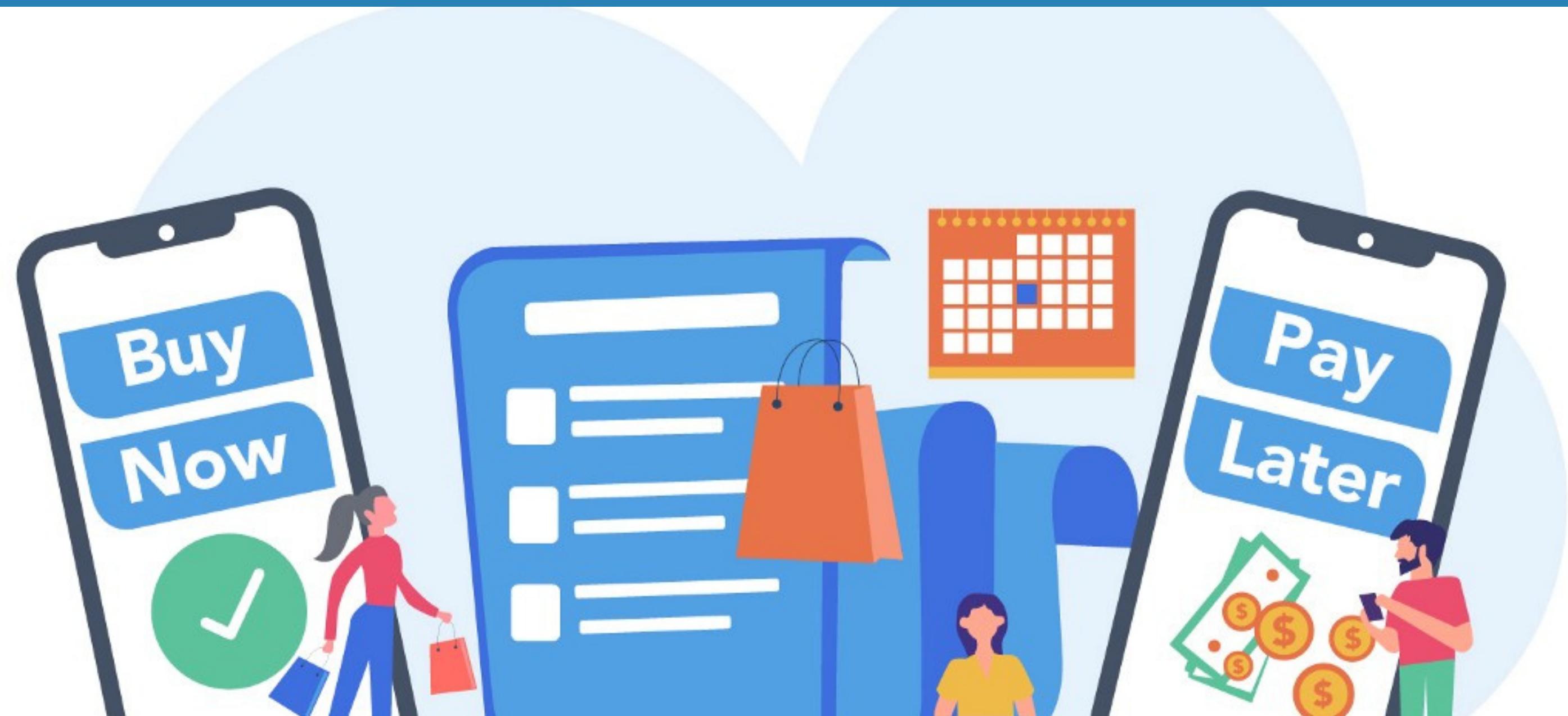


# Buy Now, Pay Later

Applied Data Science  
MAST300034



# Meet the Team



**Harshita Soni**



**Prathyush Rao**



**James Barro**



**Chai Raghuvanshi**



**Loky Wang**

# Problem Overview

**Creating a ranking system to select top 100 merchants for our BNPL firm**

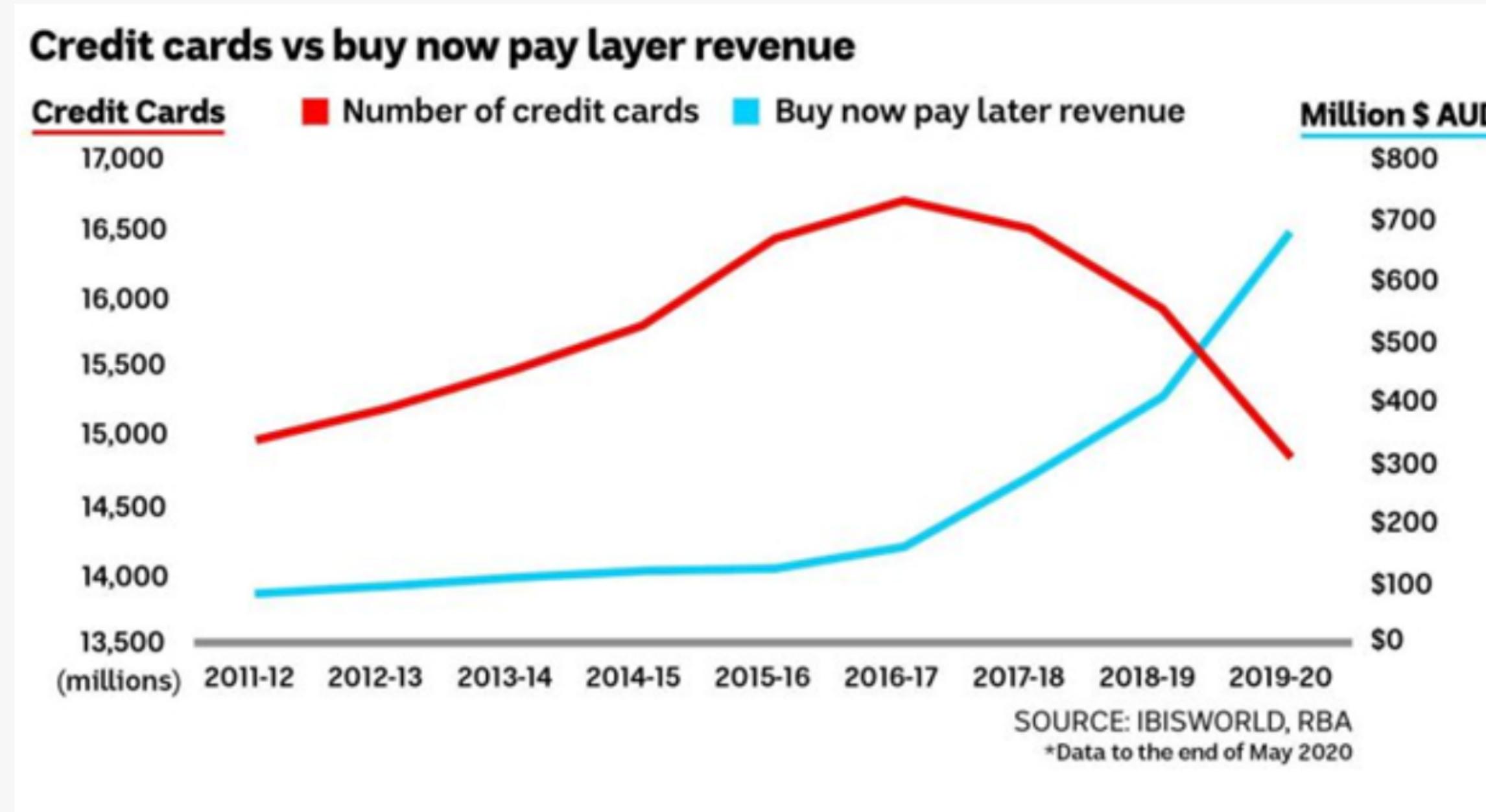
- Exploring features that help in the selection process
- Provide interesting recommendations/insights for the firm



**Creating a fraud detection algorithm to watch out for companies with major fraud transactions**

# Industry Importance

**Buy Now, Pay Later: Making payments an invisible part of the customer journey**



- **27% more people were using BNPL services in 2021 than in 2020 (The Ascent, 2021).**
- **On-boarding the right merchants is crucial not only to increase profitability but also to prevent fraud.**

# BNPL Fraud

- BNPL fraud increased by 66% between 2020 and 2021.
- Losses due to online payment fraud will exceed \$343 billion worldwide between 2023-2027 (Juniper Research).

## 'Criminals love buy now, pay later': How fraudsters exploit popular interest-free payment plans

PUBLISHED THU, NOV 18 2021 2:00 AM EST | UPDATED MON, NOV 22 2021 4:28 AM EST

 Ryan Browne

SHARE    

BUSINESS NEWS FEBRUARY 9, 2021 / 1:05 PM / UPDATED 2 YEARS AGO

### Exclusive: Australia's BNPL boom pushes identity theft to record, data shows

By Byron Kaye

5 MIN READ



Local Video Sport Lifestyle Ente

## Buy now, pay later services AfterPay and LatitudePay hit by scam texts

If you've received a message from your buy now, pay later service, this is what you need to know.



# Logistics

File	Status	Type of work	Assigned To	Last edited	Description
Merchant & Customer fraud association	Review	Prelim Analysis	H Hershey	October 7, 2022 5:13 PM	Find associations between merchant and customer fraud datasets
Download.py	Done	ETL	Prat	October 4, 2022 11:11 PM	For external datasets
Process.py	Later	ETL	Prat	October 5, 2022 4:56 PM	Finish other processing tasks
Clean.py	Review	ETL	Prat	October 5, 2022 10:45 PM	For external datasets
Analysis on merchant val...	Review	Prelim Analysis	L Loky	October 5, 2022 11:15 PM	For understanding loyalty algorithm
Loyalty score	To-do	Feature Engineering	Prat	October 5, 2022 11:15 PM	Creating a loyalty algorithm
Min, Max, SD	Later	Feature Engineering	James	October 7, 2022 5:14 PM	Add columns such as min, max, and SD for all variables in the external dataset
Seasonal/Public Holidays analysis	In progress	Prelim Analysis	Chaitanya	October 7, 2022 5:14 PM	Check for trends of fraud o...

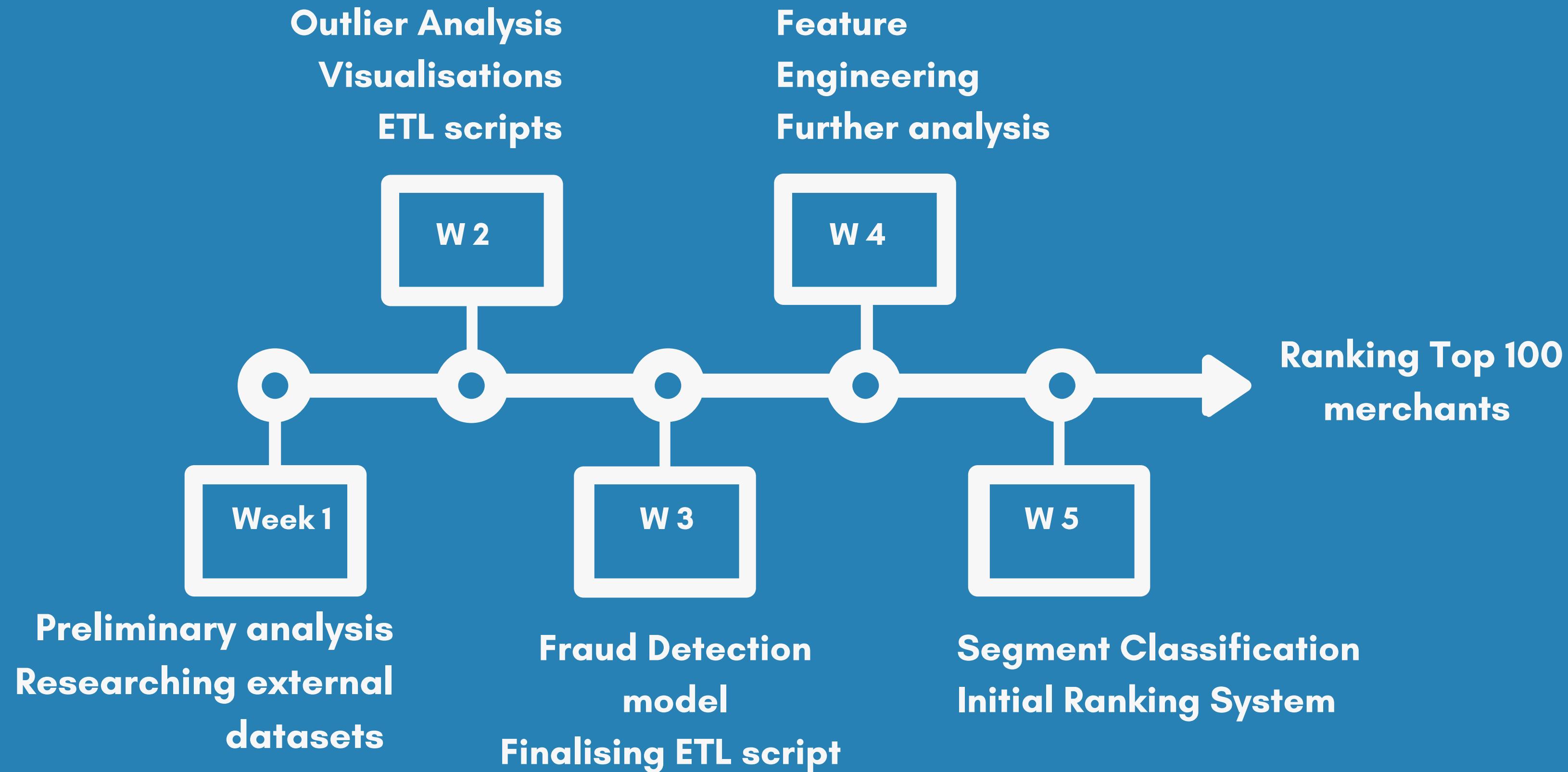
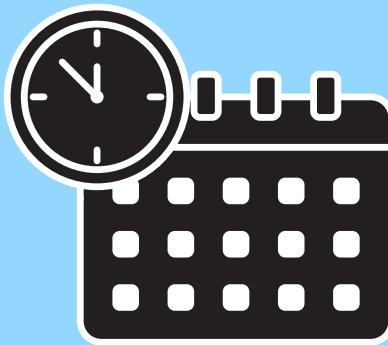
- Bi-weekly meetings (in-person and online via Zoom)

- Total commitment: 15 hours (2 meetings) per week for 6 weeks

- Project Communication via Messenger

- Project management on Notion and Confluence

# Timeline



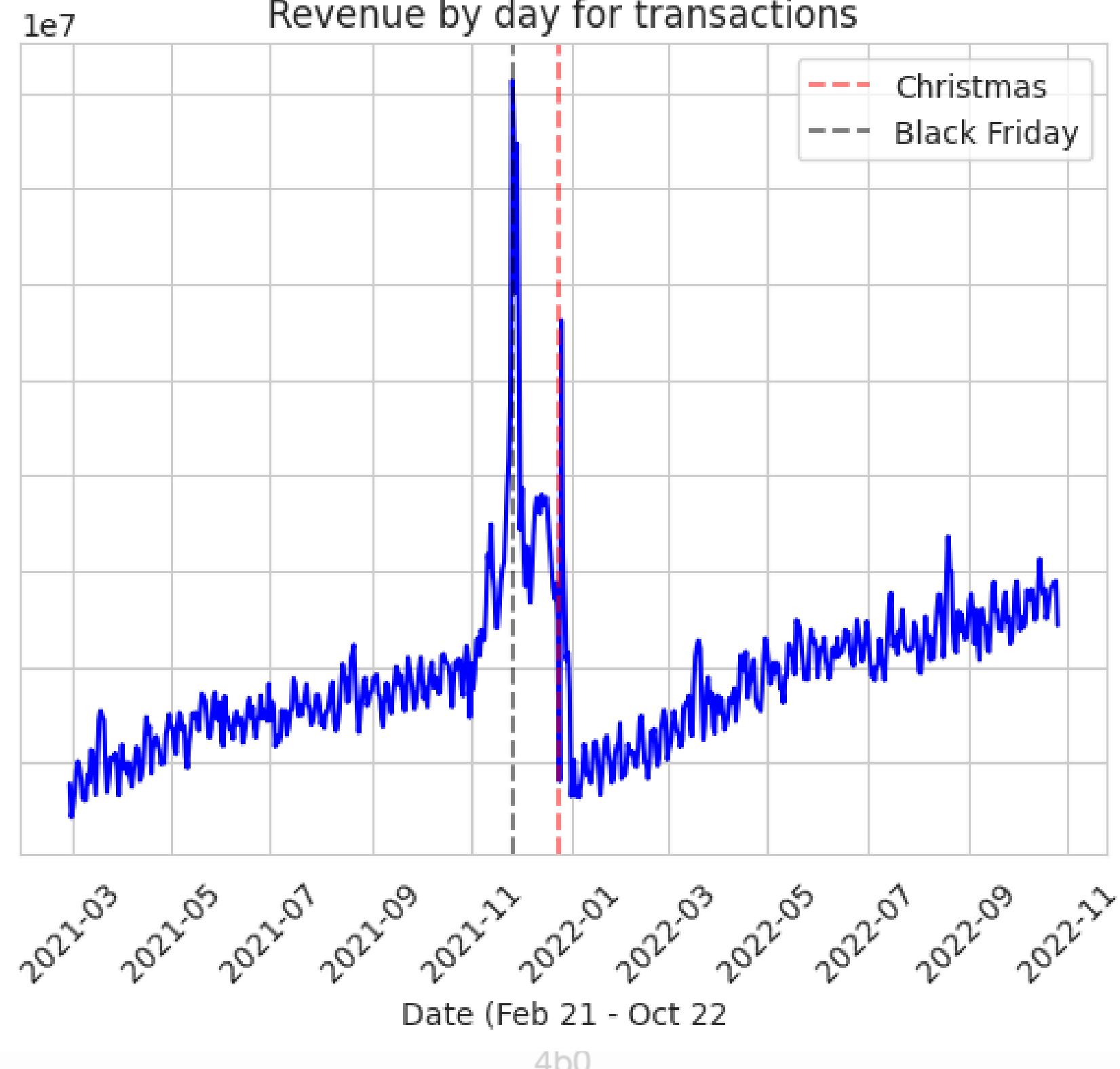


# Security and Privacy requirements

- **Private GitHub repository**
- **Only local machines used to develop and run code**
- **Private Confluence and Notion pages**

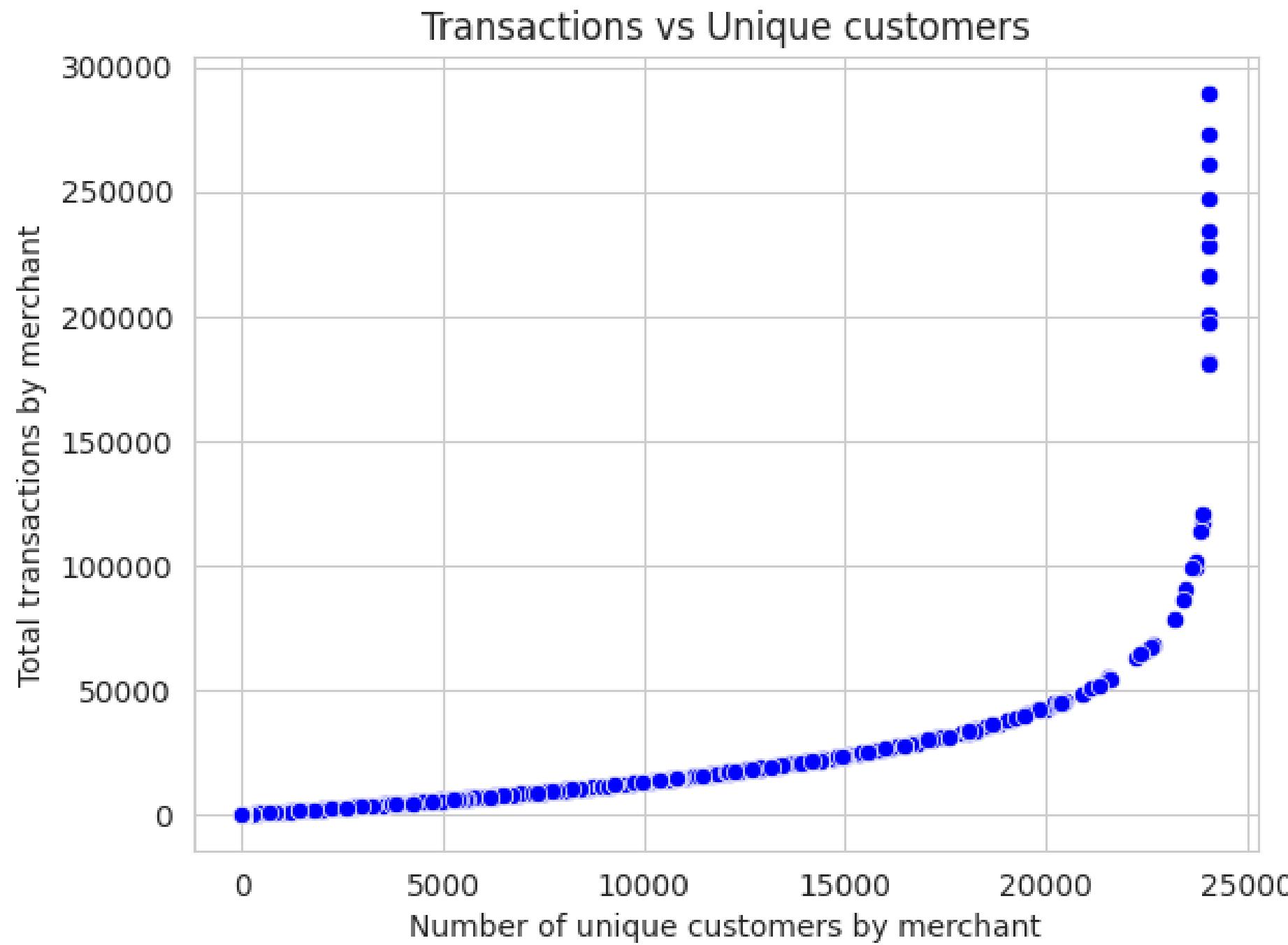
# The Data

merchant_abn	user_id	dollar_value	order_id	order_datetime	consumer_id	name	address	state	postcode	gender	company_name	Description
0410114	7	16.362338872391526	ba1c3c0b-	2021-11-26	511685	Andrea Jones	122 Brandon Cliff	QLD	4606	Female	Orci Ltd	digital go boo
49252122	49252122	1e7	Revenue by day for transactions									
00410673	1.0											
09646675	0.9											
77320845	0.8											
87473570	0.7											
12515164	0.6											
57847798	0.5											
534649627	0.4											
7	0.3											
16.362338872391526	0.2											
96a5d5f4-2f24-4c9...	0.1											
2021-11-29	0.0											
511685												



- Christmas showed the lowest revenue, black friday had the highest
  - Found 396 unregistered merchant ABNs in the transaction data; Not used in models
  - Mainly 25 distinct tags/description to label merchants (after NLP cleaning);
- Used to classify sectors

# Interesting Insights

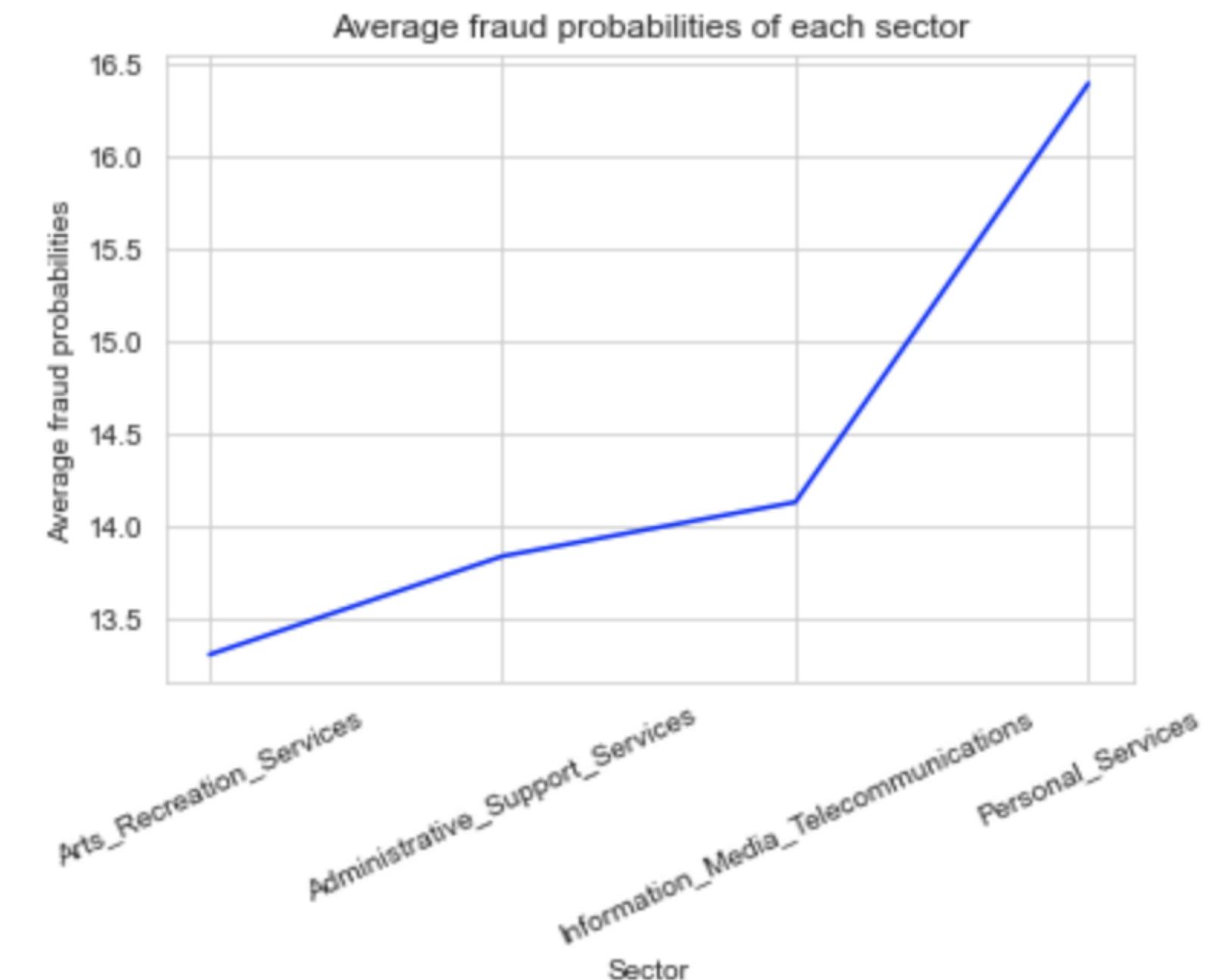
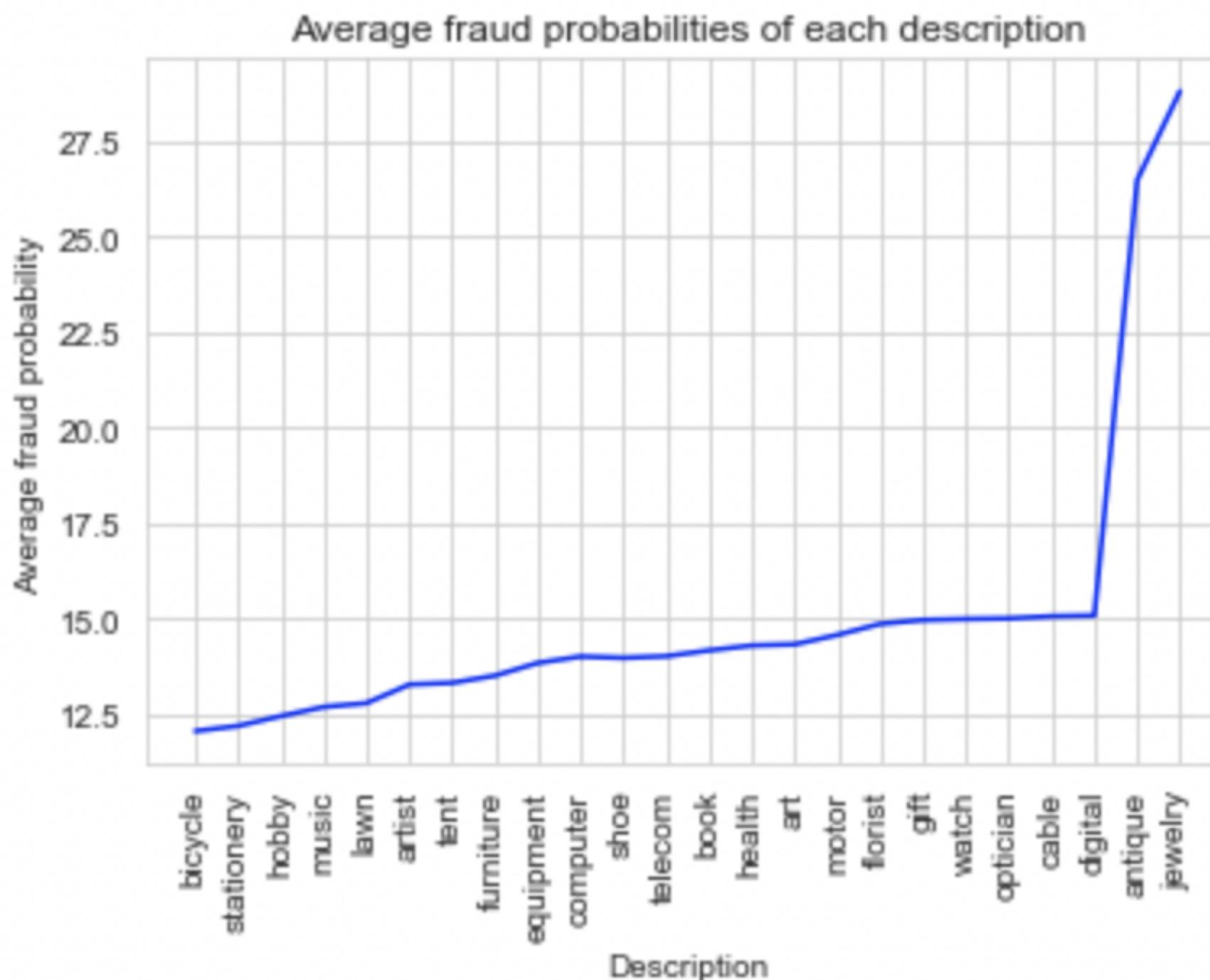


- There was an unusually large number of transactions with a specific value of \$157
- 99% of merchants do not have transactions for more than a 6-month period

# Fraud Model

# Fraud Analysis

- Personal services including jewellery and antiques have a much higher fraud probability than other industry merchants



# Fraud in the real world

## AFCA rules \$227,000 jewellery claim was fraud

22 February 2022



## Scam alert sounded for art and antiques trade as email invoice fraud tops £100m

26 March 2019

The cost to firms of cyber fraudsters' invoice scams reached nearly £100m in 2018 and small businesses, such as those in the art and antiques trade, remain vulnerable, according to the latest report.

ROB BONTA  
Attorney General

Translate Website | Traducir Sitio Web

HOME ABOUT MEDIA CAREERS REGULATIONS RESOURCES PROGRAMS APPOINTMENTS CONTACT

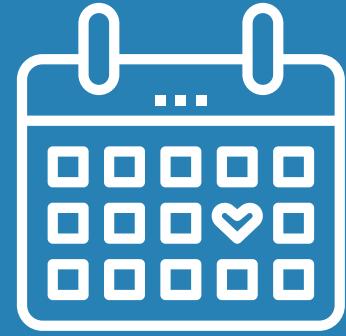
**Attorney General Bonta: CFPB Should Scrutinize Buy-Now-Pay-Later Financing to Address Compliance with Consumer Protection Laws**

# Feature Engineering



Monthly increase in:

- Customer base
- Revenue



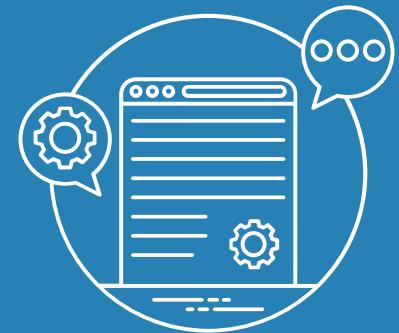
- Public Holidays
- Day of the week



- Income data
- Tax returns

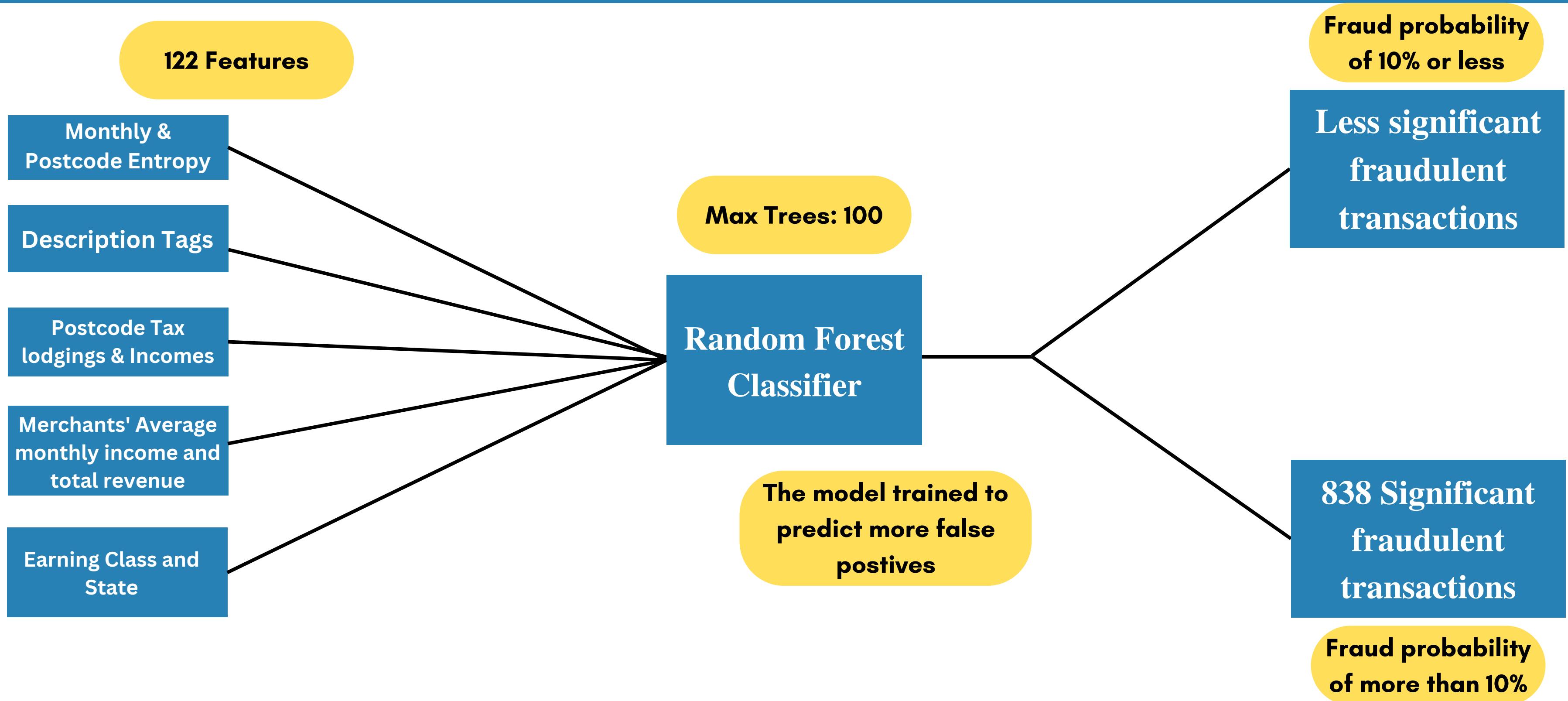


Postcode data  
features

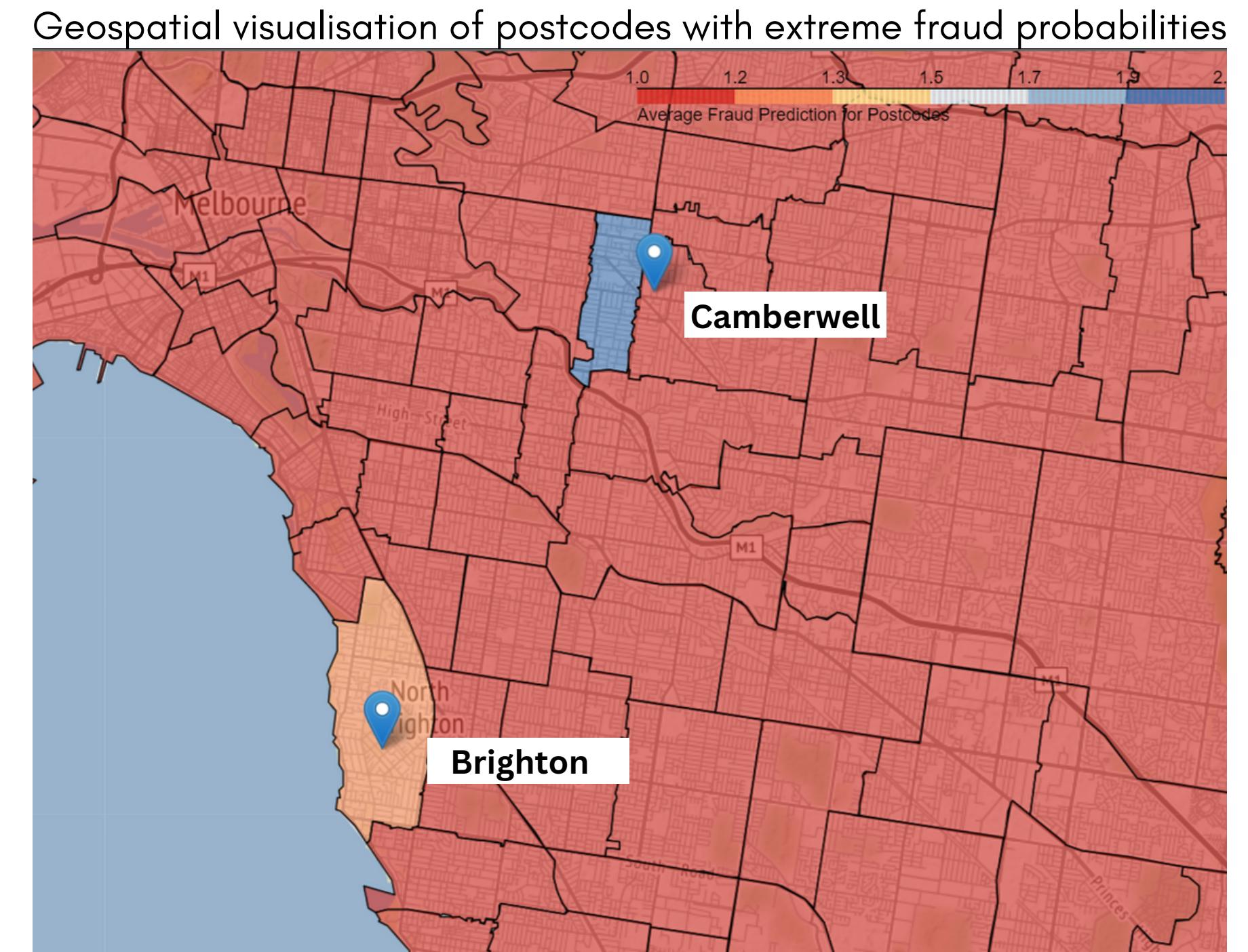
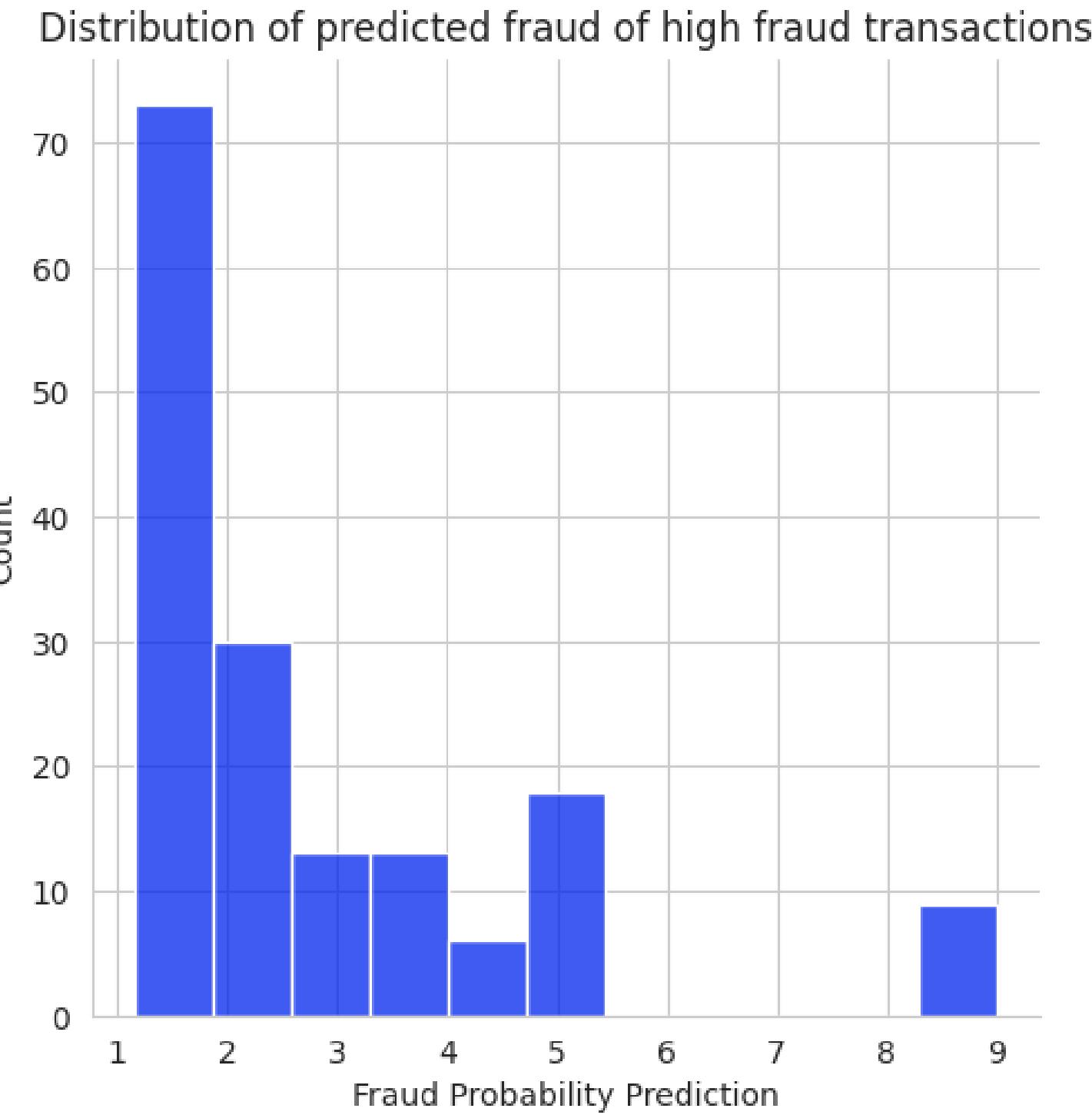


Entropy for  
merchants

# The Fraud Model



# Evaluation of Fraud Model



# Limitations of the Fraud Model

- Not enough training data (~34k rows), the prediction set is much bigger (~14 million rows).
- Data is not strong enough to truly separate between fraud quantiles.
- This model on its own is not enough to rank merchants; other factors have to be taken into consideration.



# The Ranking Model

# **Major features and their importance**

## **Financial**

- How much revenue is a merchant making?
- How much growth has happened in the last 6 months?

## **Customer Base**

- How do a merchant's customers behave?
- What is the spread like?
- Are they wealthy?

## **Sustainability**

- Has there been a stagnation of output over time?
- What is the level of experience?
- What do the current industry trends suggest?

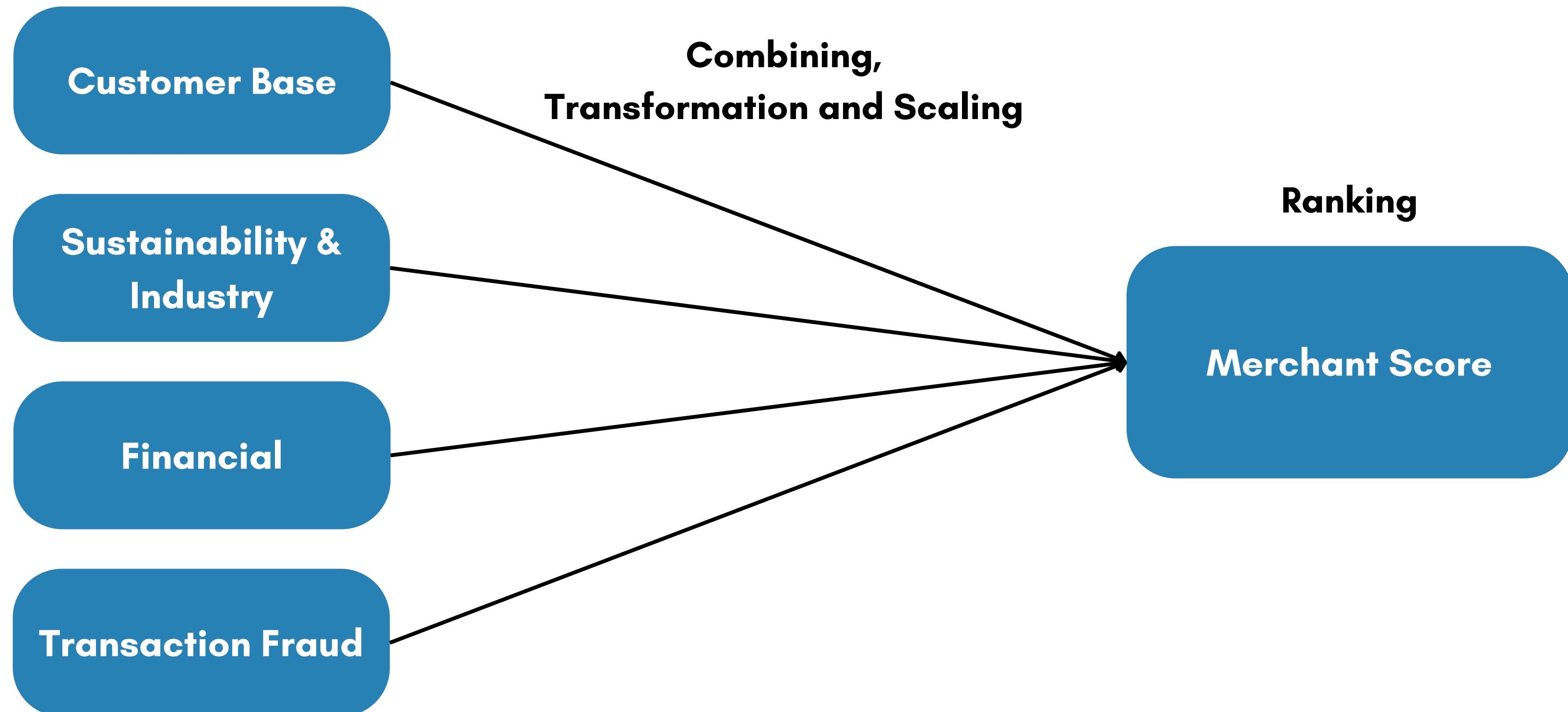
## **Transaction Fraud**

- What impact does fraud have?

# Merchant Ranking Algorithm

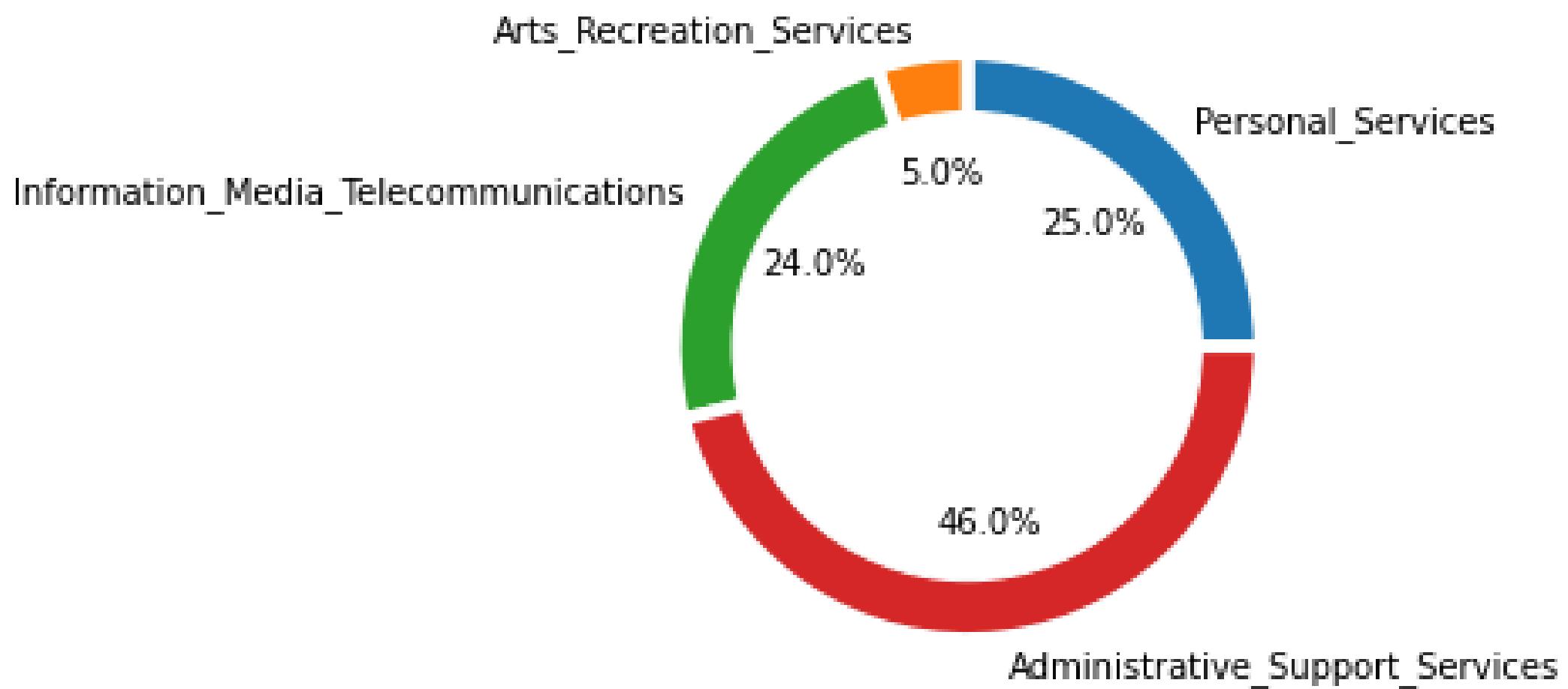
Summary of ranking model and criteria

## Identified Metrics



# Analysis of the Ranking Algorithm

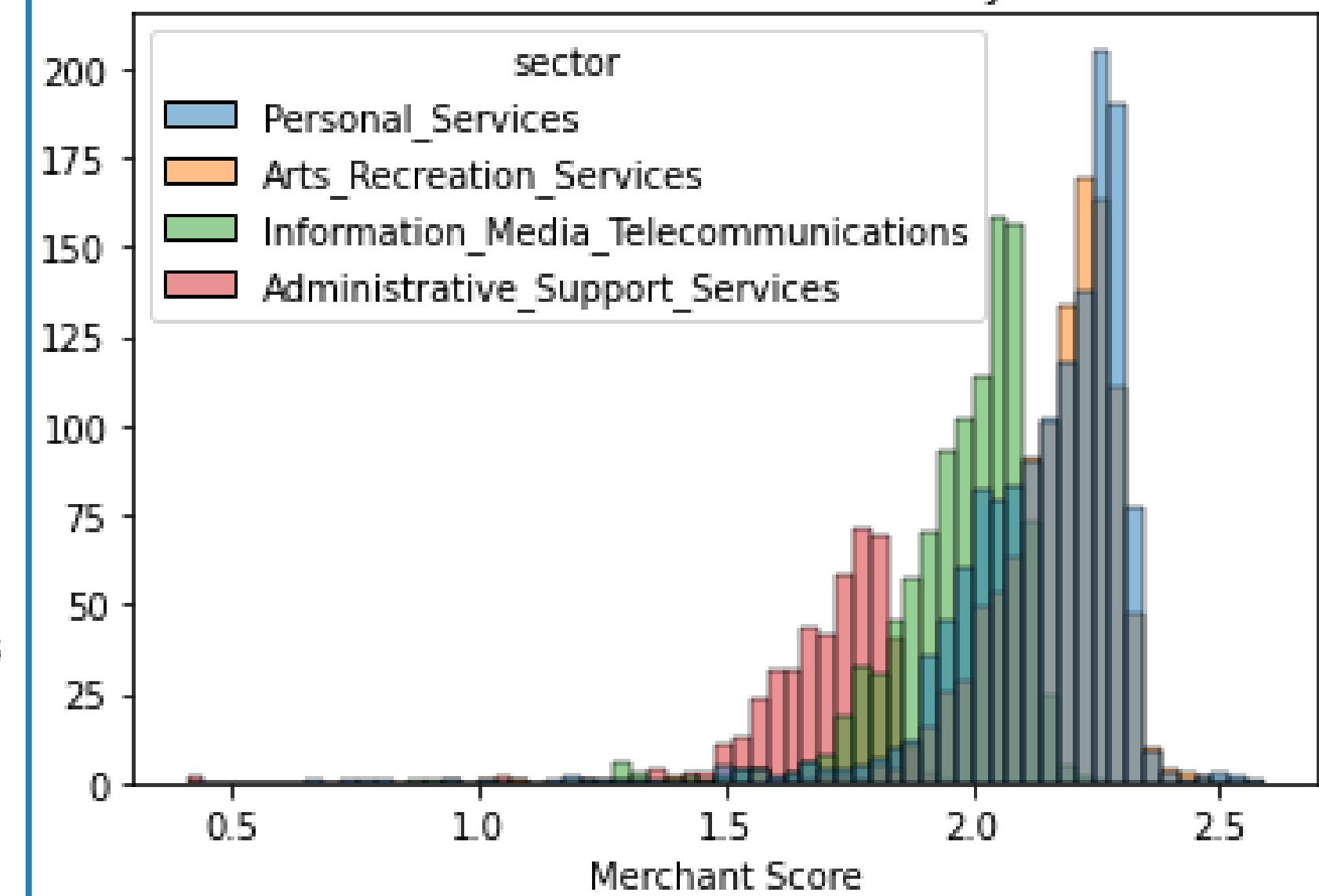
Recommendation Sector Analysis of Top 100 Merchants



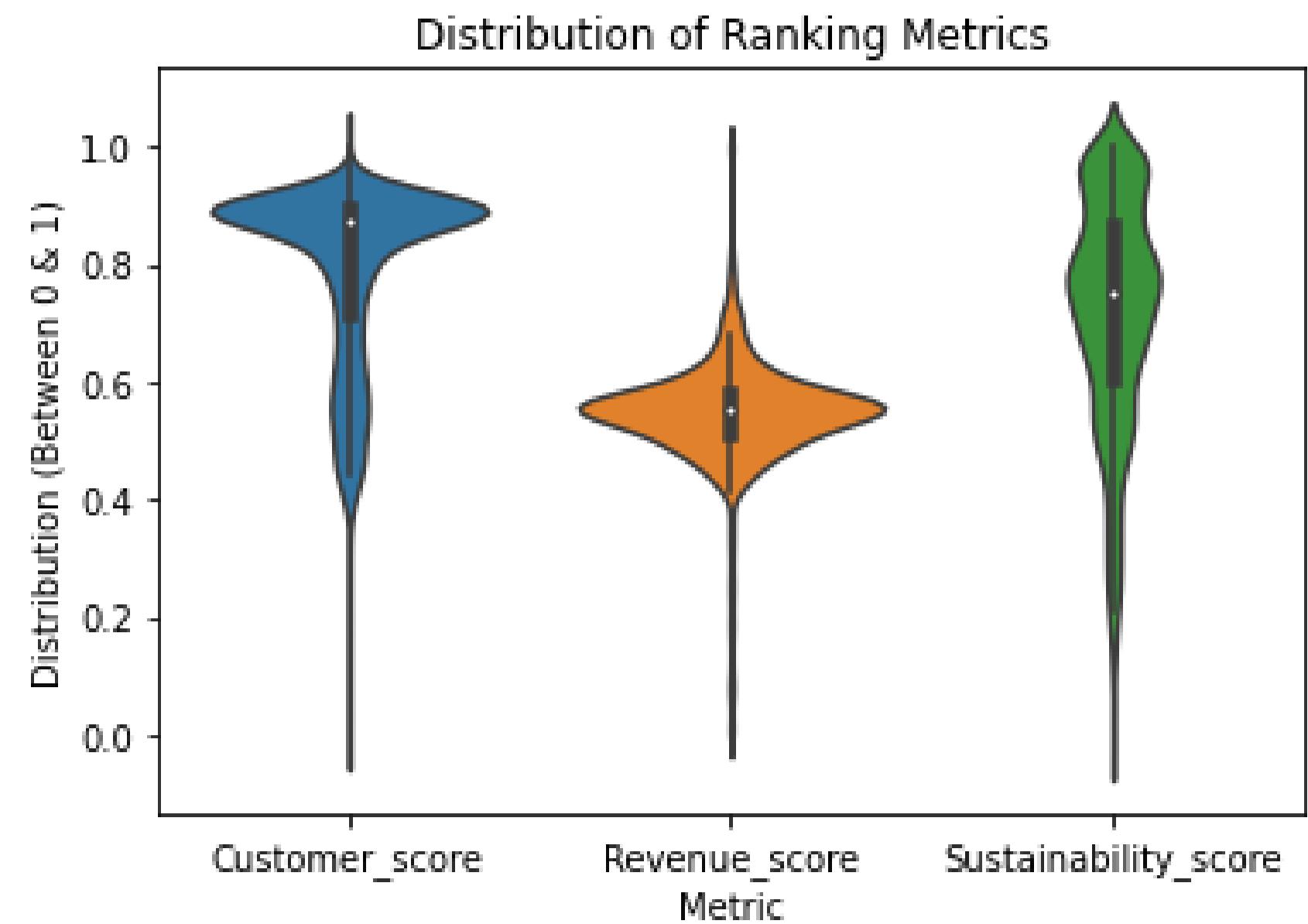
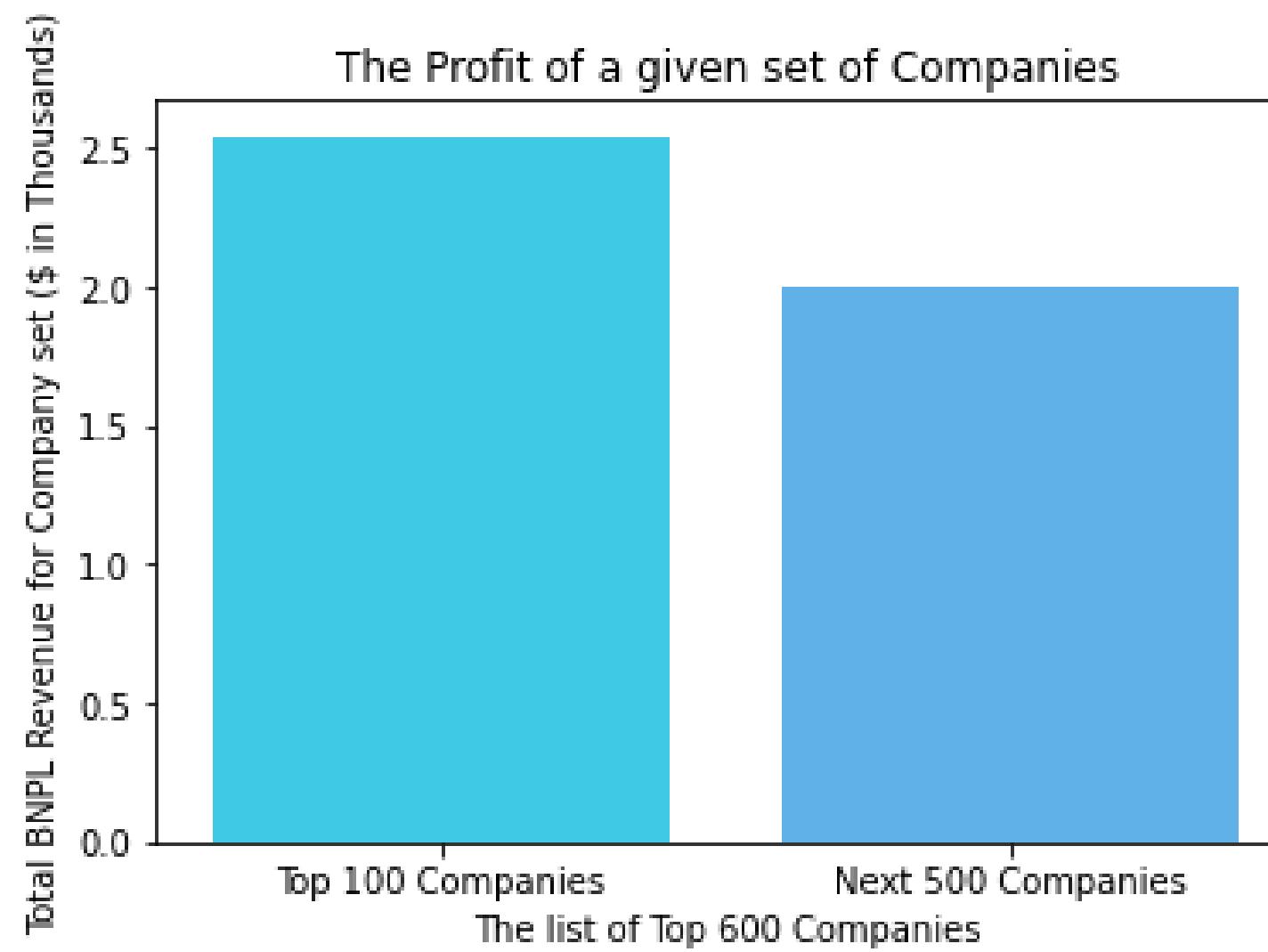
## Summary

- Strong Separation of 'best' merchants from the rest
- There is a diverse spread of Sectors in the top 100

Distribution of Merchant Score by Sector



# Further Insights



# Conclusion

# Recommendations



Avoid partnering with merchants in Arts and Recreation Services



Inform partnering merchants about customers in high-fraud areas, such as Camberwell and Brighton.



Avoid high-risk Jewellery and antique shops

# Business Aspects

Why it's worth putting in all the effort!



Generating more than \$2 million dollars of profit by onboarding 100 best merchants.



Decreasing the BNPL fraud risk by about 1%



Stronger stakeholders with a loyal customer base

# Limitations

The ranking system has flaws in its design

- The order of scaling
- Equalised feature importance
- Dealing with small values
- Unaccounted areas of a company

# Future Scope

Worth pursuing in the future

- Notify new businesses of potential fraud
- Can be feasibly implemented
- Optimises revenue
- Helps understand customer behaviour
- Allows more data-driven decisions