# Cohort Analysis for IronHack Payments

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

IRON HACK

PUERTO RICO
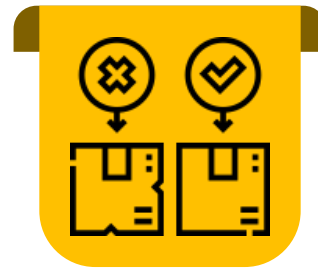
# Exploration Data Analysis

## 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).
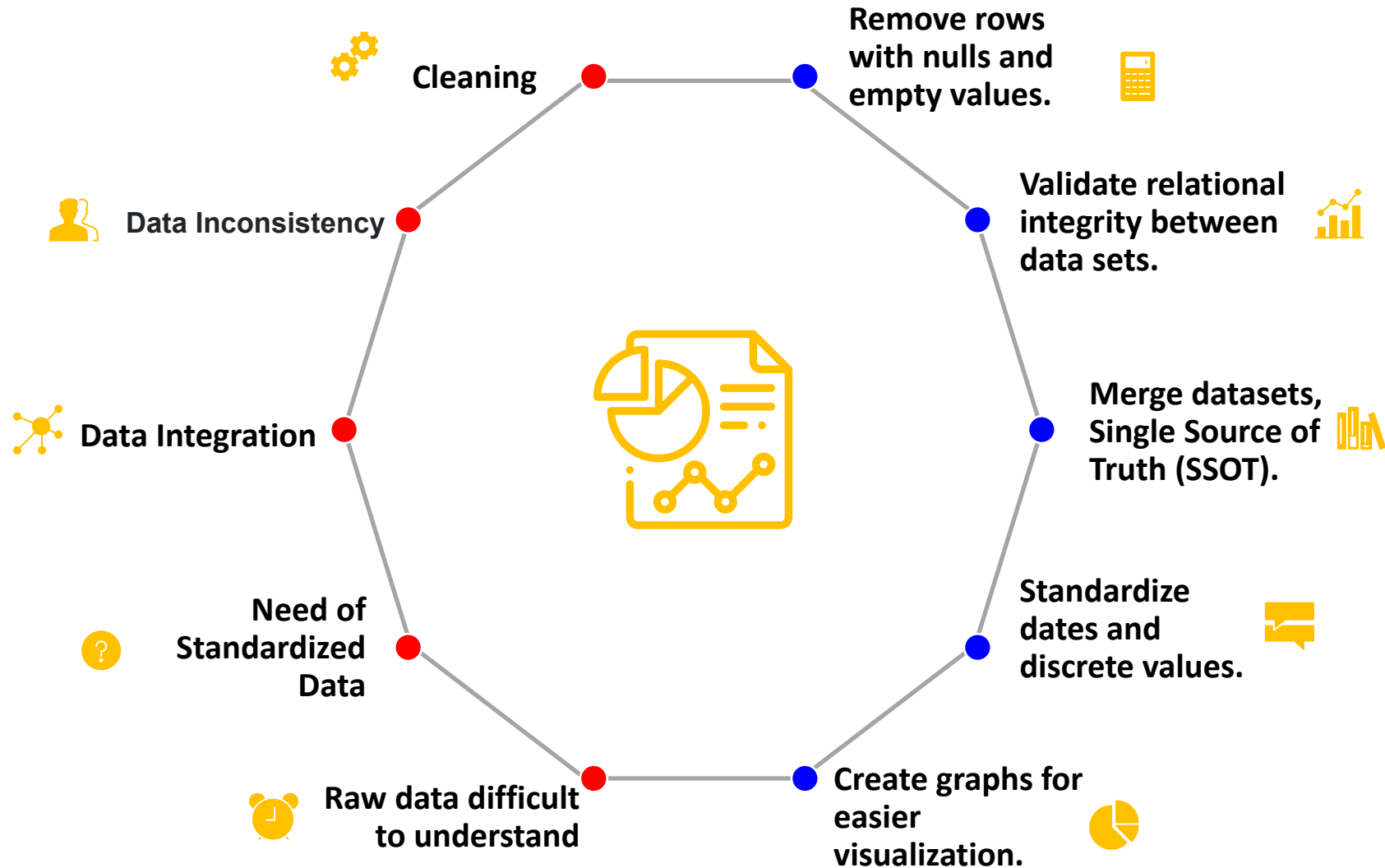
## VISUALIZATION

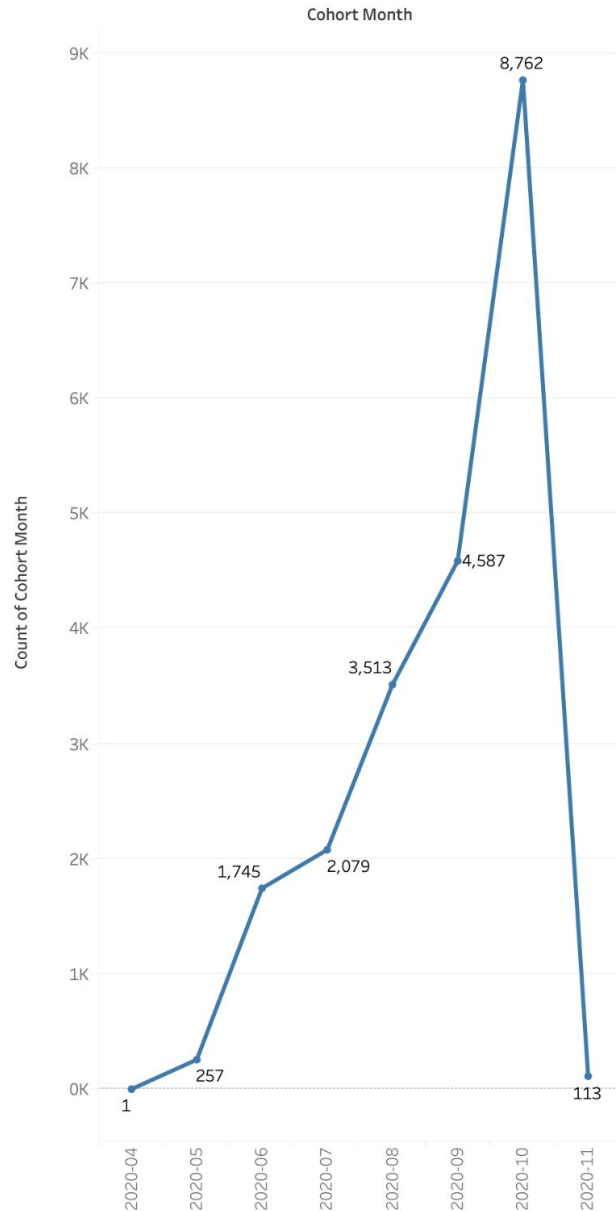Document results, observed patterns, and interpretations.

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.
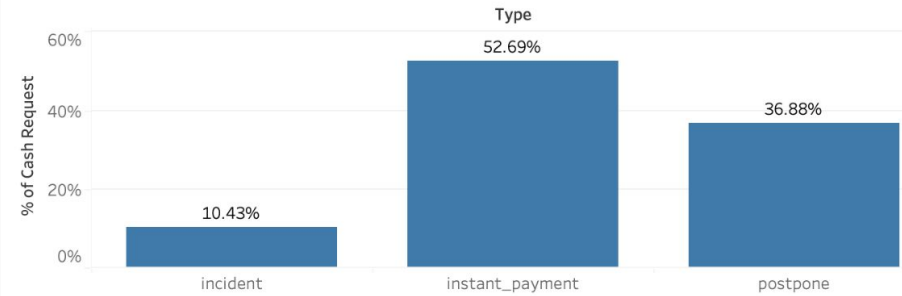
# Challenges and Solutions

Cleaning

Data Inconsistency
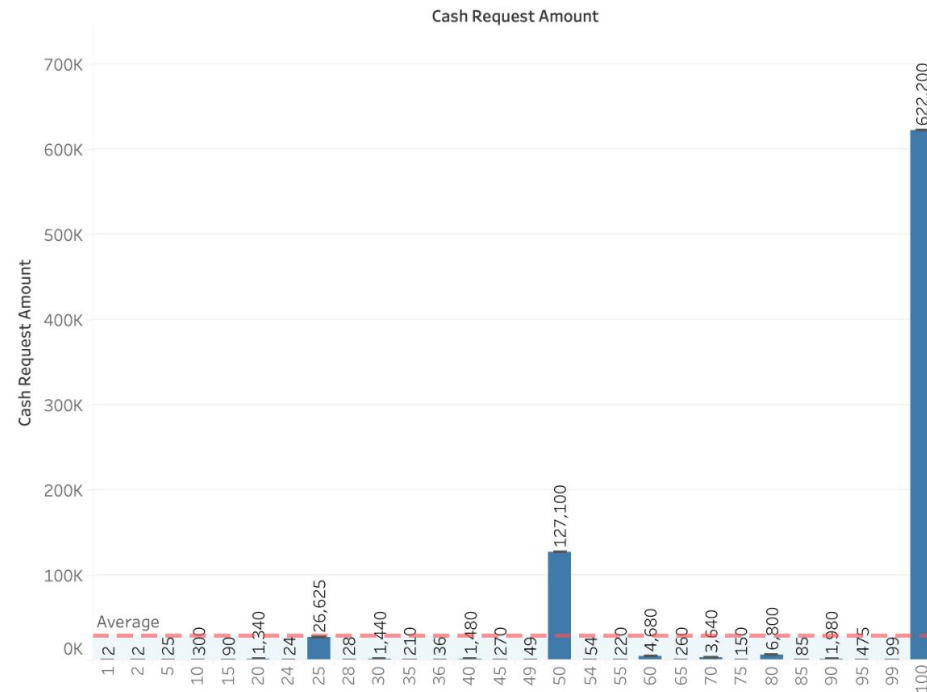
Data Integration

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.

## Success vs Effort Index

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

- Future Analysis: cost increment per transaction, bad data, problems with the fees request system, bad process…

Incidents Rates by Cohort by Cohort

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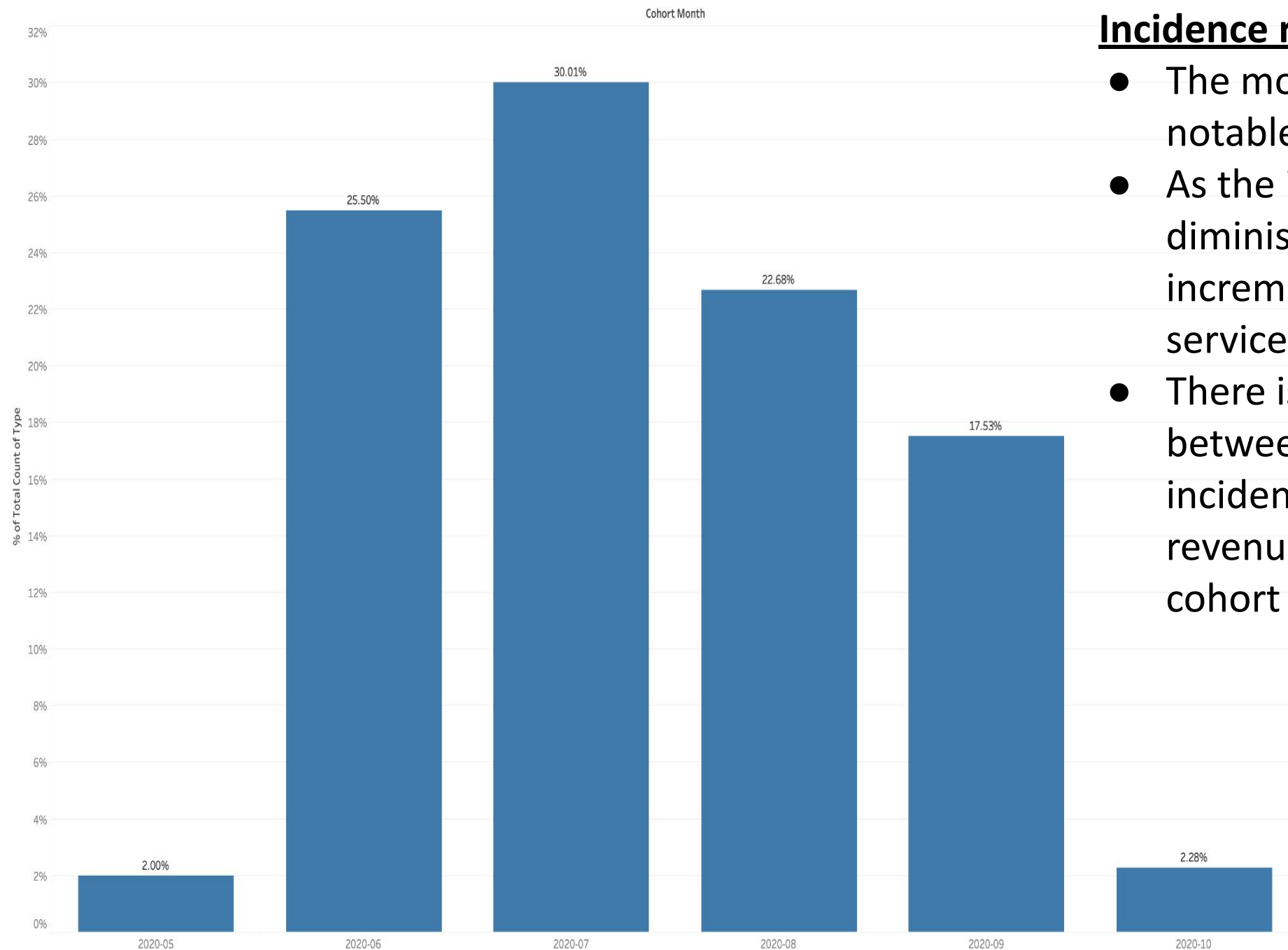
**Incidence rate**
- The month with the highest notable incident rate is July.
- As the incident rate diminished, there was an increment in the total of services and clients.
- There is a direct correlation between the increment of incidents and the reduced revenue generated in the cohort month.

# 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.
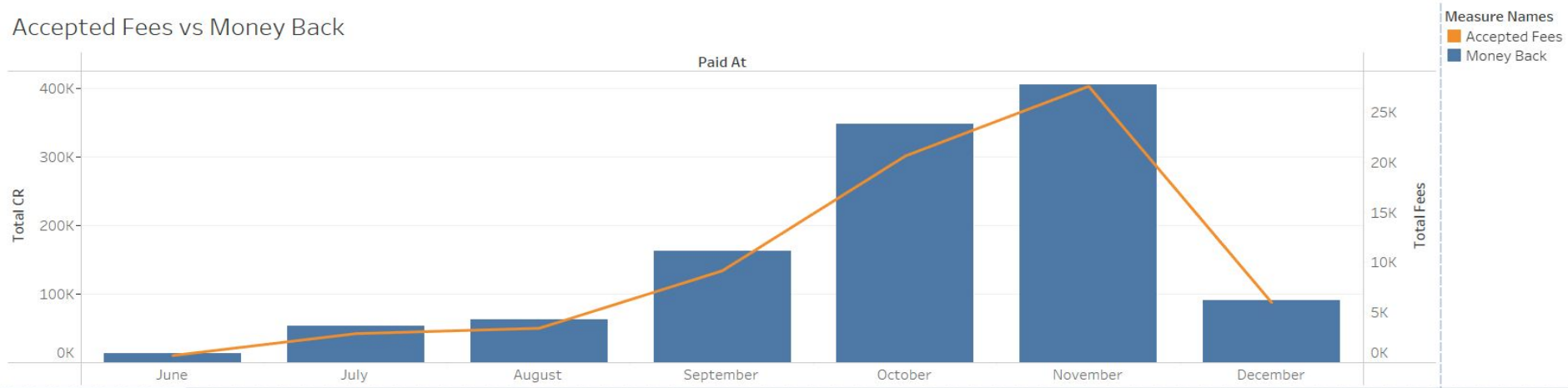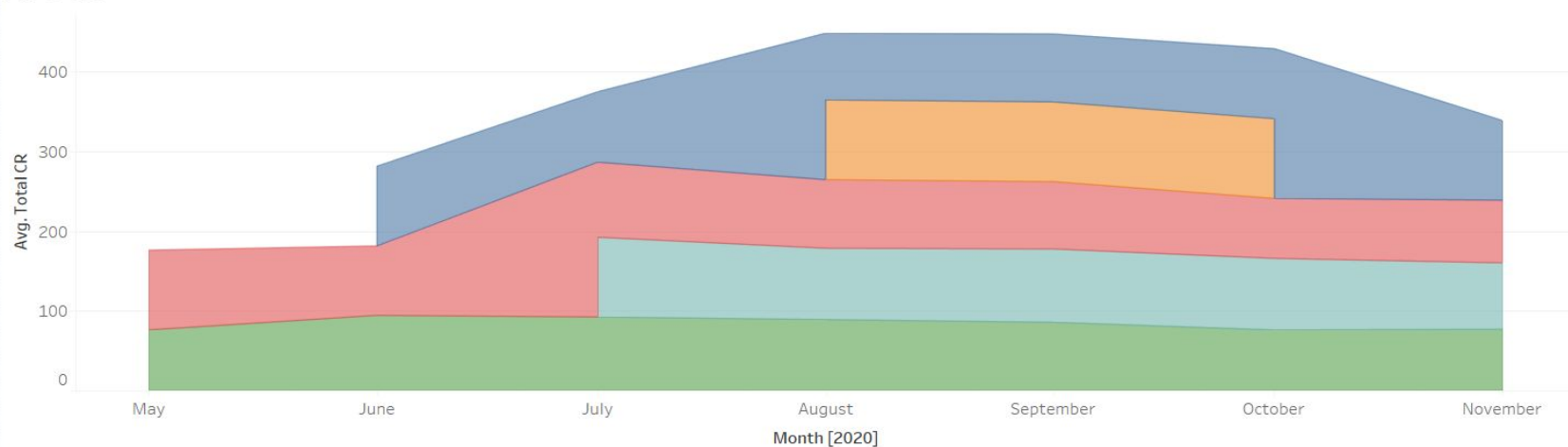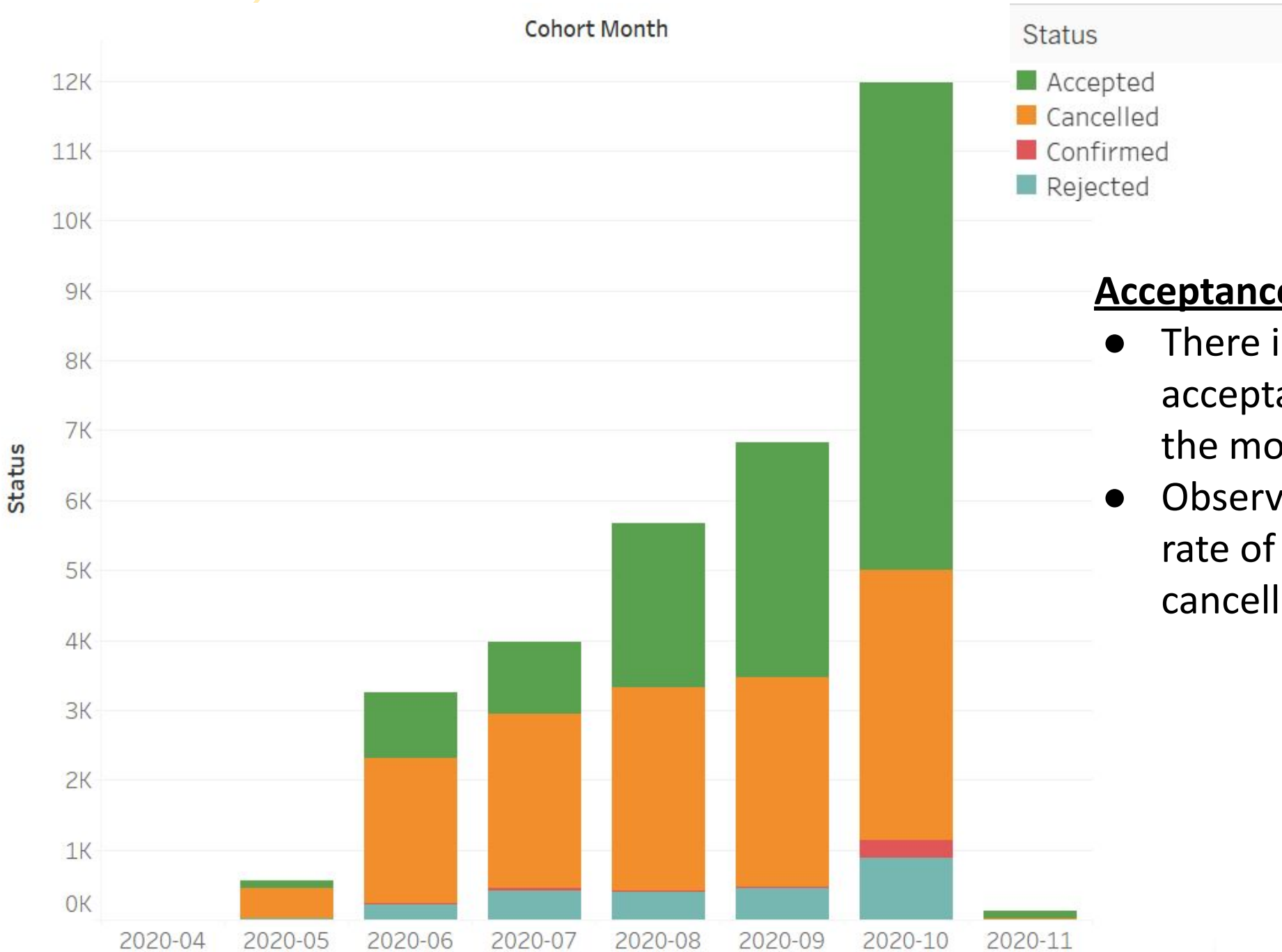
**Acceptance rate**
- There is an increment in the acceptance rate throughout the months during the cohort.
- Observable difference in the rate of acceptance and cancellation.

# Conclusion

The exploratory data analysis allowed us to clean and merge both data sets, resulting in a more comprehensive dataset. This enhanced dataset enabled the identification of patterns and the analysis of key metrics, such as service usage frequency, incident rates, and revenue generation. One of our significant findings was the development of a proposed metric focused on the efficiency of instant payments. We observed a considerable increase in the acceptance of instant payments, which directly correlated with the growth in the number of clients.

Additionally, we identified a seasonal correlation, particularly in the late-year holiday period, where service usage peaked. Based on these insights, we recommend prioritizing marketing efforts for instant payments during holiday seasons throughout the year to sustain and increase profitability. Furthermore, we suggest extracting client location data to gain deeper insights into which countries and demographics yield the highest sales and profit margins.

# 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