



# **Business Challenge: Ironhack Payments**

Project 1 Collaborators:  
Ginosca Alejandro  
Natanael Santiago  
Paola Rivera  
Dan Bigman



# Key insights

## Exploratory Data Analysis:

- Table numbers of rows, columns, and data types
- Number of missing values per column

## Data Cleaning:

- Dropped rows without id, add missing categories

## Data Processing:

- Replace unknown dates with 'NaT'
- Change date strings to pandas date format
- Add column and assign cohorts

# Exploratory Data Analysis: Cash Requests and Fees Datasets

Column	Non-Null Count	Dtype	Missing Count	Missing %
id	23,970	int64	-	0
amount	23,970	float64	-	0
status	23,970	object	-	0
created_at	23,970	object	-	0
updated_at	23,970	object	-	0
user_id	21,867	float64	2,103	8.8
moderated_at	16,035	object	7,935	33.1
deleted_account_id	2,104	float64	21,866	91.2
reimbursement_date	23,970	object	-	0.0
cash_request_received_date	16,289	object	7,681	32.0
money_back_date	16,543	object	7,427	31.0
transfer_type	23,970	object	-	0.0
send_at	16,641	object	7,329	30.6
recovery_status	3,330	object	20,640	86.1
reco_creation	3,330	object	20,640	86.1
reco_last_update	3,330	object	20,640	86.1

Status Category	Count
money_back	16397
rejected	6568
direct_debit_rejected	831
active	59
transaction_declined	48
direct_debit_sent	34
canceled	33

extract - cash request - data analyst.csv

Column	Count	Dtype	Missing Count	Missing %
id	21,061	int64	0	0.0%
cash_request_id	21,057	float64	4	0.0%
type	21,061	object	0	0.0%
status	21,061	object	0	0.0%
category	2,196	object	18,865	89.6%
total_amount	21,061	float64	0	0.0%
reason	21,061	object	0	0.0%
created_at	21,061	datetime64[ns, UTC]	0	0.0%
updated_at	21,061	datetime64[ns, UTC]	0	0.0%
paid_at	15,438	datetime64[ns, UTC]	5,623	26.7%
from_date	6,749	datetime64[ns, UTC]	14,312	68.0%
to_date	6,512	datetime64[ns, UTC]	14,549	69.1%
charge_moment	21,061	object	0	0.0%

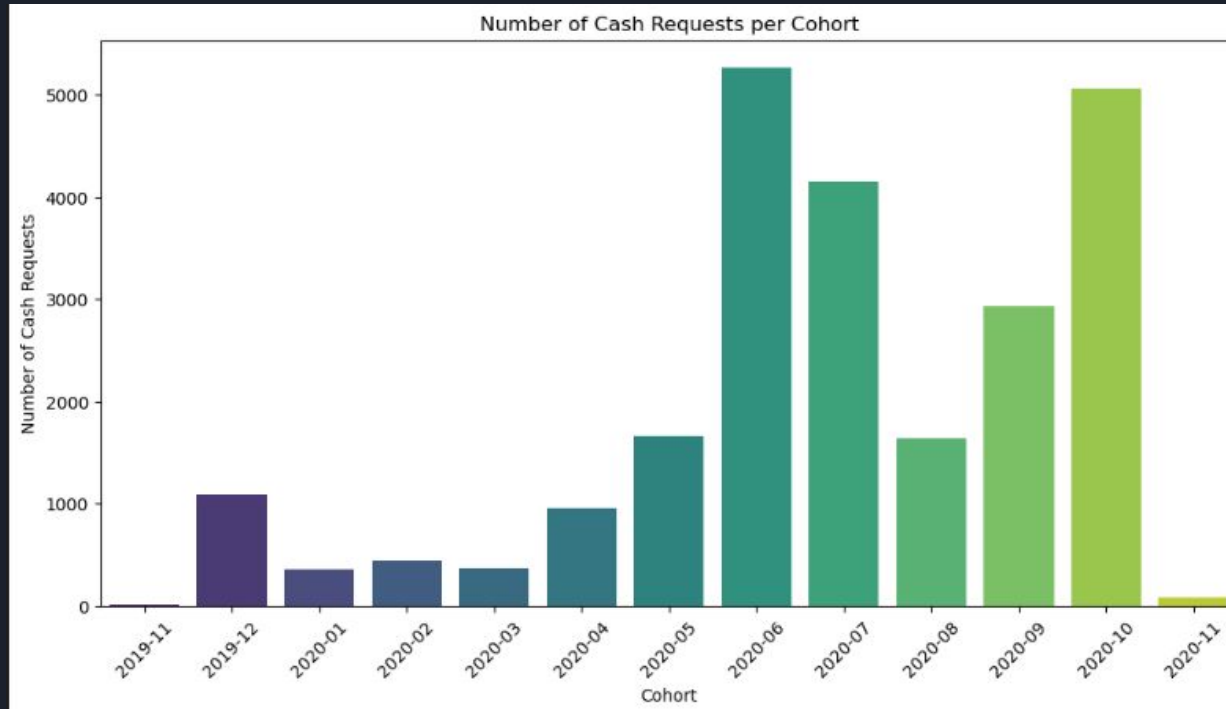
Status	Count
accepted	14,841
cancelled	4,938
rejected	1,194
confirmed	88

extract - fees - data analyst - .csv

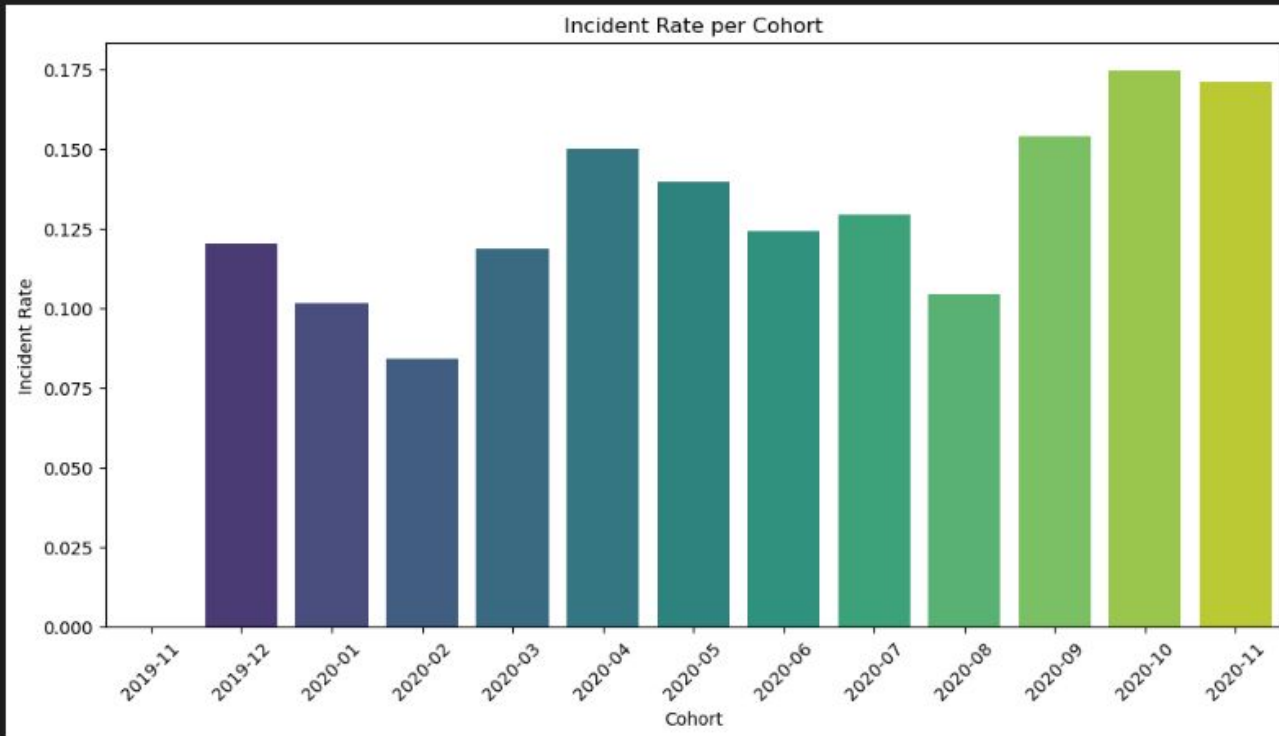
This examines the structure, completeness, and key distributions in the 'cash requests' and 'fees' datasets to prepare for downstream analysis.

- No duplicate rows in either dataset.

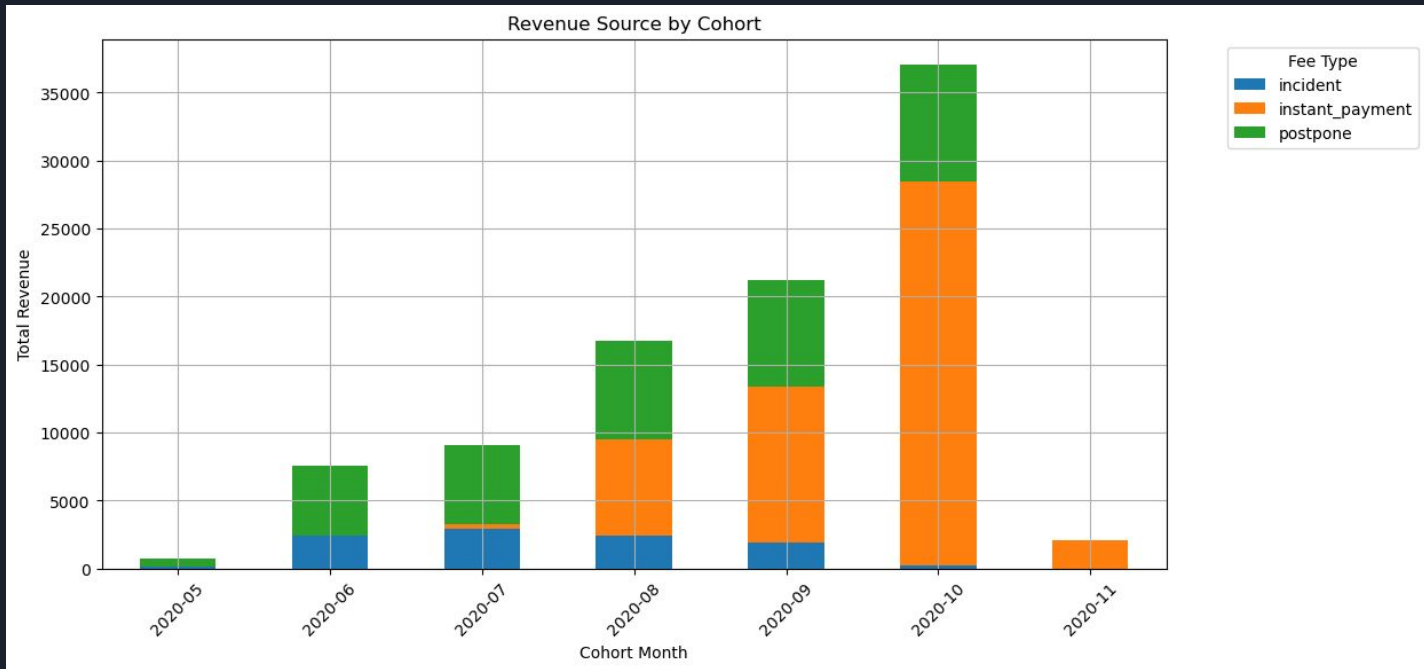
# Frequency of Service Usage



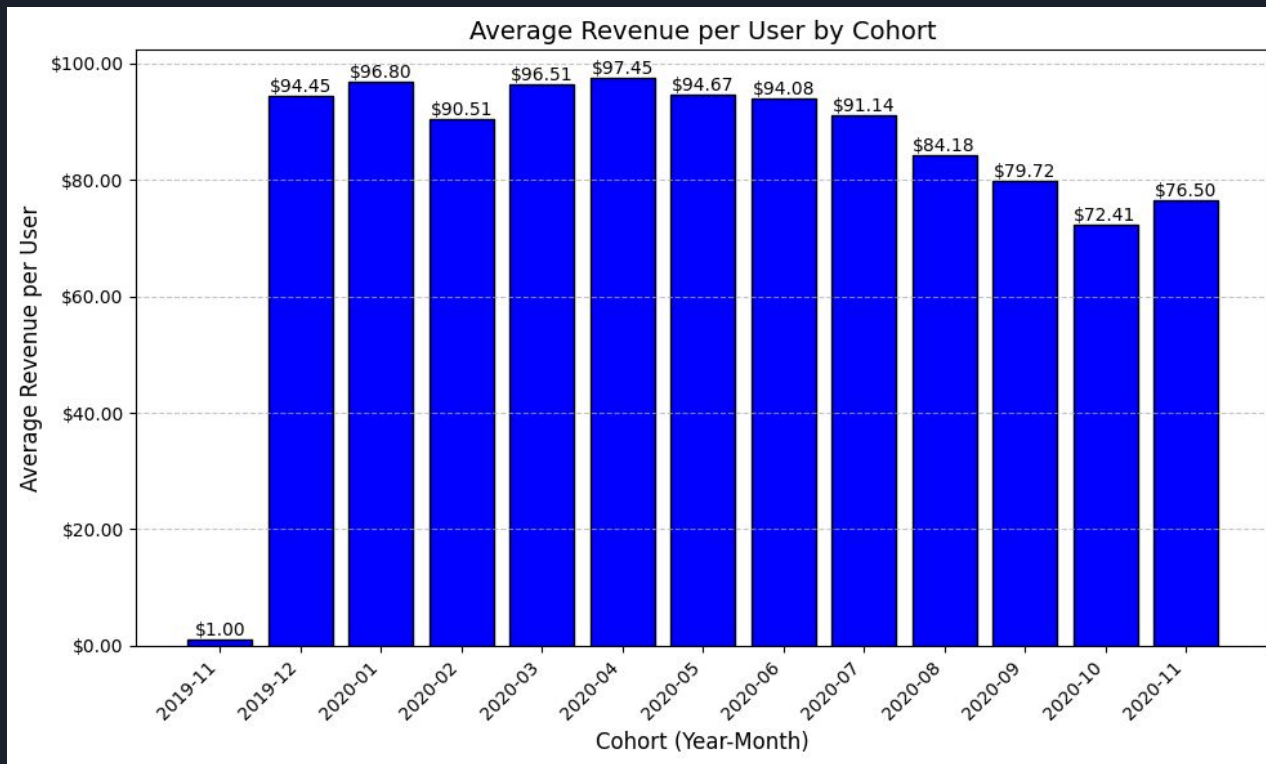
# Incident Rate



# Revenue Generated by Cohort



# New Relevant Metric: Revenue per User by Cohort

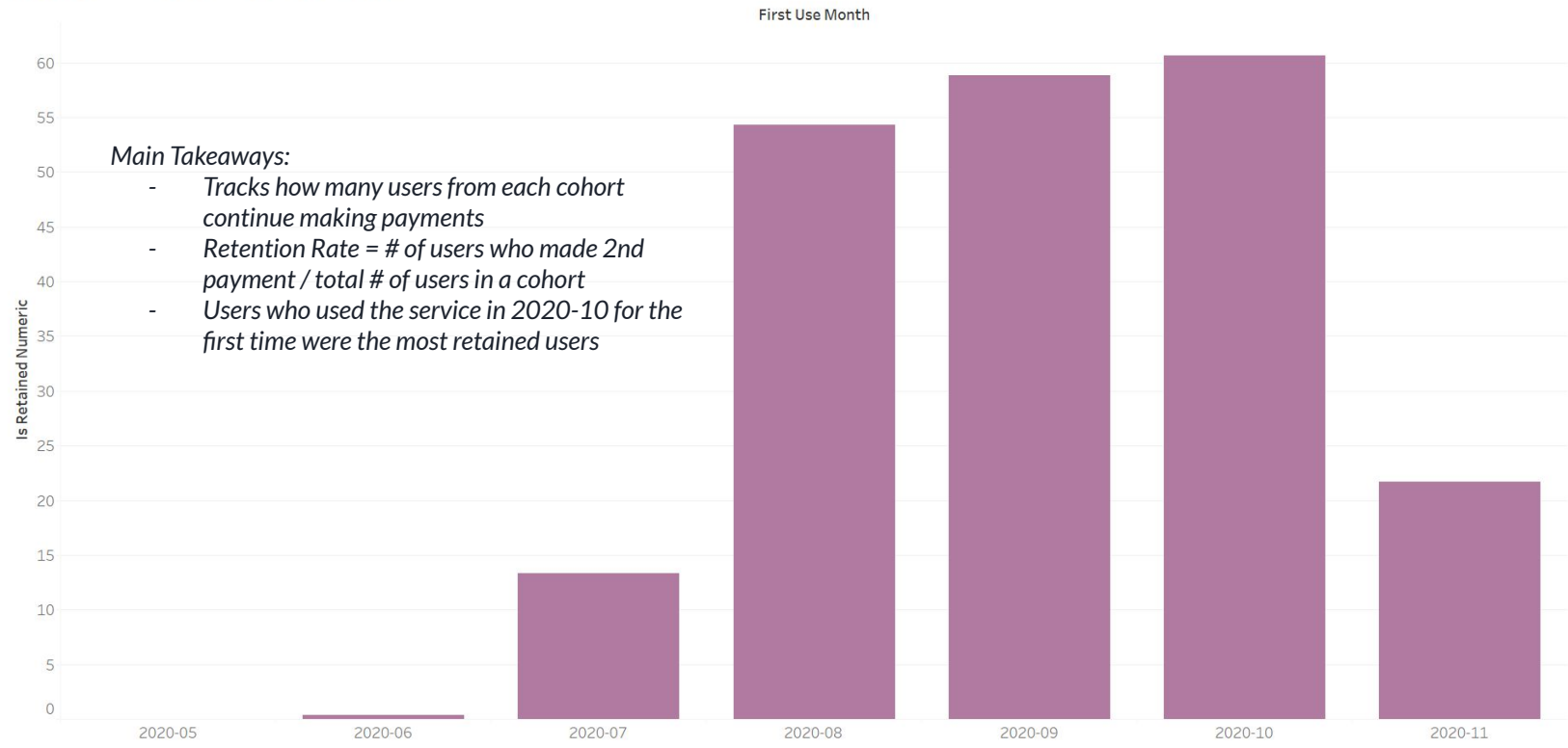


# Group by cohort and count the number of requests

```
cohort_counts = df.groupby('cohort').size().reset_index(name='cash_request_count')
```

# New Relevant Metric: Retention Rate

Retention Rate by First Use Month







# Recommendations:

As a data analyst, after performing the cohort analysis, retention rate analysis, and examining the revenue generated from cohorts, we can provide the company with several key recommendations to optimize business performance. Based on our findings, here are some possible recommendations:

1. Improve User Retention Strategies
  - a. Recommendation: If you observe significant drop-offs in retention over time (e.g., after the first month or after certain cohorts), you could recommend that the company invest in improving user engagement and retention strategies.
2. Targeted Marketing for High-Value Cohorts
  - a. Recommendation: Based on cohort revenue analysis, if certain cohorts generate more revenue over time, the company can increase marketing efforts targeting similar cohorts.
3. Address Payment Issues (Incidents)
  - a. Recommendation: If a significant number of users experience payment issues (failed or pending transactions), it's crucial to address the payment process to reduce churn and improve revenue.



# Actionable Steps:

1. Improve User Retention Strategies:
  - Implement targeted email campaigns, in-app messaging, or personalized offers to keep users engaged.
  - Identify specific touchpoints where users tend to drop off and optimize the user experience around these points (e.g., onboarding process, first payment).
  - Consider loyalty programs, discounts, or exclusive content for long-term users.
2. Targeted Marketing for High-Value Cohorts
  - Leverage user acquisition strategies focusing on characteristics of high-value cohorts (e.g., high-paying customers, frequent users).
  - Expand marketing budgets for channels that perform well with higher-value cohorts.
  - Use advanced segmentation and personalization to create tailored campaigns for these high-performing groups.
3. Address Payment Issues (Incidents)
  - Investigate the root causes of payment incidents and streamline the payment process to reduce friction.
  - Work with payment processors to ensure smoother transactions, such as better handling of failed payments and offering users alternative payment methods.
  - Provide better communication and support to users facing payment issues, offering assistance to resolve payment problems quickly.



THANK YOU!



**IRON**  
HACK