ROCKBUSTER

stealth

Data Analysis   Q3 '05 Rentals

Aaron Manzano

Table of Contents

RockBuster *stealth*

Customers

Global Movie Rentals in •**109** Countries

•**14,413** Transactions

•**$4.25** Average transaction amount:

•**599** Customers

•**$102** Avg customer spend
•(~24 rentals)

RockBuster *stealth*

Films

Film Catalog with
- **958** Titles
- **17** Categories
- **1-8** Copies per title
- **4,581** Total movies
- **$10-$30** Replacement cost
  - (avg **$20**)

Rental Duration

AVG
4.99
Day

MIN:
3
Day

MAX:
7
Day

Film Length

AVG:
115
min

MIN:
46
min

MAX:
185
min
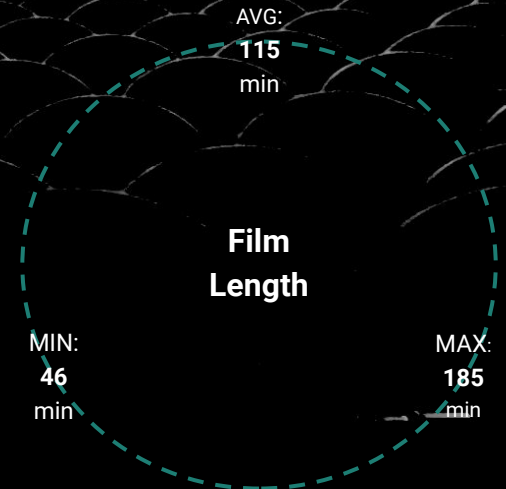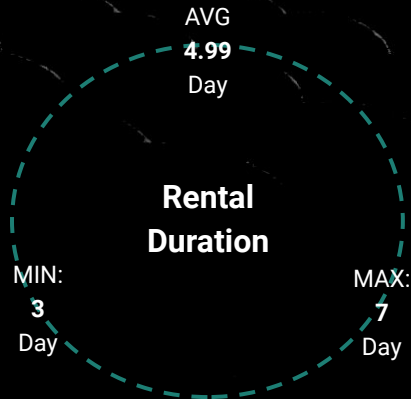
Overview

Define Key Metrics

Gross Profit Analyzes **Return on Investment**.

- Total Revenue
  - For 3 Months
- Inventory Cost
  - 1-8 film copies
- Identify **Important Variables**

Optimize Catalog

Catalog Growth Strategy:

- **Identify** weak titles
  - Low ROI Films
- **Remove** weak titles
  - Sell physical copies
  - Don't renew digital licenses

Geographic Analysis

Identify Top Markets:

- Top **Countries**
  - By Customer Base
- Top **Customers**
  - By Total Spend
- Top **Categories**
  - In Top Markets

Key Performance Metric
Gross Profit

Total Revenue

Gross Revenue from all payments with that Film. Overall, about 30% of revenue is from late fees.
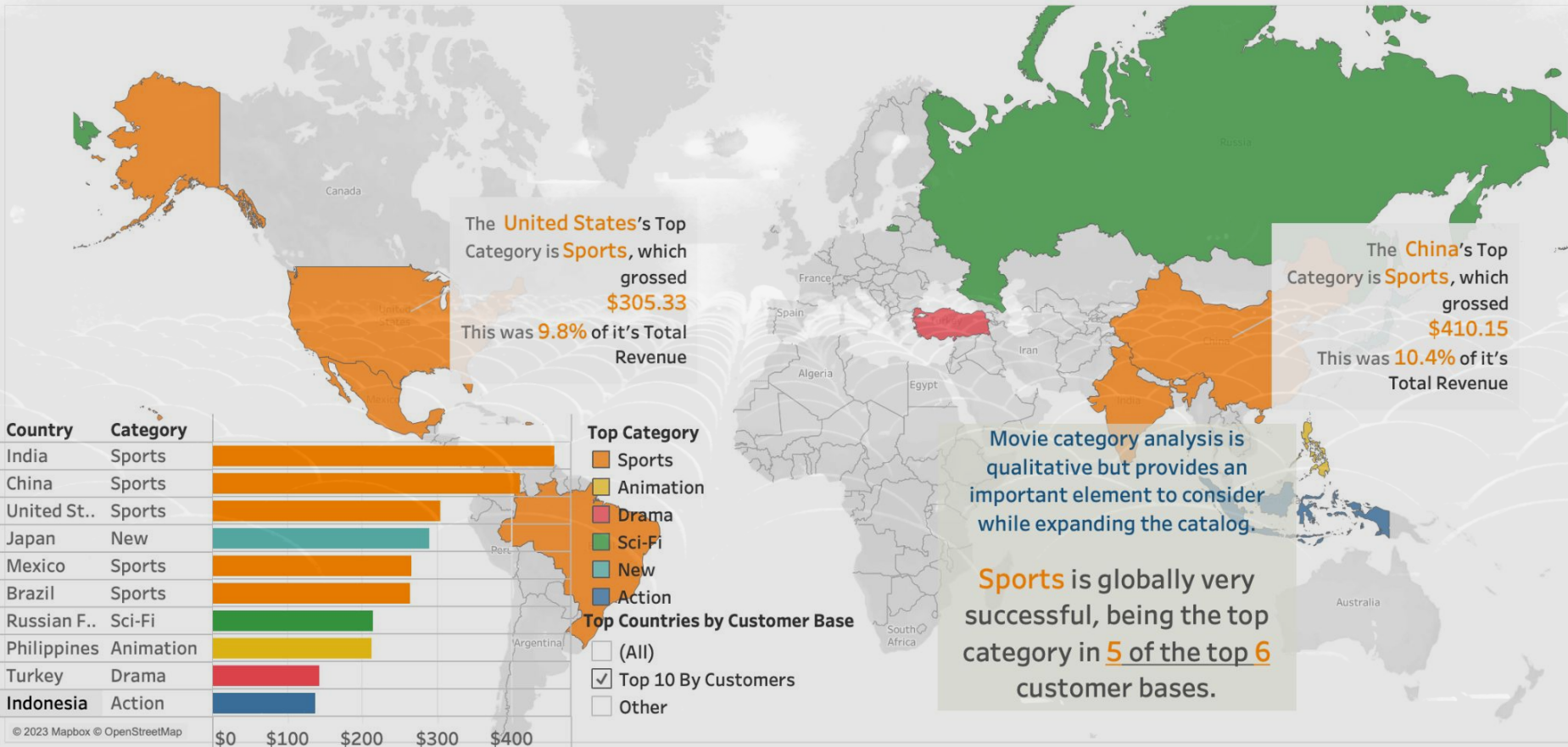
Inventory Cost

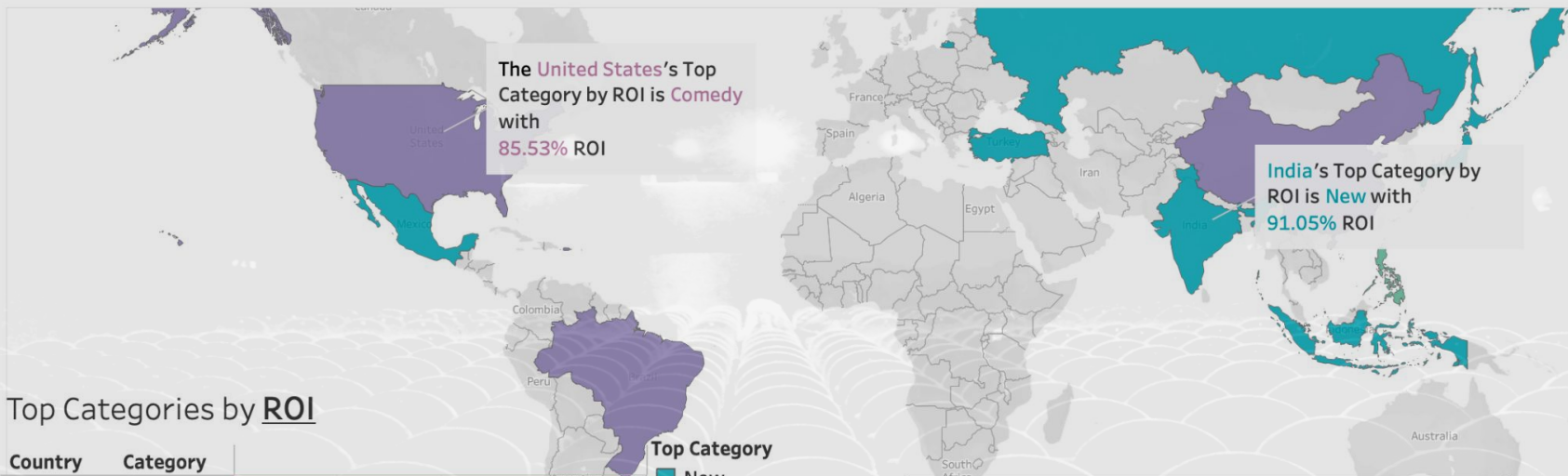**Total Cost** from multiplying the replacement cost by the number of copies in inventory.

Top Categories in Top Markets

The **United States**'s Top Category is **Sports**, which grossed **$305.33**
This was **9.8%** of it's Total Revenue

The **China**'s Top Category is **Sports**, which grossed **$410.15**
This was **10.4%** of it's Total Revenue

Country	Category
India	Sports
China	Sports
United St..	Sports
Japan	New
Mexico	Sports
Brazil	Sports
Russian F..	Sci-Fi
Philippines	Animation
Turkey	Drama
Indonesia	Action

Top Category
- Sports
- Animation
- Drama
- Sci-Fi
- New
- Action

Top Countries by Customer Base
- ☐ (All)
- ☑ Top 10 By Customers
- ☐ Other

Movie category analysis is qualitative but provides an important element to consider while expanding the catalog.

Sports is globally very successful, being the top category in **5 of the top 6** customer bases.

© 2023 Mapbox © OpenStreetMap

$0 $100 $200 $300 $400

Top Categories in Top Markets: By Return on Investment

The United States's Top Category by ROI is Comedy with 85.53% ROI

India's Top Category by ROI is New with 91.05% ROI

Top Categories by ROI

Country	Category
India	New
China	Comedy
United St..	Comedy
Japan	New
Mexico	New
Brazil	Comedy
Russian F..	New
Philippines	Foreign
Turkey	New
Indonesia	New

Top Category
- New
- Foreign
- Comedy

$0 $100 $200 $300

© 2023 Mapbox © OpenStreetMap

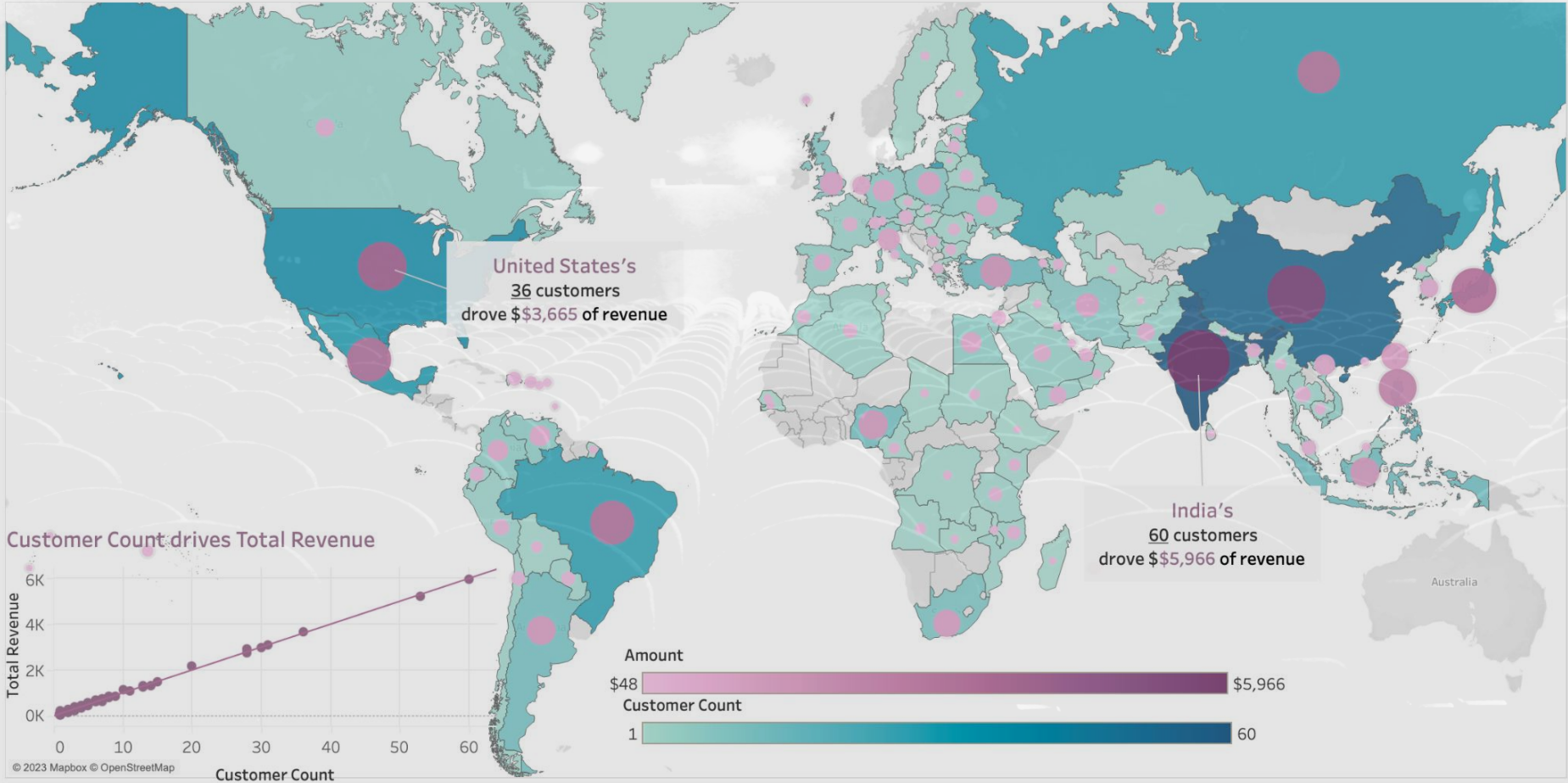**Top Countries by Customer**
- ☐ (All)
- ☑ Top 10 By Customers
- ☐ Other

Analyzing Top Category by Total Revenue results in a different view of categories to focus on.

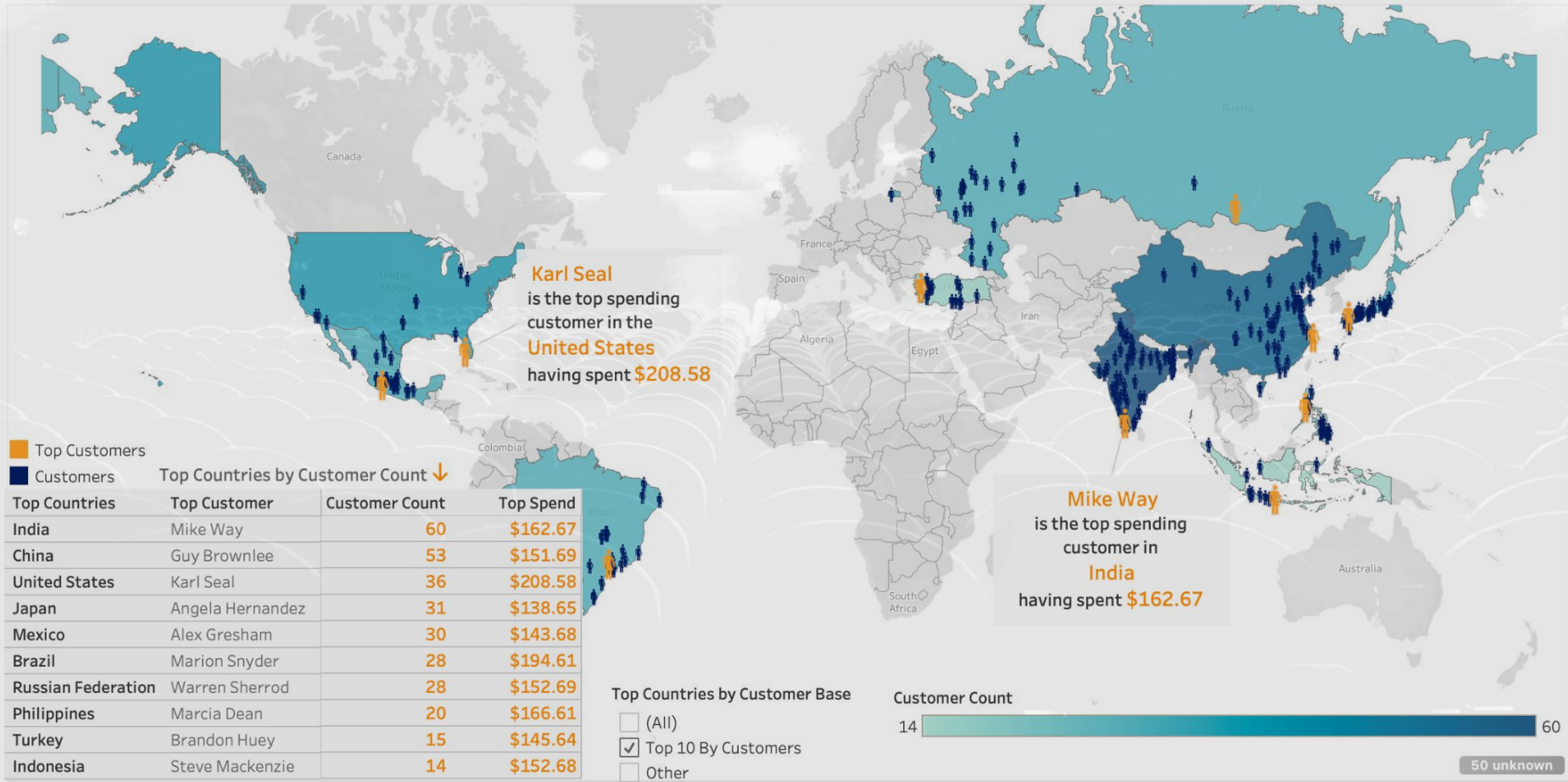New is the top category by ROI in 6 of the top 10 countries by customer base.
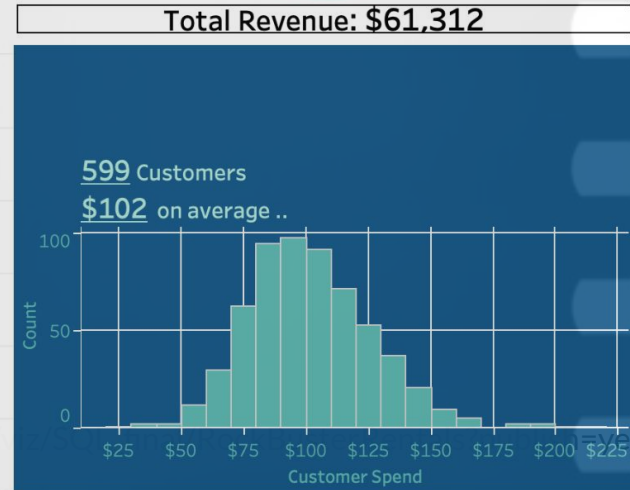
Map: Customer Base and Total Revenue

United States's
36 customers
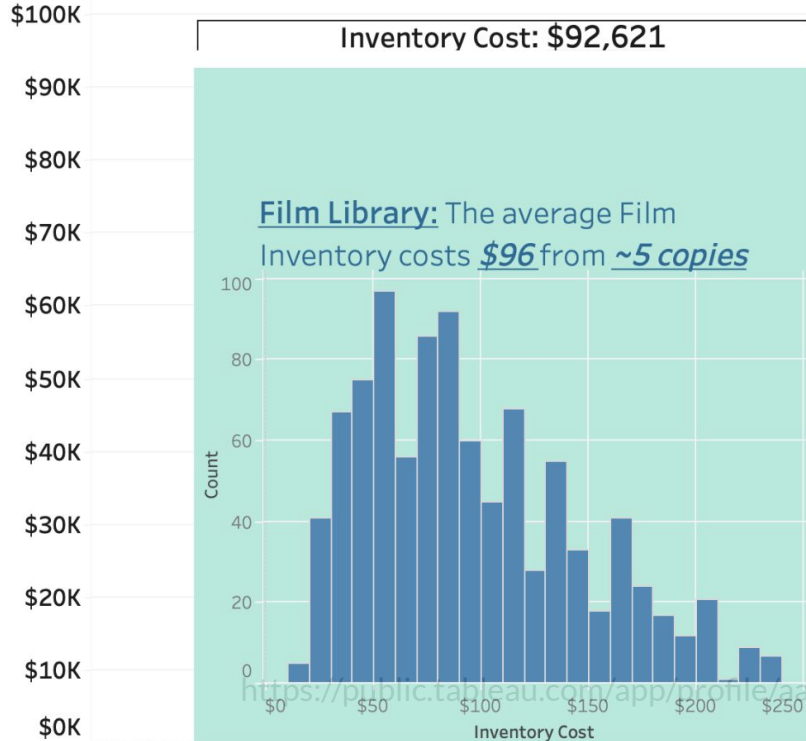drove $$3,665 of revenue

India's
60 customers
drove $$5,966 of revenue

Customer Count drives Total Revenue

Total Revenue: 6K, 4K, 2K, 0K

Customer Count: 0, 10, 20, 30, 40, 50, 60

Amount
$48 — $5,966
Customer Count
1 — 60

Australia

© 2023 Mapbox © OpenStreetMap

Map: Top Countries and Customers

Karl Seal is the top spending customer in the United States having spent $208.58

Mike Way is the top spending customer in India having spent $162.67

Top Customers
Customers

Top Countries by Customer Count ↓

Top Countries	Top Customer	Customer Count	Top Spend
India	Mike Way	60	$162.67
China	Guy Brownlee	53	$151.69
United States	Karl Seal	36	$208.58
Japan	Angela Hernandez	31	$138.65
Mexico	Alex Gresham	30	$143.68
Brazil	Marion Snyder	28	$194.61
Russian Federation	Warren Sherrod	28	$152.69
Philippines	Marcia Dean	20	$166.61
Turkey	Brandon Huey	15	$145.64
Indonesia	Steve Mackenzie	14	$152.68

Top Countries by Customer Base
☐ (All)
☑ Top 10 By Customers
☐ Other

Customer Count
14 ——————— 60

50 unknown

Cost and Revenue:
Breaking down the Key Metric: **Return on Investment**

Inventory Cost: $92,621

Film Library: The average Film Inventory costs *$96* from *~5 copies*

ROI for Q3 '05
$61,312 $92,621
Quarterly ROI: 66%
Annualized ROI: 246%

■ Inventory Cost
■ Total Revenue

Total Revenue: $61,312

599 Customers
$102 on average ..

https://public.tableau.com/app/profile/aaron.manzanowa/...

Revenue and ROI:
Visualizing the range of profitability across films
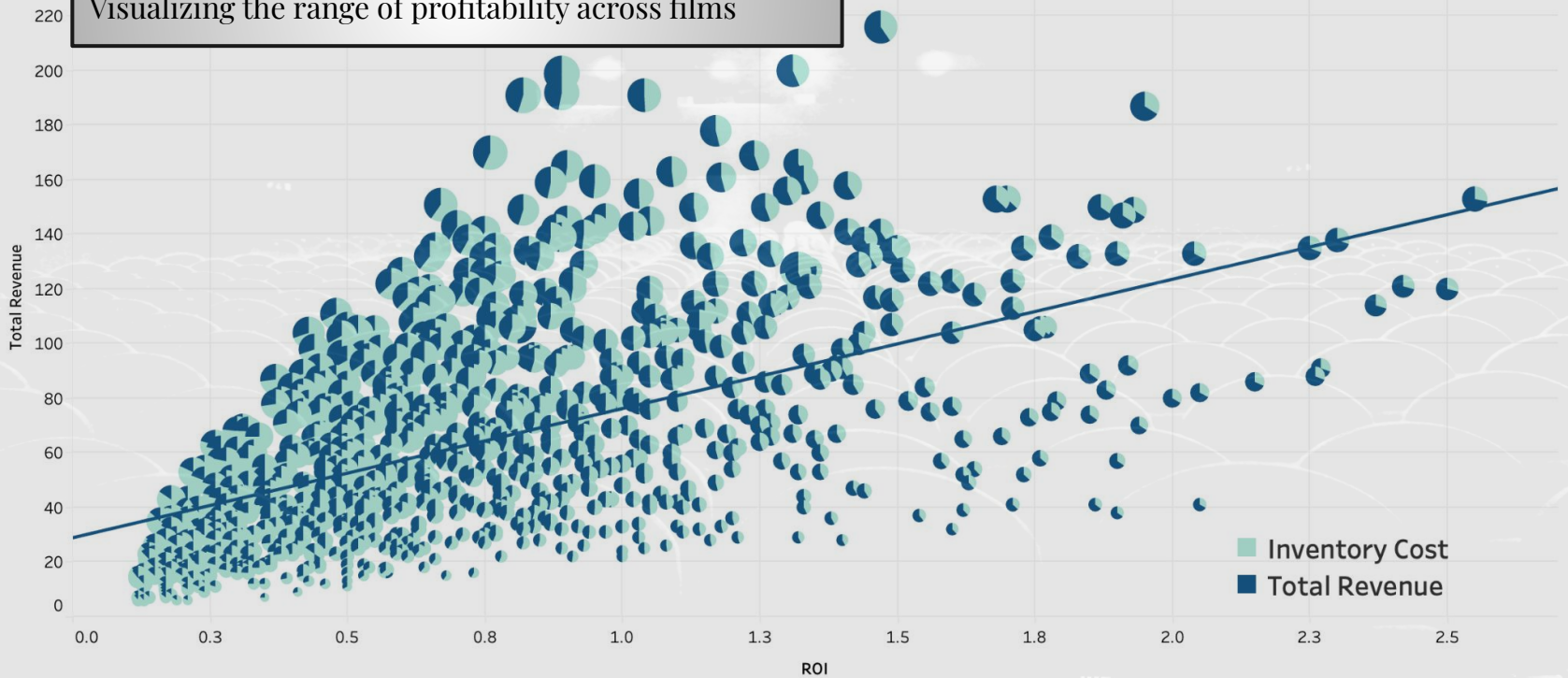
Determining the Least Profitable Titles:

The *Shapely Values* indicate that Quantitative > Qualitative.

Instead of choosing between categories or ratings, we select the bottom quartile of all categories by Gross Profit.
This measure of Inventory Cost / Total Revenue provides

This **group of 244** has less than half the median profit ratio
Removing Low Profit Titles
Total Profit Ratio: **74.5%** > **89.9%** AVG Profit ra

Gross Profit by Category Quartiles

Profitability Quartiles
- 🟥 Low Profit Titles (244)↑
- 🟩 Top Quartile + Outliers (247)↓
- ⬜ Profitable Titles (467)↑

Average: 74%

ROI axis: 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, 210%, 220%, 230%, 240%, 250%, 260%

DUDE BLINDNESS

Low Profit Titles (244)↑		Profitable Titles (467)↑		Top Quartile + Outliers (247)↓	
Cruelty Unforgiven: 12%	Japanese Run: 13%	Hills Neighbors: 33%	Chitty Lock: 39%	Maiden Home: 255%	Dude Blindness: 227%
Freedom Cleopatra: 12%	Lawrence Love: 13%	Dalmations Sweden: 35%	Midnight Westward: 39%	Trap Guys: 250%	Bright Encounters: 226%
Hollywood Anonymous: 12%	Silence Kane: 13%	Glass Dying: 36%	Army Flintstones: 40%	Kissing Dolls: 242%	Flintstones Happiness: 225%
Sassy Packer: 12%	Texas Watch: 13%	Extraordinary Conquerer: 38%	Doctor Grail: 40%	Whale Bikini: 237%	Daisy Menagerie: 215%
Clockwork Paradise: 13%	Bride Intrigue: 14%	Reunion Witches: 38%	Carrie Bunch: 41%	Fellowship Autumn: 230%	Paths Control: 205%

## Profitability Quartiles:
Identifying low profit titles
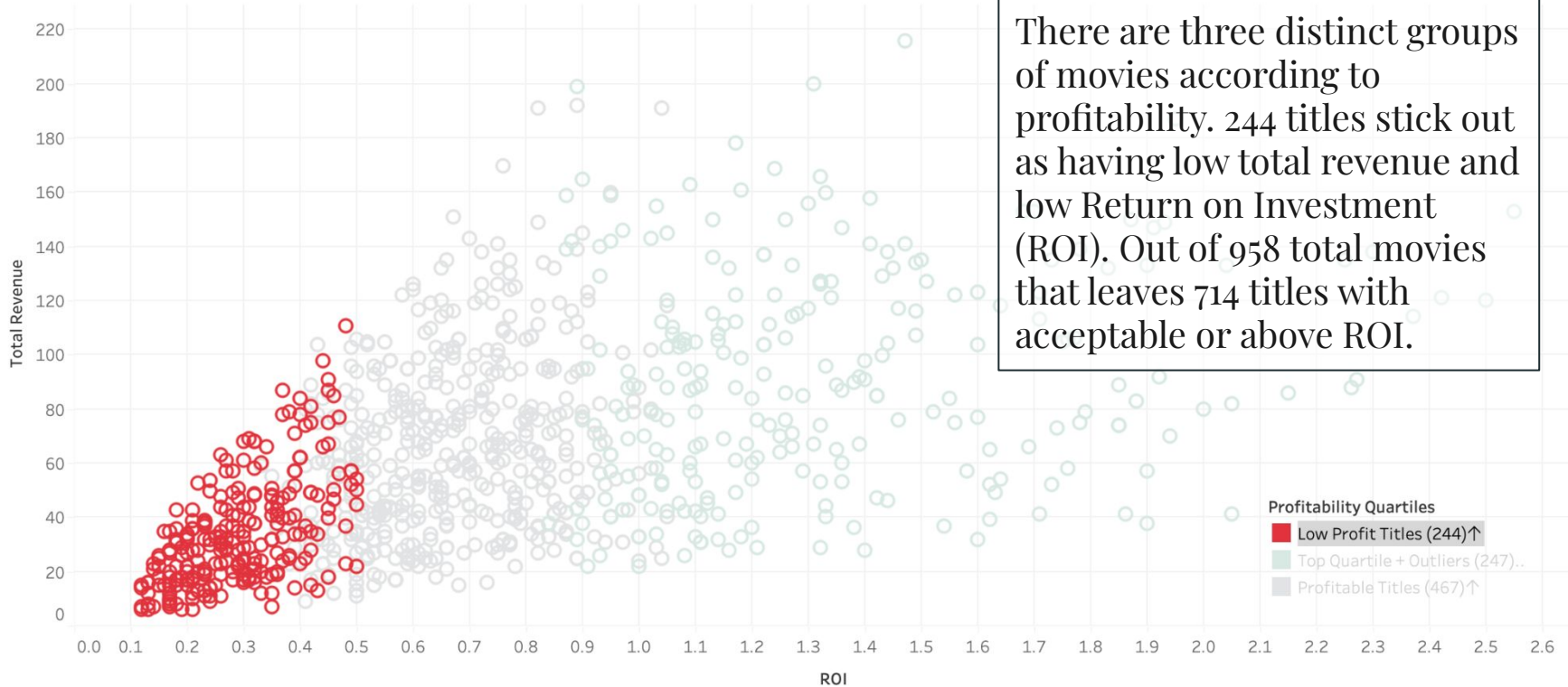
There are three distinct groups of movies according to profitability. 244 titles stick out as having low total revenue and low Return on Investment (ROI). Out of 958 total movies that leaves 714 titles with acceptable or above ROI.

Profitability Quartiles
Low Profit Titles (244)↑
Top Quartile + Outliers (247)..
Profitable Titles (467)↑

Rating Count:
Higher film count
doesnt always lead to higher ROI

5 Ratings

Average: 191.6

213

171
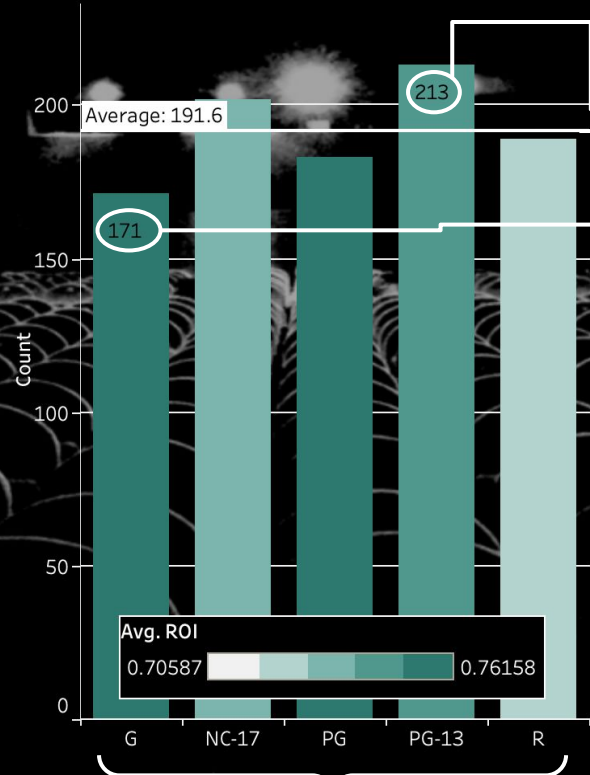
Max Count: 213
Avg Count: 191.6
Min count: 171

Avg. ROI
0.70587 — 0.76158

The **darker** the Hue, the higher the **Rating's ROI**.

Higher number of films in that Rating **don't** seem associated with higher ROI, as the **Min Count has very high ROI.**

G   NC-17   PG   PG-13   R

Count

How to determine which features are important in ***driving ROI***?

- Machine Learning is an asset

- Model interpretability is important to understand how the model is weighing different features
  - Features: Rental rate, replacement cost, categories, etc

- Train the model on predicting ROI
  - Use interpretability metrics to define a path towards higher ROI

Category

Rental Rate

Rental Rate

Prediction

Costs

Costs

Film Length
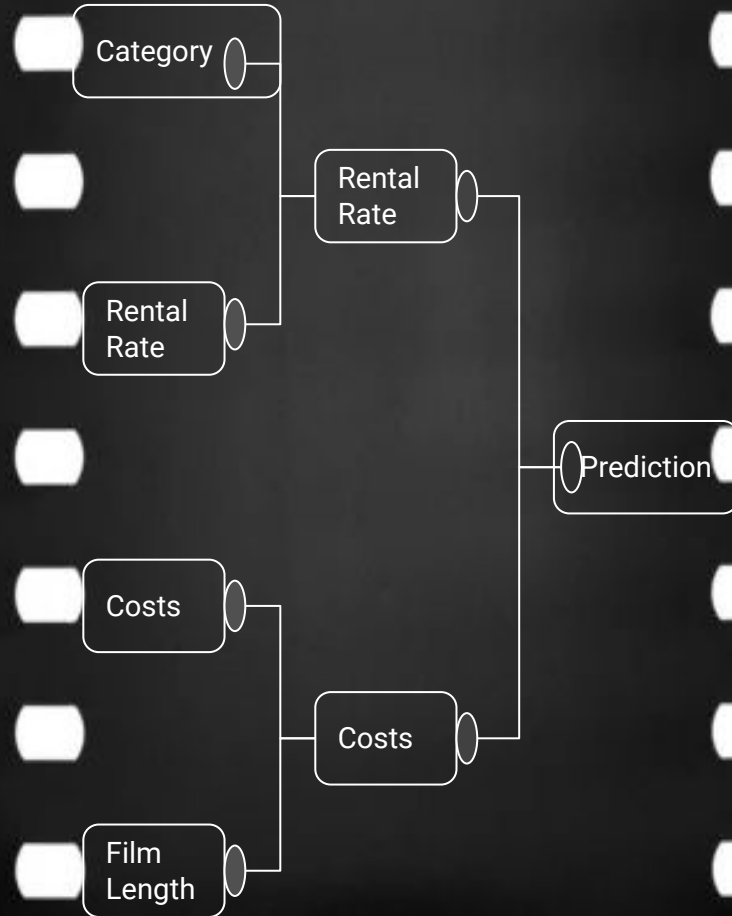
Shapely Values
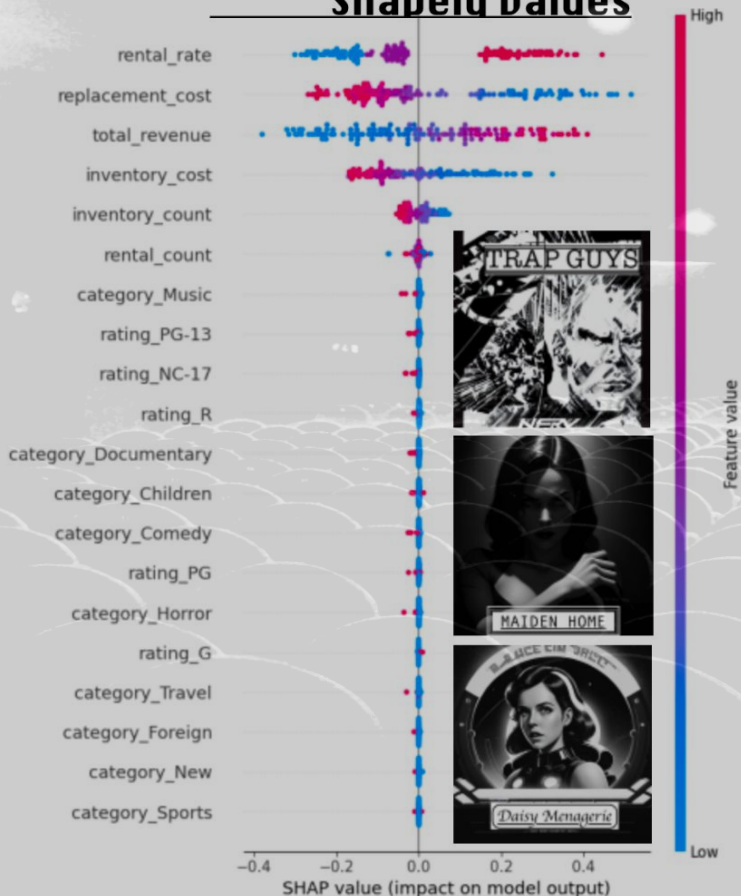
XGBoost is a preferred choice among experts for tackling intricate data challenges and generating robust predictive models.

-Decision Tree ML model trained to predict film's Profit Ratio (*total revenue / inventory cost*)
-Low MSE (0.0093) of the model indicates reliability

Shapely values offer a pragmatic framework for fairly attributing feature importance in these models, quantifying the practical significance of each variable's role in the model.

-SHAP values closer to zero indicate features with less predictive value in the model. Ratings/categories aren't predictive.
-Feature values are shown in blue (low values) to red (high values)

1) *Rental_rate*, *Replacement_cost*, *Total_revenue* are furthest from 0 and are therefore contribute most to the model.

2) *Quantitative data* is much more predictive than *qualitative data* (like categories and ratings).

2) *High rental rates* and *low replacement costs* are the key drivers of profitability.

PROFIT

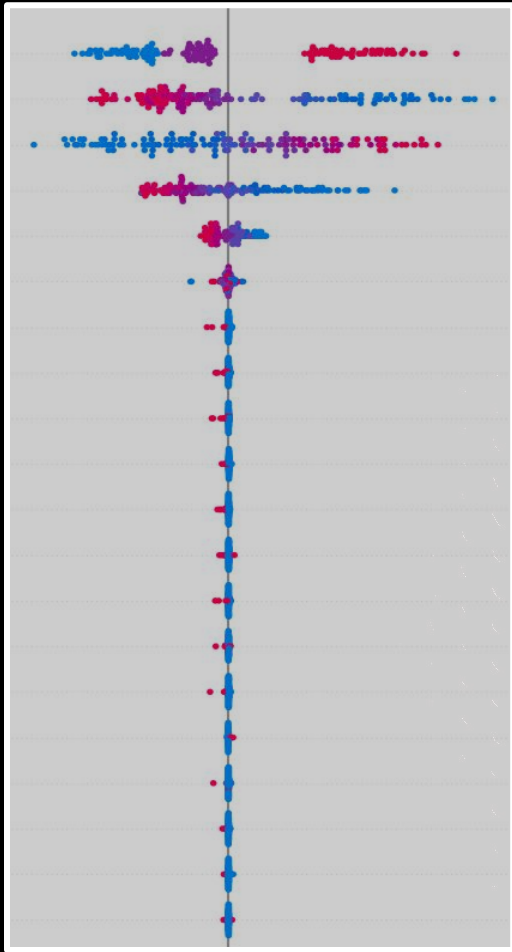Interpreting a model to determine feature importance.

Shapely Values *far from 0* are more important

Blue values are low variables, red is high.

Low rates and high costs predict *Low Profit*

High rates and low costs predict *High Profit*

rental_rate

replacement_cost

0
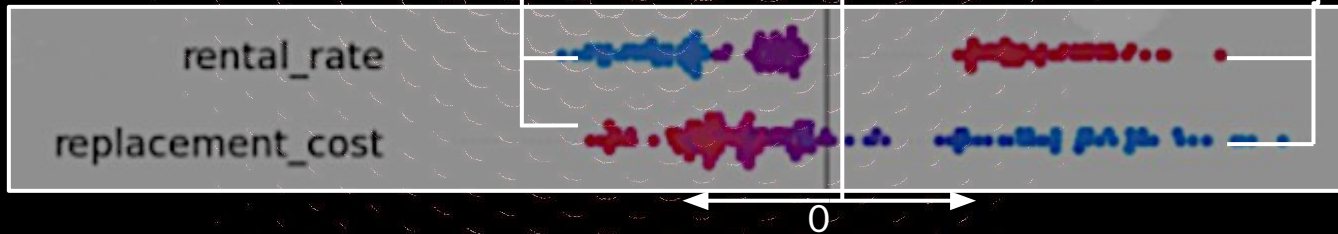
PROFIT

Categories and ratings

1 **Qualitative data is closer to zero and less useful**

Replacement Costs

2

Maintaining low inventory costs predicts profitability

Rental Rates

3 High Fees are the foundation of profitability

Feature Leakage

Feature leakage is when information from the target variable is directly or indirectly being used by the model to predict said target variable. This hinders generalization to new data and predictive capabilities of the model. Notice how the MSE degrades as leaking data highlighted in red are removed.
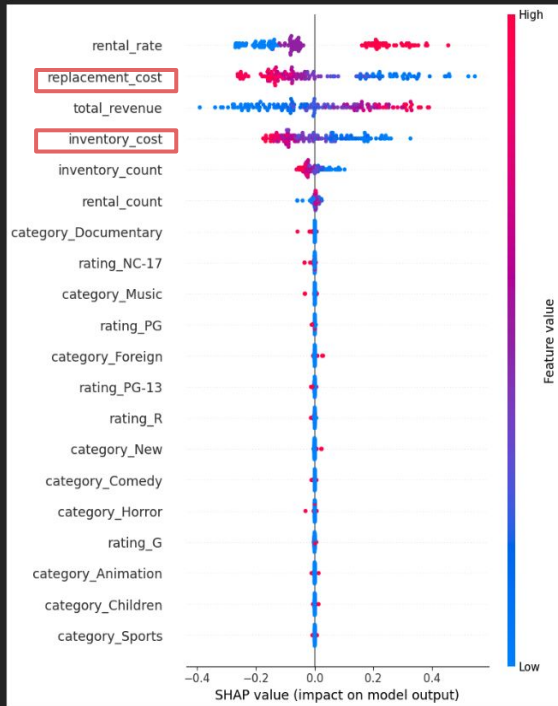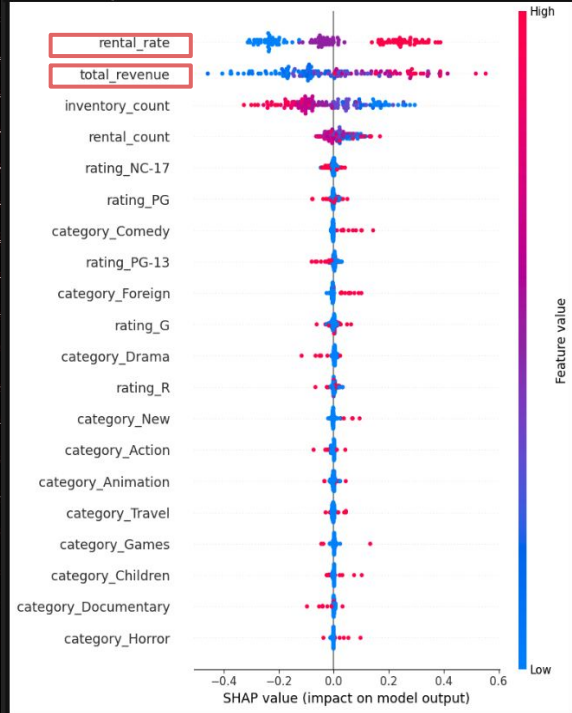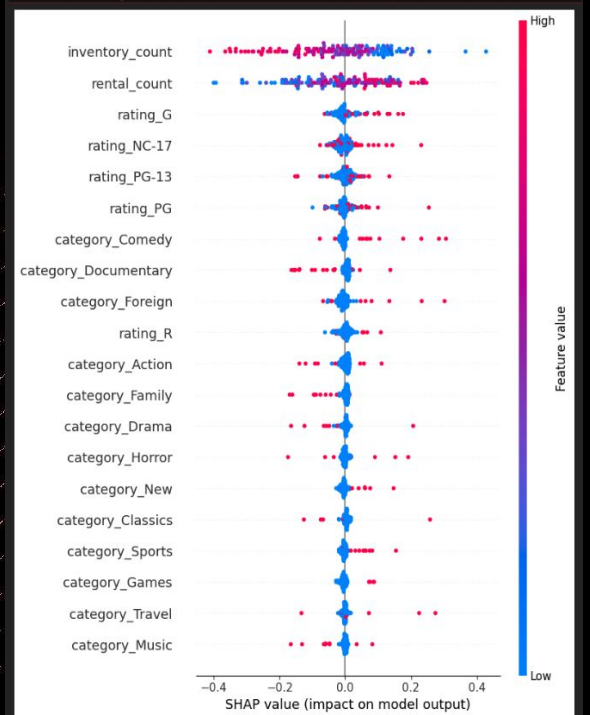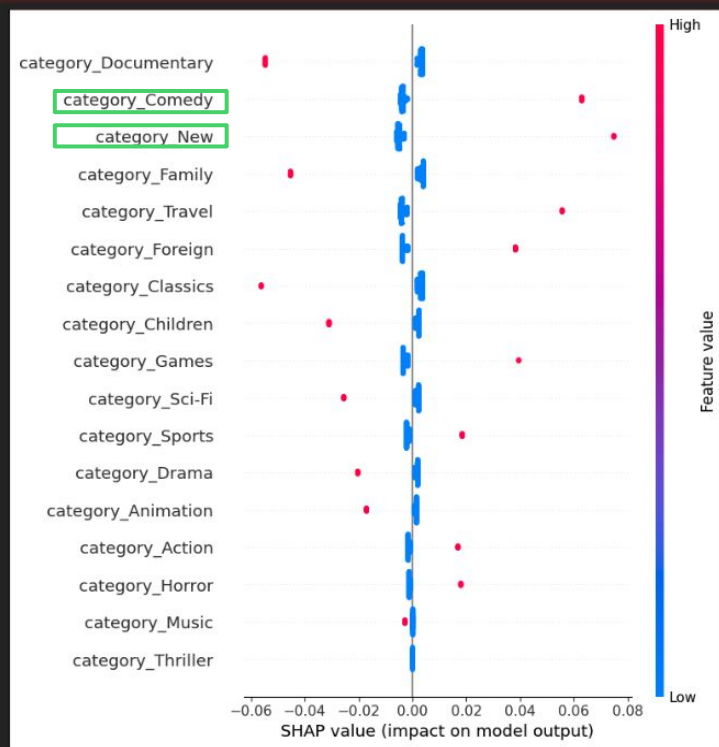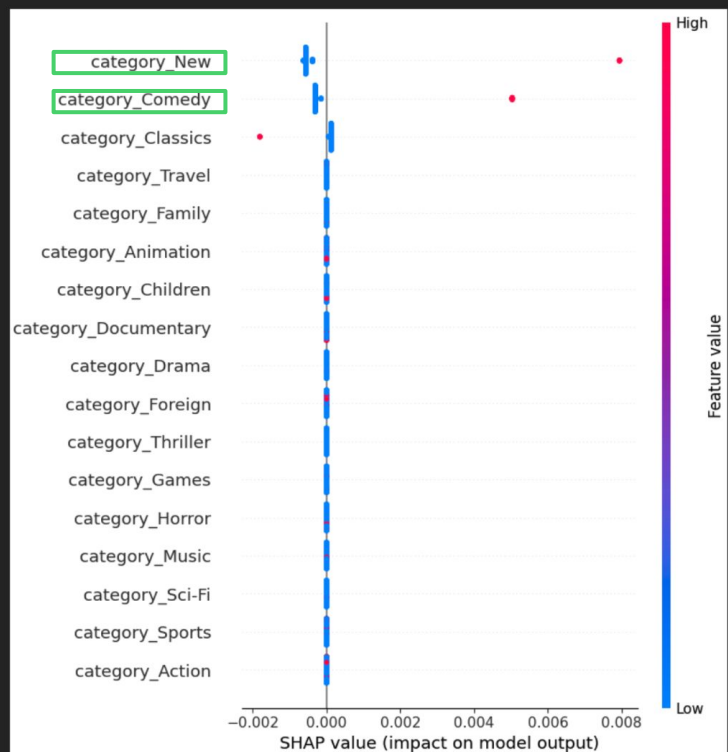
Categories | GridSearch

Here we focus only on categories and train for gross profit, then again after performing grid search for optimal hyperparameter settings. The MSE is high on both, and so predictive accuracy is questionable, but the feature importance can help separate noise from signal. Comedies and New stand out as profitable categories.
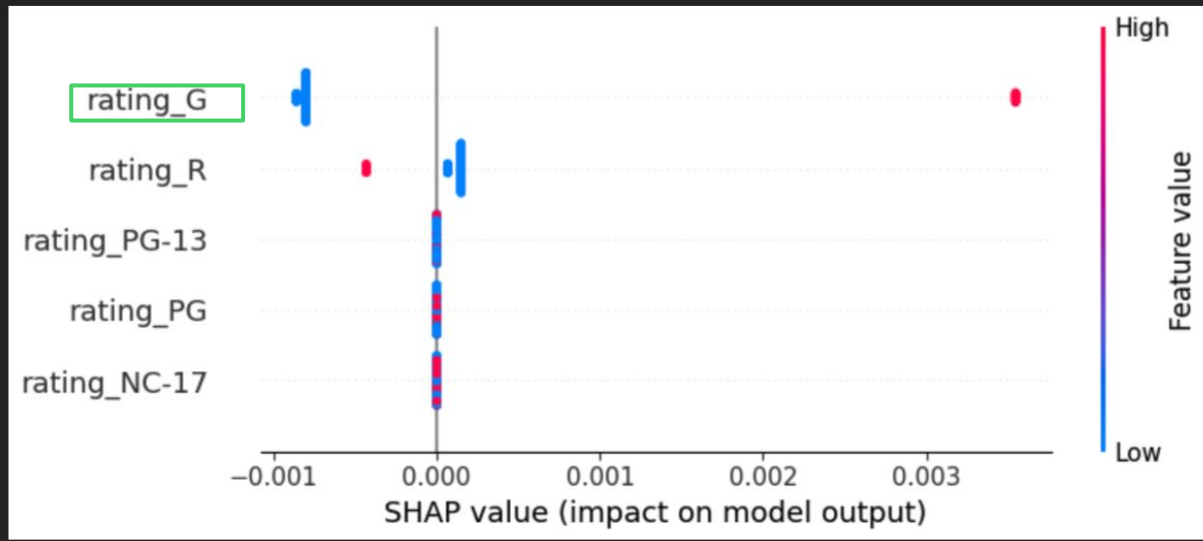
Ratings | GridSearch

Here we focus only on rating and train for gross profit with grid search for optimal hyperparameters. The MSE is not very low again, but G ratings stand out as higher impact on positive ROI than other ratings.

Rental Rates:
High importance for predicting ROI

$0.99
Low rate

$2.99
Medium rate

$4.99
High rate

Low rates are easily associated with lower ROI, as it takes more rentals to reach the same revenue as higher rates.
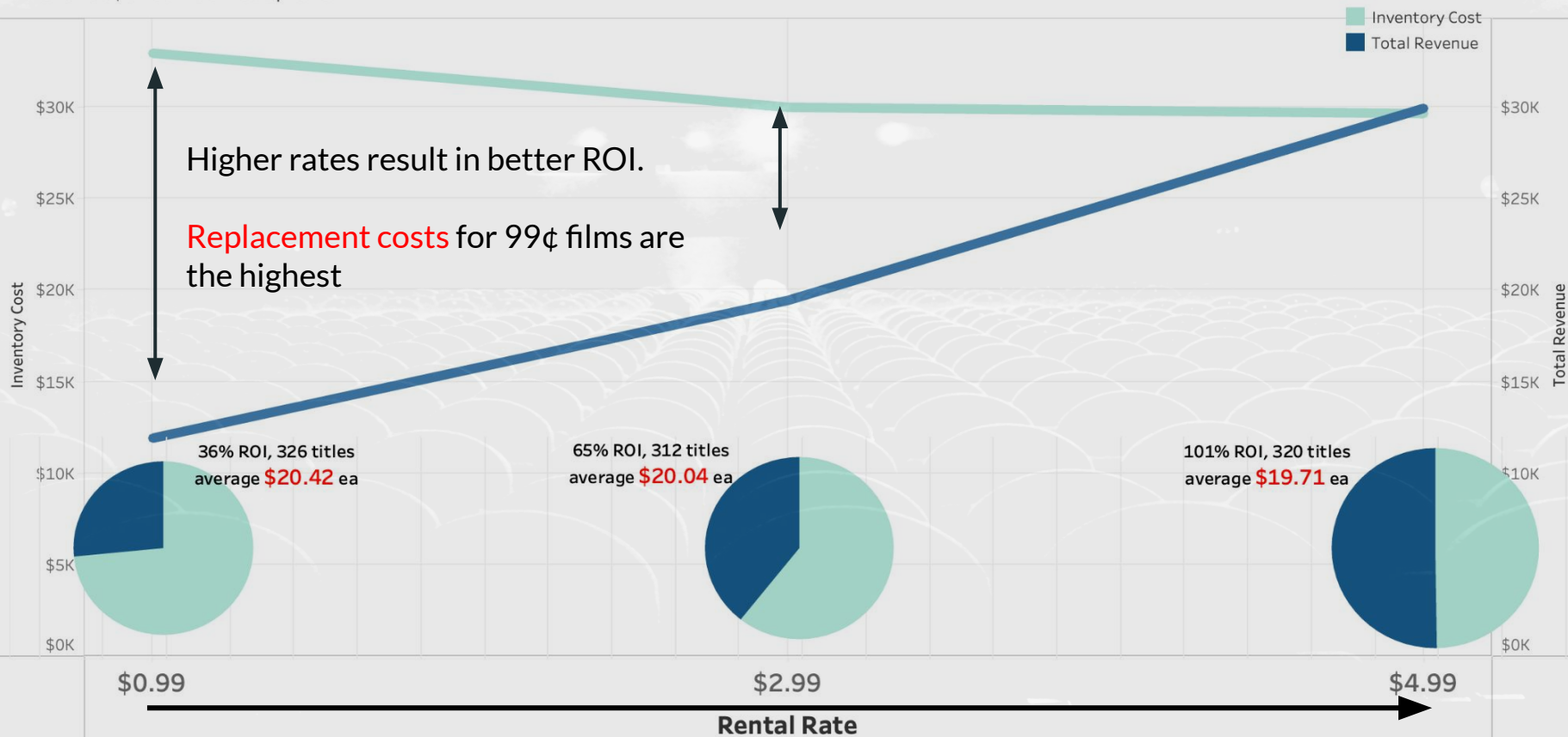
What do the rental rates look like in the context of ROI?

Do higher rates really result in more return on investment?

Rental Rates and ROI:
Films for 99¢ drive much less profit

Higher rates result in better ROI.

Replacement costs for 99¢ films are the highest

Inventory Cost
Total Revenue

36% ROI, 326 titles
average $20.42 ea

65% ROI, 312 titles
average $20.04 ea

101% ROI, 320 titles
average $19.71 ea

$0.99 $2.99 $4.99

Rental Rate

Rental Rates: Profit

Measure Names:
- Inventory Cost
- Total Revenue

$4.99

$29,944

Films for $4.99
Cost on average **$93** to stock for an average total revenue of **$93.58**

$29,660

$2.99

$19,451

Films for $2.99
Cost on average **$96** to stock our catalog and bring in an average total revenue of **$62.34**

$30,004

$0.99

$11,917

Films for $0.99
Cost on average **$101** to stock our catalog and only bring in an average total revenue of **$36.55** per quarter.

$32,957

Late Fees

Measure Names:
- late fees
- rental fees

$6,107

$23,593

$5,475

$13,787

$5,023

99¢ films:
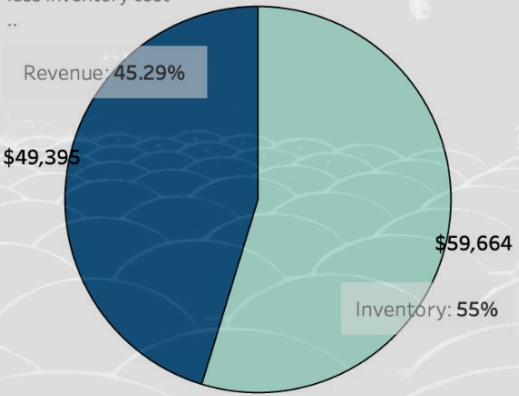
68% of total revenue from last **late fees**

This indicates a lack of ability to drive value.

$6,809

Removing $0.99 Rentals

There are **326** titles rentable for **$0.99,** removing them changes the catalog's profitability:

Total Revenue- $61,312-$49,395= **$11,917** less revenue

Inventory Cost- $92,621-$59,664= **$32,957** less inventory cost ..

Revenue: 45.29%

$49,395

$59,664

Inventory: 55%

However, cutting 33% of the catalog may be too high a threshold and can be improved with similar results.

Rate Groups
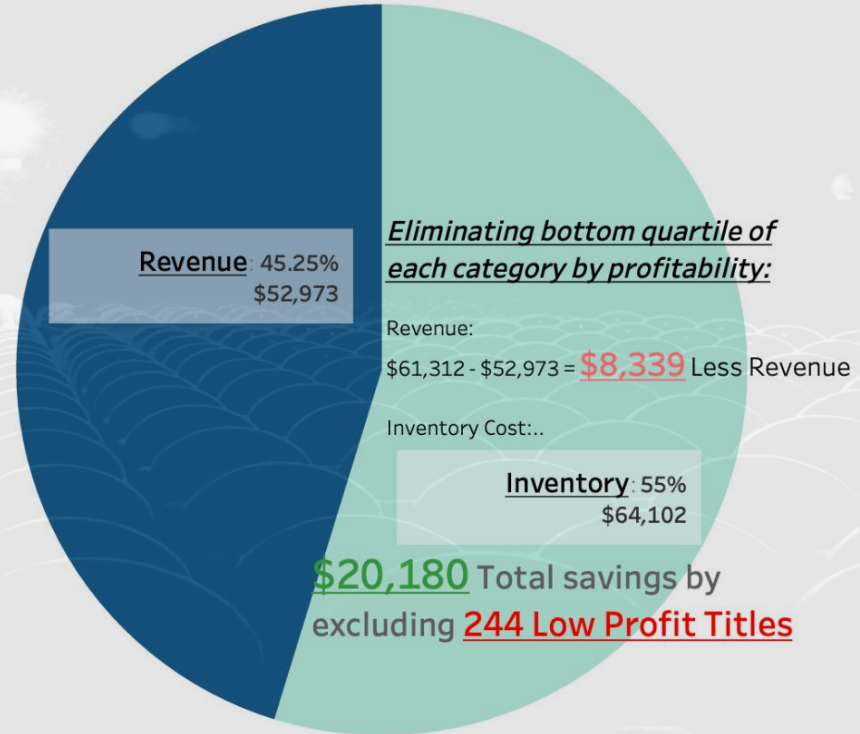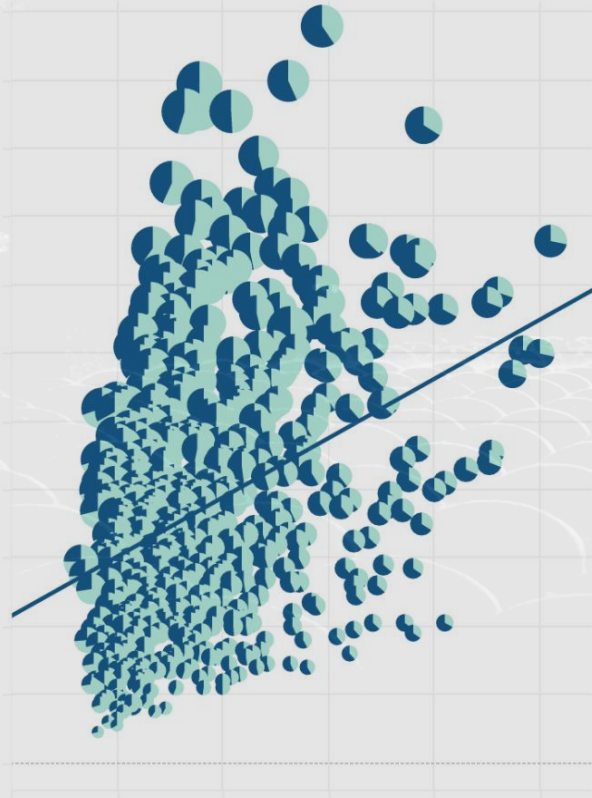- ☐ (All)
- ☐ $0.99
- ☑ $2.99 & $4.99

Profitability Quartiles

There are plenty of ways to prune the catalog, the strategy is to cut the bottom ~20% in terms of ROI each quarter by category.

This will result in less revenue but much less cost, and will continuously update the catalog to be popular and profitable.

What happens to total ROI when we remove the Low Profit Quartile by Category?

Revenue: 45.25%
$52,973

Eliminating bottom quartile of each category by profitability:

Revenue:

$61,312 - $52,973 = $8,339 Less Revenue

Inventory Cost:..

Inventory: 55%
$64,102

$20,180 Total savings by excluding **244 Low Profit Titles**

☐ (All)
☐ Low Profit Titles (244)↑
☑ Top Quartile + Outliers (247)↓

Optimize

40% ROI

45% ROI

Define Key Metrics

Goal: Improve 40% ROI

Solve for Key Metric (ROI)

Remove Low ROI Titles

Save *$20,180*

Collect data for next optimization cycle

Rates

Categories

Loyalty

Avoid 99¢ titles

Maximize $4.99 rates

Largest 10 Markets

Sports is the Highest Grossing, but New has the Highest ROI

Top customers in our largest markets

Identify and Reward

Spend *less* to get <u>more</u>

Find the *weak links*

Strengthen and **Optimize** the film catalog

Profitability Quartiles:
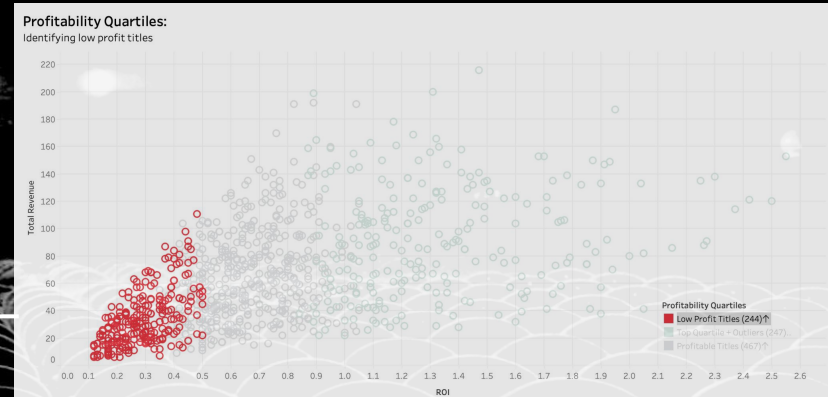Identifying low profit titles

Tableau Link:

https://public.tableau.com/app/profile/aaron.manzano/viz/SQL_final/RockBusterRentals?publish=yes