2021

Data Dictionary

ROCKBUSTER DATABASE

CAROLINA PACHON BUITRAGO

Table of Contents

ERD 0

Fact Table 1

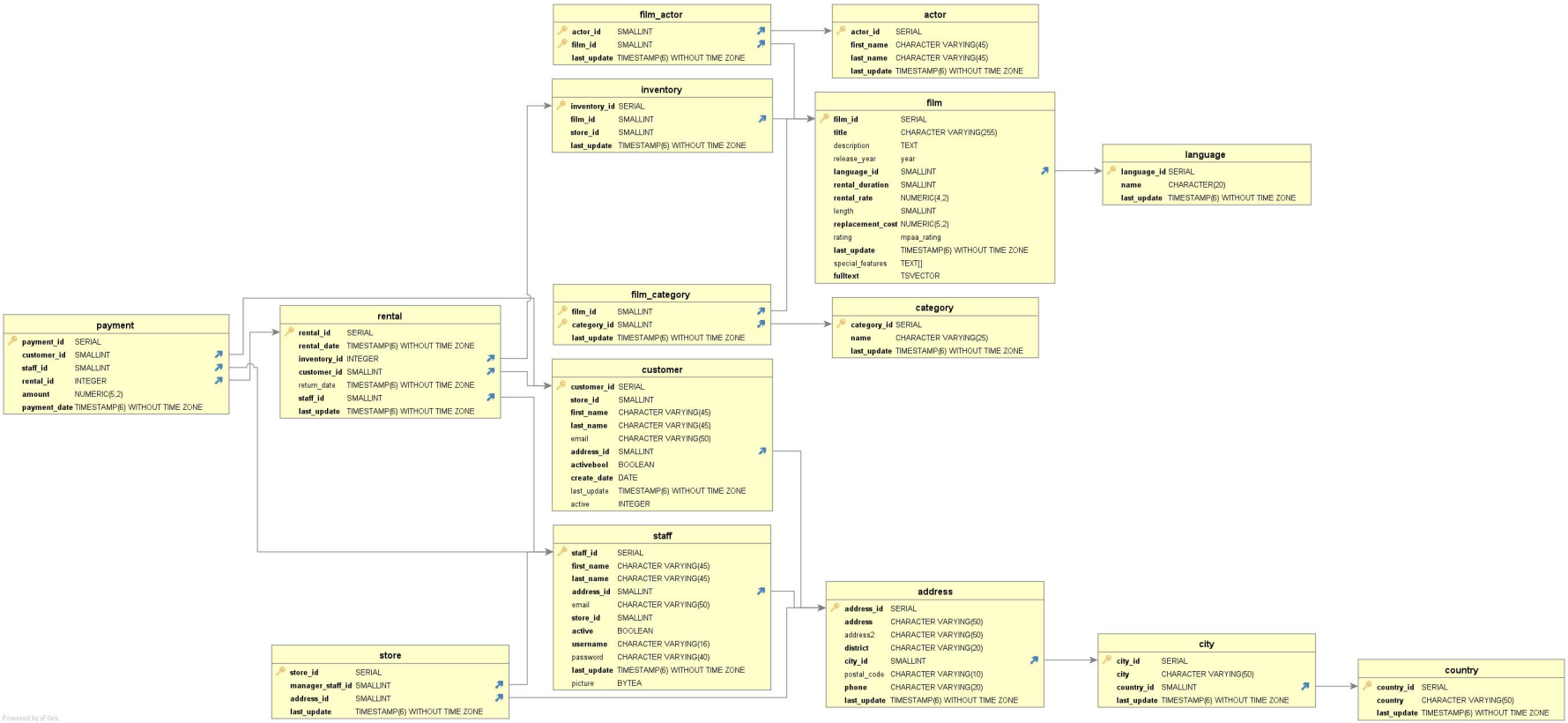
- Payment 1

Dimension tables 1

- Rental Table 1
- Film_actor 2
- Inventory 2
- Film Category 2
- Customer 3
- Staff 3
- Store 4
- Actor 4
- Film 4
- Category 5
- Address 5
- Language 5
- City 6
- Country 6

VIZ in Tableau 6

ERD


film_actor
actor_id SMALLINT
film_id SMALLINT
last_update TIMESTAMP(6) WITHOUT TIME ZONE
actor
actor_id SERIAL
first_name CHARACTER VARYING(45)
last_name CHARACTER VARYING(45)
last_update TIMESTAMP(6) WITHOUT TIME ZONE
inventory
inventory_id SERIAL
film_id SMALLINT
store_id SMALLINT
last_update TIMESTAMP(6) WITHOUT TIME ZONE
film
film_id SERIAL
title CHARACTER VARYING(255)
description TEXT
release_year year
language_id SMALLINT
rental_duration SMALLINT
rental_rate NUMERIC(4,2)
length SMALLINT
replacement_cost NUMERIC(5,2)
rating mpaa_rating
last_update TIMESTAMP(6) WITHOUT TIME ZONE
special_features TEXT[]
fulltext TSVECTOR
language
language_id SERIAL
name CHARACTER(20)
last_update TIMESTAMP(6) WITHOUT TIME ZONE
payment
payment_id SERIAL
customer_id SMALLINT
staff_id SMALLINT
rental_id INTEGER
amount NUMERIC(5,2)
payment_date TIMESTAMP(6) WITHOUT TIME ZONE
rental
rental_id SERIAL
rental_date TIMESTAMP(6) WITHOUT TIME ZONE
inventory_id INTEGER
customer_id SMALLINT
return_date TIMESTAMP(6) WITHOUT TIME ZONE
staff_id SMALLINT
last_update TIMESTAMP(6) WITHOUT TIME ZONE
film_category
film_id SMALLINT
category_id SMALLINT
last_update TIMESTAMP(6) WITHOUT TIME ZONE
category
category_id SERIAL
name CHARACTER VARYING(25)
last_update TIMESTAMP(6) WITHOUT TIME ZONE
customer
customer_id SERIAL
store_id SMALLINT
first_name CHARACTER VARYING(45)
last_name CHARACTER VARYING(45)
email CHARACTER VARYING(50)
address_id SMALLINT
activebool BOOLEAN
create_date DATE
last_update TIMESTAMP(6) WITHOUT TIME ZONE
active INTEGER
staff
staff_id SERIAL
first_name CHARACTER VARYING(45)
last_name CHARACTER VARYING(45)
address_id SMALLINT
email CHARACTER VARYING(50)
store_id SMALLINT
active BOOLEAN
username CHARACTER VARYING(16)
password CHARACTER VARYING(40)
last_update TIMESTAMP(6) WITHOUT TIME ZONE
picture BYTEA
address
address_id SERIAL
address CHARACTER VARYING(50)
address2 CHARACTER VARYING(50)
district CHARACTER VARYING(20)
city_id SMALLINT
postal_code CHARACTER VARYING(10)
phone CHARACTER VARYING(20)
last_update TIMESTAMP(6) WITHOUT TIME ZONE
city
city_id SERIAL
city CHARACTER VARYING(50)
country_id SMALLINT
last_update TIMESTAMP(6) WITHOUT TIME ZONE
country
country_id SERIAL
country CHARACTER VARYING(50)
last_update TIMESTAMP(6) WITHOUT TIME ZONE
store
store_id SERIAL
manager_staff_id SMALLINT
address_id SMALLINT
last_update TIMESTAMP(6) WITHOUT TIME ZONE
Powered by yFiles

Fact Table

Payment

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	Payment_id	SERIAL	Primary Key for the fact table. Unique identifier for the payment table
FK	Custumer_id	SMALLINT	Foreign key to customer table. Unique identifier for customer table.
FK	Staff_id	SMALLINT	Foreign key to staff table. Unique identifier for staff table.
FK	Rental_id	INTEGER	Foreign key to rental table. Unique identifier for rental table.
	Amount	NUMERIC (5,2)	Amount of money by which the transaction was made.
	Payment_date	TIMESTAMP(6) WITHOUT TIMEZONE	Date when the transaction was made

Dimension tables

Rental Table

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	rental_id	SERIAL	Primary Key for the rental table. Unique identifier for the rental of each movie
	Rental_date	TIMESTAMP(6) WITHOUT TIMEZONE	Date when a particular movie was rented
FK	Inventory_id	INTEGER	Foreign key to inventory table. Unique identifier for inventory item.
FK	customer_id	INTEGER	Foreign key to customer table. Unique identifier for each customer.
FK	Staff_id	NUMERIC (5,2)	Foreign key to staff table. Unique identifier for each staff worker.
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Film_actor

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	actor_id	SMALL INT	Foreign Key for the actor table. Unique identifier for each actor.
FK	Film_id	SMALL INT	Foreign Key for the film table. Unique identifier for each movie.
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Inventory

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	inventory_id	SERIAL	Primary key to inventory table. Unique identifier for inventory item.
FK	Film_id	SMALL INT	Foreign Key for the film table. Unique identifier for each movie.
FK	store_id	SMALL INT	Foreign Key for the store table. Unique identifier for each store.
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Film Category

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	Film_id	SMALL INT	Foreign Key for the film table. Unique identifier for each movie.
🔑	category_id	SMALL INT	Foreign Key for the category table. Unique identifier for each of the categories in movies
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Customer

	COLUMN NAME	DATA TYPE	DESCRIPTION
	Customer_id	SERIAL	Primary key to customer table. Unique identifier for each customer.
FK	store_id	SMALL INT	Foreign Key for the store table. Unique identifier for each store.
	First_name	CHARACTER VARYING (65)	First name of a customer
	last_name	CHARACTER VARYING (65)	Last name of a customer
	email	CHARACTER VARYING (50)	email of a customer
FK	Address_id	SMALL INT	Foreign Key for the address table. Unique identifier for the address of each store where the customer is renting.
	Activebool	BOOLEAN	True or False
	Create_date	Date	Date when the record for this customer was created
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated
	Active	INTEGER	1=active customer 0=inactive customer

Staff

	COLUMN NAME	DATA TYPE	DESCRIPTION
	Staff_id	SERIAL	Primary key to staff table. Unique identifier for each staff worker.
	First_name	CHARACTER VARYING (65)	First name of the staff member
	last_name	CHARACTER VARYING (65)	Last name of the staff member
FK	address_id	SMALL INT	Foreign Key for the address table. Unique identifier for the address of each store where the member staff is located.
	email	CHARACTER VARYING (50)	email of a customer
FK	store_id	SMALL INT	Foreign Key for the store table. Unique identifier for each store.
	Activebool	BOOLEAN	True or False
	Username	CHARACTER VARYING (16)	Username of the staff member
	password	CHARACTER VARYING (40)	Password related to the username
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Store

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	store_id	SERIAL	Primary Key for the store table. Unique identifier for each store.
	Manager_staff_id	SMALL INT	Unique identifier the manager of each store
FK	address_id	SMALL INT	Foreign Key for the address table. Unique identifier for the address of each store.
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Actor

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	actor_id	SERIAL	Primary Key for the actor table. Unique identifier for each actor.
	First_name	CHARACTER VARYING (45)	First name of the actor
	last_name	CHARACTER VARYING (45)	Last name of the actor
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Film

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	Film_id	SERIAL	Primary Key for the film table. Unique identifier for each movie.
	Title	CHARACTER VARYING (255)	Title of the movie
	Description	TEXT	Description of what the movie is about
	Release_year	Year	Year when the movie was released
FK	Language_id	SMALLINT	Foreign Key for the language table. Unique identifier for each language.
	rental_duration	SMALLINT	Duration of time for a rental of a movie
	Rental_rate	NUMERIC (4,2)	Price for renting certain movie
	length	SMALLINT	Duration in time of a film
	Replacement_cost	NUMERIC (5,2)	Cost of the film in case is not returned or gets lost
	Rating	Mpaa_rating	Scale of rating according to the audience of the film
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated
	Special_features	TEXT[]	Extra material as “Deleted Scenes” or trailers
	Fulltext	TSVECTOR	Other information

Category

	COLUMN NAME	DATA TYPE	DESCRIPTION
	category_id	SMALL INT	Primary Key for the category table. Unique identifier for each of the categories in movies
	name	CHARACTER VARYING (45)	Name of the category
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Address

	COLUMN NAME	DATA TYPE	DESCRIPTION
	Address_id	SMALL INT	Primary Key for the address table. Unique identifier for the address of store.
	Address	CHARACTER VARYING (50)	Address of the store
	Address2	CHARACTER VARYING (50)	Another address
	District	CHARACTER VARYING (20)	District where the address is located
FK	City_id	SMALLINT	Foreign Key for the city table. Code of the city where the address is located
	Postal_code	CHARACTER VARYING (10)	Postal code of the address
	Phone	CHARACTER VARYING (10)	Phone corresponding to that particular address
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Language

	COLUMN NAME	DATA TYPE	DESCRIPTION
	Language_id	SERIAL	Primary Key for the language table. Unique identifier for each language.
	name	CHARACTER (20)	Name of the language
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

City

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	City_id	SERIAL	Foreign Key for the city table. Code of the city where the address is located
	City	CHARACTER VARYING (50)	Name of the city
FK	Country_id	SMALLINT	Foreign Key for the country table. Code of the country where the address is located
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

Country

	COLUMN NAME	DATA TYPE	DESCRIPTION
🔑	Country_id	SERIAL	Primary Key for the country table. Code of the country where the address is located
	Country	CHARACTER VARYING (50)	Name of the country
	last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date when this table was last updated

VIZ in Tableau

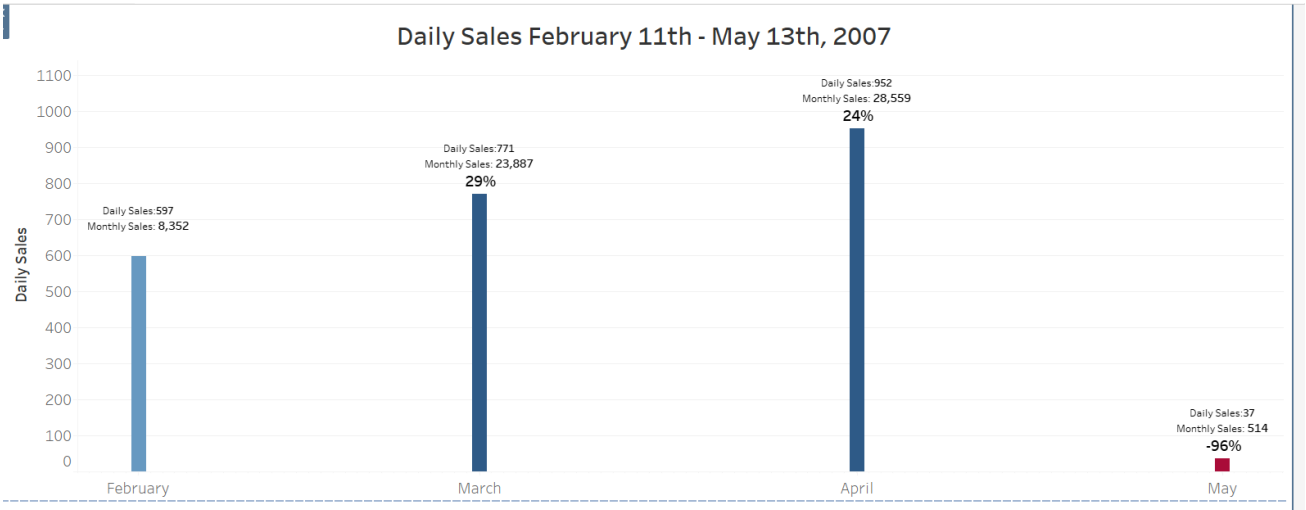
https://public.tableau.com/app/profile/carolina.pachon/viz/RockbusterAnalysis/Story1

In case the storyboard is displayed in a messy way I added the screenshots of the dashboards.

Rockbuster Stealth LLC is a movie rental company that used to have stores around the world. Facing stiff competition from streaming services such as Netflix and Amazon Prime, the Rockbuster Stealth management team is planning to use its existing movie licenses to launch an online video rental service in order to stay competitive.

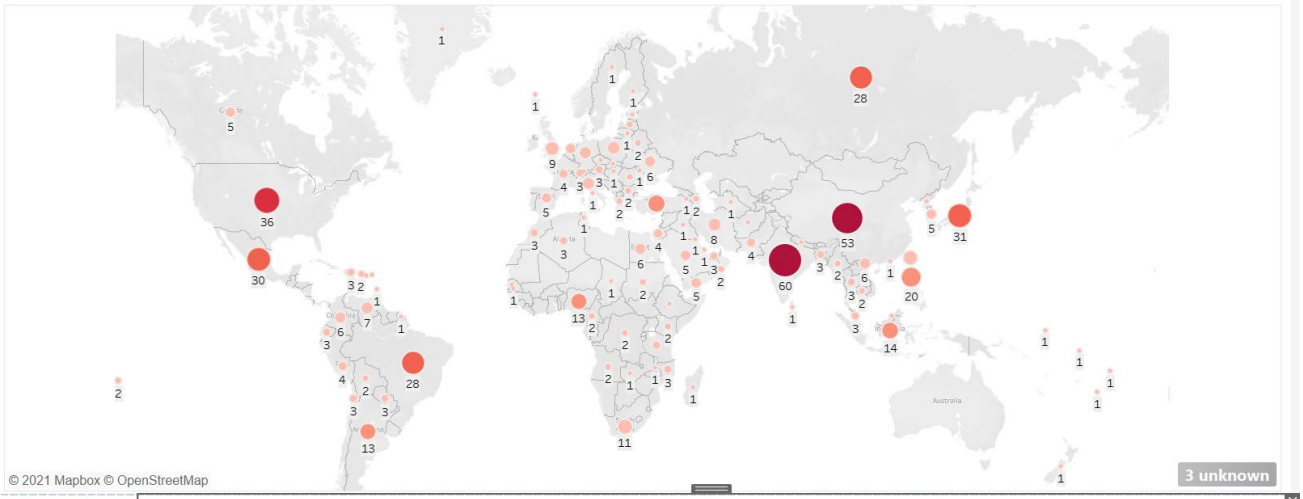
Key Questions and Objectives

- Which movies contributed the most/least to revenue gain?
- What was the average rental duration for all videos?
- Which countries are Rockbuster customers based i n?
- Where are customers with a high l ifetime value based?
- Do sales figures vary between geographic regions?



Total Sales account for $61,312 for a period of 13 weeks.
When compared the behavior of weekly sales, there is a percentage change with respect to the previous month's daily sales of -96% , which raises a red flag.

Sales and Count of Customers by Country



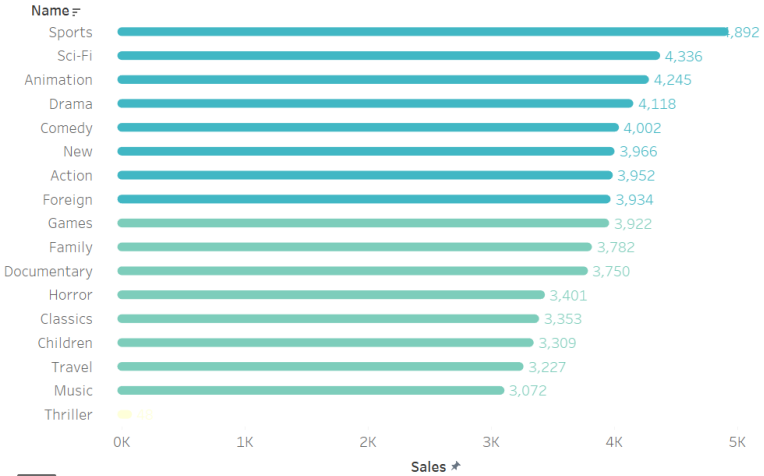
Country's Sales
48 6,035

Rockbuster has a **customer base of 599** customers all around the world with its highest concentration in India, China, United States, Japan, Mexico, Russia and Brasil.
As well most of its sales concentate around these countries.

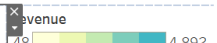
10 Best and 10 Worst Film Contribution to Revenue

Title	Number Cop..	Rental Count	Revenue
Telegraph Voyage	7.0	25.0	216.0
Zorro Ark	8.0	28.0	200.0
Wife Turn	8.0	27.0	199.0
Innocent Usual	8.0	26.0	192.0
Saturday Lambs	8.0	26.0	191.0
Hustler Party	8.0	22.0	191.0
Titans Jerk	8.0	27.0	187.0
Harry Idaho	8.0	27.0	178.0
Torque Bound	8.0	23.0	170.0
Dogma Family	8.0	28.0	169.0
Lights Deer	2.0	7.0	8.0
Japanese Run	2.0	6.0	8.0
Young Language	2.0	7.0	7.0
Treatment Jekyll	2.0	6.0	7.0
Rebel Airport	2.0	7.0	7.0
Cruelty Unforgiv..	2.0	6.0	7.0
Texas Watch	2.0	6.0	6.0
Oklahoma Juma..	2.0	6.0	6.0
Freedom Cleopat..	2.0	5.0	6.0
Duffel Apocalypse	2.0	6.0	6.0

Revenue by Category



In general most of the categories share the same popularity except for the thriller genre.
The most popular movies had a rental count at least three times superior than the movies that performed the worst. In terms of revenue, the best movies outperformed in sales at least 25 times. This is crucial when Rockbusters decides on what will be the future portfolio it will offer to its clients.



After seeing the whole picture of the business and before defining which customers matter the most, let's understand the retention rate and analyse how many customers kept on purchasing throughtout the 13 weeks of reported sales.

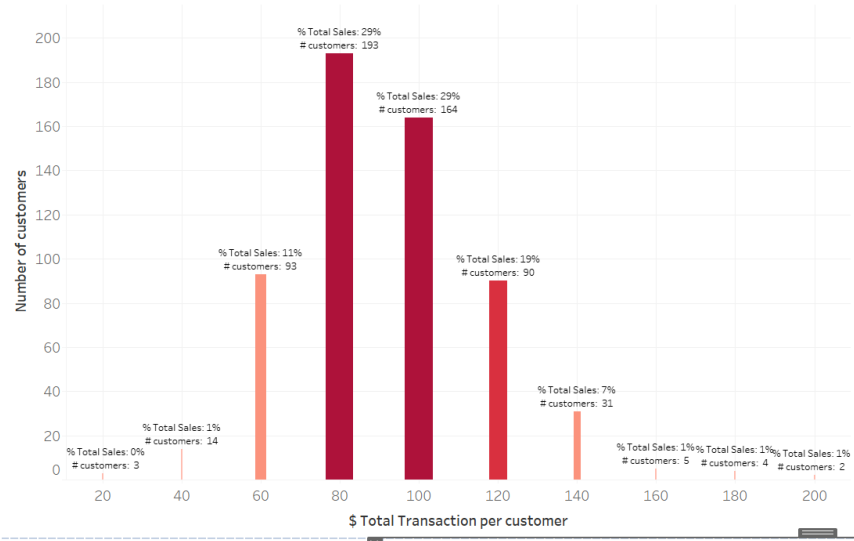
Cohort Analysis

Week of Cohort	Week Number 0	1	2	3	4	5	6	7	8	9	10	11	12	13
11 February, 2007	433	381	381		286	431		342	428		385	415		115
18 February, 2007	85	78		59	85		70	84		78	80		23	
25 February, 2007	69		52	69		57	67		62	67		16		
11 March, 2007	9	9		9	9		7	9		3				
18 March, 2007	3		2	2		3	2		1					

From this Cohort Analysis we can see how the customers who reported sales from the first week were the ones who generated most of the revenue. On the other hand, customers who were acquired later did not came back to repurchase.

Interesting to see how in week 5, 8 and 11 most of the sales activity is concentrated.

Customer segmentation by total purchase



In this graph we can see the distribution of total sales by customer's purchase power that ranges from $20 to $200. We can also see the number of customers that are concentated in each range of purchase.

Loyal customers who made purchases over$140 only account for 10% of the revenue(7% +1%+1%+1%), while the gross of the revenue is among those customers whose spending was between $80 and $120 (29% +29%+19%).

Instead of rewarding the most loyal customers, Rockbuster should focus where most of its business condense (customers with a higher spending of $80), meaning the PARETO CUSTOMERS.

Percentage of Sales: 0% 10% 20%

Percentage of Sales: 0% 29%

Sales and Count of the pareto Customers(85% sales) by Country

From the 599 paying customers, Rockbuster should focus on **489** clients.
This map can give an idea of where marketing and sales efforts should be focused in case future campaigns and discounts cannot be deployed worldwide.

Sales Country: 81 5,275

Count Customer: 1 10 20 30 40 49

Recommendations and Conclusions

Rockbuster should redirect it's efforts towards the customers that represent most of its revenue, Pareto Customers. Just focusing on the best 10 customers won't make the business prosper or reach sales target easily. For this, implementing strategies of acquisition to expand the customer base will payoff better than encouraging loyal customers to rent more movies.

A deeper analysis on what was done during weeks 5,8, and 11 may shed light into why customers did most of the renting during these weeks.

Keeping on track on what cusotmers rent what categories they enjoy the most will ensure to keep the content relevant to viewers and will extend their chance to repurchase again.

For Rockbuster to transition from rental-based to a suscription-based businesss it needs to define a flat rate princing model. These transition will imply changing its business model, for instance closing physical stores and developing strong capabilities to stream online content. All of these changes suggest an initial heavy investmen and a complete disrupt on how Rockbuster operates.