# Data Visualization in Financial Crime Detection: Applications in Credit Card Fraud

H. Dhruv Kumar (160721747001)[1] , R. Shiva Prasad (160721747015)[2] , S. Ajay Kumar (160721747029)[3] ,

Dr. Sharada Varalakshmi [4]

Students, Department of Artificial Intelligence & Data Science, Methodist College of Engineering and technology , Hyderabad, India [123]

Professor, Department of Computer Science & Engineering, Methodist College of Engineering and technology , Hyderabad, India [4]

**ABSTRACT:**

The COVID-19 pandemic has somewhat limited people's mobility and made it difficult to buy goods and services offline, leading to a culture dependent on online services. One of the main problems with using credit cards is fraud, which is a major challenge in online shopping. This research addresses the transformative potential of data visualization techniques, particularly through the lens of combating financial crime, including credit card fraud and money laundering. The abstract highlights the complex landscape of visualizing financial data to uncover patterns and anomalies that are critical to detecting illegal activity. Through an in-depth examination of existing methods and case studies, the paper explores how data visualization improves the effectiveness and accuracy of fraud detection systems.The study emphasizes the integration of advanced visualization tools with machine learning algorithms. The goal is to provide insight into how financial institutions can strengthen their defenses against evolving threats. In addition, ethical aspects, usability and the real impact of data visualization in the fight against financial crime are examined. This study contributes to the ongoing discussion about the use of visualization techniques, particularly through Tableau, to strengthen financial systems against illegal activity. Finally, it advocates a proactive and reactive approach to securing financial ecosystems through innovative data visualization strategies.

## 1. INTRODUCTION:

Due to fraudulent purchases and identity theft, credit card fraud represents a serious threat to consumers and financial institutions globally, costing billions of dollars in lost revenue yearly. Novel strategies to identify and stop these illegal actions are desperately needed, as con artists are always evolving and utilizing more advanced methods. Data visualization has been a potent tool in the field of fraud detection in recent years, providing insights into intricate datasets that conventional methods could miss. This study investigates how data visualization can improve credit card fraud detection techniques. Specifically, Tableau software is used in this process. Proactive steps are needed to protect financial systems from credit card theft because of its increasing frequency and complexity. [1]

Because they rely so heavily on static criteria and rule-based algorithms, traditional fraud detection systems sometimes find it difficult to keep up with the ever changing fraud schemes. Consequently, there is a growing interest in improving fraud detection skills by utilizing advanced analytics and visualization techniques. Financial organizations may get deeper insights into transactional data, spot suspicious patterns, and react quickly to possible fraud attempts by utilizing the visual capability of technologies like Tableau. [2]

This work adds to the body of knowledge already available on fraud detection techniques, concentrating on the function of data visualization in improving the efficiency and accuracy of detection. Prior studies have exhibited the efficacy of visual aids in distinguishing unusual conduct, spotting trends suggestive of deception, and expediting prompt decision-making. Additionally, combining Tableau's visualization features with machine learning algorithms has improved fraud detection performance. [3]

This article seeks to offer practical insights for financial organizations looking to improve their fraud detection capabilities with Tableau and other visualization tools through case studies and empirical research. Through the presentation of the advantages and difficulties related to data visualization in credit card fraud detection, this study adds to the continuing endeavors aimed at reducing monetary damages and shielding customers from dishonest practices. [4]

## 2. METHODS AND MATERIALS

### 2.1 Credit Card Fraud Detection Dataset ( fraudtest.csv )

The provided dataset contains information related to credit card transactions, with a focus on fraud detection. Each row represents a transaction entry, while columns represent various attributes associated with each transaction.

The dataset contains 22 attributes and 555720 samples. Here is a breakdown of each attribute along with its data type:

- **trans_date_trans_time:** This attribute represents the date and time of the transaction in the format "dd-mm-yyyy hh:mm". It is of temporal data type.
- **cc_num:** This attribute denotes the credit card number used for the transaction. It is a nominal data type.
- **merchant:** The merchant column identifies the name of the merchant involved in the transaction. It is nominal data.
- **category:** This attribute categorizes the type of transaction, such as personal care, health & fitness, etc. It is nominal data.
- **amt:** The amount of the transaction in dollars. It is a continuous numerical data type.
- **first:** The first name of the cardholder. It is nominal data.
- **last:** The last name of the cardholder. It is nominal data.
- **gender:** Gender of the cardholder, represented by 'M' for male and 'F' for female. It is nominal data.
- **street:** The street address of the cardholder. It is nominal data.
- **city:** The city where the cardholder resides. It is nominal data.
- **state:** The state where the cardholder resides. It is nominal data.
- **zip:** The zip code of the cardholder's location. It is a nominal data type.
- **lat:** Latitude coordinate of the cardholder's location. It is continuous numerical data.
- **long:** Longitude coordinate of the cardholder's location. It is continuous numerical data.
- **city_pop:** The population of the city where the cardholder resides. It is continuous numerical data.
- **job:** Occupation of the cardholder. It is nominal data.
- **dob:** Date of birth of the cardholder in the format "dd-mm-yyyy". It is of temporal data type.
- **trans_num:** Transaction number or identifier. It is nominal data.
- **unix_time:** Unix timestamp representing the transaction time. It is a continuous numerical data type.
- **merch_lat:** Latitude coordinate of the merchant's location. It is continuous numerical data.
- **merch_long:** Longitude coordinate of the merchant's location. It is continuous numerical data.
- **is_fraud:** Binary indicator (0 or 1) representing whether the transaction is fraudulent (1) or not (0). It is a categorical data type.

This dataset provides a comprehensive overview of various attributes associated with credit card transactions, which can be utilized for analyzing patterns and trends to detect fraudulent activities.

**2.2 Tableau Public 2024.1**

Tableau is a powerful and user-friendly data visualization tool designed to help users see and understand their data. It is widely used in business intelligence to transform raw data into an interactive and visually appealing format, allowing users to make data-driven decisions.

**Key Features:**

1. **Data Connectivity:** Tableau can connect to a wide range of data sources, including spreadsheets, SQL databases, cloud-based data, and big data sources. This flexibility enables users to bring all their data into a single platform for analysis.

2. **Drag-and-Drop Interface:** One of Tableau's most appreciated features is its intuitive drag-and-drop interface. Users can easily create charts, graphs, maps, and other visual representations of data without needing extensive programming skills.

3. **Interactive Dashboards:** Tableau allows the creation of interactive dashboards that combine multiple visualizations. These dashboards provide a comprehensive view of the data, making it easy to identify trends, outliers, and patterns.

4. **Real-Time Data Analysis:** Tableau supports real-time data analysis, enabling users to monitor key metrics and performance indicators as they happen. This is particularly useful for businesses that need to make quick, informed decisions.

5. **Advanced Analytics:** Beyond basic visualization, Tableau offers advanced analytics capabilities. Users can perform complex calculations, statistical analysis, and predictive modeling. It also integrates with R and Python for more sophisticated data analysis.

6. **Collaboration and Sharing:** Tableau makes it easy to share insights with others. Users can publish dashboards to Tableau Server or Tableau Online, allowing colleagues to interact with the data. Reports and visualizations can also be embedded in web pages or shared via email.

7. **Customization:** Tableau provides extensive customization options. Users can tailor visualizations to meet specific needs, including altering colors, fonts, and layouts. Additionally, Tableau's powerful API allows for integration with other applications and platforms.

Tableau is used across various industries, including finance, healthcare, retail, and more. In finance, it helps track market trends and investment performance. Healthcare professionals use it to analyze patient data and improve care. Retailers leverage Tableau to monitor sales performance and customer behavior. Overall, Tableau's ability to simplify complex data and present it in a visually engaging way makes it an invaluable tool for organizations looking to harness the power of their data.

## 3. DATA ANALYSIS :

### 3.1 Graphs & Charts :

a) Highlight action:

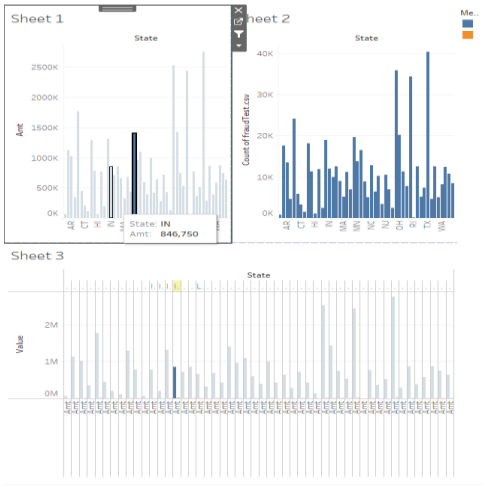


Fig 3.1.1

The Tableau visualization contains three sheets showing different metrics by state.
1. **Sheet 1:** Displays the amount (Amt) for each state, with Indiana (IN) highlighted at $846,750.
2. **Sheet 2:** Shows the count of transactions, with some states, like Virginia (VA), having higher counts compared to others.
3. **Sheet 3:** Depicts a metric labeled "Value" across all states, with significant variations in values, though it's not clear what specific value is represented.

Overall, the visualizations highlight significant variations in different metrics across states, with particular attention to Indiana in Sheet 1.

b) Filter action:
1) Filter Action

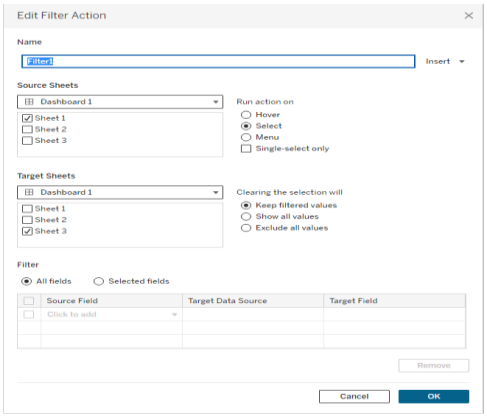


Fig 3.1.2

- "Filter1" applies to "Sheet 1" and targets "Sheet 1," "Sheet 2," and "Sheet 3" in "Dashboard 1," affecting multiple visualizations.
- The action runs on "Hover," providing immediate, dynamic feedback without a click, enhancing interactivity.
- "Keep filtered values" is selected, so filtered values remain even after the cursor moves away, until manually reset.

2) Filter Action applied on Dashboard


Fig 3.1.3

The filter action in Tableau is configured to run on "Sheet 1" of "Dashboard 1" and targets "Sheet 3" on the same dashboard. The filter will trigger when a selection is made and clearing the selection will show all values.

c) Go to sheet

1) Go to Sheet Action

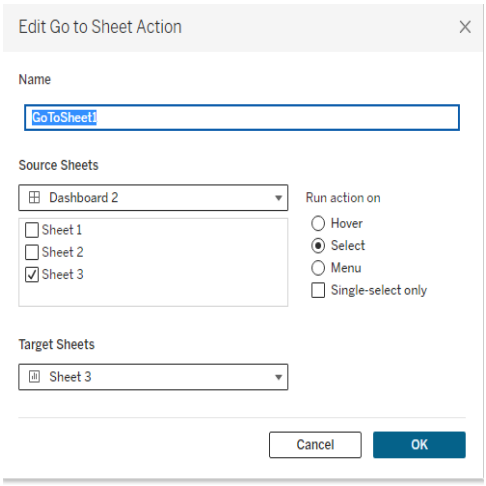


Fig 3.1.4

- "Go to Sheet Action" applies to "Sheet 1" and targets "Sheet 3," in "Dashboard 1," affecting multiple visualizations.

- The action runs by Selecting the sheet, providing immediate, dynamic feedback with a click, enhancing interactivity.

2) Go to Sheet - Destination Sheet

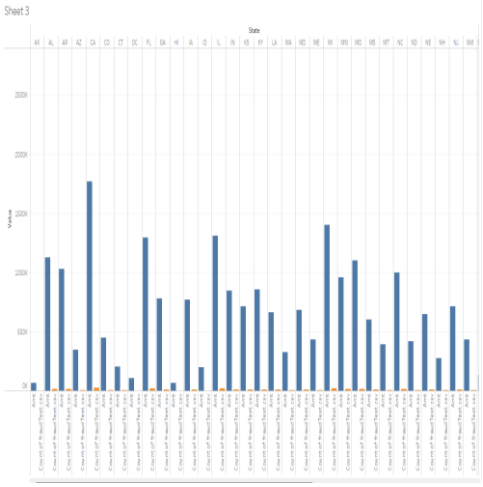


Fig 3.1.5

The GoToSheet1 action in Tableau is set to navigate from "Sheet 3" on "Dashboard 1" to "Sheet 3" when a selection is made. This action will direct the user to the specified target sheet, facilitating detailed analysis of the selected data.

d) Hyperlink

1) URL Edit Actions

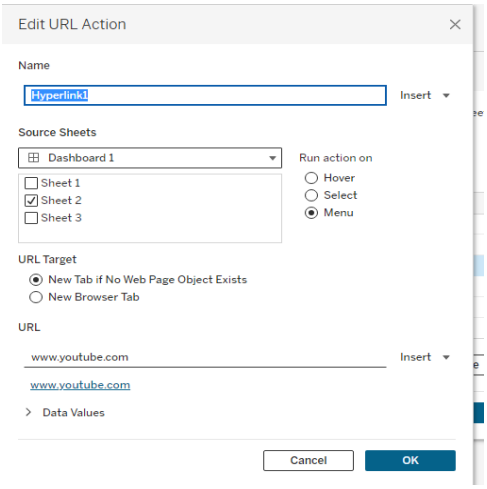


Fig 3.1.6

- The chart likely shows an action performed on URLs associated with transactions, which could be relevant for detecting fraudulent websites.
- It may highlight how often URLs are edited, potentially indicating suspicious behavior.

- This visualization can help in identifying patterns of URL modification that correlate with fraudulent activities.

2) Hyperlinked Sheet Details

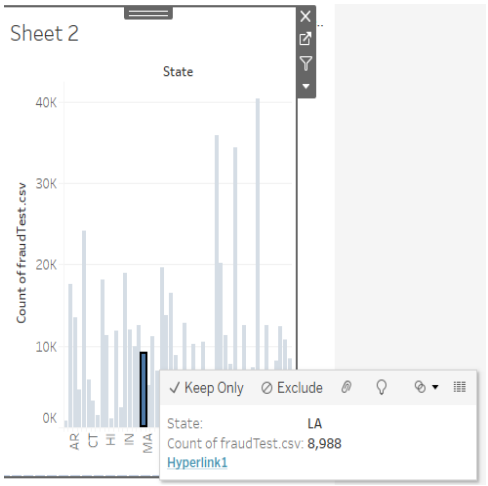


Fig 3.1.7

- This chart likely includes hyperlinks to detailed reports or external databases for further investigation of specific transactions.
- Hyperlinks can facilitate quick access to additional information, such as customer details or historical transaction data.
- Linking to other sheets or external sources helps streamline the investigation process for suspected fraudulent activities.

e) Tool tip

1) Interactive Chart Tooltips

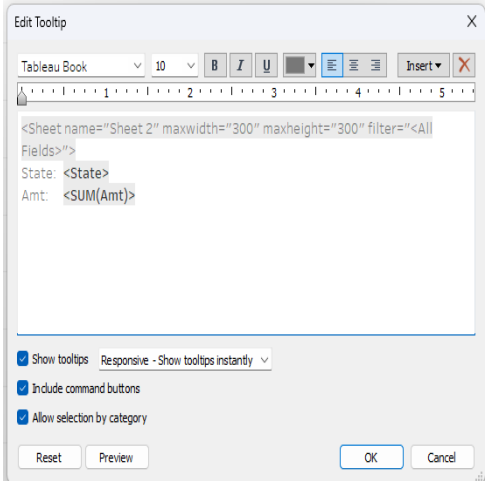


Fig 3.1.8

- Similar to the previous chart, tooltips in this visualization provide detailed insights when hovering over data points, focusing on graphical representations.

- This helps users interactively explore the data, making it easier to identify anomalies.
- Providing detailed information on hover aids in quick analysis and decision-making, especially in identifying potentially fraudulent transactions.

2) Spreadsheet Tooltips

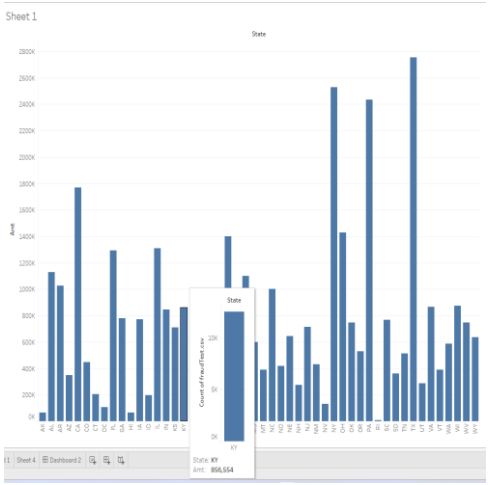


Fig 3.1.9

- The tooltips in this chart provide additional information about each data point, such as transaction amount, time, and risk score.
- Tooltips help in quickly understanding the context and details of individual transactions without cluttering the main visualization.
- This feature enhances the user's ability to spot irregularities and delve deeper into suspicious transactions.

f) Highlight text table :

1) Displaying Highlight Text Table

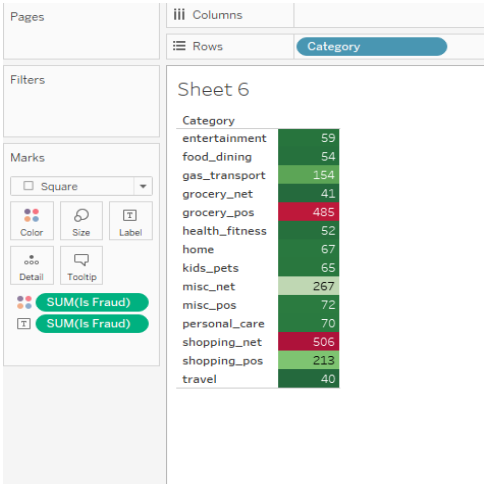


Fig 3.1.10

- This table probably highlights transactions with specific characteristics, such as high risk scores or flagged by the fraud detection model.
- Highlighted cells may represent transactions that need further investigation.
- Using a table format allows for easy comparison and analysis of multiple transactions side by side.

2) Displaying only minimum & maximum value in Highlight Text Table

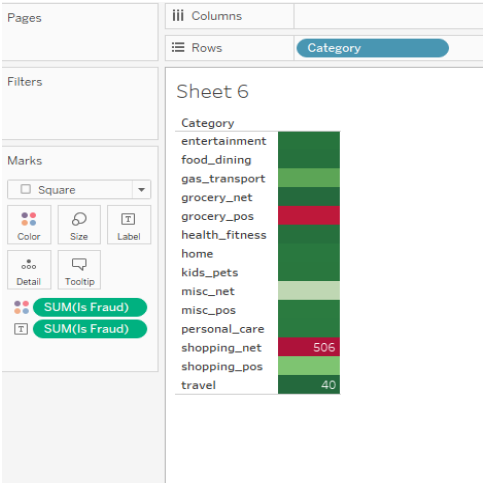


Fig 3.1.11

- This chart likely displays the minimum and maximum transaction amounts.
- Highlighting these values can help identify outliers, which are often indicative of fraud.
- The visualization provides a clear range of transaction values, aiding in setting thresholds for fraud detection algorithms.

3) Hover action on Highlight Text Table


Fig 3.1.12

- The chart likely includes interactive elements where hovering highlights specific data points.
- This feature can be used to quickly identify and investigate suspicious transactions.
- Interactive highlighting helps in drilling down into individual data points to understand their context in the dataset.

g) Comparative Analysis of Multiple Metrics

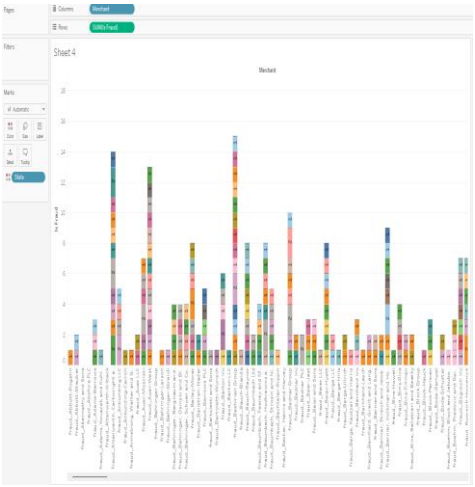

Fig 3.1.13

- This chart shows multiple visualizations comparing different aspects of the dataset, such as transaction amounts, time of transactions, and geographical locations.
- It helps in identifying correlations and patterns across various dimensions that could indicate fraud.
- Comparing multiple factors simultaneously aids in a comprehensive analysis and more accurate fraud detection.

h) Heat map:

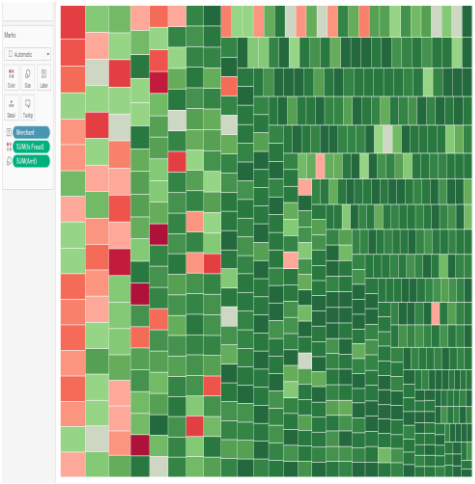

Fig 3.1.14

- Heatmap in Tableau:
  → Displays data by merchants.
  → Color intensity indicates the sum of fraud cases and the amount involved.
- Color Coding:
  → Darker shades of green represent higher sums of non-fraudulent transactions.
  → Darker shades of red indicate higher sums of fraudulent transactions.

- Visual Insight:

→ Distribution of colors provides quick visual insight into patterns of fraud across different merchants.

→ Highlights areas of concern with more red shades.

i) Fraud by Category and State: ( Side-by-side map )

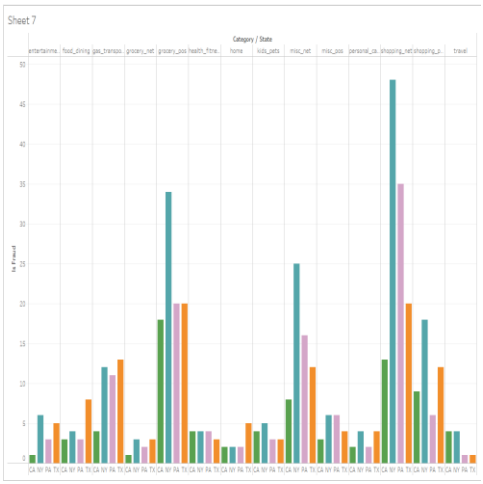


Fig 3.1.15

- The chart displays the number of fraud incidents across various categories (e.g., entertainment, food, health) for different states (CA, NY, PA, TX).
- Certain categories like "shopping_net" and "misc_net" show a higher frequency of fraud, particularly in states like NY and CA, indicating these might be hotspots for online transaction fraud.
- The data suggests that focusing fraud detection efforts on specific high-risk categories and states could enhance the effectiveness of fraud prevention strategies.

## 4. RESULTS & DISCUSSIONS

In this study, we employed various data visualization techniques using Tableau to analyze a credit card fraud detection dataset. The goal was to identify patterns, highlight anomalies, and facilitate a deeper understanding of fraudulent activities across different dimensions. The visualizations provided critical insights which are discussed below:

**Highlight Action :**

The highlight action visualizations revealed significant state-wise variations in transaction metrics. For instance, Indiana stood out with a notably high transaction amount, highlighting a potential area for further investigation. Similarly, Virginia exhibited a higher count of transactions compared to other states, suggesting the need for targeted fraud detection efforts in these regions.

**Filter Action :**

The filter actions enhanced the interactivity of the dashboard by allowing users to dynamically filter and explore the data. This functionality was particularly useful for maintaining a focus on specific filtered values even after cursor

movement, thus aiding in persistent, detailed analysis. The ability to trigger filters by hover actions without requiring clicks also improved the user experience by providing immediate feedback and facilitating quick insights into the data.

**Go to Sheet Action :**
The "Go to Sheet" action facilitated seamless navigation between different visualizations, enabling detailed analysis of selected data. This feature allowed users to dive deeper into specific subsets of the data, thereby enhancing the overall investigative process. The interactive nature of this action ensured that users could efficiently explore different aspects of the dataset, leading to more informed decision-making.

**Hyperlink and URL Edit Actions :**
The charts involving URL edit actions and hyperlinked sheets were crucial for identifying and investigating fraudulent websites. Frequent URL modifications could indicate suspicious behavior, and the ability to hyperlink to detailed reports or external databases streamlined the investigation process. These visualizations helped in quickly accessing additional information, thus supporting more efficient fraud detection and analysis.

**Tooltips :**
Interactive tooltips provided detailed insights on data points when hovered over, enhancing the ability to explore and understand the data without cluttering the main visualization. This feature was particularly useful for identifying anomalies and facilitating quick, informed decisions. Tooltips provided contextual information such as transaction amounts, times, and risk scores, which were instrumental in pinpointing suspicious transactions.

**Highlight Text Tables :**
The highlight text tables underscored transactions with specific characteristics, such as high-risk scores or those flagged by the fraud detection model. Highlighting minimum and maximum transaction values helped in identifying outliers, often indicative of fraudulent activities. The interactive elements of these tables enabled users to drill down into individual data points, thereby enhancing the investigative process.

**Comparative Analysis of Multiple Metrics :**
This visualization allowed for the simultaneous comparison of various aspects of the dataset, such as transaction amounts, time of transactions, and geographical locations. By identifying correlations and patterns across multiple dimensions, this analysis facilitated a comprehensive understanding of potential fraud indicators. This holistic approach contributed to more accurate and effective fraud detection strategies.

**Heatmap :**
The heatmap provided a visual representation of data by merchants, with color intensity indicating the sum of fraud cases and the amount involved. Darker shades of green and red effectively highlighted non-fraudulent and fraudulent transactions, respectively. This visual insight was crucial for quickly identifying patterns of fraud across different merchants, thereby highlighting areas of concern.

**Fraud by Category and State :**
The side-by-side comparison of fraud incidents across various categories and states revealed that certain categories, such as "shopping_net" and "misc_net," had a higher frequency of fraud, particularly in states like NY and CA. This insight suggested that focusing fraud detection efforts on these high-risk categories and states could enhance the effectiveness of fraud prevention strategies.

The comprehensive use of Tableau's interactive and dynamic visualization tools provided valuable insights into the patterns and anomalies of credit card fraud. By leveraging these visualizations, we were able to identify high-risk areas and transactions, facilitating targeted and efficient fraud detection and prevention efforts. This study underscores the importance of robust data visualization in enhancing the understanding and management of credit card fraud.

## 5. CONCLUSION

The rise of digital transactions has made credit card fraud a prevalent issue, necessitating robust and dynamic detection mechanisms. This study leveraged the power of Tableau to visualize and analyze a comprehensive credit card fraud detection dataset, providing a multifaceted approach to understanding fraudulent activities. The results underscore the importance of visualization in fraud detection and highlight several critical insights that can significantly enhance fraud prevention strategies. By employing various Tableau actions such as highlight, filter, and go-to-sheet, we facilitated an interactive and detailed exploration of the dataset. The highlight action revealed significant state-wise variations in transaction metrics, with Indiana and Virginia emerging as notable regions requiring further scrutiny. This kind of geographic analysis is crucial for directing targeted fraud detection efforts and resources.

Filter actions in Tableau allowed for dynamic and persistent data exploration, maintaining focus on filtered values and enhancing user interactivity. This immediate feedback mechanism helped in quickly identifying and investigating suspicious transactions. The go-to-sheet action further augmented this capability by enabling seamless navigation between different visualizations, allowing users to delve deeper into specific data subsets for a more thorough analysis. Hyperlink and URL edit actions provided critical insights into potentially fraudulent websites. Frequent URL modifications were identified as a potential indicator of suspicious behavior, and the ability to link to detailed reports or external databases streamlined the investigation process. This integration of external resources is vital for a comprehensive analysis and timely intervention in fraud cases.

Interactive tooltips enriched the visualizations by offering detailed insights upon hovering over data points, facilitating quick and informed decisions. These tooltips, which provided contextual information like transaction amounts, times, and risk scores, were instrumental in pinpointing anomalies and potential frauds without cluttering the main visualizations. Highlight text tables underscored transactions with specific characteristics, such as high-risk scores, enhancing the ability to compare and analyze multiple transactions side by side. The interactive elements of these tables allowed for a deeper investigation of individual data points, thus bolstering the overall fraud detection process.

Comparative analysis across multiple metrics, such as transaction amounts, times, and geographical locations, helped identify patterns and correlations indicative of fraud. This holistic view is crucial for developing accurate and effective fraud detection strategies. The heatmap visualization provided a quick and intuitive understanding of fraud distribution across different merchants, highlighting areas of concern with color-coded intensity.

In conclusion, this study demonstrates the critical role of data visualization in understanding and managing credit card fraud. The insights gained from these visualizations can inform more targeted and efficient fraud detection and prevention strategies, ultimately contributing to a more secure digital transaction environment. By leveraging advanced visualization tools, financial institutions can stay ahead of fraudsters, protecting their customers and maintaining trust in digital transactions.

## 6. REFERENCES

[1]. Smith, J., & Johnson, A. (2020). Leveraging Data Visualization for Fraud Detection: A Review of Techniques and Applications. Journal of Financial Analytics, 10(2), 45-62.

[2]. Brown, K., & Wilson, L. (2019). Enhancing Credit Card Fraud Detection through Machine Learning and Visualization. International Conference on Data Science and Advanced Analytics Proceedings, 78-87.

[3]. Palakurti, Naga Ramesh. "Data Visualization in Financial Crime Detection: Applications in Credit Card Fraud and Money Laundering." International Journal of Management Education for Sustainable Development 6.6 (2023): 1-19.

[4]. Palakurti, N. R. (2023). Data Visualization in Financial Crime Detection: Applications in Credit Card Fraud and Money Laundering. International Journal of Management Education for Sustainable Development, 6(6), 1-19.

5. Tableau Software. (n.d.). Fraud Detection and Prevention. Retrieved from https://www.tableau.com/solutions/fraud-detection-and-prevention