Cohort Analysis for IronHack Payments

by:
Jarian N. Del Valle
Diego Rosa Paz
Natalia Torres
Dylan Rodriguez
Javier A. Dastas

IRON
HACK

PUERTO RICO
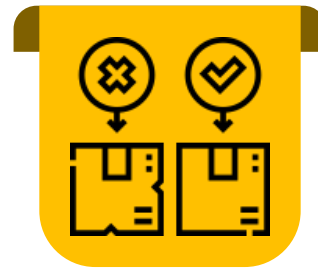
Exploration Data Analysis

EXPLORATION

Examine the structure, data types, and relationships between datasets.

Analyze distributions and key descriptive statistics.

Identify missing values, duplicates, and inconsistent formats.

DATA CLEAN & QUALITY

Correct inconsistent or missing values.

Secure standardize formats (e.g., dates, types, status).

Create new columns derived from existing data (e.g., cohorts based on creation month).
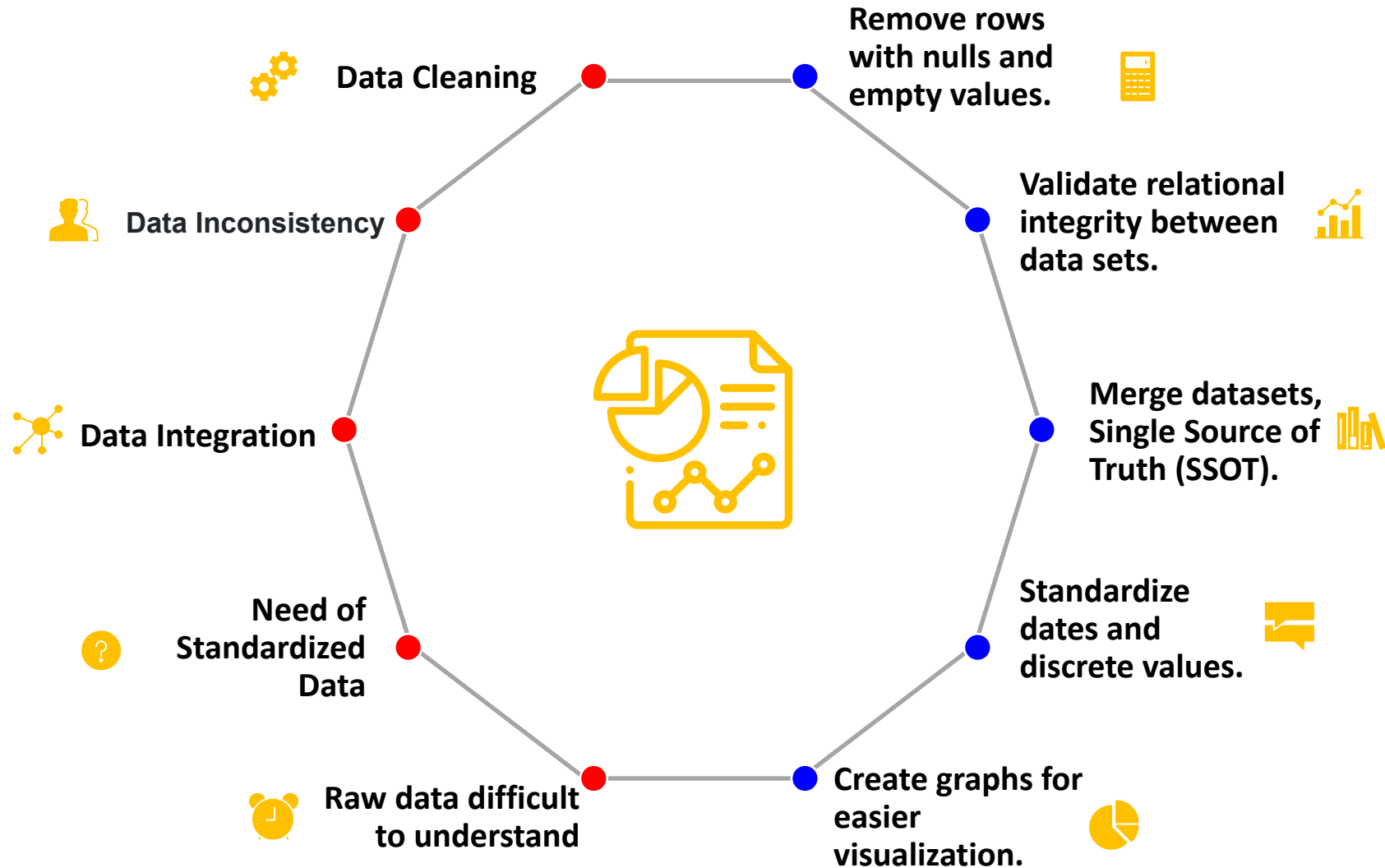
VISUALIZATION

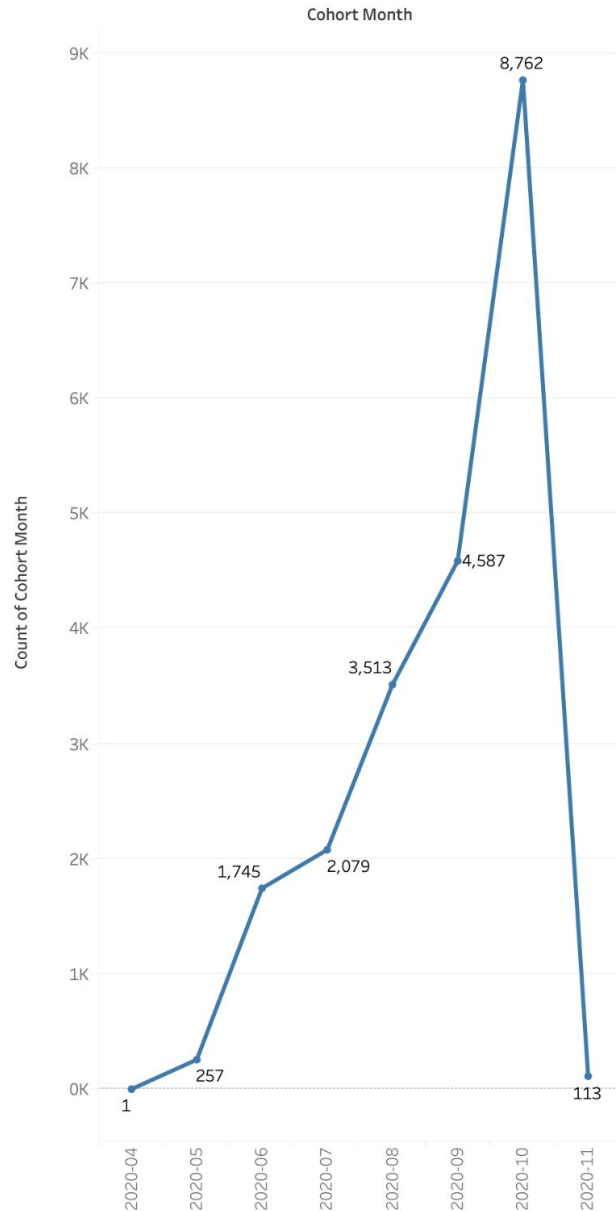Document results, observed patterns, and interpretations.

Create visualizations to explore trends and patterns (e.g., histogram of amounts, usage frequency by month).

Create charts to showcase the monthly evolution of metrics.
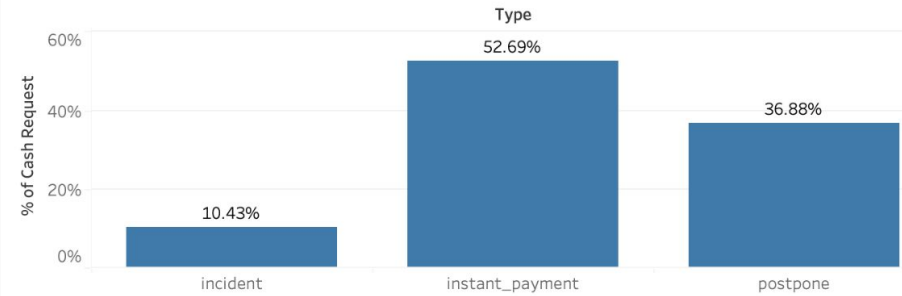
Challenges and Solutions

Data Cleaning

Data Inconsistency

Data Integration

Need of Standardized Data

Raw data difficult to understand

Remove rows with nulls and empty values.

Validate relational integrity between data sets.

Merge datasets, Single Source of Truth (SSOT).

Standardize dates and discrete values.

Create graphs for easier visualization.

FREQUENCY OF REQUEST BY COHORT
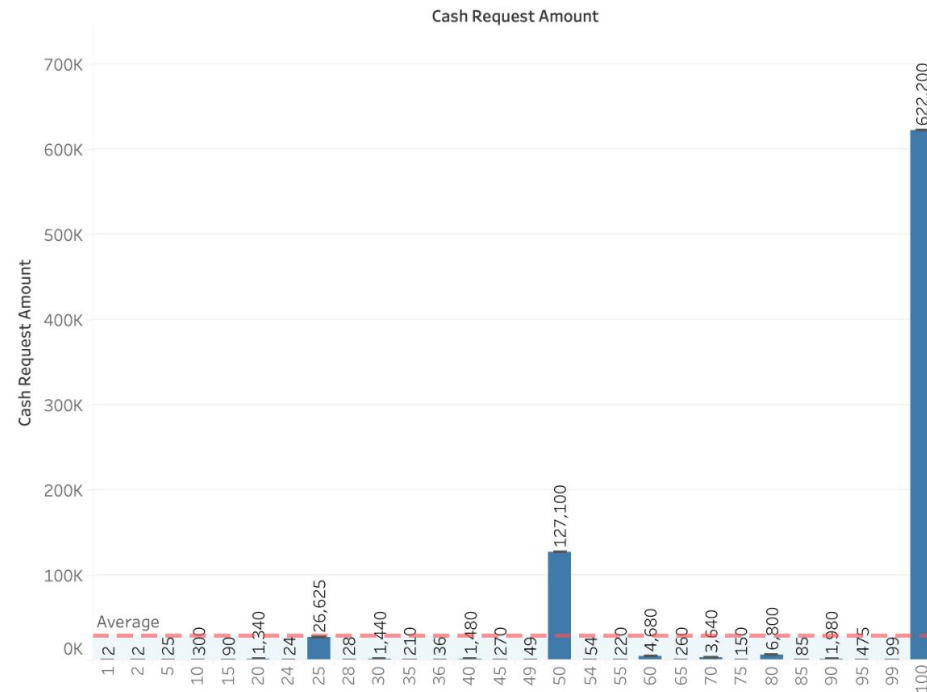
Cohort Month
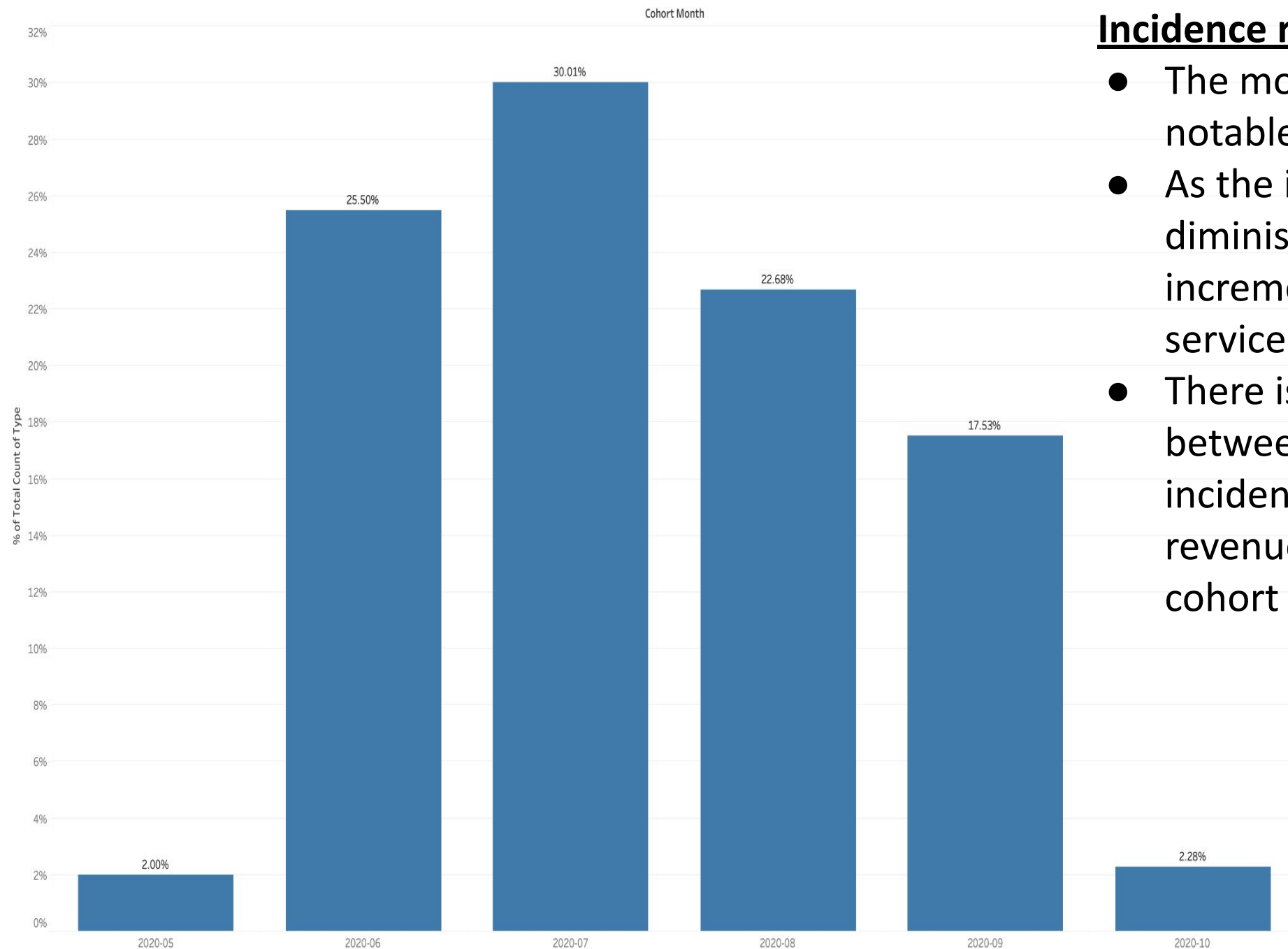
TOTALS

Rate of Services Generate (All)

Instant Payment and Accepted Income: $799,664
Effort Index: **2.045**   Success Index: **0.4890**

Cash Frequency Distribution

Success vs Effort Index

- More effort than necessary is being invested in transactions, but an adequate completion rate is not being achieved.

- Future Analysis: check the cost per transaction, possible bad data, problems with the fees request system, bad process...

Incidents Rates by Cohort by Cohort

Cohort Month

Incidence rate

- The month with the highest notable incident rate is July.
- As the incident rate diminished, there was an increment in the total of services and clients.
- There is a direct correlation between the increment of incidents and the reduced revenue generated in the cohort month.

Highly Seasonal Pattern

- Most CRs are grouped nearest to end of year holidays, peaking significantly in October and November.
- The yearly average on CRs of all status does peak nearer to August, suggesting a prognosis of the peak months may be possible if some external factors come into consideration.
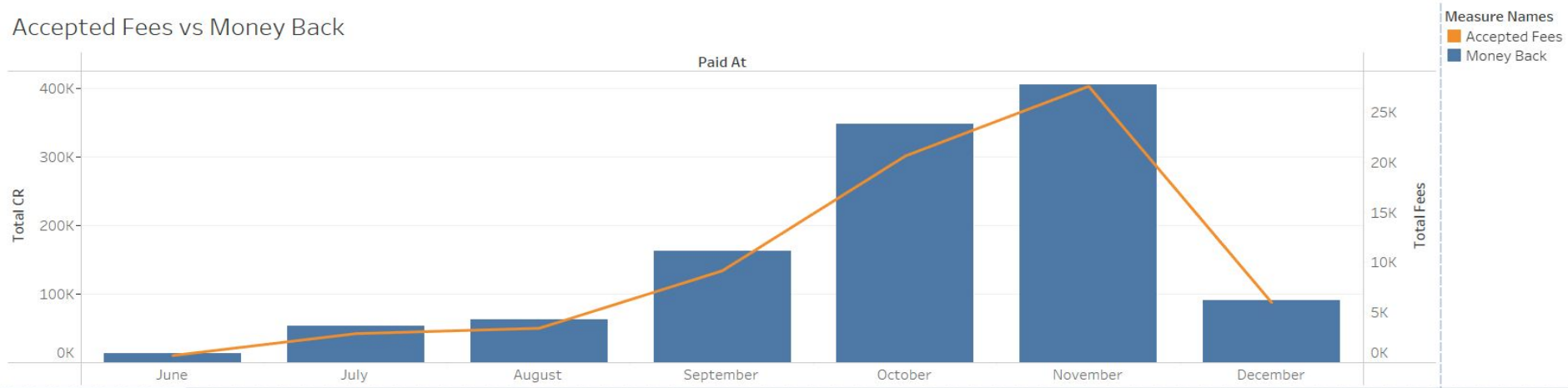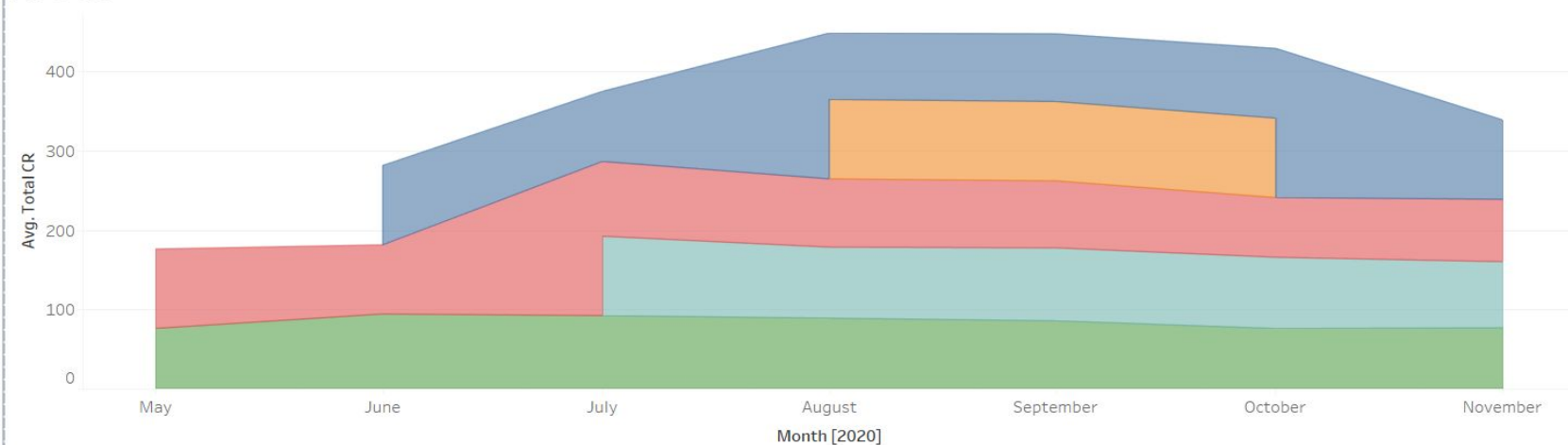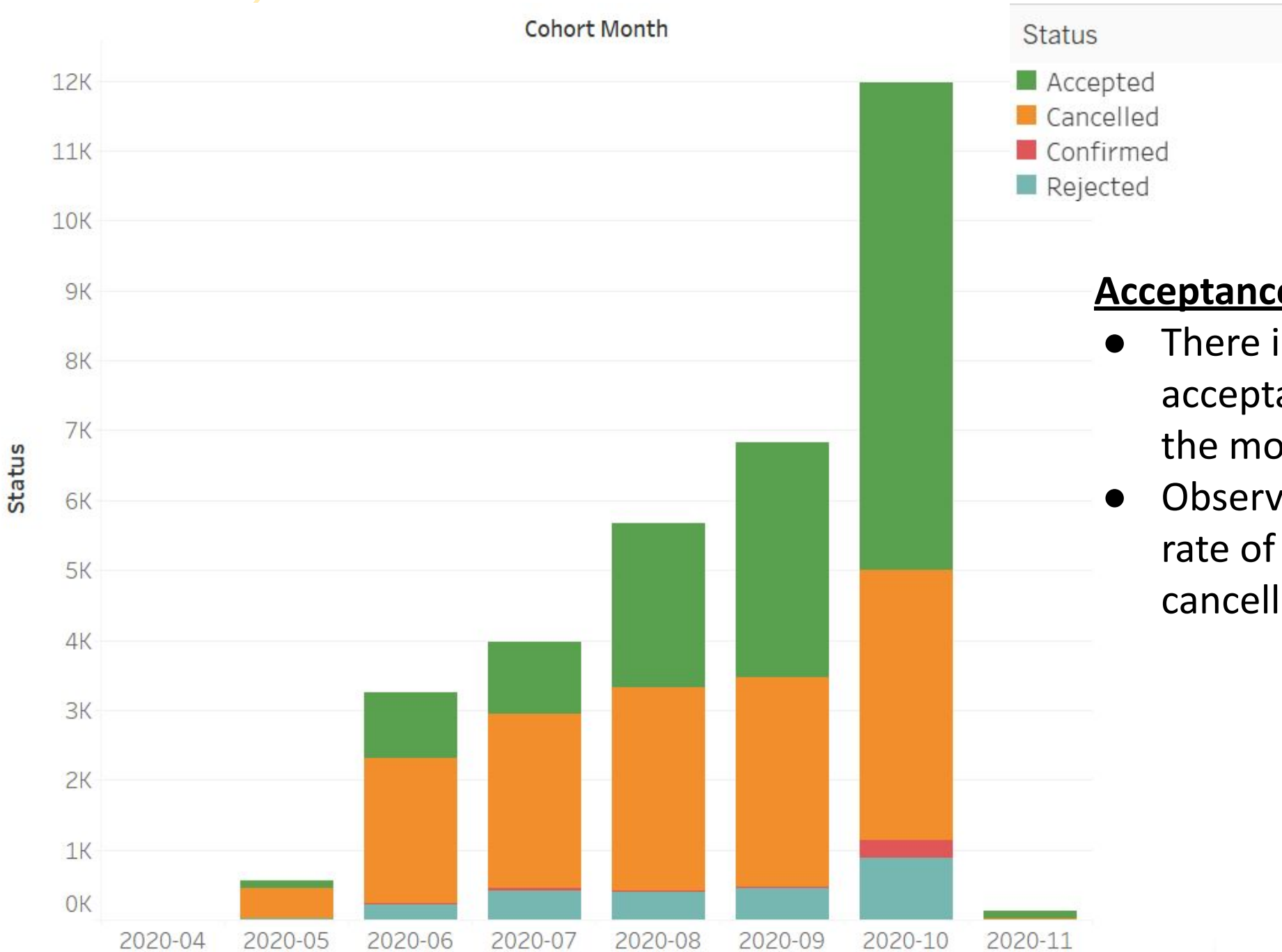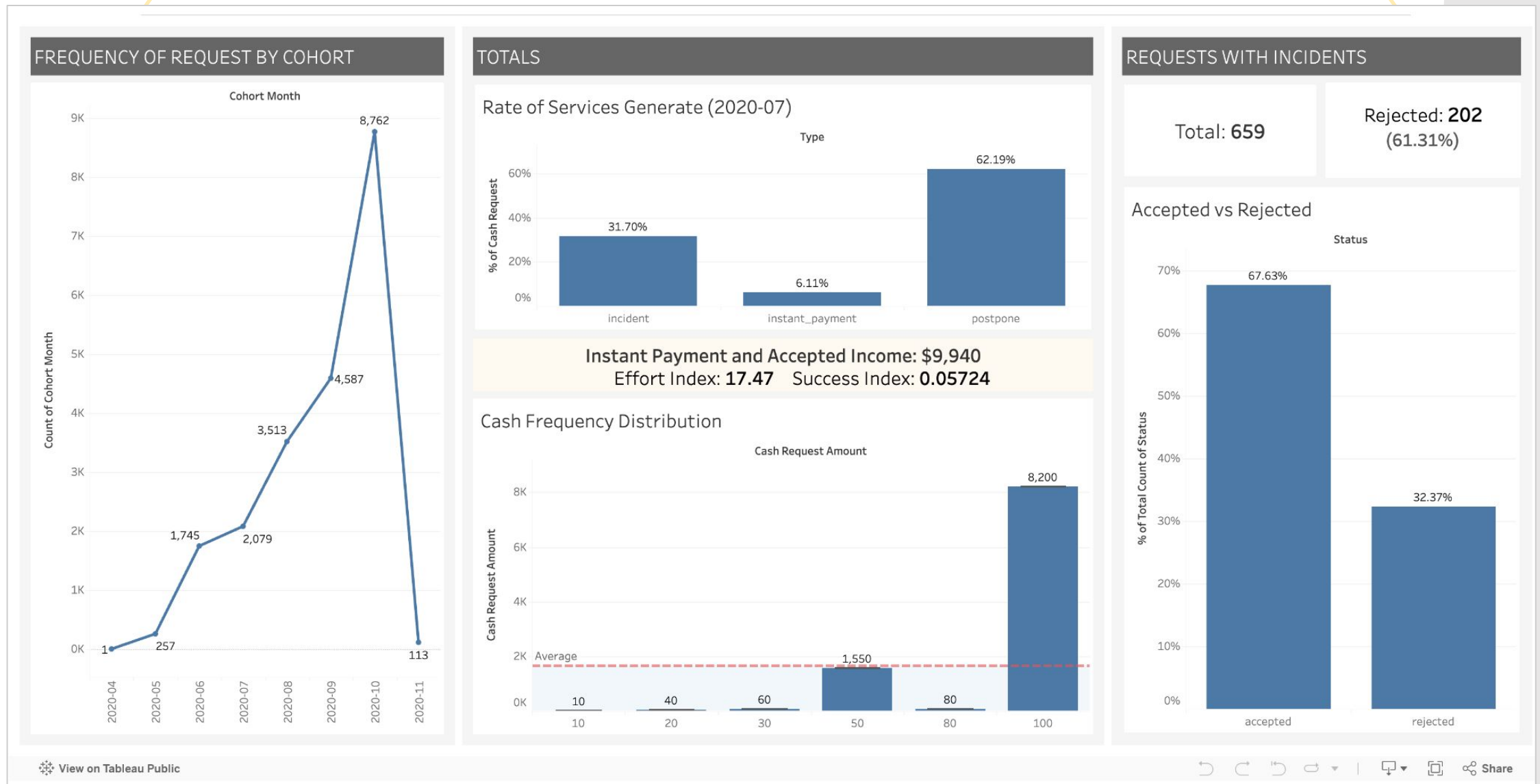
Acceptance rate
- There is an increment in the acceptance rate throughout the months during the cohort.
- Observable difference in the rate of acceptance and cancellation.

Dashboard Publication in Tableau Public

https://public.tableau.com/views/DSMLBootcampProjectI-DemoData/Dashboard1?:language=en-US&publish=yes&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

Strelit - Analysis Dashboard

http://dsmlbootcamp.org/

Conclusion

- Cleaned and merged datasets for a comprehensive analysis.

- Identified patterns in service usage, incident rates, and revenue generation.

  - Developed a new metric focused on instant payment efficiency and monthly cohorts.

  - Observed a rise in instant payment acceptance, correlating with client growth.

  - Detected seasonal peaks in service usage during late-year holidays.

  - Detected more effort than necessary is being invested in transaction.

- Recommended prioritizing holiday marketing or incentive programs for instant payments to boost profitability.

- Suggested extracting client location data to identify top-performing regions and demographics.

Cohort Analysis for IronHack Payments

THANK YOU

by:
Jarian N. Del Valle
Diego Rosa Paz
Natalia Torres
Dylan Rodriguez
Javier A. Dastas

IRON HACK

PUERTO RICO