Understanding Financial Fraud Through Data Visualization

1. Project Goal:

This project aims to explore and uncover hidden patterns in financial fraud by using interactive and informative data visualizations. Leveraging adata set of 50,000 labeled financial transactions, we aim to analyze how fraudulent activities vary across transaction types, periods, geographic locations, and user behaviors. By translating these complex data relationships into intuitive visuals, we hope to make fraud detection more understandable and accessible — not just to data scientists, but also to financial analysts, cybersecurity professionals, and even the general public.

2. Key Questions Addressed

• Do frauds vary by time of day/week?

• What user behaviors correlate with fraud?

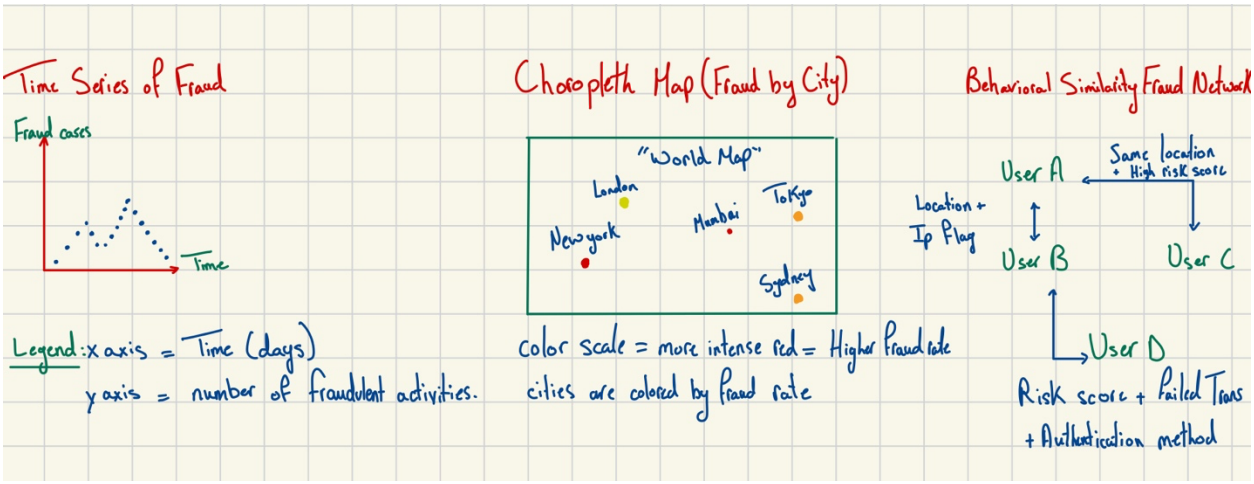
• Is there a spatial/geographic pattern?

• Can we identify clusters of suspicious users?

3. Technologies and Tools

Visualization	Technology/Library	Related Lectures
Time Series (Fraud Over Time)	D3.js (Line chart)	Lecture 4_2: D3 Basics, 5_2: Scales & Axes
Choropleth Map (Fraud by City)	D3.js Geo or Leaflet.js	Lecture 8_1: Maps, 8_2: Practical Maps
Similarity Network Graph	D3.js Force Layout	Lecture 10: Graphs
Webpage Layout	HTML, CSS, JavaScript	Lecture 1_2: Web Dev, 2: JavaScript
Color/Design Choices	Color scales, Perception-based	Lecture 6_1: Perception & Colors

4. Sketches of Visualizations



5. Core Visualization (MVP) and Optional Features

We have defined a set of core components that make up our minimum viable product (MVP). First, we'll build a static time series chart to show how fraudulent transactions evolve, helping us spot trends, spikes, or unusual activity tied to specific dates or periods. To make this more informative, we propose adding a simple selector that allows the user to break down the graph by transaction type or merchant category, so that patterns specific to certain behaviors or sectors can be explored more easily. Next, we'll create a static choropleth map that visualizes fraud distribution by location to help identify how hotspots change across different regions over time. In addition to these, we plan to include a few straightforward and informative plots, such as a bar chart showing the fraud rate by authentication method or a distribution plot of risk scores, which provide quick, valuable insights to the viewer. All these visualizations will be brought together in a scroll-based web page layout, allowing users to explore the data in a clear, structured flow. These elements will form the project's foundation and will be prioritized for completion.

If time allows, we'd like to explore a few additional features to enhance the interactivity and depth of the project. One idea is to include a behavioral similarity network graph, where users are connected based on shared suspicious traits like location, risk score, or authentication method. This could help uncover potential clusters of fraudulent behavior. We're also considering using calendar-style heatmaps to show patterns in fraud activity over days or weeks, highlighting specific timeframes with unusual spikes. We're also considering adding simple filters and toggles that allow users to explore specific data slices, such as focusing on a particular merchant category or zooming into a certain risk score range. Finally, subtle animated transitions between sections or chart states could make the storytelling experience more dynamic and intuitive.