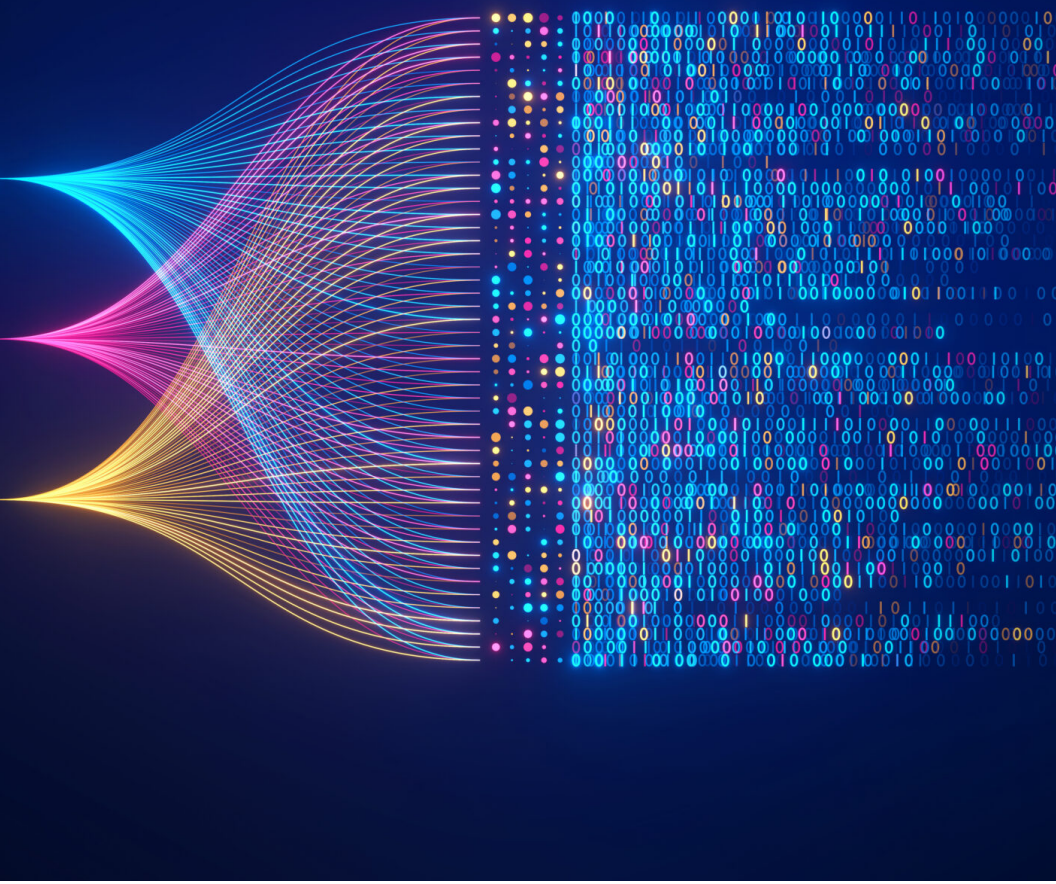
UNSW Fintech Society
presents



FRAUD FORENSICS

A POWER BI CHALLENGE

DATASET

The dataset can be accessed at transactional_data.csv

KEY DATES

Case Release	11/09/2023 (Monday)
Workshop 1	13/09/2023 (Wednesday) Held online - Zoom link will be emailed
Submission	11:59PM 17/09/2023 (Sunday)
Announcement of Finalists	19/09/2023 (Tuesday)
Finals	22/09/2023 (Friday) Held in-person 6-8pm on campus (MCIC)

CASE BRIEF

You are part of an exciting FinTech startup specialising in end-to-end fraud detection and analytics for financial institutions with retail credit products. A major credit card provider has approached the firm with their customers' transactional data from 2022, wanting to better understand customer behaviour and factors that influence fraudulent transactions. The client has provided a large but messy dataset of a total 1000 customers. It's your job to interpret it and achieve their following objective:

Which combination of demographics lead to higher fraudulent transaction rates?

Based on the provided dataset, the client would like to learn more about the characteristics of the fraudulent transactions. To help guide you with your analysis, they have provided you with some questions you may want to answer:

- What category of purchases had the highest rate of fraud?
- Is there a common timeframe in which fraudulent transactions occur?

- Which regions have the highest rate of fraud?

- How significant are the personal characteristics of the customers to fraud rate?

Feel free to also answer your own questions if you find something insightful or interesting. However for all questions, you must be able to answer questions on HOW you arrived at your final insights. Be prepared to describe your process/methodology.

Like all client-facing roles, the client requests these insights to be delivered in the form of a presentation. You are free to use whatever tools to analyse the data and draw insights (e.g. Excel, Python, R, etc.) HOWEVER, you MUST use Power BI to visualise these results in the final presentation.

DATA DESCRIPTION

ssn: The individual's Social Security Number
cc_num: The individual's credit card number
first: First name
last: Last name
gender: Gender
street: Street number and suburb
city: City name
state: State initials
zip: Postcode
lat: Latitude
long: Longitude
city_pop: Total population of the city
job: The individual's occupation
dob: The individual's date of birth
acct_num: The individual's bank account number
trans_num: Transactional number of the purchase
trans_date: Date of the transaction
trans_time: Time of the transaction
unix_time: Time of the transaction in Unix time
category: Category of purchase item/service
amt: Amount spent on transaction, in dollars ($)
is_fraud: Flag for fraud (0 = No; 1 = Yes)
merchant: Name of merchant
merch_lat: Merchant's latitude
merch_long: Merchant's longitude

GETTING STARTED WITH POWER BI

Microsoft's Self-learn courses https://powerbi.microsoft.com/en-au/learning/

Workshop 1 Video Recording will be uploaded to UNSW Fintech Society's Social Media Accounts. Follow the event on Facebook.

EVALUATION CRITERIA

Methodology: Does the team have an appropriate methodology for data analysis. Do they appropriately clean data, deal with missing values, appropriately munge data, etc.

Presentation: Does the team present their slide deck smoothly and effectively? Do the visuals complement the oral presentation?

Use of Power BI: Is the team able to effectively use Power BI to analyse and present data in an easily understood and appropriate format?

Q&A: Ability to justify their methodology and visualisations in front of the judging panel, demonstrating high competence in understanding the data set provided.