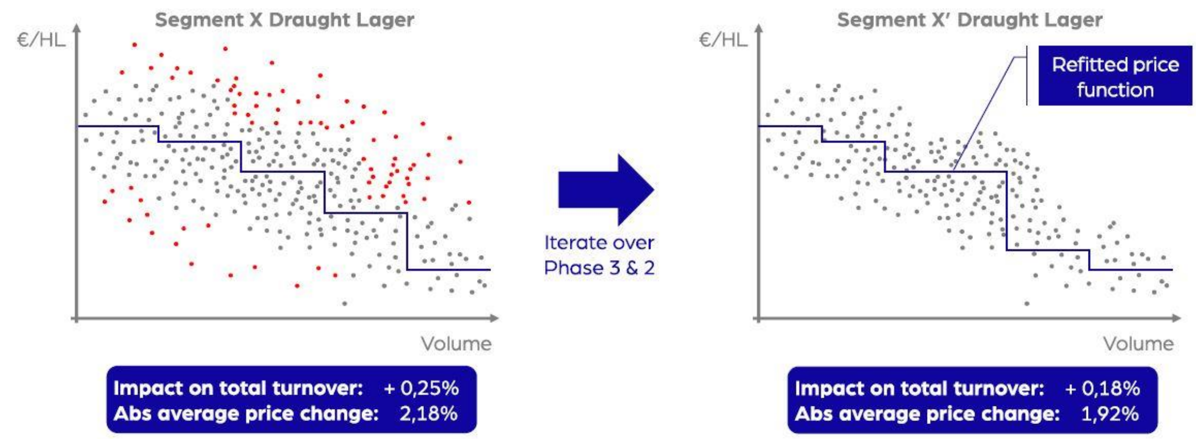
The heart of this procedure is an evaluation of which price changes would be the most undesirable (i.e. which are the most likely to result in customer defection).

The image on the left shows a very simple evaluation where the absolute distance (i.e. the absolute magnitude of the price change) is interpreted as a metric for this.

However, more complex variations considering the "value at risk" can also be designed. An example could be the square of the price change, multiplied by the volume sold to the specific client:

$$\min(P'_c - P_c)^2 \cdot V_c$$

At every iteration, consisting of both the adjustment of the segments and refitting of the tiered price functions, the "entropy" surrounding the stepwise linear function is reduced.

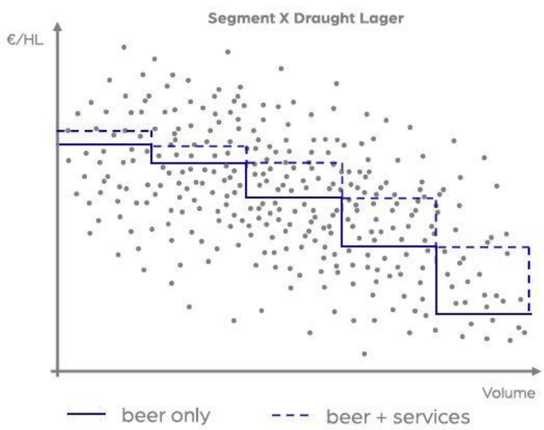


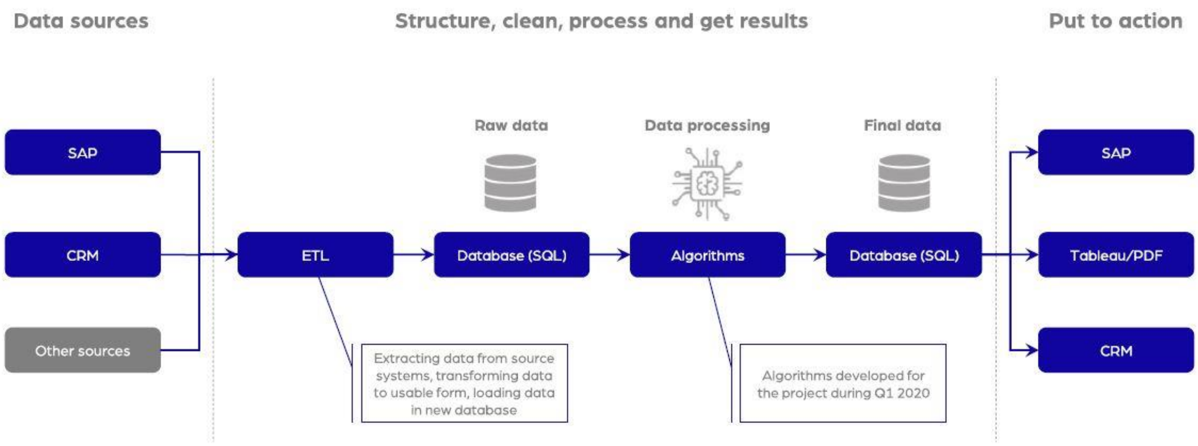
Iterate over phase 2 and 3 and see if you can remove any outliers.

For the sake of simplicity, the preceding analysis has made abstraction of the effect of adding prices for services, since this also changes the overall price point for clients.

The final methodology will deal with this aspect in much the same manner as the volume, using the amount of services as an explanatory variable in the first prediction model, but excluding this from the logic of building segments.

As such, the piecewise linear pricing function is extended with price discrimination tactics for the services as well as the volume.

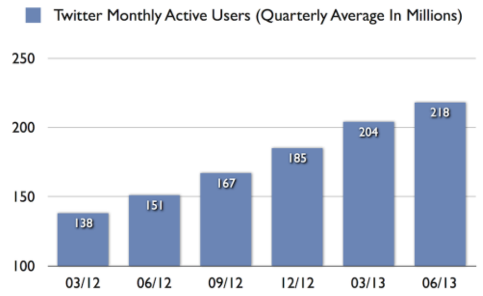




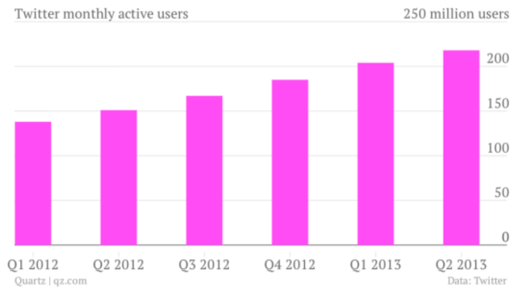
This diagram represents a simplified version of the data processing flow.

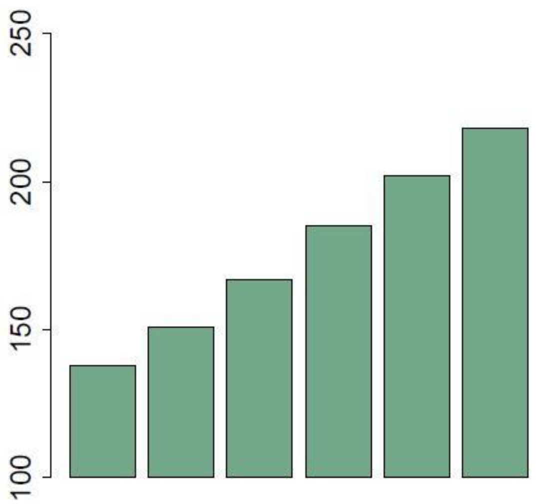
Data visualization

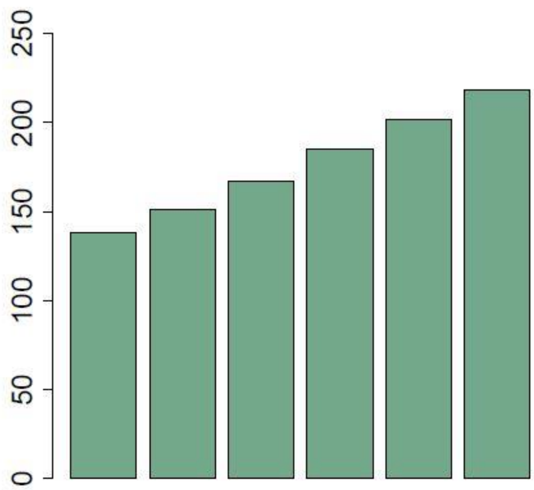
The focus lays on what the ideas are, the theory behind data visualization. Not so much how to do it in practice, but what the difference is between a good and a bad visualization.



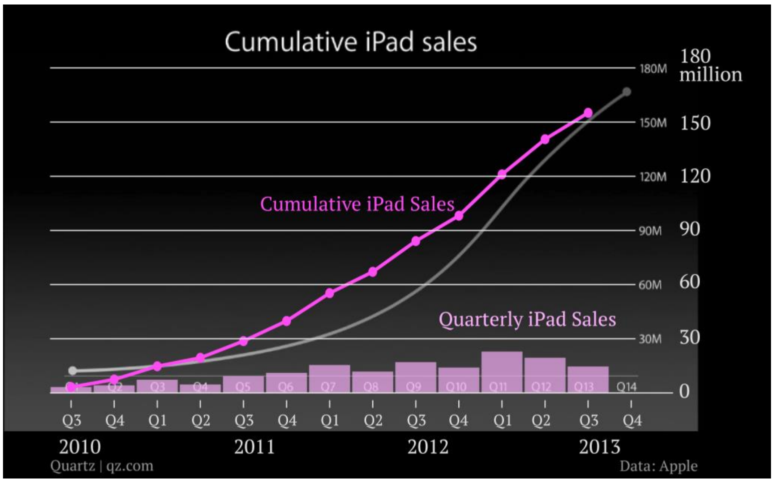
This is a figure that appeared in IPO (initial public offering) of Twitter in 2013. They pot the share on the stock market. In the report, Twitter shows why it is a good idea to buy stocks from them. It shows that from march 2012 until June 2013 the number of active users had increased dramatically. But they got criticized for that! Because **the Y-axis does not start at 0**, the slope of the imaginary line, drawn at the top of the histogram, seems much more dramatic than it actually is! Here is the adjusted figure on the right. So not starting the Y-axis at 0 can really change the point of view of the figure, and this happens very often.







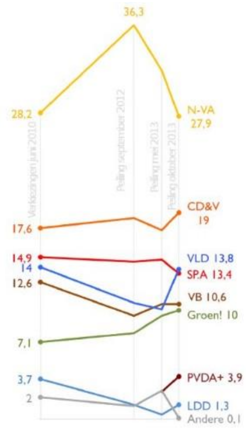
The next figure is about Apple. They have a big conference every year about new products. In this figure they talked about the iPad sales, they wanted to motivate their story by showing data visualizations. But what is misleading here? They used **the cumulative sales**: those sales can never go down! So the graph was misleading. It seems as if the sales are skyrocketing, but what would the actual sales look like? In the last quarters, the sales actually went down! Also, they **truncated the Y-axis** a little bit.



Visualization is the power or process of forming a mental picture or vision of something not actually present to the sight. A picture thus formed. Visualization is also the construction of images which represent important aspects of some situation or process. Such as a plot, graph, diagram or picture.

Partij	Verkiezingen 06/10	Peiling 9/12	Peiling 5/13	Peiling 10/13
N-VA	28,2	36,3	32,1	27,9
CD&V	17,6	18,5	17,4	19
SP.A	14,9	14,5	14,7	13,4
VB	12,6	9,5	10,6	10,6
VLD	14	10,7	10,1	13,8
Groen!	7,1	7,9	9,5	10
LDD	3,7	1,3	0,4	1,3
PVDA+	5	5	2,5	3,9
Andere	2	1,2	2,6	0,1

Now why should we visualize things? **Our brain excels in interpreting visual things.** For example: there are different parties and election on different time periods (see picture). Now what is the information in this table? That's very difficult to tell! We have to read all the numbers and then we still have to try to figure out what they want to tell. But when they used visuals, you can see it immediately.



Video also shows the power of visualization. For example, how counties evolve through the last 200 years, you cannot understand what is going on in the data if you don't visualize it.

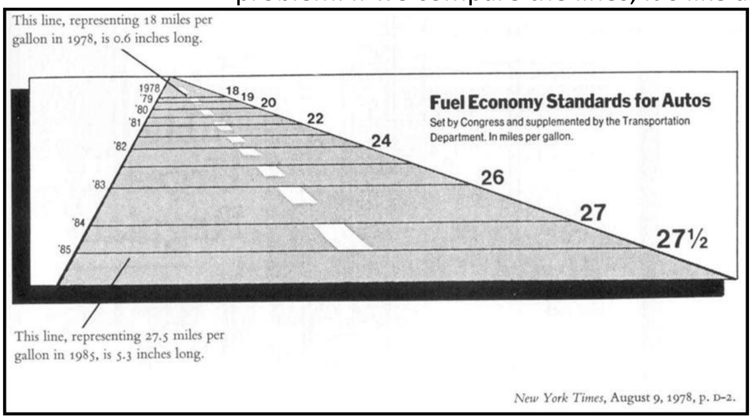
So now again, why do we visualize?

- ***Communication.*** Visualization provides a quick way to communicate a very rich message.
- ***Discovery.*** To discover certain patterns or trends (always first do some descriptive analysis + make some plots. Humans are better in absorbing information when it is visual). Visualization provides a way of displaying a large amount of data so we can uncover new facts and relationships.
- ***Insight.*** We work with formulas in this course, this is information that is provided to us in a logical way but for most of us difficult to really understand but once you draw that function, it becomes much more easy to understand what the formula is about. Visualization provides a way to obtain better insight into things we already know.

Theory

Tufte's principles:

1. **Data integrity**, what does this mean?
   You see miles per gallon that cars could drive through the years. More miles per gallon means that the cars are more energy efficient. So here: more is better. The message that we get from the graph: in 1998 they could only drive X miles, and in 1995 so much. But there is a problem with data integrity here. The horizontal line at the end represents 18, the one in the beginning 27. The proportional increase when we go from 18 to 27 (that means, not even times 2) is the problem. If we compare the lines, it's like a tenfold increase!

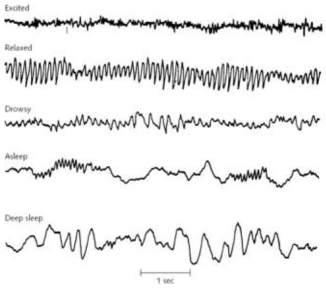


Data integrity: the visual difference should reflect the real data, which is not the case here. To quantify this a bit, Tufte developed the lie factor: the size of the effect shown in the graph, divided by the size of the effect in the data. So you should aim for a lie factor of 1: perfect resemblance between the visual information and the real information in data.

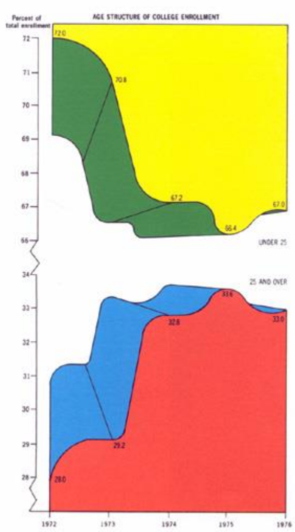
Lie factor = $\frac{size\ of\ effect\ shown\ in\ graph}{size\ of\ effect\ in\ data}$

Size of effect = $\frac{|second\ value - first\ value|}{first\ value}$

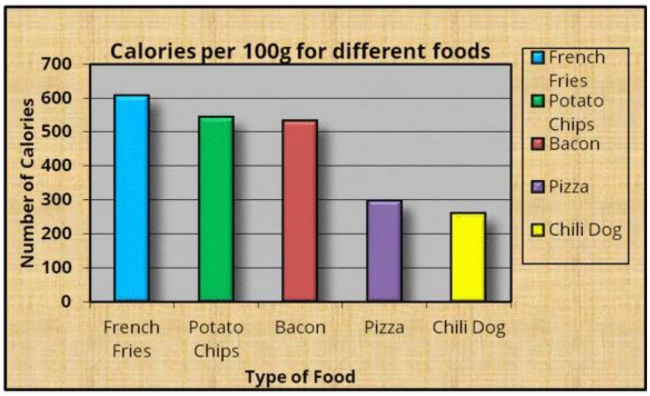
2. Data-Ink



A good visualization should have a **very high data-Ink ratio**: every point that you put on the figure, should be related to a datapoint, don't spill too much ink in the figure that is not data-based. These visualizations have a very high data-ink ratio because everywhere you see ink, we have a data point. but don't throw it overboard!



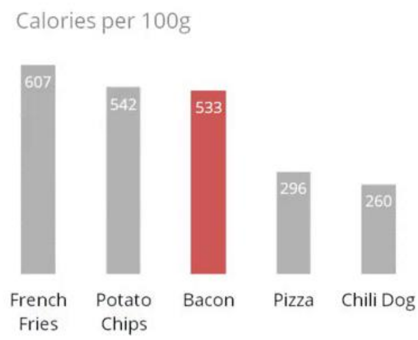
The next picture is the percentage of students older than 25 when they start studying at university. In the beginning it was 28%. What is wrong with this visualization? First of all, way to much colours. The upper and down part contain the exact same information. One is percentage over 25, the other is percentage under 25, they are just complements. So we should get rid of the top part of the graph. There is some 3D effect, but we only have 2 dimensions? The 3D effect does not relate to any data: **chart junk!** They also make very strange interpolations between the points. We have just 5 data points (there were we have a number) and between the points they don't use straight lines, they make some curved line through it, very weird, it does not reflect any data at all!



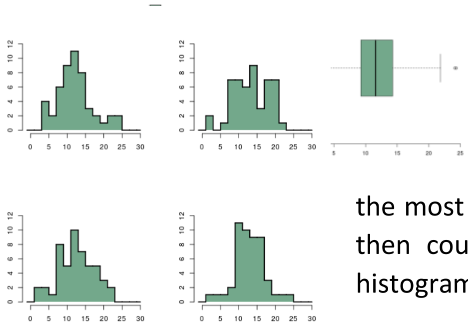
How could we improve the graph on the left?

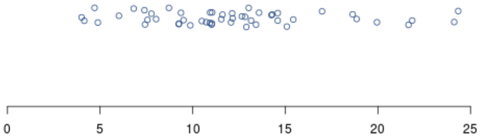
- Change background colour.
- Legend is not adding anything so get rid of it.
- Shading in the bar chart, doesn't add anything (not everyone has the same taste however).
- Calories "for different foods" can be removed.
- Axes do not need to be labled
- Remove borders
- Reduce number of colours (you should only use colours when you want to highlight something.
- Remove special effects
- Remove bolding
- Lighten the labels so that bars become more visual
- Remove lines in the background
- Do direct labelling: but data on top of bar instead of on Y-axis.

LESS IS MORE! End result →



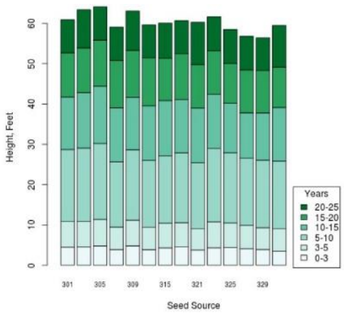
Visual summary of one variable. We want to visualize age. It is not good to represent it in a line graph: you want to know how many people are in the same age group, what the typical age is, how young the youngest person is, so more the distribution of the age of the customers. One of the most obvious ways is by creating a histogram. It creates bins in data and then count how many are in each bin. You could also smooth out the histogram: density plot. Another way is a boxplot.



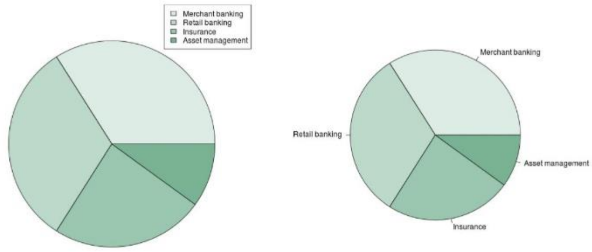


Compare distributions. What with different distributions that you want to compare with each other? We have 4 variables, product categories and you want to see the distribution of the age in that product category. So we create the same type of plot as before, but now several together: small multiples. One plot that contains of several subplots. We have a histogram with the 4 categories next to each other. Or all boxplots next to eachother.

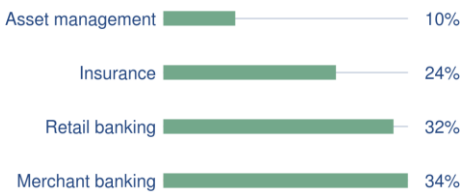
Stacked bar charts/histograms are not a good idea: it becomes very difficult to compare different values. If you are interested in the category of 15-20, and you want to know where the highest frequency lays, so we are looking for the longest green bar, that is very difficult because they all start at different points. Alternative version: take all the light greens, then all the dark greens and so on...



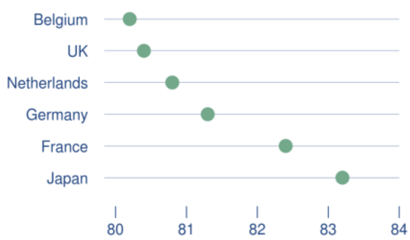
Compare single values. In the field of visualization they say that pie charts are a suboptimal way. Most people would suggest not to use pie charts. If you do want to use pie charts, you can improve it by removing the legend and just add direct labelling.



Why are pie charts not optimal? Well our brain is much better in comparing length than it is in comparing surfaces. Here the surface represents the number, but that is difficult for us. How much larger is merchant banking than retail banking? If you use a bar chart, you can see the difference immediately.

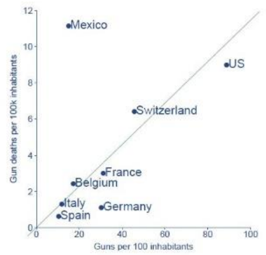


Next we want to compare the life expectancy in different countries. We are interested in the difference between the countries so it is not optimal if you show it in a bar chart. We could make the mistake to not truncate the Y-axis, if we do that we focus on the differences. But actually this is unfair: the heights of the bars are not representative for the data. Solution: dot charts! Why is truncating dishonest in bar charts and not in dot charts? A bar chart focuses our attention to the length of the bar so if you then truncate, that is dishonest because the truncated bars do not represent the underlaying data. This is not the case for dot charts because here the focus is put on the



location, on the dot. So you could use bar charts but if you want to zoom in on the differences, you should use dot charts. This allows you to maximize the effect without putting the attention on the length.

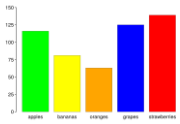
Compare 2 continuous variables. For example, gun ownership and gun deaths. We can visualize this in a scatter plot, then you can see quicker what is going on in the data compared to the bar chart. US scores high on both dimensions, there is no other country with more guns and also many gun deaths. You can always add a regression line, this shows us that US has the expected gun death for the number of guns they have. In Germany there are relatively a lot guns but not many gun deaths. In Mexico it's the other way around. There are strikingly many gun deaths in comparison with the amount of guns they have there.

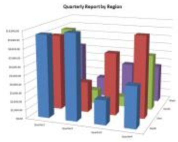


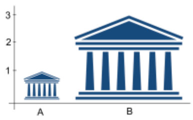
Showing evolution is best done with a line chart. A bar chart is not optimal because when the focus lays on showing evolution, a line chart is better. But what is wrong with these plots? There is missing data for the month November! So we go from October to December. So when you use bar charts, the bars are put next to each other, which is kind of misleading. Bar charts are more suited for discrete information, for different categories, for grouping data. If the X data is time, use a line chart. It is also very important that you should always indicate where your data is and where you have interpolations. The circles in the line chart are the data points and the lines in between are interpolation.

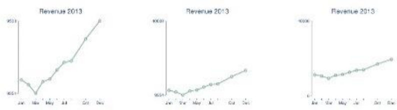
Don'ts

Don't overdo colours. Only use colors if you want to highlight something. Also don't do 3D. First of all there is no 3th dimension but it gives also an additional problem: it leads to confusion, comparing the surfaces is even more difficult now. Watch out with scaled pictures. B is 3 times bigger than A but if you look at the image, that is not really the case! A can fit 9 times in B, even though you have the numbers of the Y-axis. The goal of data visualization is that when you look at the graph, you immediately see what is going on and what the data behind the graph is. Do not create a nice visualization that you then have to interpret by looking at the axes/at the data. Instead stack A 3 times to get B. Also be careful about how you scale your axes. And when you show only 1 piece of data, make sure you show the full data too (improper extraction).

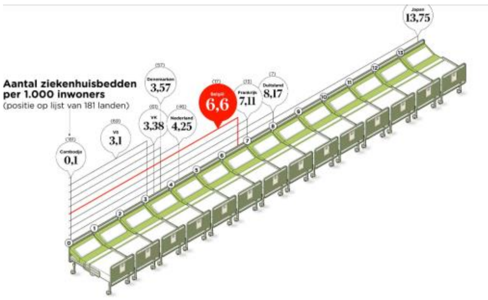








EXERCISE



The labels are misleading in this image, you get the impression that Belgium is doing a bit worse than France but if you look at the height of the bars, they are almost the same, very close to each other. By putting the label further, Belgium seems to do worse than they actually do. This is even more the case for the US. The label is 1/3th away from the actual top of the bar, so you get a wrong impression.

<u>EXERCISE</u>

Soccer red cards. The legend takes up as much space as the graph. Direct labelling is better but then it would be necessary to rotate the bars horizontally, otherwise because of the length of the label, you would push the bars away from each other. Bad practice would be to rotate the labels. It is also not a good title and the Y-axis label is rotated. On the exam: cut data visualization and then we have to relate it to the do's and don'ts and the principles we saw.

<u>EXERCISE</u>

They violated the principle of truncated Y-axis. it starts at 100 million and by doing so it looks as if there are 8 times more people on welfare than people with a full time job.

Take-Aways

Do's:

- Summarize single variable → histogram or boxplot
- Compare distributions → histogram or boxplot
- Compare single value → bar chart or dot chart
- Compare 2 continuous variables → scatterplot
- Compare > 2 variables → small multiples
- Show evolution → line chart
- Scale your axes "fair"

Don'ts:

- Avoid using a legend → explicit labelling
- Avoid using many colours → colour important datapoints
- Avoid using pie charts → use bar charts or dot charts
- Avoid using 3D → no more dimensions than variables
- Avoid using surface or volume to compare values → use length
- Avoid "unfair" extraction of data