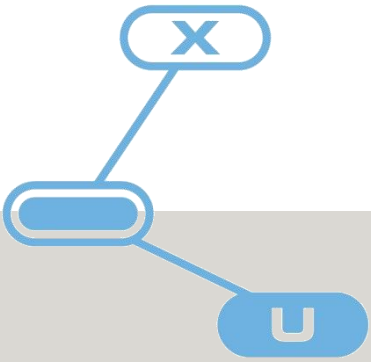


White Paper – Straight from Customer Insights

Wieke Selten, Peter Stehouwer and Roel Moradi

To survive in the 21st century as a company you have to continuously listen to, and be able to understand and respond to, customer feedback. Such word-of-mouth data can originate from a variety of sources, such as internet reviews, complaint files and social media. In this white paper we take the example of airlines to show you how to make sense of all this available input and come up with actionable insights.

Reviews & Ratings

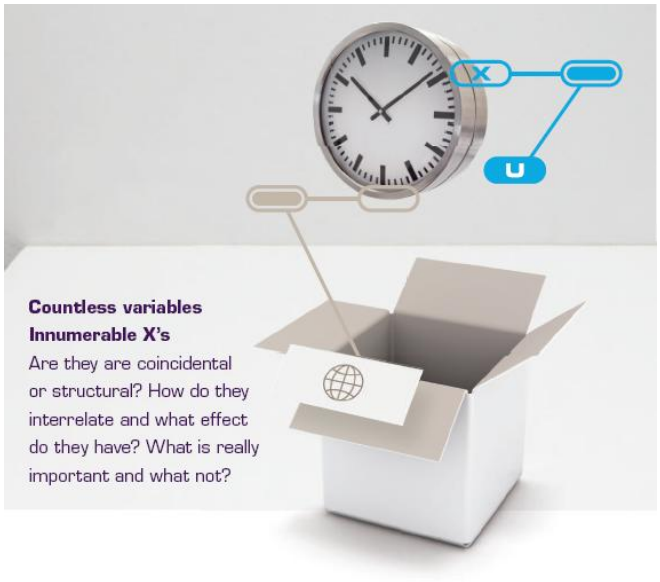
Have you ever written a review on a product you bought? Or have you, while looking for a good restaurant to eat, read such reviews? Did reading them influence your decision in your restaurant choice?

Reviews are just one example of unstructured word-of-mouth data. Increasingly, people give their opinion about products and services on the internet, from sharing their opinion about their experiences with a general practitioner to writing book or tool reviews on sites like Bol.com and Amazon.com. Besides a review in free text format, customers are often also asked to assign an overall rating for their experiences with a product or service. Some reviews can contain even more information, when customers also assign sub-scores to various aspects related to the product/service and provide some personal characteristics such as their gender, age or residence.

Reviews not only serve as an outlet for the writer, but are also an important driver in the online decision-making of potential buyers, also called the 'zero moment of truth'. According to Econsultancy.com, 61% of customers read online reviews before making a purchase decision[1]. Consumer-to-consumer communication is considered to be the most important factor

in purchasing decisions[2]. It is therefore no surprise that an increase in rating is related to an increase in sales[3].

Often, online reviews are quite extensive and available in large numbers. Moreover, they are free to pick up from the internet, as are the reviews of your competitors' products and services. Online reviews and ratings generally contain high quality information offering valuable insights. In short, they can have a high impact on organizations today.



The sky's the limit

Many holidaymakers write about their experiences with their accommodation, destination or the airline that took them there. Immediately after returning home, holidaymakers are asked for their opinion on these things, resulting in thousands upon thousands of reviews. The Davy Crockett bungalows near Disney World in Paris have had more than 1,000 reviews and the airline, KLM, has been reviewed by over 4,000 people.

Let's take a case study of airlines. In this paper we study reviews of the 'Big Five' airlines in the Netherlands: KLM, Transavia, Easyjet, Arkefly and Ryanair. As well as giving an opinion in free-text format and an overall rating (ranging from 1-10), holidaymakers also assessed some or all of the subcategories: leg space, flight times reliability, entertainment & facilities, food & drinks, hygiene, child friendliness, price/quality and staff service. Hereafter, we shall refer to these subcategories as 'drivers'.

Airline	*# of reviews*	*Average overall rating*
KLM airlines	4.456	7.9
Transavia	2.733	7.3
ArkeFly	1.519	6.8
EasyJet	354	7.0
RyanAir	172	6.9

Table 1. Overview airlines, number of reviews and average rating

What do reviewers write?

An example of a review included in the analyses is one written by a 72 year-old about their experience with KLM. While this reviewer complains about paying extra for a decent amount of leg space, they still give KLM an overall rating of 9. Except for the driver Reliable Flight Times, the subcategories are also rewarded with good ratings of at least an 8.

Actionable insights

Using statistical techniques, one can link the driver ratings to the overall rating given by customers. Figure 1 illustrates a selection of the results of these analyses.

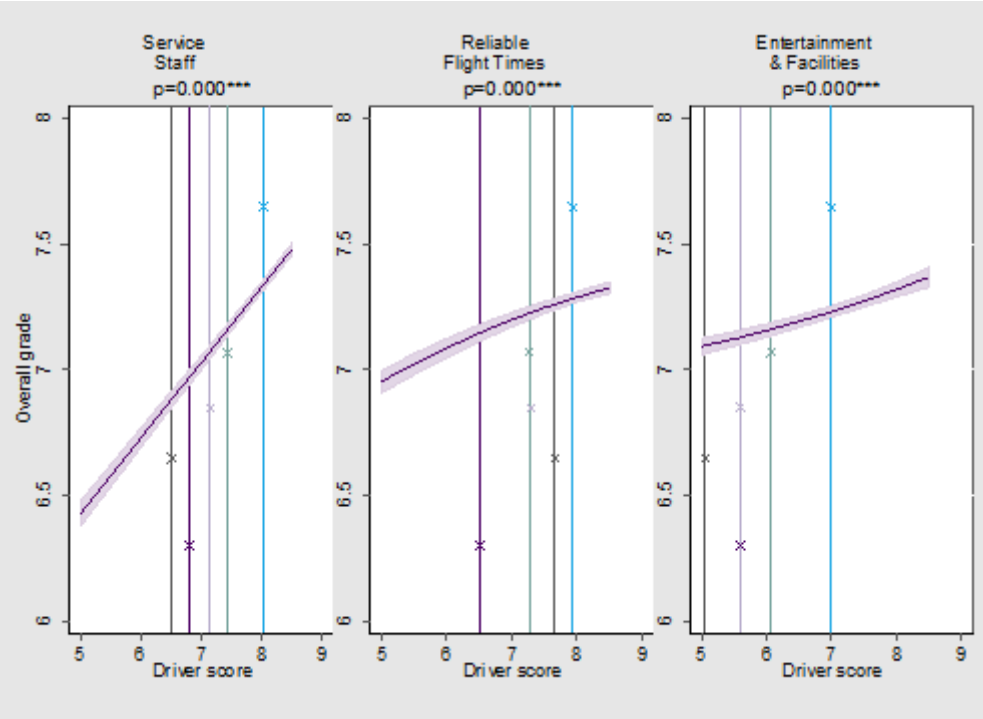


Figure 1. Effect plots

How to read these graphs?

For each driver a separate effect plot is drawn, representing the relation between the driver and the expected overall rating (controlling for the other subcategories).
The y axes show the overall rating, ranging from 6-8. The x axes represent the scores for each driver. The purple lines indicate the best estimate of the relation between the driver score and the expected overall rating.
The 95% confidence interval for this estimated relation is shown by the light purple area around the line. The vertical lines are the average driver scores per airline and the crosses on them represent that airline's average overall rating.
If the p-value at the top of each graph is less than 0.05, the driver is said to have a significant effect on the overall rating.

Must

At a glance, we can immediately see that the graph of the driver Service Staff has the steepest line, meaning that this driver has the largest effect on the overall rating. The line should be interpreted as follows: an increase of one rating point on Service Staff results in an increase of about 0.3 points in the overall rating. Using the terminology of the Kano model (fig. 3)[4], this driver appears to be a 'performance need': the more positive customers are about Service Staff, the higher the overall rating.

Want...?

The relationship between the driver Reliable Flight Times and the overall rating is a 'basic need' in Kano terms: the impact of this driver on the overall rating is greater for lower driver ratings than it is for higher ones. Customers expect an aircraft to depart on time, so if this basic need is not met, they will be dissatisfied and give the airline a lower overall rating. However, when customers' expectations on Reliable Flight Times are met, they do not generally reward this by giving a higher overall rating.

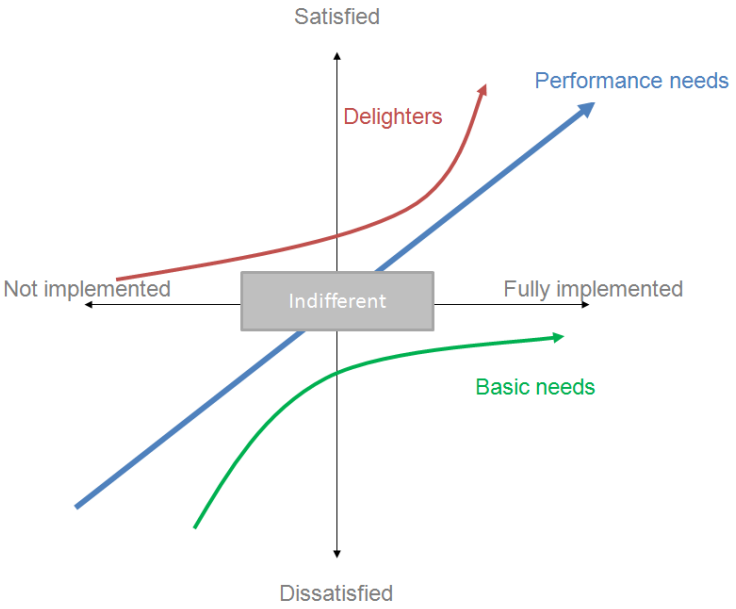


Figure 3. Kano-model

...Or Wow?

In Kano model terms, the driver Entertainment & Facilities most closely resembles an 'excitement need'. However this resemblance is not very convincing based on the numerical data. Let's have a look at the textual data: quotes from review texts show that the driver can act as a wow factor. These customers seem to be surprised by the airline's entertainment facilities, which is also reflected in the high overall rating they give.

'*Nice monitors in the plane on which you can see where you are and videos are played in between* (10)'

'*And what I really liked is the possibility to watch movies all the time, I love it!* (10)'

'*The 11 hour flight went very quickly because of the distraction provided by the video screens with very nice movies in the chairs in front of me* (9)'

Figure 2. Word cloud of review texts (overall rating presented in parentheses).

So, where to invest?

Even though these effect lines already give valuable insights, for decision-making on where to make improvements the airline should also consider their average score on that driver. For instance, as the average reliable flight times score is 6.5 for Arkefly and about 8 for KLM, and this driver is a basic need, it makes no sense for KLM to invest in the reliability of its flight times, as it will no further increase their star rating. For Arkefly, however, there is much more to be gained from an investment in this driver in terms of star rating and competition with other airlines.

Differentiate yourself

Combining the effect lines with the average driver scores also gives insight into possible relevant differences between a company and its competitors. For the driver Entertainment & Facilities, none of the airlines has an average driver score higher than 7. Take out KLM and the highest average score is only 6. This offers opportunities for differentiation: investing in Entertainment & Facilities enables airlines to differentiate themselves from the competition.

Be careful!

Creating trustworthy and actionable insights from unstructured word-of-mouth data can be tricky and requires a combination of data science skills and domain knowledge of a company's products and markets. Furthermore, there are several potential pitfalls that should be properly addressed when analyzing unstructured data:

- Studies based on freely available review texts involve observational data rather than properly designed experiments. This usually means unbalanced datasets in which the observations are unequally distributed over all the possible combinations of drivers. With unbalanced data, you should always check for multicollinearity and be careful in drawing conclusions. The main risks are: unjustified extrapolation and/or unjustified causation.

- Review datasets typically involve a lot of missing values. If a reviewer doesn't write about or rate a given driver, a missing value occurs. In building a model, you should think very carefully about how to deal with such missing values.

- Also bear in mind that the data is not a representative reflection of the users in general. It only contains data from customers writing reviews[5].

In addition to more traditional statistical techniques, machine learning techniques like deep learning[6] and Bayesian Networks[7] are gaining popularity in the analysis of unstructured textual data. It's important to remember, however, that no technique can solve the intrinsic issues listed above. This is why, for this kind of study, at CQM we combine domain knowledge with data science techniques. To avoid making unfounded decisions, we always calculate 95% confidence intervals into our predictions. A thorough understanding of both what you're doing and the application domain are vital in order to reach trustworthy conclusions.

What next?

The analysis in this case study is just a sneak preview of the kind of analysis that can be performed with unstructured textual data. Such rich datasets can provide many more interesting insights. To list a few of the possibilities:

- Text analytics: So far in this paper we have only taken into account the ratings given on the drivers. The reviews that customers write on review sites contain even more valuable information, as holidaymakers also give their opinions in a free-text format. Employing natural language processing (NLP) and sentiment analysis, one can map the voice of the customer using textual data to gain actionable insights.

- Segmentation: some review sites ask reviewers, for example, their age, place of residence, tour group (and age range of the group) and travel period. Analyzing whether the relationship between the drivers and the overall rating differs per subgroup can be very revealing. For instance, are other aspects important in an airline for people travelling alone than for people travelling with a partner? Are there differences between those travelling in high season and low season?

- Differences between companies' customers: are Ryanair customers comparable to KLM customers, or do they have completely different expectations regarding a flight? Presumably customers of a discount airline like Ryan-air are less likely to expect high-quality service on board than KLM customers. And one wouldn't expect Ryanair customers to be as dissatisfied when their flight is delayed as KLM customers are. Studying different groups of customers using the kind of analyses discussed in this paper yields additional valuable insights.

- Competitor analysis: as well as giving insight into where the competition stands regarding the market's drivers, it is also interesting to dive deeper in the differences between the airlines. Why are there differences between e.g. KLM and Transavia? Which are the main drivers explaining this difference?

Whatever business you are in, we hope it is clear that word-of-mouth data is everywhere and that it is possible to extract valuable decision information from it. Don't wait too long, because your competitor has probably already started!

[1] *Charlton, G. (2015). Ecommerce consumer reviews: Why you need them and how to use them. Retrieved from econsultancy.com/blog/9366-ecommerce-consumer-reviews-why-you-need-them-and-how-to-use-them*

[2] *Bughin, J., Doogan, J. and Vetvik, O.J. (2010). A new way to measure word-of-mouth marketing. Assessing its impact as well as its volume will help companies take better advantage of buzz. Markting & Sales practice. McKinsey Quarterly (April).*

[3] *Coulson, A. (2013). Negative online reviews about your business are costing you money. Retrieved from acuwebservices.com/internet-marketing/negative-online-reviews-about-your-business-are-costing-you-money/.*

[4] www.kanomodel.com

[5] www.hadeninteractive.com/consumers-write-reviews/

[6] http://en.wikipedia.org/wiki/Deep_learning

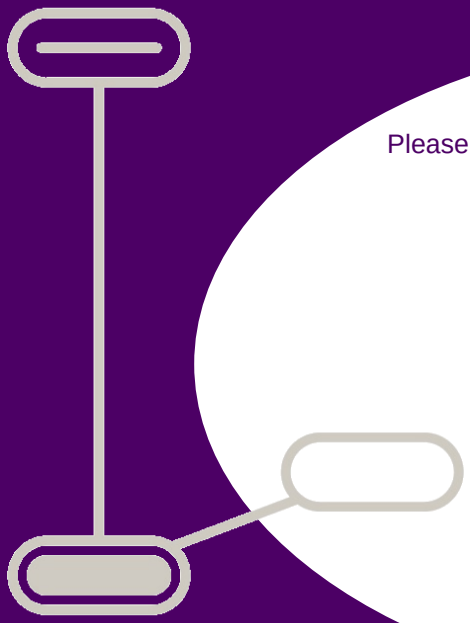
[7] http://en.wikipedia.org/wiki/Bayesian_network

CQM helps organizations make complex processes transparent. Using quantitative models, we create a framework to analyze processes and make decisions based on the facts, enabling you to optimize your planning and logistics, and improve your product and process innovation. Intelligence, that takes your organization to a lasting, higher level. We analyze and clarify, with a genuine understanding of the issues you face. That is the way we do things. **From x to u**

Text analytics and CQM

CQM has been in the data science business for more than 35 years. We are constantly developing our methodology on text analytics and can help you find trustworthy, actionable insights from the growing amount of unstructured word-of-mouth data collected by your company.

Are you curious what your customers write online about your products and services? Do you have other textual or numerical (big) databases (e.g. the voice of the customer) you would like to analyze for actionable insights? Or would you simply like to discuss the opportunities for your organization regarding text analytics in general and/or the structured mapping of the customer's voice specifically.



Please do not hesitate to contact us via mail or phone (040 750 23 23) or visit our website www.cqm.nl for more information.

Wieke Selten - Junior Consultant
selten@cqm.nl

Peter Stehouwer - Senior Principal Consultant
stehouwer@cqm.nl

More ?

Would you like to receive future white papers?
Subscribe here!

This paper is loosely based on a presentation on airline data given by Senior Consultant Roel Moradi at the Data Science Meetup in Amsterdam in March 2016.