Leveraging Big Data to make your customers happy

How the Dutch National railways (NS) uses data to monitor and make smart decisions about the experience of its customers

Prof. Peter Verhoef
Professor of Marketing
University of Groningen

Joost Bosma
Head of Customer
& Market Insights
Nederlandse Spoorwegen

Thijs Urlings
Customer Experience
Consultant
Nederlandse Spoorwegen

Hanneke van der Boog
Customer Experience
Consultant
Nederlandse Spoorwegen

Martin Heijnsbroek
Managing Partner
MIcompany

Helen Rijkes
Program Manager
MIcompany

MI COMPANY

Leveraging Big Data to make your customers happy

How the Dutch National railways (NS) uses data to monitor and make smart decisions about the experience of its customers

In January 2011 the customers of the NS had suffered from serious waiting times and cancelled trains. This was now the second winter in a row, where management was forced to embark on a series of crisis interventions that would limit the damage done to their customers. Besides launching a task-force to improve the winter-resistance of the rail network, the NS asked itself the question: “how can we more structurally monitor and improve on the drivers of the experience of our customers?”. Management was convinced that the answer to that question was hidden in the very large sets of operational and marketing data available. And that through uncovering the answer, the NS would be able to continuously optimize the investments for the experience of the customer base. The approach that was developed lead to considerable impact and was selected as one of the five finalists of the global Gary Lillien Marketing Prize in 2016.

Context. Improving Customer Experience suffered from multiple measures and insufficient insights on key drivers.

While many companies have put Customer Satisfaction or the Net Promotor Score as a critical KPI, and even some have defined their mission and purpose around serving customers in the best possible manner, there are few companies whose license to operate depends on how they serve their customers. As a state owned company, and regulated monopoly, optimizing on profits or revenues does not make sense. At the NS, the railway services need to be offered at the highest possible satisfaction against the lowest price. Meeting therefore the ambitions and targets on customer experience is of critical importance for the NS role in society. As many companies suffer from multiple lenses, interpretations, metrics and often even datasets to manage performance, this was also the case at the NS for customer experience. Historically the NS would use the NPS metric, a Reputation tracker and a Customer Satisfaction score. While there can be a logic for having multiple metrics for different purposes, the NS realized that they lacked one single view on their customer performance, with the risk of management cherry-picking the right metric to defend or promote a specific customer initiative.

At the same time, the NS had an ambition to become more data-driven in the way its organisation operates, and chooses the best initiative to delight their customers. Instead of making these decisions based on opinions of the stakeholders, management or media, the NS had a desire to move towards a data-driven way of working that would allow them to monitor and improve the actual drivers of customer experience. That data-driven way of working was a major ambition, given the fragmentation involved in measurement and the decision processes and strong opinions in the organisation.

> Hans Peters, Commercial Director of the NS: "In the past, different departments in the organisation aimed to evaluate their own performance, as a result of which nobody took ownership to achieve overall targets or to prioritize initiatives across the departments."

Approach. Leveraging Big Data to understand, monitor, predict and improve the experience of customers.

Given the importance and sensitivity in developing a more structural and more granular perspective on managing customer experience, the NS had decided to develop a data-driven approach. The central idea evolved around 4 steps:

1. Aligning on a single metric: By adopting one measurable metric of customer experience, the NS would be able to steer the company by fact-based decisions based on this single KPI.

2. Estimating the driver based model for customer experience: By combining millions of records of operational variables, customer variables, marketing variables, external variables (e.g. weather conditions) and the – to be defined – single measure of customer experience, Big Data Analyses would then reveal the most important drivers of customer experience. Instead of trying to improve and maintain all possible drivers of experience, the NS concluded that identification of the top drivers would focus management initiatives. This would enable building a granular monitoring system that allows continuous focus on the top drivers, on specific routes or customer segments where performance was lacking.

3. Developing a prediction model for customer experience: Moreover, such a driver-based model would become the core of a forecasting capability, that would allow the NS to monitor their performance versus year-end ambitions, and correct for typical trends of seasonality or expected improvement initiatives.

4. Optimizing investments on contribution to customer experience: Even more exciting, such a driver-based model could become the engine for estimating the incremental effect of an initiative on customer experience. More specifically, quantifying what an additional euro investment to a specific customer initiative would contribute to the experience of its customers. Allowing the NS to allocate its customer investments to optimize the customer experience of its customer base.

> Joost Bosma, Head of Customer & Market Insights NS: "Although there were multiple opinions on the value of using hard operational data versus softer marketing research data, we all united on the mission to start optimizing one single version of customer experience, and become more data-driven in the process. We had asked MIcompany to support us with this complex and data-intensive challenge, and invited professor Peter Verhoef, from the Rijksuniversiteit Groningen, to guide and challenge the team based on his extensive knowledge and research experience on the topic of customer experience. Although we had assembled a team of independent, but also stubborn thinkers, in the process we managed to unite on the excitement of contributing to a more fact-based view on how to improve the experience of the Dutch railway passenger."

Aligning on a single metric.

Although there is probably not a single metric that is the silver bullet for all purposes, from a firm governance perspective it was critical to avoid confusion and even cherry picking the right metric for the right argument at the right moment. In search of the 'one' customer experience KPI, three different metrics were investigated:

- The Customer Satisfaction score, a large survey which is measured on a 10-point scale on a daily basis in trains, which was also the input for the government target;
- the Net Promoter Score to evaluate customer processes in more detail;
- the Reputation Tracker, used to understand feedback of customers mainly from a corporate communication level.

To make sure the metric would meet the ambitions for a single customer experience instrument it was important to set the right selection criteria. Based on Criteria by Ailawadi et al. (2003), we came up with a solid framework for assessing the strengths and weaknesses of each of these three metrics. We summarize the evaluation in Table 1. The framework provided a strong grip on the different angles that are important for a metric. Key angles included the theoretical foundation for the metric, the robustness in the development over time, the possibility to link it to service crises, the variations over scores and its implication, and the volume of data present. Based on this systematic scoring of the metrics on these criteria the customer satisfaction score (% 7 or higher) proved to be the strongest. The board unanimously accepted the customer satisfaction metric as their key-customer metric, and with that decision stopped the internal time-consuming debates on which customer metric to use. Moreover, the focus on this KPI was used in the upcoming discussion with the Dutch government and resulted in concrete agreements.

Estimating the driver based model for customer experience.

One of the first aims of the management was to understand how they could influence customer satisfaction. As noted in the past the NS executed multiple satisfaction studies, but the impact of these studies was limited. In sum, our approach was to measure these drivers, through a multi-method approach. This multi-method approach consisted out of three different modelling techniques:

1. Big Data driver model: finding patterns hidden in the existing data-sets of the NS.

2. survey-based driver model: measuring drivers through interviewing customers on an integral set of customer experience drivers.

metric / criteron	CSAT	NPS	REP-TRACK
Theory-based	+	-	-
Complete: multi-dimensional	-	-	-
Diagnostic for crises	+	+	-
Robust and reliable	++	+/-	+/-
Single number	+	+	+
Intuitive and trustworthy for top management	++	++	++
Intuitive and trustworthy for stakeholders	++	-	-
Validated with outcome measure: customer value development	++	+	?
Based on existing data	++	+	+

Figure 1. Metric evaluation

3. driver importance using input from the firm: letting experts score the most important drivers based on their customer experience.

Combining these three methods, we landed on a single approach to measure the importance of the considered driver. And because findings were consistent across the three models, the approach was strongly supported by the management and experts in the NS.

Big Data driver model

The first challenge we had to address was to collect the data for this approach. Since the NS serves more than 9 million unique customers each year and has multiple products and measuring systems at place, we knew there was a lot of data available. But the question was; Which data should we use and in what way can we link it to a customer satisfaction score?

In our first model we gathered data from several sources: the survey (Our sample consisted of around 120.000 respondents), observations of the interviewer of the survey on some operational issues (i.e. train cleanness of train), internal operational data on punctuality, stations etc., and internal marketing data and external data on (social) media presence. One of the challenges was that we could not track customers over time to see how their experience evolved. Given that we had a dataset that consisted of multiple snapshots of performance (repeated cross-sectional data), we tried to link the data as close as possible to the experience of the customer. Therefore we integrated the data using two linking variables: (1) the travelled trajectory (from train station A to train station B) and the timing of the travel. Another challenge was the large number of explaining variables, that could also be correlated strongly and would mess-up our statistical model. To correct for potential correlation between explaining variables, we first executed a principal components analysis. This analysis produced a number of new components that would replace correlated variables (mostly regarding the cleanness of the train, cleanness of the station, exposure in media, marketing efforts, presence of shops and service points on stations, presence of parking and taxi on station). The overall results are shown in Figure 2.

Figure 2. The Driver Model contains individual variables as well as factors

This first Big Data driver model showed that almost all included train variables were significant. Punctuality has the expected positive effect, whereas delays, the fullness of the train and the number of complaints had expected negative effects. For the station especially, the presence of good station facilities had a relatively strong positive effect. Interestingly the service variables were not very significant. Finally, we found effects of social media mentions. As one can observe the train factor is the most important driver, followed by the station and service.

With a very large database we were able to build a driver model that included many potential drivers of customer satisfaction. However, the outcome of this phase also revealed three problems: First, the integration of the data was not perfect, given that we were confronted with missing values. Second, the predictive quality of the model to estimate the satisfaction for an individual passenger was still relatively low, partly because it was hard to link back performance data to the individual customer level. Third, although we were able to gather data on major drivers, we had to use existing data which had all sort of issues (for example the surveyor in the train registered a lot information only in a binary way, so either the train seats are clean or they are not, with nothing in between).

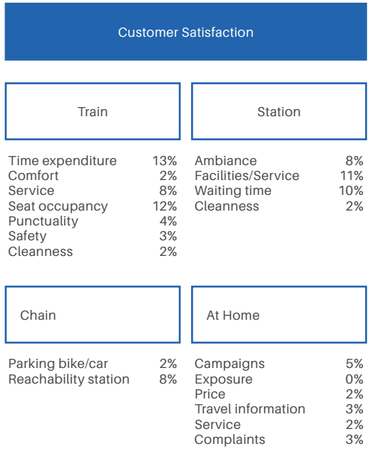

Figure 3. Importance of Drivers as result of the Multi-method approach

Survey based model and expert opinions

In this model, we critically evaluated the model as developed in the Big Data approach, with its limitations from using existing measured data only. Based on the learnings of the Big Data driver model, we now explicitly included variables that accounted for the profile of customers. The main data source in this part of the project was a survey, which was executed during 8 weeks among train travellers in 2014.

At the same time, we also looked at input from experts. We considered two target groups, that we could use to further fine tune our model: (1) Experts within the marketing, market research and intelligence departments and (2) service employees on the train. The results of this Multi-method approach are shown in figure 3.

> Hanneke van der Boog, Customer Experience Consultant NS: "Involving the Service employees on the train proved to be a very interesting and valuable step in the project. We realized that this group has a good sense on what is important for customers since they work in this context day in day out. Good example was when they told us how trains can be very dirty when they are deployed during the night and then continue directly to a morning shift, causing people to step into a train compartment full of beer and hamburgers. These sessions turned out to be valuable both ways, since the involved Service employees also got a broader understanding of all factors that influence the Customer Satisfaction (which goes further than just the train being on time)".

Overall, each method confirmed the importance of key aspects such as punctuality, seating availability, and the use/presence of facilities. The new survey-based driver model added more insights on the importance of the accessibility of the station. Overall, the input from the management and employees from the NS relatively confirmed the results of the model. The model results provided directions for how to improve customer satisfaction and specifically what the most important drivers were. In determining the priority drivers, we did not only consider the effect size (how strong a driver influences customer satisfaction), but also the potential to improve on the driver. We defined and measured the improvement potential as the extent to which a current performance differs from the maximum performance. More specifically, we calculated a potential improvement score by dividing the max score by the current realized performance. Combining the potential improvement score with the effect size, we created a new management framework (see Figure 4), which provided direction on which drivers should be targeted with customer initiatives.

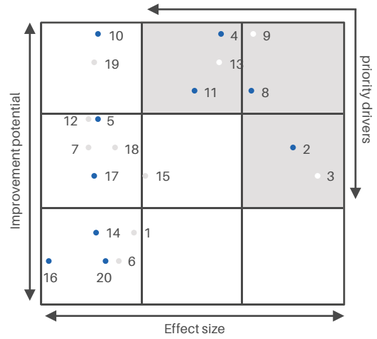

Figure 4. Combining Effect Size and Improvement Potential

Hans Peters: "Now we really understand how CSAT moves, we have much more confidence in our improvement measures and performance management, and dare to invest in a focused manner... We have set clear priorities on those fact-based drivers that matter most and can be influenced by our team. As part of this strategy we have even evaluated initiatives on the revenue and CSAT contribution"

Prediction model for customer experience

Where the first aim of the management had been to understand how they could influence customer satisfaction, a second important steering instrument was a prediction model for aggregate customer satisfaction scores. This prediction model would allow the management of NS to assess whether satisfaction score ambitions were realistic, given the expected operational service performance (i.e. punctuality).

For this phase we used four years of data from December 2006 to December 2011 which contained a total of 430.000 surveys. Based on the experiences in the prior phase and specifically the Big Data driver model, we again critically considered which variables to include in our prediction model and focused on these variables that had a strong influence, as our focus is on prediction.

The prediction model was immediately used to assess whether targets on the percentage of customers giving a 7 or higher could be achieved. This target was set at 77% for 2012. However, early 2012 there was again a strong winter problem, resulting in decreases in satisfaction. Hence, the management were keen to understand how feasible it was to still reach the target satisfaction levels towards January 2013. We used our prediction model to predict satisfaction levels under three scenario's: (1) the operational performance is the average of the last six years, (2) the operational performance is as it was in the year before, (3) the operational performance is the max of the past six years. Thereby we assumed that some other factors (i.e. the ratio between travel motives) follow the normal average. The predictions of these scenarios are given in Figure 5. As one can observe the target is unlikely to be achieved under all 3 scenarios with average achieved satisfaction levels of 74,3%, 75,0% and 76,2%. So even under the very positive scenario the target was unlikely to be achieved. Overall, the NS learned that they should set more realistic targets, that are achievable with reasonable improvements in service operations.

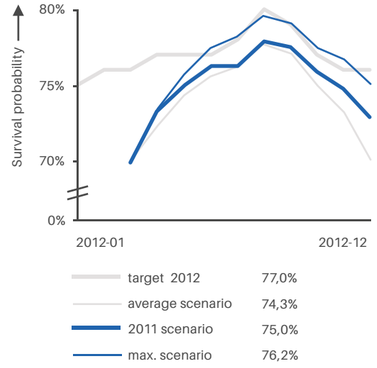

Figure 5. Predictions for 2012 after early Year Winter Problem

Thijs Urlings, Customer Experience Consultant NS: "With one single metric, an aligned set of drivers and a capability to forecast we can now centrally steer our business. With the models as starting point we have developed dashboards to monitor the scores, set realistic ambitions and involve the business through a monthly update. Now when a peak or dip occurs in the score we can quickly respond to with insight in the underlying causes."

Optimizing investments on contribution to customer experience

The fact-based approach lead to a strong support for the identified priority drivers of customers satisfaction. Interestingly, the same model could quantify the contribution of a specific customer investment to the customer satisfaction. Given that an organisation like the NS will always have limited budget, the key question is how to allocate the budget in such manner that overall customer experience uplift is maximal.

The calculation is straightforward, if you want to optimize the allocation of the budget, you should start with those initiatives that have the highest customer experience uplift per euro investment. We labelled this metric as CSAT ROI, the customer satisfaction return on investment. The following figure 6. illustrates the importance of this approach, as CSAT ROIs vary immensely across initiatives.

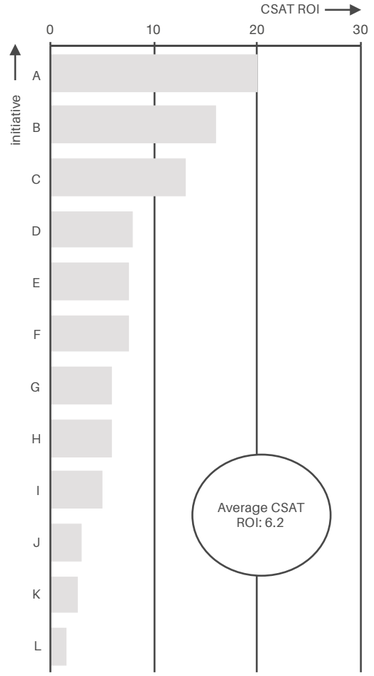

Figure 6. CSAT basis points per million euros

Hans Peters: "Although there might be always be dips during extreme winters or operational crises, we aim to remain focused on how we can make the difference in Marketing. For example, we identified the importance of available seating in the train, prioritized the operational issues with increasing seat capacity, and improved our customer communications through our seating app. We also prioritized rebalancing our social media conversations on positive versus negative topics with our customers, since that had a strong contribution to more satisfied customers. Overall we now also understand the contribution of softer emotional elements, which allows us to trade-off concrete CSAT improvement initiatives and those that manage the perception of the customer"

NS reflections on impact.

The initiative has developed into the corner stone for the NS customer-centered company strategy. More specifically the impact of the data-driven approach was experienced in four areas:

- Customer Strategy sharpened. First and foremost, the insights from this initiative have centered the current company strategy, appropriately named "traveller at number 1,2, and 3", around improving customer satisfaction whilst improving our services.
- Customer Satisfaction centrally managed. If there is debate about what the right metric is, if there is uncertainty about what the drivers are, and if you cannot forecast how CSAT could evolve over time, then there is no firm ground to centrally manage the customer satisfaction.
- Marketing Intelligence Consolidated. Through the multi-method approach the NS experienced the critical importance of collaboration between marketing research, customer analytics, and business intelligence.

Prioritization of (Marketing) initiatives. With the developed insights and tools, the NS can now isolate those drivers our organisation cannot influence, such as weather conditions and infrastructural disturbances. Instead the NS now focuses on translating those drivers that they can influence with commercial initiatives - such as wifi-in-train, better atmosphere in stations, traveller communication.

In summary, the initiative has taken out the noise from managing CSAT within the organisation of the NS, simplified the conversation with stakeholders, while focusing us on contributing to those drivers that matter most to the valued customers. As such, it has been the corner stone for putting the NS passenger on number 1, 2 and 3.