



# Data Analysis for Online Video Service

---



Rockbuster Stealth LLC | Austin Tam



# Table of Contents

---

*01*

## *Data Overview*

Descriptive statistics and data limitations

*02*

## *Content*

Most popular kinds of movies by category

*03*

## *Geographical Factors*

Countries and cities with largest customer base

*04*

## *Next Steps*

What else we can do to promote the platform



01

# Data Overview

---



# Customer & Payment Table

Serving 600 Customers

Two Retail Locations

605 different addresses

\$4.2 is the Average Spent

\$11.99 is the Most Spent

customer	
customer_id	serial
store_id	smallint
first_name	varchar(45)
last_name	varchar(45)
email	varchar(50)
address_id	smallint
activebool	boolean
create_date	date
last_update	timestamp(without timezone)
active	integer

payment	
payment_id	serial
customer_id	smallint
staff_id	smallint
rental_id	integer(5,2)
amount	numeric
payment_date	timestamp(without timezone)

## Data limitations

- We only have 600 customers to study, so not enough to conduct any statistical analysis without bias
- No other real descriptive information to draw insights from
- All our customers rented from stores, so much of their activity will not inform us how they will use an online video service

97% of Customers Still Active

## Data limitations

- Payment dates and rental dates do not align whatsoever
- There is no further data that is collected like payment method, or overdue fees
- All rentals happened in May of 2005, so we cannot do any significant temporal analysis



# Film Table



1000 Different Films

All Films are in English

Average Rental Duration is 4.86

Average Replacement Cost is \$19.98

Most Common Rating is PG-13

film	
film_id	serial
title	varchar(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(without timezone)
special_features	text
fulltext	tsvector

## Data limitations

- We are missing some data that would be ideal for analysis e.g. the cost of acquiring the film for our inventory, the production company of the film
- We do not carry any tv shows, which in todays age is prevalent amongst any online video service
- Some of the data does not have units

All Movies were Released in 2006

Average Rental Rate is \$2.98

Most Common Features are Trailers, Commentary, and Behind the Scenes

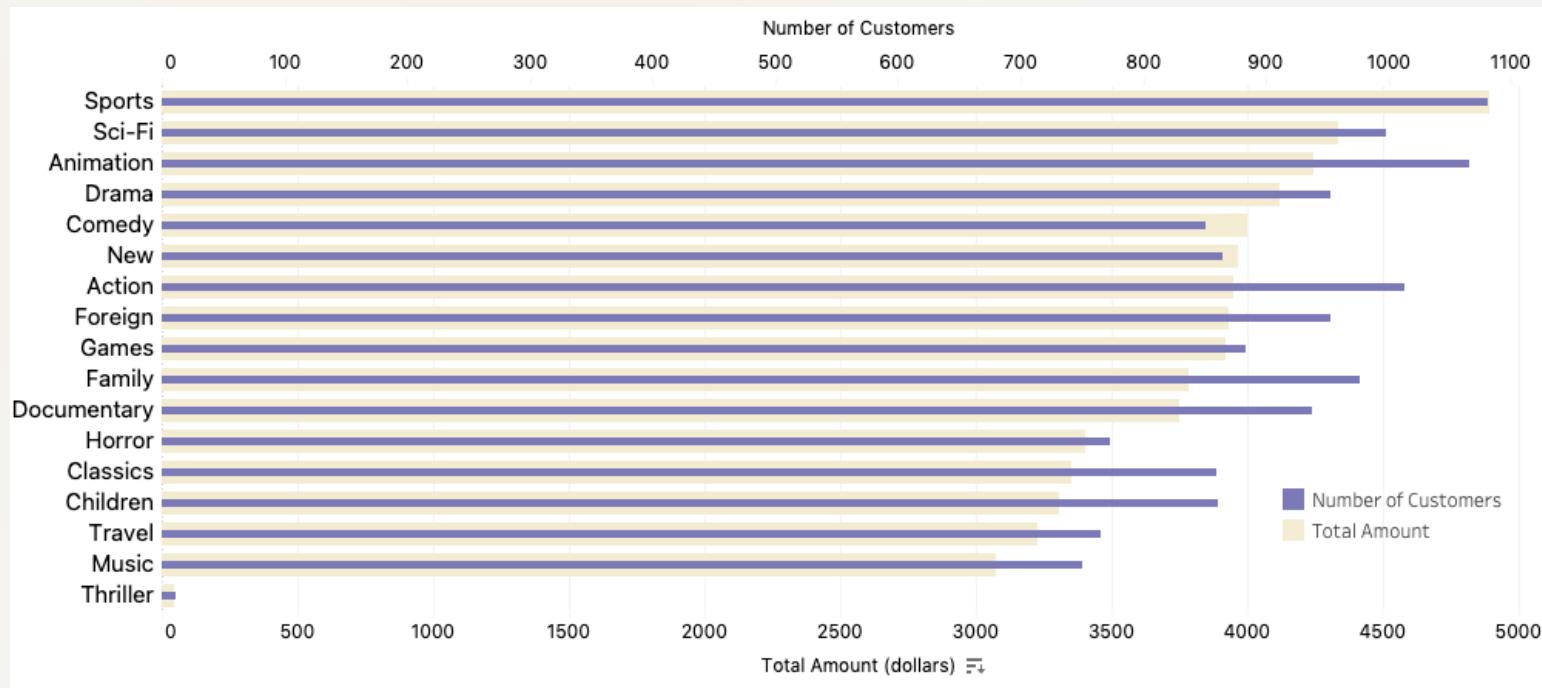


02

# Content

---

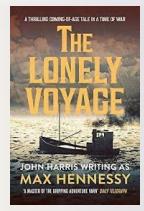




Here, we can see the most popular Movie Category by Revenue. However, as shown through the number of customers, many of the less revenue generating categories are just as popular.



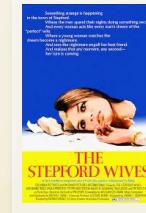
# Highest Grossing



Telegraph Voyage

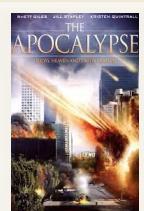


Zorro Ark



Wife Turn

# & Lowest Grossing



Duffel Apocalypse



Texas Watch



Oklahoma Jumanji



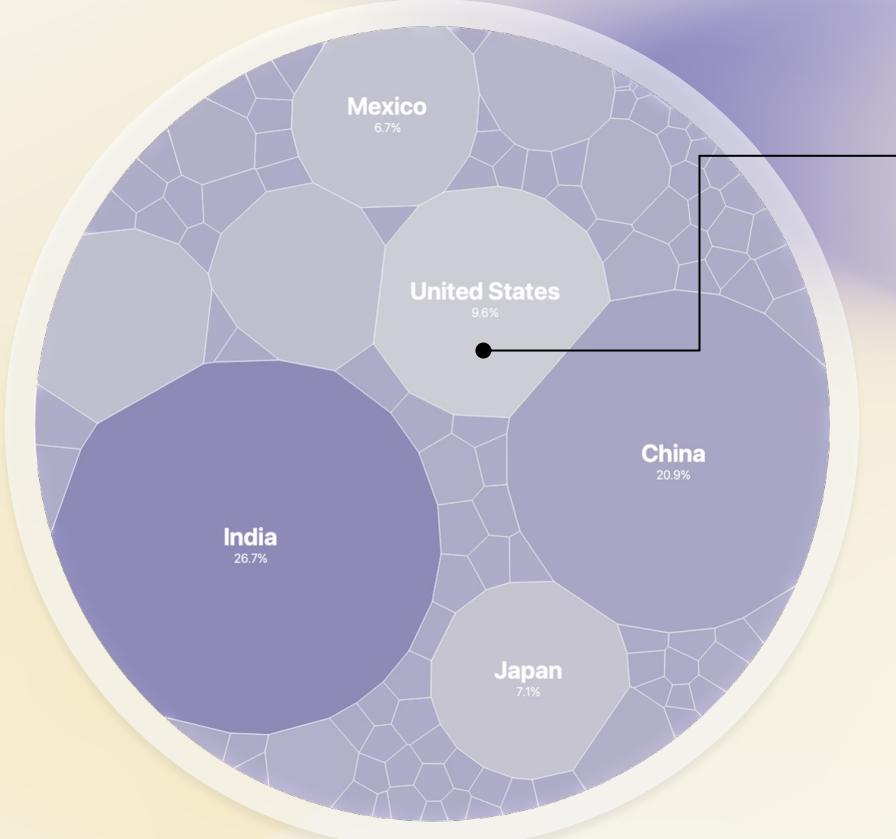
03

# Geographical Factors

---



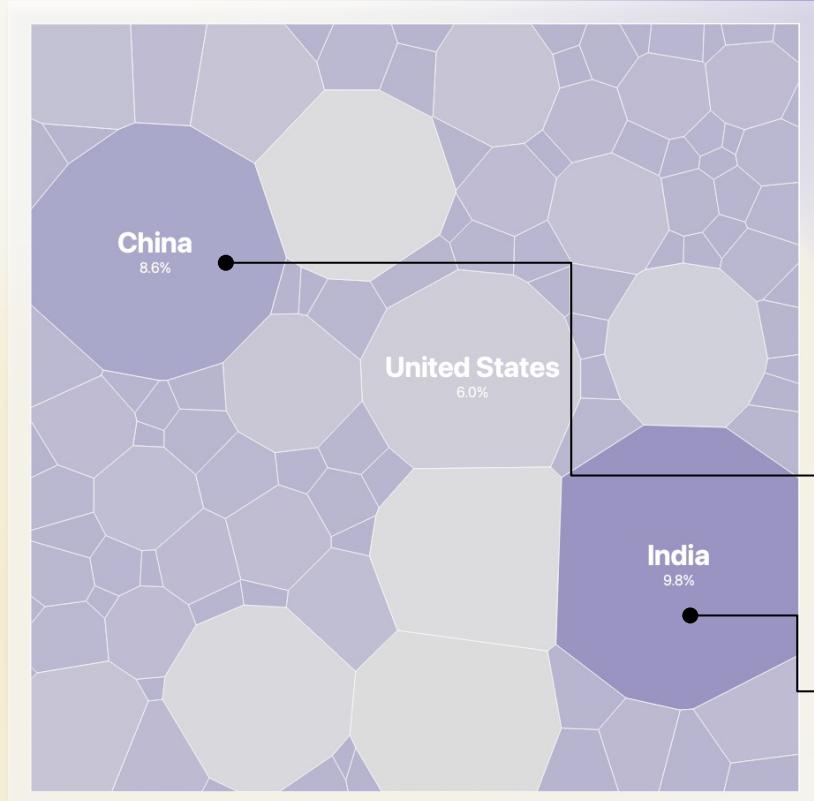
# The Top Countries & Cities by Number of Customers



*Aurora, Colorado and London, England had the Highest Number of Customers at 2*

*The Rest of the Cities in All the Other Countries had One Customer*

# The Top Countries & Cities by Amount of Revenue



*The Top Three Cities are: (Sant-Denis, Runion), (Cape Coral, US), (Santa Barbara, Brazil)*

*The Top City in China was Zhoushan*

*The Top City in India was Valparai*



04

# Next Steps

---



# Top 5 High Lifetime Value Customers

---



*United States*

---

Aurora | Clinton Buford

Aurora | Scott Shelley



*United Kingdom*

---

London | Cecil Vines



*Algeria*

---

Bchar | Judy Gray

Skikda | June Carroll





# Most Popular Actors/Actresses

---



01

---

Susan  
Davis



02

---

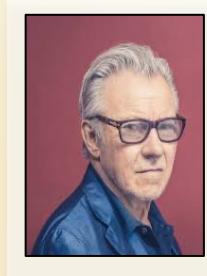
Gina  
DeGeneres



03

---

Matthew  
Carrey



04

---

Mary  
Keitel



05

---

Scarlett  
Damon



# Thank You

---



The visualizations in this presentation can be found [here](#). The data dictionary which includes the data used in this presentation can be found [here](#)