



StripeConnect

STRIPE CONNECT

PAYOUT ANALYSIS

Presented by: Shalini Dutta



AGENDA

01

Data Overview and
Cleaning

02

EDA

03

Question 1

04

Question 2

05

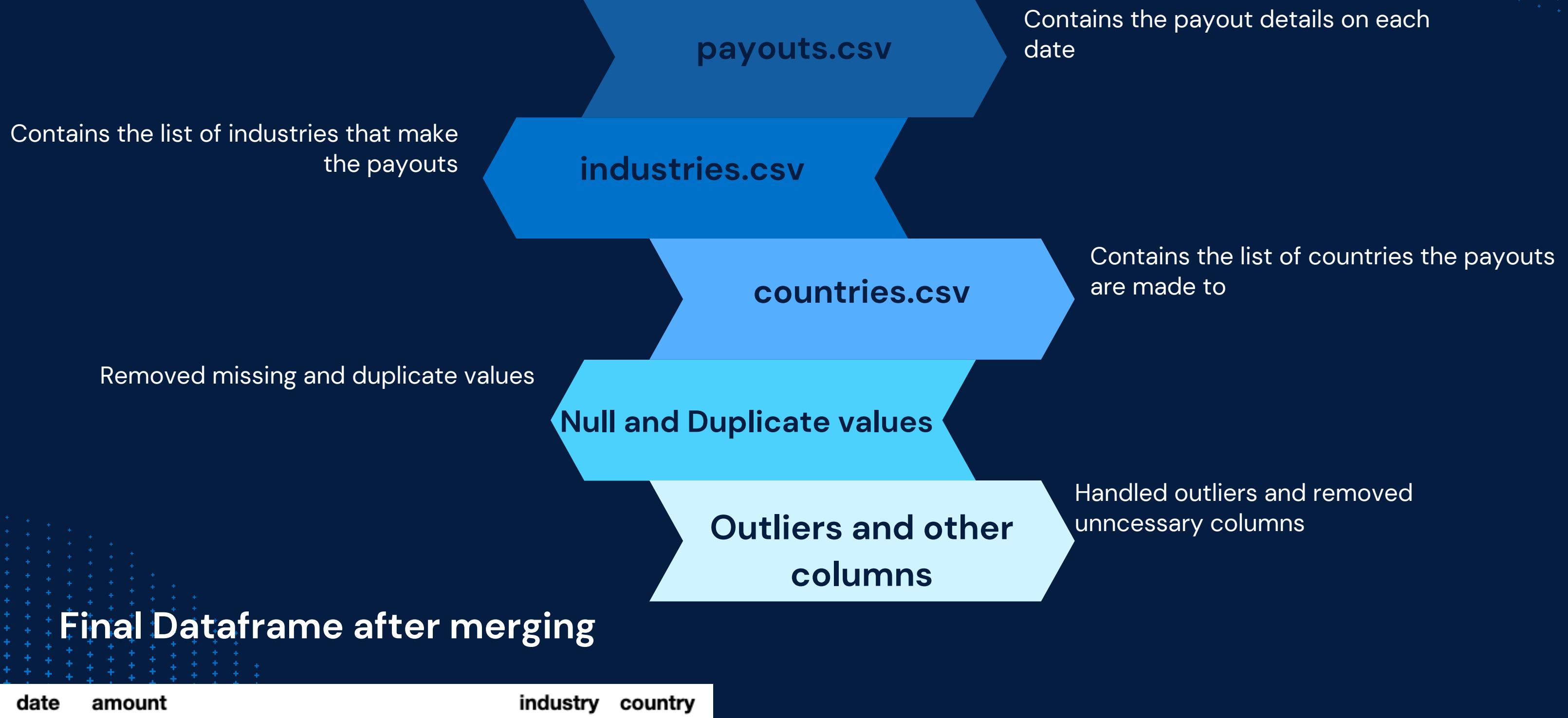
Question 3

06

Future Work

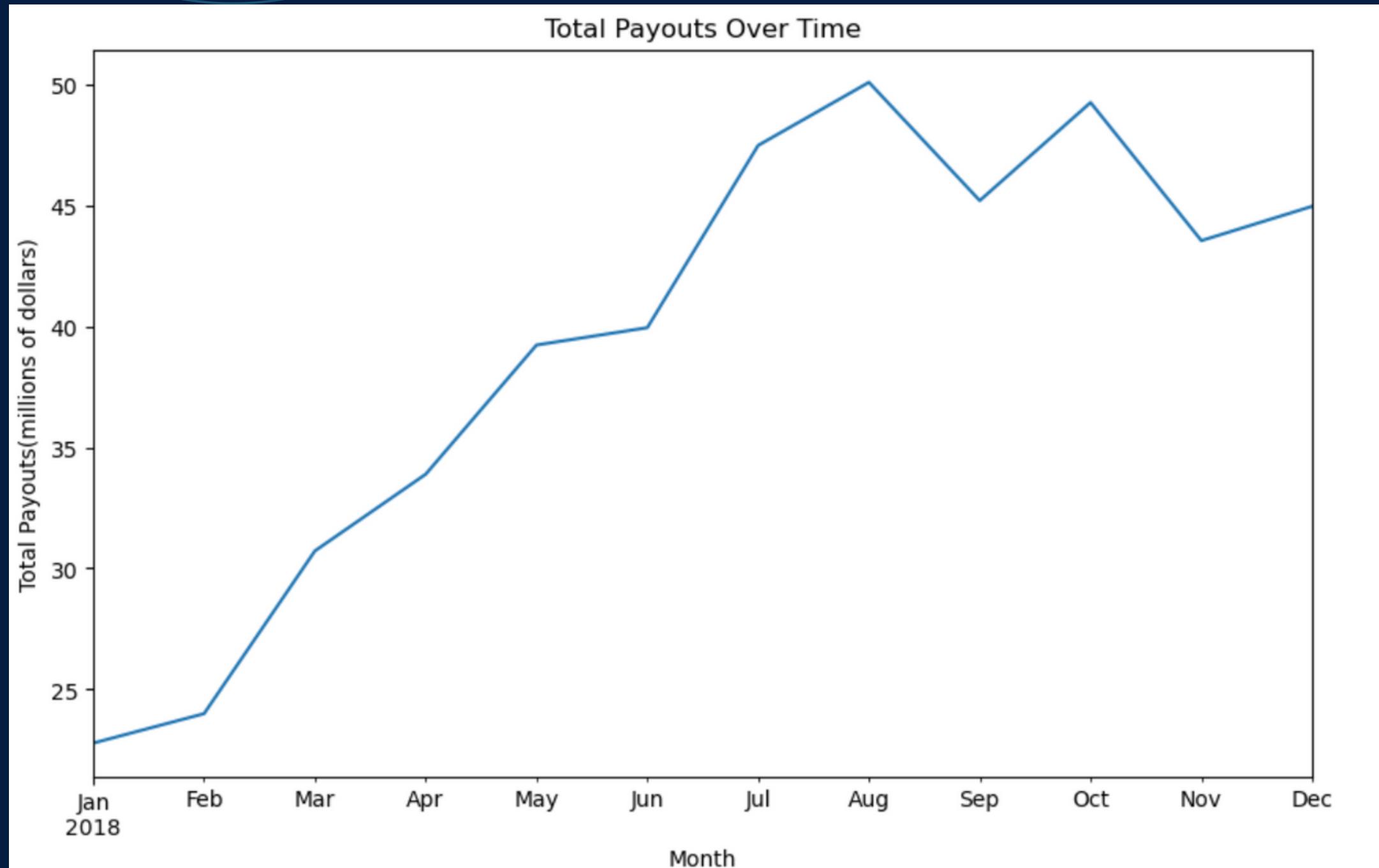


Dataset Overview and Cleaning



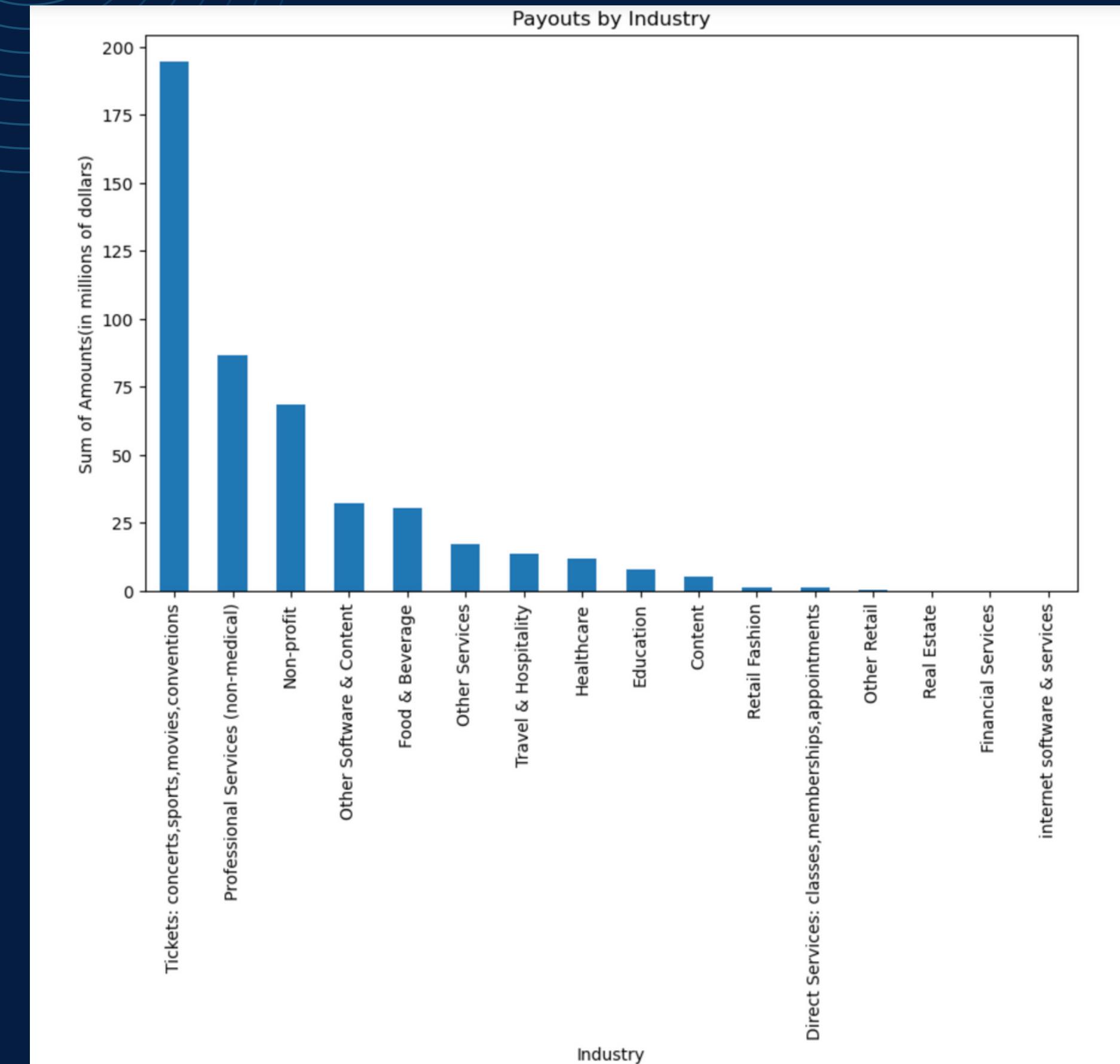
Exploratory Data Analysis

- **General Upward Trend:** Total payouts show a consistent increase throughout the year, indicating overall growth.
- **Seasonal Variations:** Clear cyclic patterns observed:
 - Significant rise from March to May.
 - Peak in August, the highest point in the year.
 - Decline post-August, with a notable dip in September.
 - Rise in October followed by a decline in November and subsequent increase in December.
- **End-of-Year Behavior:** Decrease in total payouts in November and December, possibly indicating seasonal decrease in activity or policy changes.
- **Volatility:** Monthly changes suggest volatility, possibly due to market conditions, business cycles, or specific events.

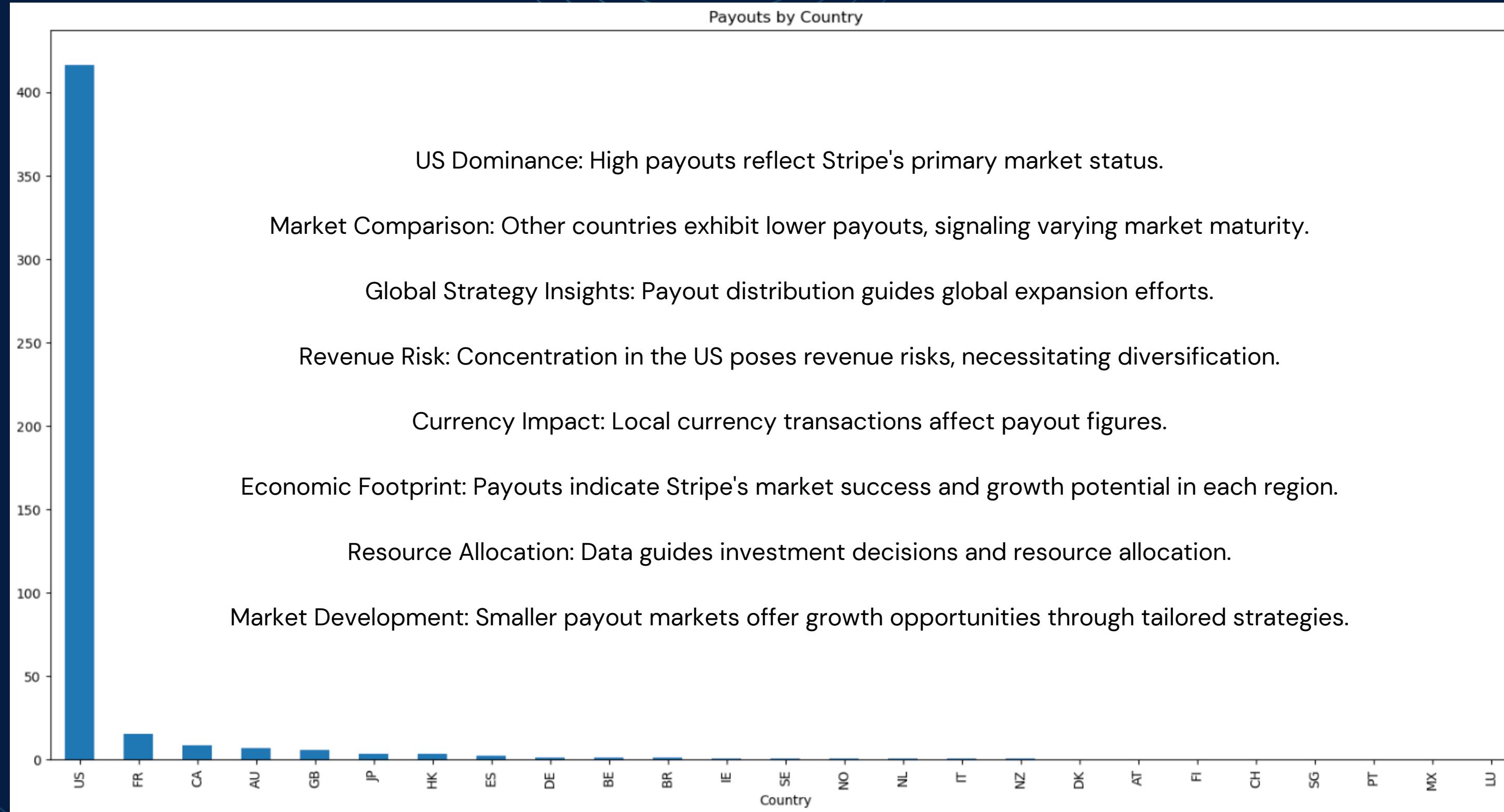


Total Payouts by Industry

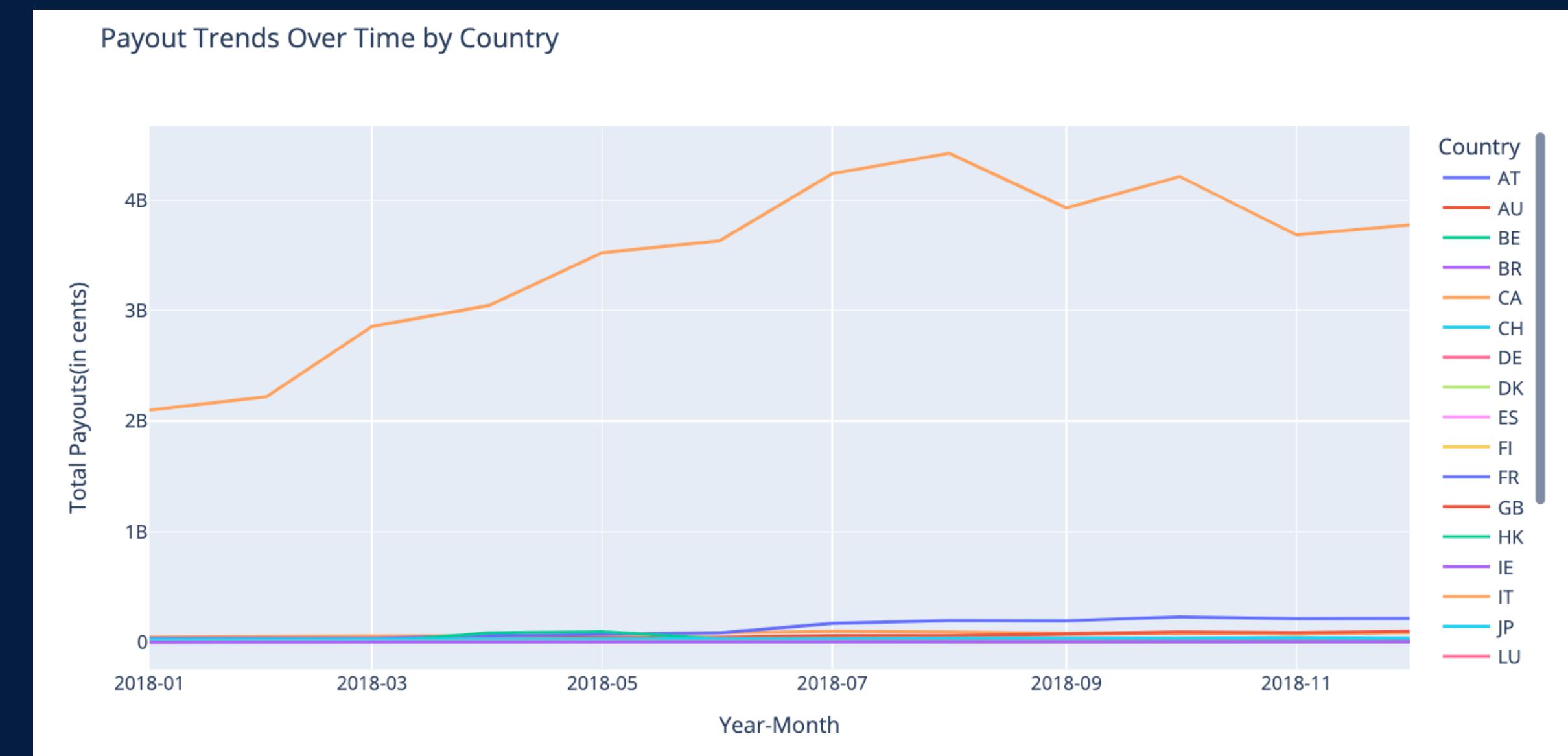
- High Activity Sectors: Ticket, Concerts, Sports, and Movies dominate payouts, nearing \$200 million.
- Notable Recipients: Professional Services and Non-profits receive \$50 to \$100 million, indicating significant funding or grants.
- Moderate Payouts Across Diverse Industries: Sectors like "Other Services" & "Content," "Food & Beverage," "Travel & Hospitality," "Healthcare," and "Education" show moderate payouts around \$25 to \$50 million.
- Lower End Payouts: Industries like "Direct Services," "Other Retail," "Real Estate," "Financial Services," and "Internet Software & Services" receive less than \$25 million.
- Variation in Payouts: Influenced by industry size, external factors, and operational differences.
- Economic Indicators: Higher payouts may signal robustness or compensation issues, while lower payouts could indicate less economic activity or smaller scale.
- Snapshot of Activity: Offers insights into economic health across sectors for stakeholders and analysts.



Total Payouts by Country



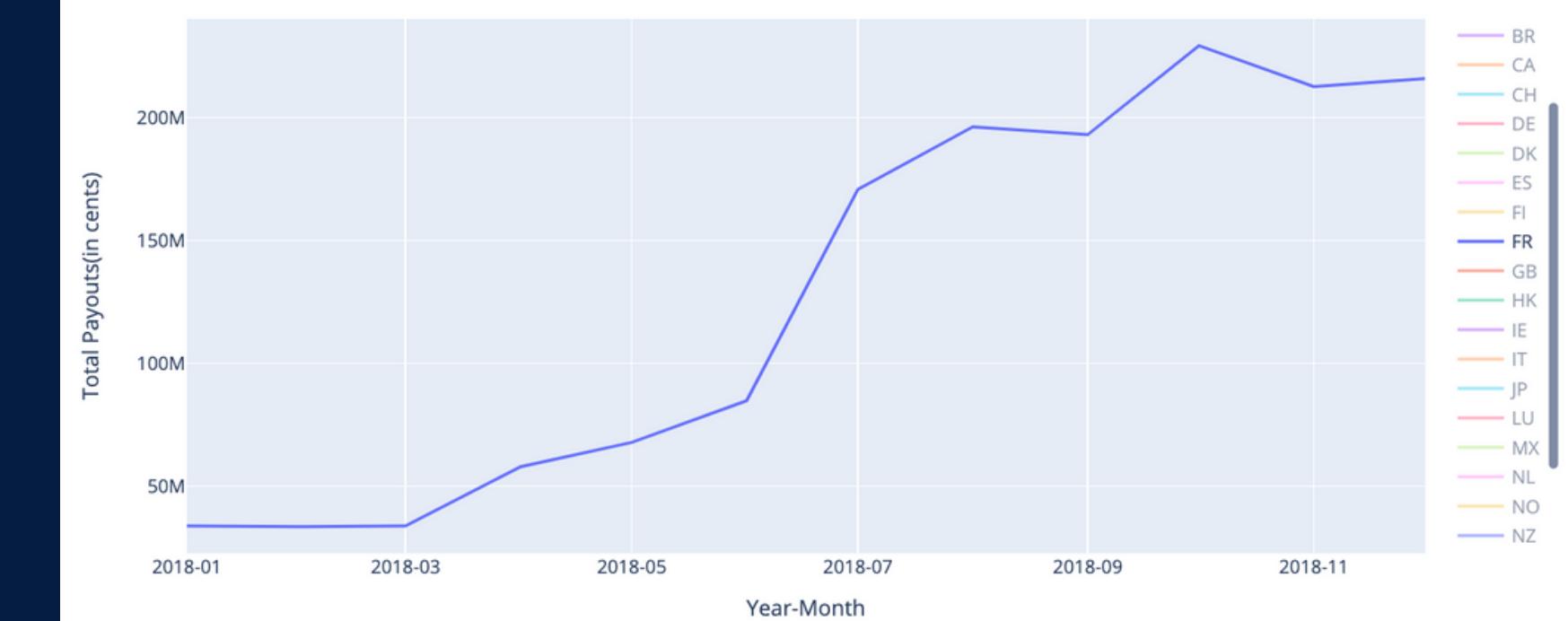
Interactive country-wise total payout trends over time (2018)



UNITED STATES



FRANCE





QUESTION 1

Using the data provided, please provide an estimate for the amount of money we should expect to be paid out to each country on Jan. 1, 2019 (the day after the last day in the dataset).



Estimation using average payout projection per country

- Data Preparation:

Convert 'date' column to datetime format.

Filter dataset to include data up to December 31, 2018.

- Calculation of Averages:

Group historical data by country.

Calculate average payout for each country based on data up to end of 2018.

- Estimation for January 1, 2019:

Create empty dictionary for estimated payouts.

Loop through each country in average payout data.

Filter historical data for each country.

Recalculate average payout for specific country.

Add entry to dictionary with country code as key and calculated average payout as value.

- Output:

Print estimated payouts for each country on January 1, 2019, based on historical averages up to December 31, 2018.

Estimated payout for AT on January 1, 2019: 10187.973089035213
Estimated payout for AU on January 1, 2019: 27490.672099906744
Estimated payout for BE on January 1, 2019: 10033.307322485207
Estimated payout for BR on January 1, 2019: 4517.70773574758
Estimated payout for CA on January 1, 2019: 29264.170085925532
Estimated payout for CH on January 1, 2019: 13802.21639108555
Estimated payout for DE on January 1, 2019: 14262.592722448593
Estimated payout for DK on January 1, 2019: 21865.59841179807
Estimated payout for ES on January 1, 2019: 11735.575507580223
Estimated payout for FI on January 1, 2019: 52139.25811688312
Estimated payout for FR on January 1, 2019: 12982.005211604734
Estimated payout for GB on January 1, 2019: 20416.848813905184
Estimated payout for HK on January 1, 2019: 6729.562694959369
Estimated payout for IE on January 1, 2019: 26681.610078828828
Estimated payout for IT on January 1, 2019: 18091.337721102427
Estimated payout for JP on January 1, 2019: 11733.38207858287
Estimated payout for LU on January 1, 2019: 14825.802816901409
Estimated payout for MX on January 1, 2019: 44187.1746031746
Estimated payout for NL on January 1, 2019: 29354.23180940116
Estimated payout for NO on January 1, 2019: 19081.77555816686
Estimated payout for NZ on January 1, 2019: 21077.373260437376
Estimated payout for PT on January 1, 2019: 26539.898496240603
Estimated payout for SE on January 1, 2019: 28772.145266594125
Estimated payout for SG on January 1, 2019: 39948.200902934535
Estimated payout for US on January 1, 2019: 39224.5372213672

Estimation using Linear Regression

- Data Preparation:

Convert 'date' column to datetime format and localize to UTC timezone.

Filter dataset to include data up to December 31, 2018.

- Linear Regression Model:

Calculate average payout per country.

For each country, filter historical data.

Prepare data for linear regression: X as day of the year, y as payout amount.

- Model Training and Prediction:

Fit linear regression model for each country.

Predict payout for January 1, 2019, using fitted model.

- Result Storage and Output:

Store prediction in dictionary with country as key.

Print estimated payouts for each country on January 1, 2019, based on model predictions.

```
Estimated payout for AT on January 1, 2019: 11900.128239685124
Estimated payout for AU on January 1, 2019: 29291.170893305647
Estimated payout for BE on January 1, 2019: 11180.6460323712
Estimated payout for BR on January 1, 2019: 4913.970268421807
Estimated payout for CA on January 1, 2019: 30635.538995989584
Estimated payout for CH on January 1, 2019: 17380.37366757656
Estimated payout for DE on January 1, 2019: 16862.980148595787
Estimated payout for DK on January 1, 2019: 23392.95333972881
Estimated payout for ES on January 1, 2019: 15259.601906902708
Estimated payout for FI on January 1, 2019: 53566.3364268343
Estimated payout for FR on January 1, 2019: 11084.054131142948
Estimated payout for GB on January 1, 2019: 22364.231124804224
Estimated payout for HK on January 1, 2019: 1851.2506432769696
Estimated payout for IE on January 1, 2019: 24425.14418670752
Estimated payout for IT on January 1, 2019: 15756.753661847979
Estimated payout for JP on January 1, 2019: 10678.153836364698
Estimated payout for LU on January 1, 2019: 22746.245694694764
Estimated payout for MX on January 1, 2019: 41525.06819874432
Estimated payout for NL on January 1, 2019: 22575.17750686259
Estimated payout for NO on January 1, 2019: 20732.922897790228
Estimated payout for NZ on January 1, 2019: 28651.15673691267
Estimated payout for PT on January 1, 2019: 28587.157094920447
Estimated payout for SE on January 1, 2019: 31445.98016410094
Estimated payout for SG on January 1, 2019: 58314.33205279428
Estimated payout for US on January 1, 2019: 39950.75106205013
```

Estimation using Random Forest Algorithm

- Data Preparation:

Convert 'date' column to datetime format and localize to UTC timezone.

Filter dataset to include data up to December 31, 2018.

- Modeling Preparation:

Calculate average payout per country (for reference).

Extract country-specific data.

Prepare inputs for Random Forest model: X as day of the year, y as payout amounts.

- Random Forest Regression:

Instantiate Random Forest regression model for each country with specified parameters.

Train the model using country-specific data.

- Prediction:

Calculate day number for January 1, 2019, from minimum date in dataset.

Predict payout for January 1, 2019, using trained model.

- Result Storage and Output:

Store predictions in dictionary with country as key.

Print estimated payouts for each country on January 1, 2019.

Estimated payout for AT on January 1, 2019:	16259.999549119872
Estimated payout for AU on January 1, 2019:	30774.272579654218
Estimated payout for BE on January 1, 2019:	12972.34985930787
Estimated payout for BR on January 1, 2019:	3738.6484155101907
Estimated payout for CA on January 1, 2019:	31480.394849765315
Estimated payout for CH on January 1, 2019:	5565.779118409365
Estimated payout for DE on January 1, 2019:	13383.203801784002
Estimated payout for DK on January 1, 2019:	23491.9092950938
Estimated payout for ES on January 1, 2019:	11654.206751419588
Estimated payout for FI on January 1, 2019:	51316.97948917749
Estimated payout for FR on January 1, 2019:	13194.363082312637
Estimated payout for GB on January 1, 2019:	17933.050785545864
Estimated payout for HK on January 1, 2019:	46149.96233333334
Estimated payout for IE on January 1, 2019:	19711.502215052657
Estimated payout for IT on January 1, 2019:	33543.28187662338
Estimated payout for JP on January 1, 2019:	12971.83567728118
Estimated payout for LU on January 1, 2019:	27847.869454394866
Estimated payout for MX on January 1, 2019:	42958.61855555555
Estimated payout for NL on January 1, 2019:	25027.665505772
Estimated payout for NO on January 1, 2019:	20990.887708102357
Estimated payout for NZ on January 1, 2019:	36739.53431007882
Estimated payout for PT on January 1, 2019:	33155.60828571428
Estimated payout for SE on January 1, 2019:	35485.42913095238
Estimated payout for SG on January 1, 2019:	45900.27911904761
Estimated payout for US on January 1, 2019:	56037.81327702984

- **Varying Estimates:**

Noticeable variation in estimated payouts among different models suggests each captures different aspects of data or has different sensitivities.

- **Model Agreement:**

Consensus in some countries indicates clear trends or less volatility in historical data.

- **High Variability:**

Significant discrepancies in certain countries could stem from outliers, noise, or non-linear patterns.

- **Peak Predictions:**

Some countries show peak predictions, suggesting higher payout activities or seasonal spikes.

- **Low Predictions:**

Other countries exhibit consistently lower estimated payouts, possibly indicating lower activity levels.

- **Country-Specific Trends:**

Random Forest model shows more fluctuation, indicating sensitivity to non-linear trends, while Linear Regression provides smoother estimates.

- **Model Preference:**

Consistent accuracy against actual payouts can indicate superiority or suitability of a particular model.

- **Actionable Insights:**

Insights help identify target countries for business development or risk mitigation strategies.

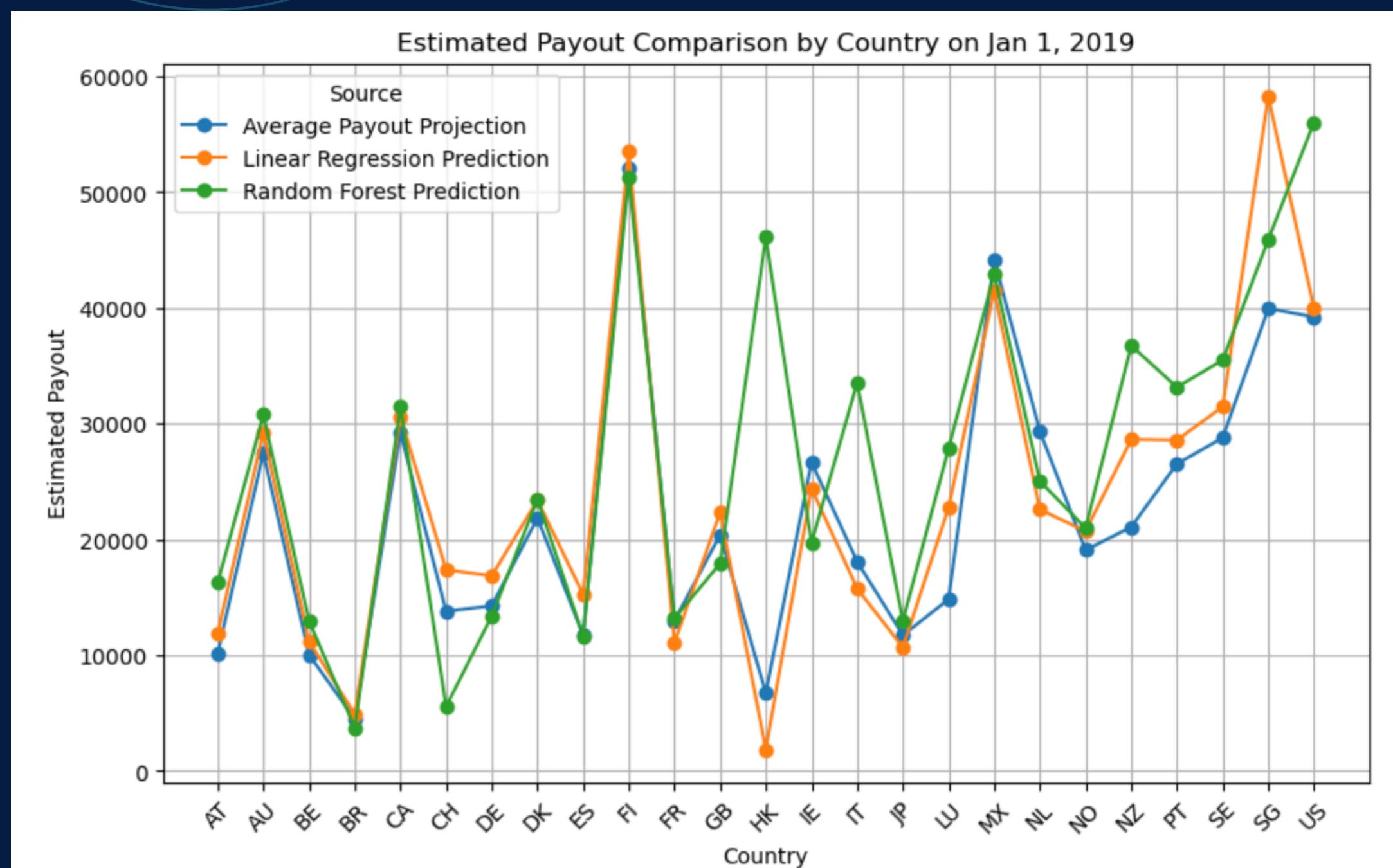
- **Model Improvement:**

Differences between predictions can inform model refinement, such as incorporating seasonal effects or market-specific factors.

- **Strategic Decisions:**

Consensus on growth in certain countries suggests capitalizing on increasing markets, while lower predictions may require stimulating growth strategies.

Comparison of the estimation





QUESTION 2

Assume that one year from the end of the dataset, we will have 15 platforms from the Education industry, 5 from Hotels, Restaurants & Leisure and 40 from Food & Beverage. How much total payout volume do you expect to see on a typical day in 2019, from these three industries?



Define platforms_2019 dictionary indicating expected platforms for three industries.

Sum all expected daily payout volumes across industries to compute total_expected_payout_2019.

Expected total payout volume on a typical day in 2019:
Education: 123242808.88579386
Hotels, Restaurants & Leisure: 0
Food & Beverage: 460854030.5753425
Total across the three industries: 584096839.46

Calculate Daily Totals and Averages

- Compute daily_totals by summing 'amount' for each industry and date group.
- Calculate daily_averages by taking the mean of daily_totals grouped by industry.

Expected Number of Platforms in 2019

Estimate Expected Daily Payout Volume

Calculate Total Expected Payout Volume

Display the Results

- Initialize expected_payouts_2019 dictionary.
- Iterate over platforms_2019, retrieving historical daily average payout for each industry.
- Calculate expected_payout by multiplying historical average payout by expected platforms in 2019.
- Store calculated expected_payout in expected_payouts_2019 dictionary.

Print expected daily payout volume for each industry and total for all industries combined, formatting total to two decimal places.



QUESTION 3

Given this data, what metrics would you define and track in order to help product and business teams understand the progress and behavior of Stripe Connect?



METRICS TO TRACK

Volume Metrics:

Total Payout Volume
Average Payout Size
Growth Rate



Temporal Metrics:

Seasonality
Time to Payout



Operational Metrics:

Error Rate
Customer Support Queries
Innovation and Growth



Metrics:

New Market Penetration
Platform Retention Rate

Distribution Metrics:

Country-wise Distribution
Industry-wise Distribution
Recipient Concentration

Operational Metrics:

Error Rate
Customer Support Queries

Innovation and Growth Metrics:

New Market Penetration
Platform Retention Rate

Visualization and Reporting:

Dashboards
Trend Analysis
Geographical Analysis
Industry Comparison



FUTURE WORK

Advanced Techniques

Employ advanced analytical techniques including time series forecasting, machine learning models, clustering analysis, anomaly detection, causal inference models, sentiment analysis, and cross-industry/country analysis.

Growth and Success

These approaches enable better prediction of payout behaviors, understanding of market dynamics, optimization of operations, and strategic expansion, contributing to Stripe's growth and success in the fintech industry.