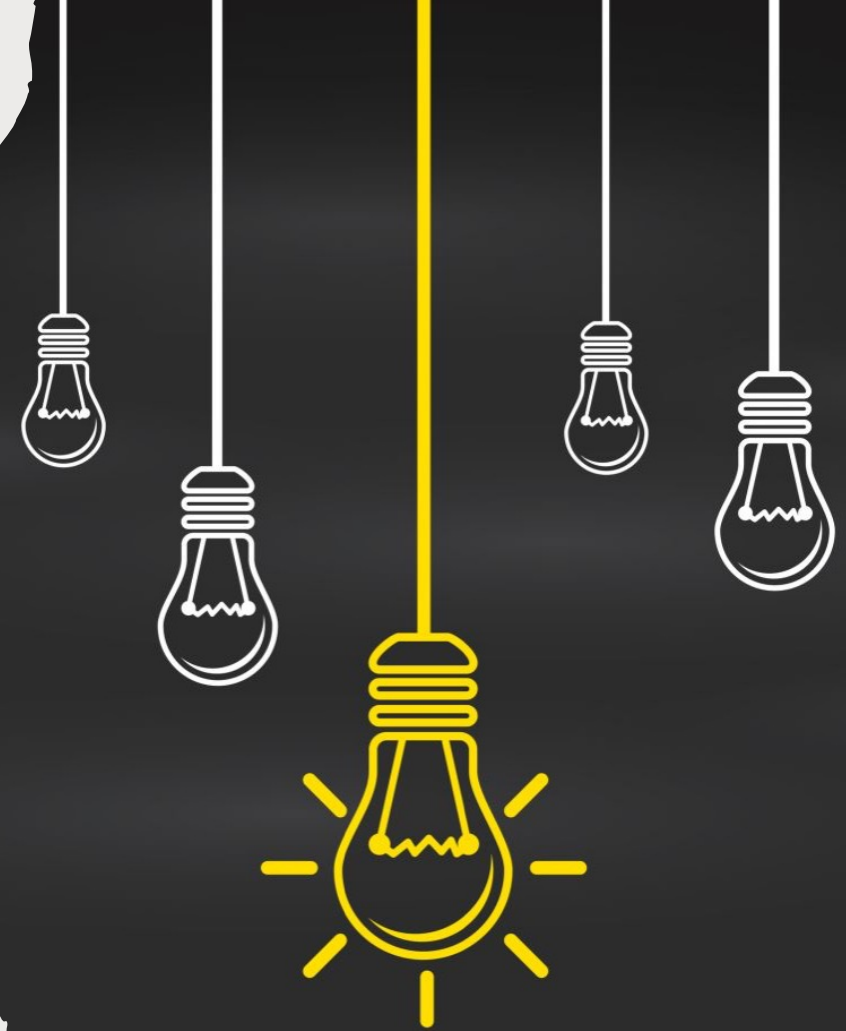# ANSWERING BUSINESS QUESTIONS FOR AN ONLINE VIDEO RENTAL COMPANY

## A DATA ANALYSIS CASE STUDY

Dena Coduri

# THE PROJECT

## Company

Rockbuster Stealth LLC is a fictional movie rental company that used to have stores around the world. The Rockbuster Stealth management team is planning to use its existing movie licenses to launch an online video rental service in order to compete with Netflix and Amazon Prime.

## Objective

Help the business intelligence department with the launch strategy for Rockbuster's new online streaming service and answer any ad-hoc questions from other departments along the way.

# THE QUESTIONS

The Rockbuster Stealth Management Board has a series of business questions, and they expect data-driven answers that they can use for their strategy. Here are the main questions they would like to answer:
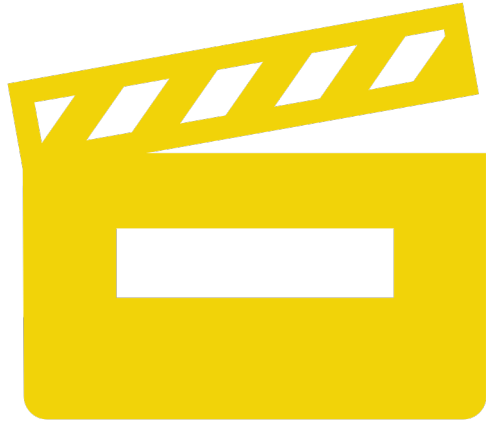
- Which movies contributed the most/least to revenue gain?

- What was the average rental duration for all videos?

- Which countries are Rockbuster customers based in?

- Do sales figures vary between geographic regions?

They have also asked for a data dictionary to be created for their SQL database.

# THE PROCESS

The data used for this analysis was provided by [Career Foundry](Career Foundry).

Software used – SQL, Excel, Google Slides, Tableau

Minimal cleaning was required for this dataset. Consistency issues were fixed.

## Steps

- Fix consistency issues in the data
- Explore the data using SQL queries and answer ad-hoc business questions
- Prepare an Excel file with important SQL queries/outputs for technical colleagues
- Create an interactive map in Tableau
- Create a data dictionary in Excel
- Create a presentation in Google Slides that answers the questions from the business

# THE OBSTACLE

The biggest challenge for this project was technical! At the outset of the project, the Rockbuster data would not load into PostgreSQL on my computer. Without the data, the work could not begin.

I contacted my mentor. After hours of video calls with my mentor, she was able to guide me through the process to reinstall pgAdmin4 on my Macbook through the command line. Prior to that, I had thought the command line was something only hackers used, and I found it intimidating.

The project was set back by a week because of this issue, but I learned that technical problems are common, and it is important to budget more time than you think you will need.

pgAdmin 4
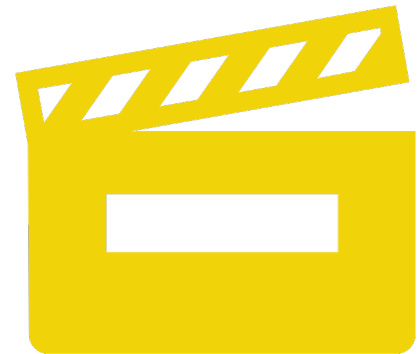
# THE DELIVERABLE

The 14-page data dictionary took some time to create, but the process was straight-forward. I had to inspect every single table in the Rockbuster database, describe every row, and show which other tables were linked to and from the table at hand.

## Dimension Table: city

Lists the name of the city associated with the city id number.

Columns

| Key | Columns | Data Type | Description | To |
|-----|---------|-----------|-------------|-----|
| 🔑 | city_id | SERIAL | Unique identification number for the city. | |
| | city | CHARACTER VARYING(50) | Name of the city. | |
| | country_id | SMALLINT | Unique identification number for the country. | ↗ |
| | last_update | TIMESTAMP(6) WITHOUT TIME ZONE | Date and time the record was last updated. | |

Links to

| Table | Join |
|-------|------|
| country | city.country_id = country.country_id |

Links from

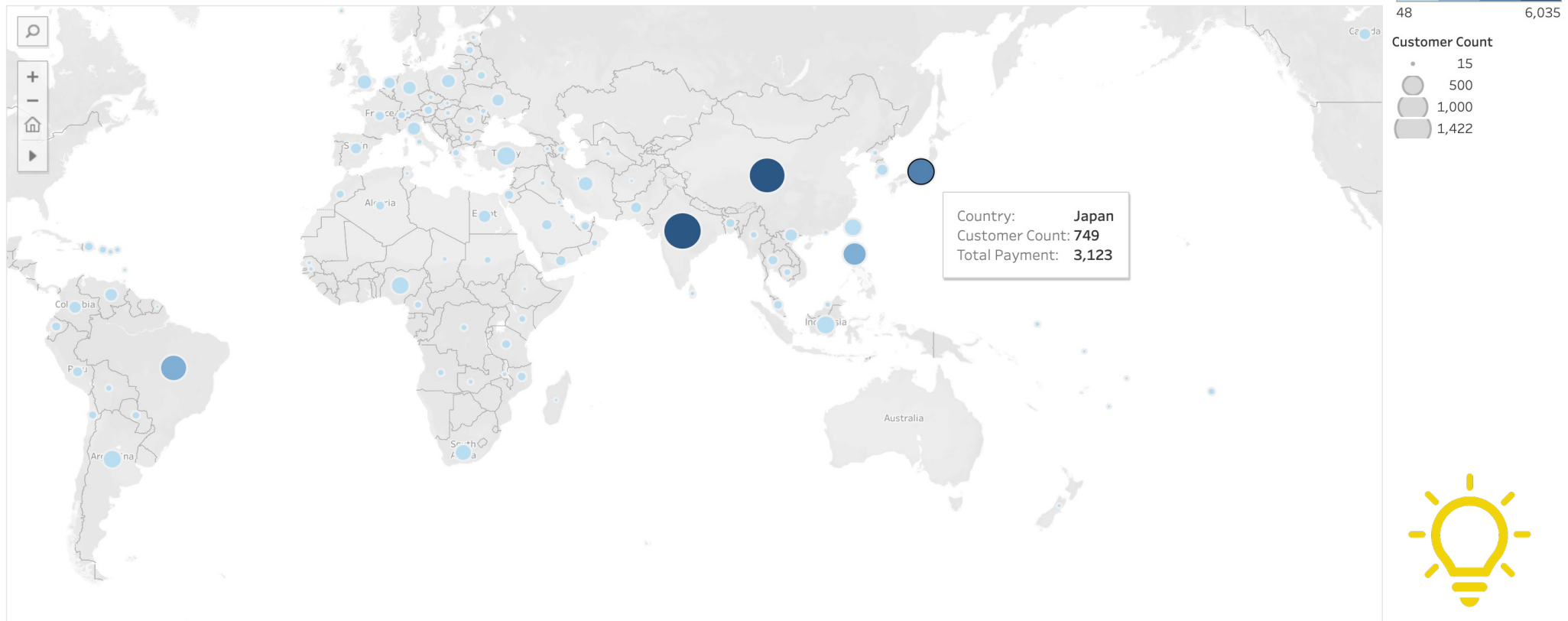| Table | Join |
|-------|------|
| address | city.city_id = address.city_id |

Unique Keys
city_id
country_id

# THE VISUALS



**Rockbuster Project Example** by [Dena](#)

Total Payments and Customer Counts by Country

**Total Payment**
48        6,035

**Customer Count**
- 15
- 500
- 1,000
- 1,422

Country: **Japan**
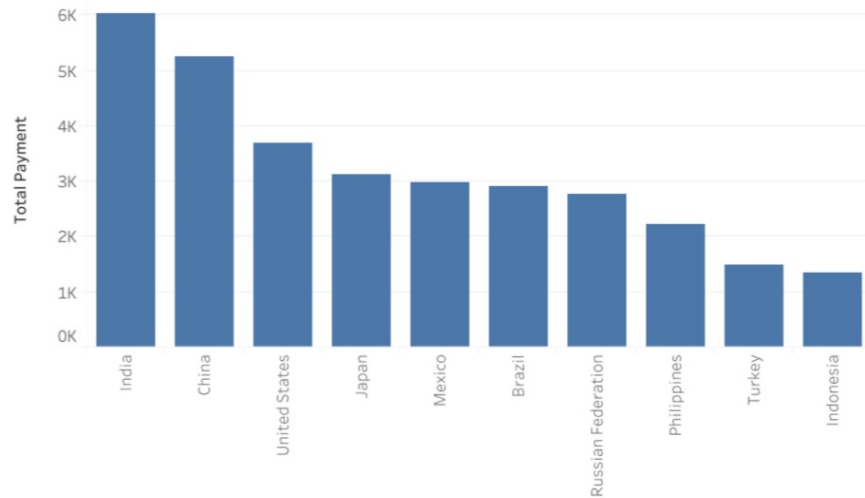Customer Count: **749**
Total Payment: **3,123**

An interactive map was created in Tableau so the business intelligence department could easily see where most of the customers are located and the total payments for each country.
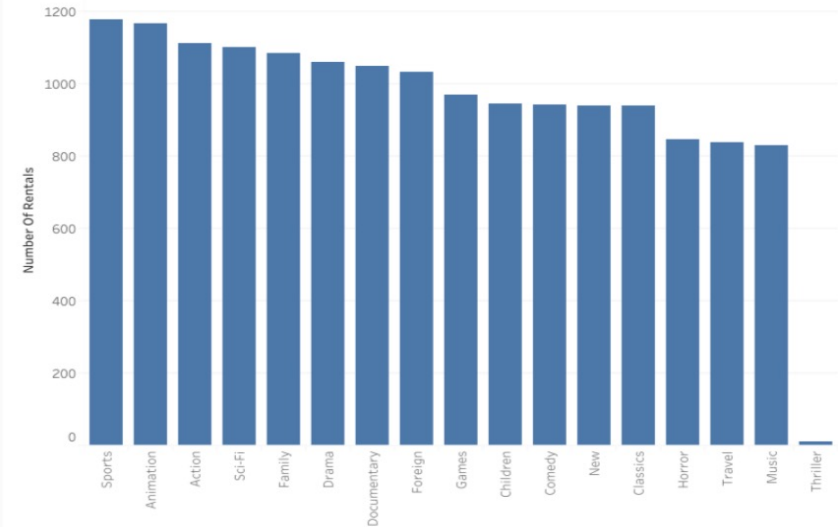
[Link to interactive Tableau map](#)

# THE INSIGHTS



**Which countries are Rockbuster customers based in?**

Total Payments for the top 10 Countries



**Which film categories are the most popular?**

Number of Rentals for each Film Category

Bar charts were created in Tableau to show the top countries and film categories.

- Rockbuster's new online streaming service should launch first in the top two countries: India and China
- The most popular film categories should be released first: Sports and Animation

# THE FEEDBACK

My original SQL queries for the top 10 cities and top 5 customers were hard-coded, meaning I had included specific country/city names in my queries. (See the query to the right).

Although the queries were correct, my mentor challenged me to rewrite them so that as the top countries/cities changed over time, the queries would still be valid.

```sql
SELECT A.customer_id, B.first_name, B.last_name,
    E.country, D.city, SUM(A.amount) AS amount_paid

FROM payment A
    INNER JOIN customer B on A.customer_id = B.customer_id
    INNER JOIN address C ON B.address_id = C.address_id
    INNER JOIN city D ON C.city_id = D.city_id
    INNER JOIN country E ON D.country_ID = E.country_ID
    WHERE city IN ('Aurora', 'Acua', 'Adana', 'Adoni',
                   'Ahmadnagar', 'Akishima', 'Akron',
                   'Allappuzha (Alleppey)', 'Allende', 'Alvorada')

GROUP BY A.customer_id, B.first_name, B.last_name,
        E.country, D.city, A.amount

ORDER BY amount_paid DESC
LIMIT 5
```

Link to GitHub Repository with SQL queries and outputs

# THE RETROSPECTIVE

- This project was my first introduction to SQL, and my first experience analyzing using a program other than Excel.

- I enjoyed answering questions using SQL, though I felt a bit disconnected from the data, because in a SQL database the information is scattered across multiple tables and harder to imagine.

- I feel more confident with analyzing using Excel and Python, so going forward, I would like to practice my SQL skills even more and become comfortable with data that is organized in this manner, preferably using a non-fictional data set.

- I plan to avoid hardcoding my SQL queries in the future, so that they remain useful as the data changes. Reusing old queries makes more sense than rewriting them every time.

- Although the assignment was to find the top 5 cities within the top 10 countries, and the top 5 customers within those top 5 cities, I would like to add to this project by looking up the top 5 cites and customers overall.