Sakila Database Analysis

DECEMBER 17

Databases and Data Visualisation Group 21
Ricardo Ramgolam 11418567
Florian Burger 12172812
Gabriella Veda
Bruno Sotic 11353473

Contents

Case Description

Business problem

For this assignment, we must look at the Sakila database, find, and then solve issues within the company's business model. For this, we first must investigate the structure of the database, construct KPIs based on the Balanced Scorecard, create a Data Warehouse and the according schema; visualize and analyze the data with regards to our KPIs, and finally summarize/discuss our findings which can then hopefully be used to benefit the company in the form of strategic advice.

Dataset

The Sakila dataset is a fictitious dataset about a DVD rental company. The company has two stores, in Australia and Canada; and rents out DVDs globally. For each store, there is only one employee that handles all aspects of the store from categorizing DVDs in the inventory to sending out the DVDs to the customers. Based on this, we assume that this is an online business operating from these two stores but not offering in-store rental. The dataset includes information about the company's revenue, movie actors, customers, and multiple internal business processes such as inventory.

Database

The original database is shown in figure 1 below and it is a normalized schema modeling the DVD rental company. It is split into four different parts which are Customer Data, Inventory, Business, and Views.

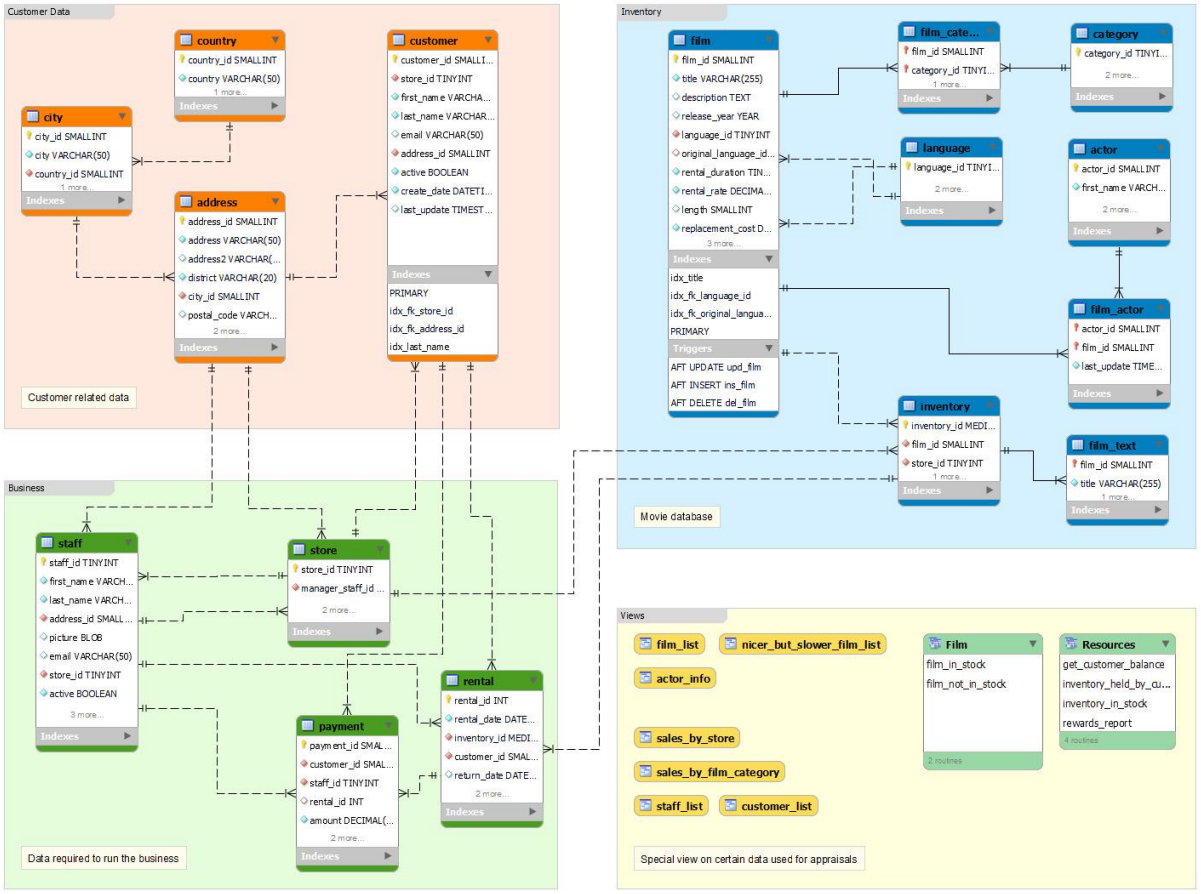


Figure 1. The schema of the original database

The table below shows the most relevant measurements in our dataset. Some values appear in multiple tables but were only described once as the description for the measurement would not change.

Measurement	Description
Actor ID	A number given to each actor that plays in a movie within our inventory
Category ID	A number given to each category of movies within our inventory
Name (Category)	Name of the category
Address ID (Customer)	A number given to each address within the system

Customer ID	A number given to each customer within our system
First Name (Customer)	The first name of the customer
Last Name (Customer)	The last name of the customer
Country, City	The location of each customer specified per country and city
Postal Code	Postal code per customer
Description (Film)	A description of each movie within our stock
Film ID	A number given to each film within our inventory
Title	The name of each film
Rating (Film)	The age rating for each film
Release Year (Film)	The year the film was released
Length (Film)	The length of each film in minutes
Rental Duration	The length of the rental duration in days
Rental Rate	The cost of renting the film for the rental duration
Replacement Cost	The cost charged to the customer in case a DVD is destroyed/damaged
Payment ID	A number given to each payment conducted
Payment Date	The date the payment was processed
Amount (Payment)	The amount that was paid per payment
Inventory ID	A number given to the item being rented
Rental Date	Start date of when the film was rented
Return Date	Date when the film was returned

Rental ID	A number that is given to each rental
Rental Days	The number of days a movie was rented
Staff ID	A number identifying each staff member
First Name (Staff)	First name of the staff member
Last Name (Staff)	Last name of the staff member
Address ID (Staff)	The address of each staff member
Store ID	A number identifying each store

Balanced Scorecard

We are using the balanced scorecard approach (Kaplan & Norton, 2005) to analyze how well the business is performing. The balanced scorecard approach combines strategic strategies with financial aspects to return a holistic view of a business. It includes four different perspectives which should be used to create relevant key performance indicators (KPI). These four perspective are:

- Financial
- Customer & Stakeholder
- Internal Business Processes
- Learning & Growth

Using this approach will allow us to focus on the company as a whole which means that we also have more areas which can potentially be improved.

KPIs

Financial Dimension

Since the data provided only contains a few months of usable data, we have opted to analyse the financial structure instead of the financial forecast of the company.

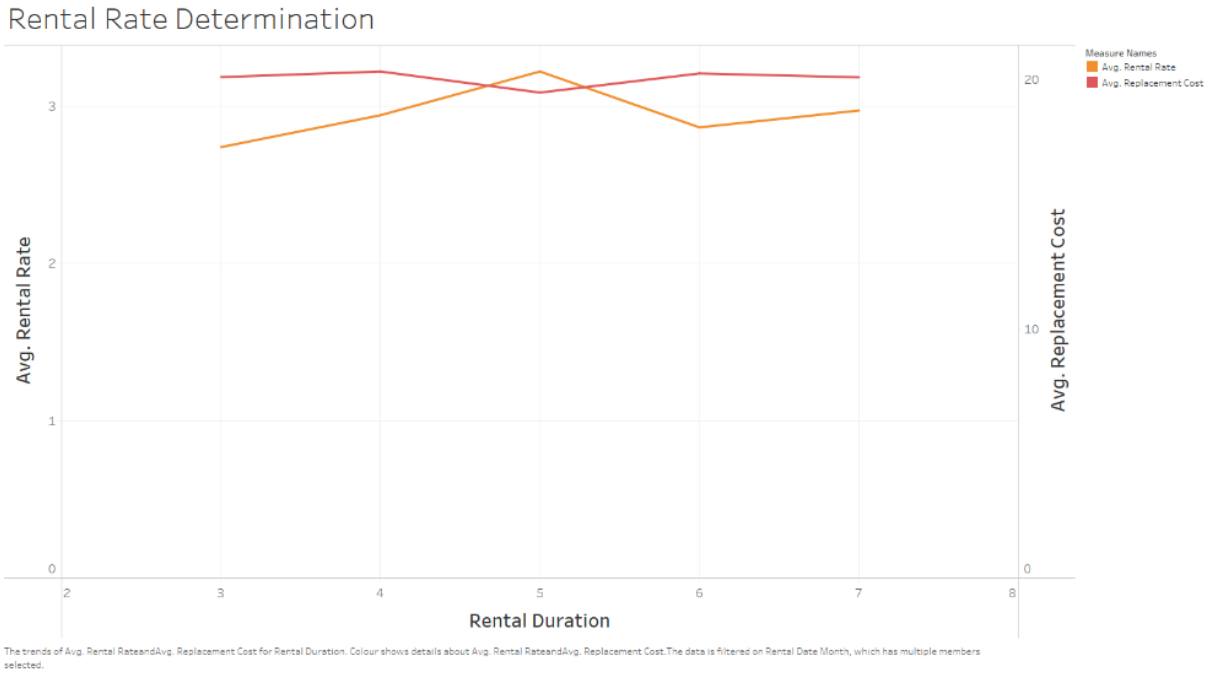
Geographical Revenue

Geographical Revenue indicates how much revenue is generated per country in which the company operates. Most revenue comes from the South Asian region, i.e., from India and China. Those two countries alone account for roughly 10% of the total revenue. The second biggest markets are the United States and Russia, followed by Brazil. European and African countries make up for a minor amount of the revenue respectively. Finally, the continent of Australia contributes almost nothing to the company's revenue.

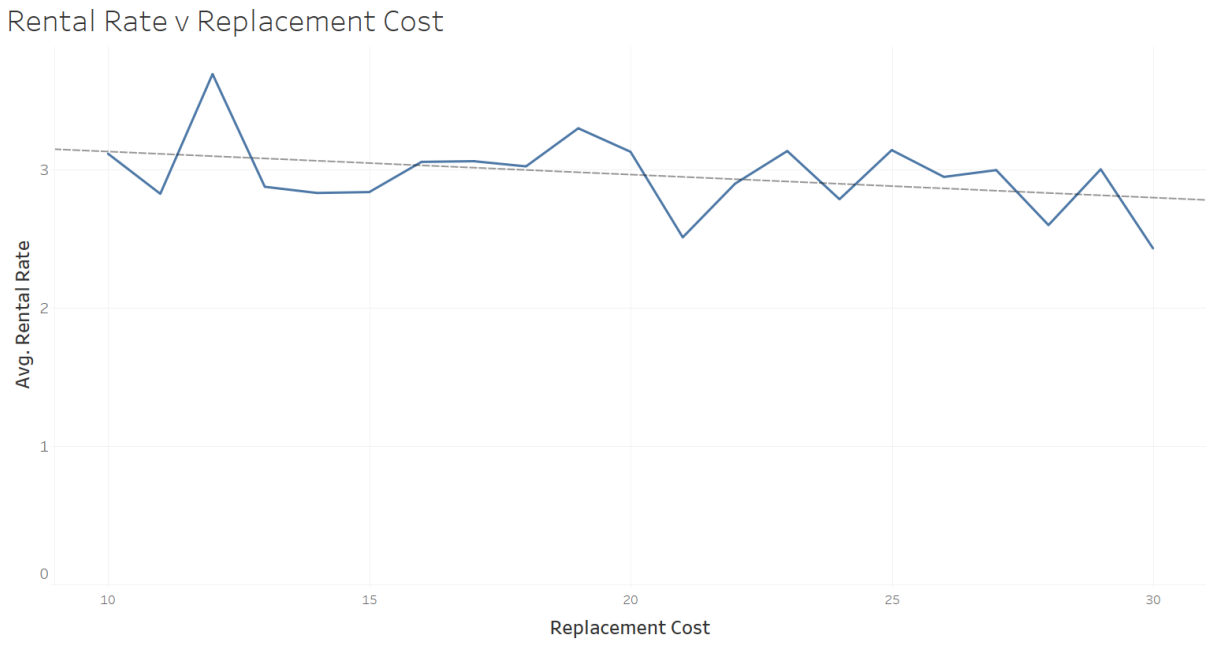
Penalty Rate

We analysed the sources of revenue to be able to provide the company with insights on how to improve its financial performances. The main source of revenue for the company is the payment for DVD rentals. The amount a customer must pay for its rental is determined when it returns its rental to the company. This amount is made up of 1) the rental rate, and 2) the late fee (when applicable). We will further examine these two elements to properly analyse revenue.

The company charges a rental rate per DVD, which is pre-determined for each film. Each DVD is presumably purchased at replacement cost. With the rental duration set per film basis, it would make sense to see a positive relationship between rental rate and rental duration, since you'd most likely charge more the longer you rent your film. However, this is not necessarily the case when we analysed the data. What was more surprising, is that rental rate seems to be unrelated to replacement cost as well. We were expecting a positive relationship between rental rate and replacement cost since the company would want to manage for the contribution margin.



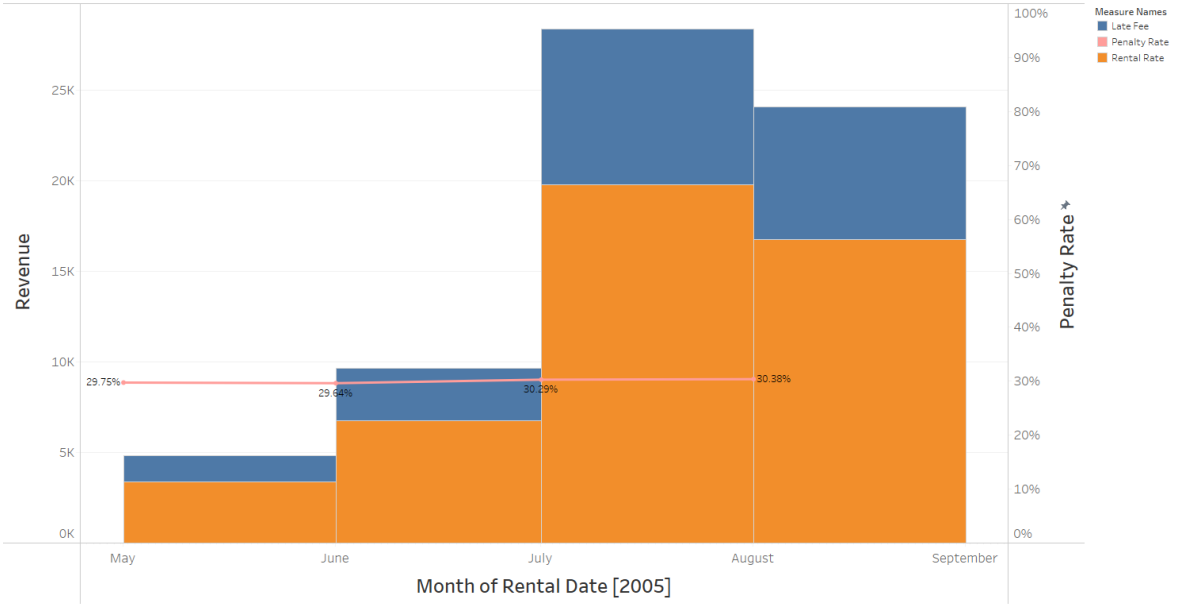
The graph above shows the relationship between rental duration, average rental rate and average replacement cost. As shown by the graph, the relationship between average replacement cost and rental duration is almost flat which means the rental duration set by the company is independent from the replacement cost of a DVD. Also, the average rental rate doesn't seem to differ much when controlling for rental duration. This would suggest the company does not use replacement cost and rental duration as determinants of the rental rate it charges. As a result, this would indicate mismanagement of assets since there does not seem to be any business strategy involving the determination of the rental rate. To further analyse this, we graphed average rental rate against replacement cost to see if there is any relationship.



There appears to be a slightly negative relationship between average rental rate and replacement cost, however, it's not a significant one. This finding goes against any financial logic, since you would expect rental rate to be higher when the cost of replacement goes up or the duration of the rental increases.

The second part of revenue from rentals appears in the form of late fees. Late fees are charged when customers exceed the number of days set as the rental duration by the company. To determine the effect of these fees on the company's financials, we examine its rate compared to monthly revenue.

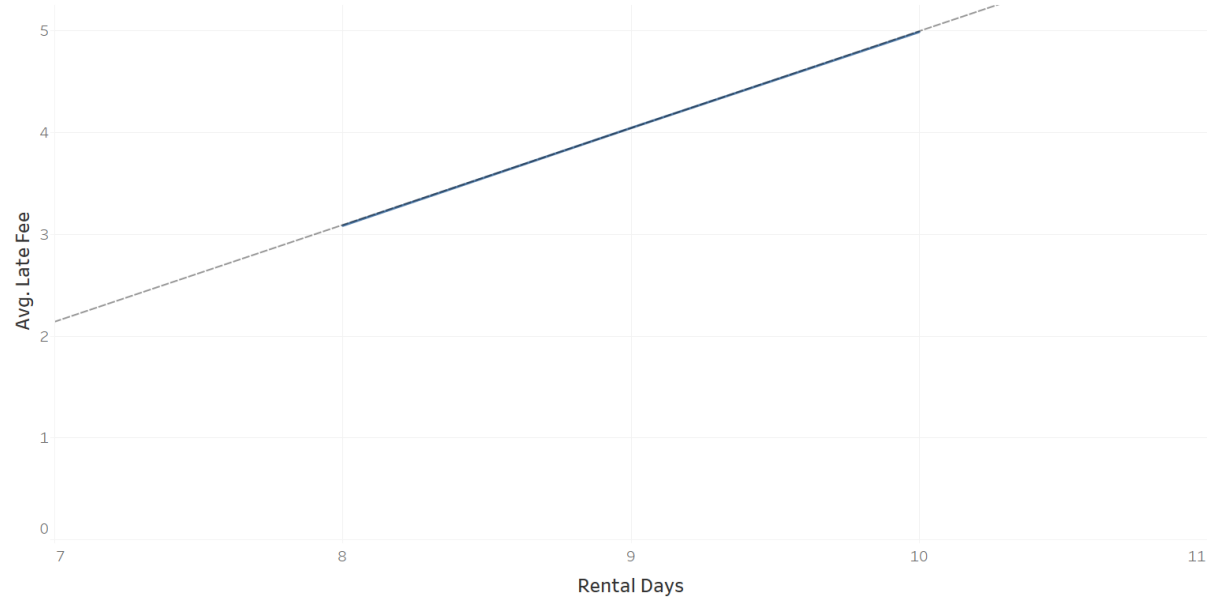
Effect of Penalty on Revenue



The trends of Late Fee,Rental Rate,Penalty RateandPenalty Rate for Rental Date Month. Colour shows details about Late Fee,Rental RateandPenalty Rate. For pane Penalty Rate: The marks are labelled by Penalty Rate.The data is filtered on Rental Date Month, which has multiple members selected.

On the graph, the average penalty rate is about 30% and stays steady over the analysed period. We can conclude that the penalty fees paid by customers have a big impact on revenue and, therefore, it is important to analyse the dynamics of late fees. Firstly, we want to analyse how the late fee is determined. Accordingly, we plot the average late fee against rental days to find a possible relationship between them.

Late Fee

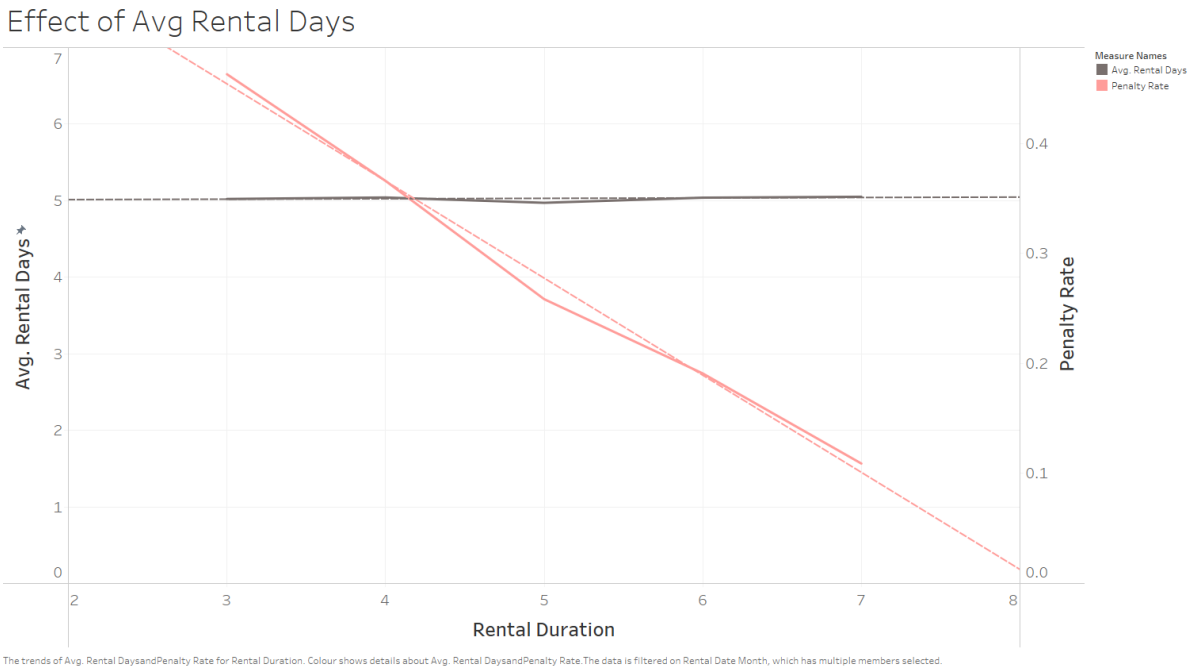


The trend of average of Late Fee for Rental Days.The data is filtered on Rental Date MonthandPenalty as an attribute.The Rental Date Month filter has multiple members selected.The Penalty as an attribute filter ranges from 1 to 1.

For this graph, we filtered out the rentals who did not have a penalty in order to properly estimate the relationship between the variables. It is evident that the penalty is set at a flat 1 unit per day over the set rental duration. Thus, revenue for the company can be formulated as: $R = \sum P_r + D_{R-D}$, where R is the revenue, P_r the rental rate set by the company and D_{R-D} are the days over the rental duration. The equation presents another important factor in the determination of the revenue for the company: the average rental days.

Average Rental Days

Since late fees dependent on the number of rental days over the original rental duration, it is important to analyse the average rental duration. We present this analysis by comparing the average rental days against the rental duration to get an understanding of their relationship.



According to the figure above, the average rental duration is 5 days, independent from the original rental duration. Also, the penalty rate appears to have a near perfect negative linear relationship with the rental duration. Intuitively this makes sense: the number of rentals who are past due when the rental duration is 3, is greater than when rental duration is 7, because the average of rental days is equal for every rental duration. For this reason, the average rental days is a way to control for the penalty rate and therefore a necessary KPI for management.

Internal Business Dimension

Inventory

Performance per category vs. DVDs available shows the performance of each film category versus the number of DVDs available in the inventory per category. The purpose of this graph is to compare and assess whether the company has a proportional number of DVDs to the customer demand. Looking at the overall trend of the chart, the number of DVDs available in the inventory is proportionate to the performance of each film category. There are a few exceptions such as with the Comedy category where customer demand is slightly higher than the inventory stock, but the proportional difference is not significant in general.

Turnover Rate

Replacement Cost vs. Rental Rate shows the cost of replacing DVDs versus the rental rate of the DVDs. One would expect that as the value of the DVD (replacement cost) increases, so would the rental rate but this graph shows the opposite trend. While this trend was not significant (p=.15) one can clearly tell that the trend is against the expected direction. Based on this, increasing the rental price proportionally with the replacement cost would be a logical step. However, there are limitations to this as we currently do not possess the data necessary to support this decision further. We do not possess data showing whether there is a significant loss of DVDs. Additionally, we do not know whether customers are billed automatically for the loss of a DVD. Thus, while it would make sense to increase the rental price in line with the replacement cost, we are missing additional data that could further support this decision.

Customer Dimension

Customer Preferences

Customer Preferences in terms of film categories can be used to identify patterns across film categories throughout the year. An interesting finding is that across all categories is that the month of July scores the highest number of rentals. A reason for this occurrence might be because the month of July is right in the middle of the summer break when the summer vacation is in full swing, so people have more free time to watch movies. It is also interesting to note that

the film category with the largest number of rentals in July is Sports, followed by Animation, Action, Sci-fi, Drama, and Family movies. These top categories in July also make sense because most of them can be considered as blockbusters, which are usually released and/ or are popular during the summer. August is the second most popular month across all film categories, followed by June, May, and February in the last place. The top 3 months July, August, and June are all summer vacation months. Considering this observation, it is advisable for the company to promote and advertise more during the months outside of the summer vacation, in the hope that the amount of film rentals can be relatively stable and high throughout the year, leading to a more stable income.

Growth Dimension

The learning & growth perspective is mainly concerned with factors that increase employee performance. As of yet, the only available data for this perspective are the ‘sales per employee’, which is also reflected in ‘sales per store’ (since the company only has 2 employees, one for each store). As discussed earlier, the revenue generated by the respective employees differs only slightly; and this dataset is missing information on factors that can be used to assess employee performance more accurately. Since the company only gathers basic personal information with regards to its employees, this report emphasizes that the company should also keep track of other metrics that could be used to shine a brighter picture on employee performance and satisfaction. This will become especially relevant in the near future, given that the company operates on a global scale and sales are rising by the month.

Recommendations

In today’s increasingly competitive market, it is essential for companies to innovate in order to keep up with evolving demand. Despite earning acceptable levels of revenue, there are a number of changes that Sakila can implement to improve their position in the global market. After analyzing the balanced scorecard perspectives through key performance indicators and gaining insight of the company’s current business model, we propose a new and improved business model.

Our first suggestion concerns the financial aspect of the business model. Currently, approximately 30% of Sakila's annual revenue comes from penalty rates, i.e., late fees for going over the specified rental duration. Although penalty rates generate income, we propose a better way to earn revenue in regard to rental duration. We recommend the company to change the rental procedure from the current pay-by-movie approach into a monthly subscription system. With a monthly subscription system, customers can rent an unlimited number of DVDs, and the company could impose a certain number of DVDs as a maximum per batch of shipment. Customers are able to create a watchlist where they can queue the movies that they want to watch next, and these DVDs that are next in line will be delivered in the next shipment batch after the customers return the DVDs from their previous batch. A monthly subscription system will benefit both the customers and the company. It is a lot more convenient and accessible for customers to rent movies, while for the company it is cost-effective and allows for revenue scaling and forecasting. The subscription system will also address the penalty issue, since customers pay for Sakila's services up front every month instead of paying for each DVD they rent and being subjected to a fee every time they return late DVDs.

A second observation of the financial aspect that we identified is that the continent of Australia contributes almost nothing to the company's revenue. We suggest the company to either further investigate why this is the case or change the distribution channel to a more beneficial location instead.

Next, for the customer aspect, we focused on customer preferences of film categories and on the customers with most rentals. Our analysis yielded results that suggest the company to promote and advertise more during the months outside of the summer vacation and advertise less popular categories, in the hope that the amount of film rentals can be relatively stable and high throughout the year, leading to a more stable income. Another thing we identified is the top paying customers, or in other words, customers with most rentals. This information can give us insight as to which customers can be rewarded with special promotions and also to which customers we should increase promotions and advertisements to get them to rent more DVDs.

Another suggestion that could improve customer experience is to implement personalization where customers receive movie recommendations based on their rental history. Recent business research and practices have shown that

personalization leads to customer loyalty because it builds upon trust, commitment, and reciprocity – fundamentals of customer-brand relationship (Nham, 2021). Therefore, generating personalized recommendations will improve the customer experience and make customers stay engaged to Sakila for their movie rental needs.

This leads us to the aspect of internal business processes. We recommend Sakila to launch a website where customers can view their personalized recommendations and easily browse through Sakila's offers. The interactive website should be user-friendly by having clear and helpful tabs such as film genres, watchlists, and new arrivals. Another observation concerning the internal business process is that there is no data on whether there is a significant loss of DVDs or whether customers are billed automatically for the loss of a DVD. We suggest the company collect data on missing DVDs and specify the regulations and procedure when losses occur. This information will help the company decide to increase the rental price in line with the replacement cost or not. Our analysis also finds that actors playing in many movies bring in less revenue. Thus, we suggest Sakila to focus on actors with less movies instead.

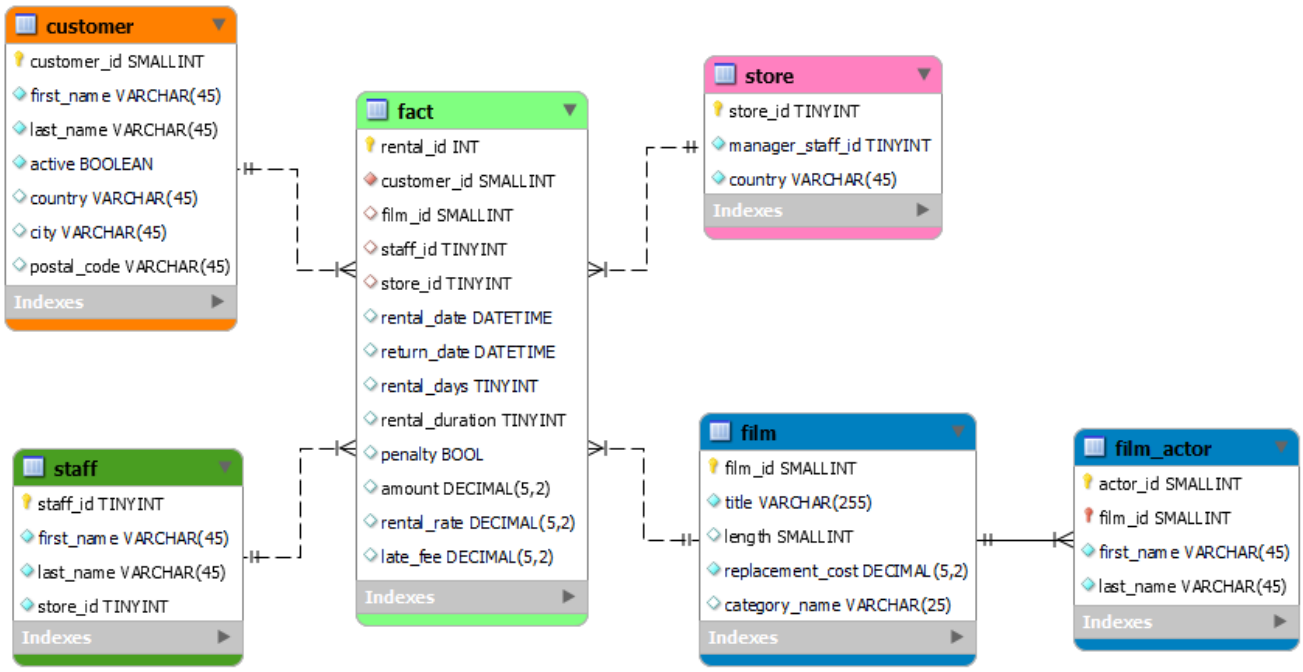
Regarding the learning and growth perspectives, we analyzed the dimensions of staff and store. It is a necessity for companies to pay attention to not only employee performance in stores, but also to their satisfaction. In fact, research has shown that companies with high numbers of satisfied employees have increasing annual profits (Marin, 2021). Thus, we recommend Sakila to implement an employee rewards system to boost both employee performance in stores and employee satisfaction.

Going forward, we recommend Sakila to begin shifting towards the business model we propose in this paper. It is also important to note that streaming services are the most popular way for customers to enjoy movies nowadays. This is a fact that Sakila should take into account and adapt accordingly as their business grows into the future. As a final recommendation, Sakila should adopt the new proposed business model and grow a loyal customer base, then utilize the gained resources to venture into the streaming service market. It is essential that Sakila implements the proposed improvements first instead of directly switching to a streaming platform because it will be difficult to compete with both

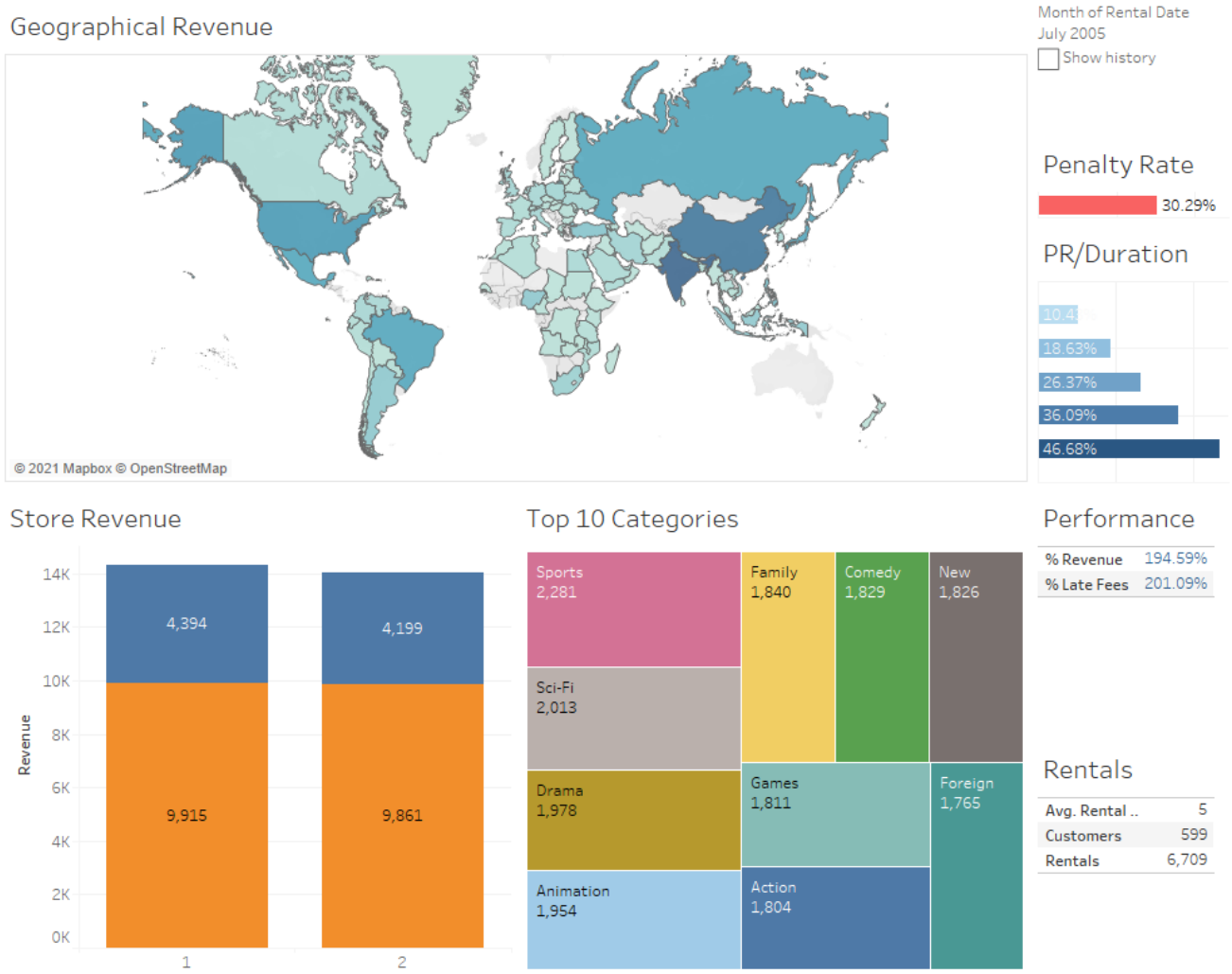
existing streaming services and in the DVD rental market with insufficient resources to do so.

Data Warehouse Design

We designed a data warehouse in order to create a dashboard and conduct our analysis. The figure below shows our finale diagram which we used in MySQL to create our schema. Relevant SQL scripts to create the data warehouse from the original Sakila database are provided with this report in separate files.



Dashboard



In order to help manage the different KPIs, we have created a dashboard which shows all relevant information in a timely way. The dashboard helps management to quickly identify key issues and control measures. The Tableau dashboard is included in a separate file with this report.

References

Kaplan, R. S., & Norton, D. P. (2005). The balanced scorecard: measures that drive performance. *Harvard business review*, *83*(7), 172.

Marin, M.J. (2021). Employee Satisfaction and Company Success. *Gateway Gives Back.* https://gatewaycfs.com/education/employee-satisfaction/

Nham, L. (2021, March 25). How To Effectively Use Personalization To Boost Customer Loyalty. *3 tier logic*. https://www.3tl.com/blog/personalization-and-customer-loyalty

Kaplan, R. S., & Norton, D. P. (2005). The balanced scorecard: measures that drive performance. *Harvard business review*, *83*(7), 172.