**PRE – REPORT FOR THE PROBLEM STATEMENT :**

**Chip vs. Swipe Performance: Do chip-based transactions result in fewer errors or discrepancies compared to swipe transactions?**

**STRATEGIC DECISIONS USING POWER BI – ASSIGNMENT 1**

**Name :**

**Devata Sri Sai Shravya**

**Registration Number:**

**2023JULB01107**

Contents

[1. Problem Statement 2](#_Toc183092341)

[1.1 Why is it important to solve this problem. 2](#_Toc183092342)

[1.2 Insights that can be gained from solving the problem 2](#_Toc183092343)

[2. Data Requirement 2](#_Toc183092344)

[2.1 Transaction Data 2](#_Toc183092345)

[2.2 Error Data: 3](#_Toc183092346)

[2.3 Possible Data Sources: 3](#_Toc183092347)

[3. Data Collection and Data Validation 3](#_Toc183092348)

[4. Data Cleaning 4](#_Toc183092349)

[4.1 Steps to clean: 4](#_Toc183092350)

[5. Tools used for Visualization of the problem - Power BI 5](#_Toc183092351)

[5.1 Power BI 5](#_Toc183092352)

[5.2 Why visualization is effective 5](#_Toc183092353)

[5.3 Charts implemented in our dashboard 6](#_Toc183092354)

[6. Dashboard 6](#_Toc183092355)

[7. Description of the Dashboard 7](#_Toc183092356)

[7.1 Bar Graph 7](#_Toc183092357)

[7.2 Pie Chart 7](#_Toc183092358)

[7.3 Funnel Chart 8](#_Toc183092359)

[7.4 Stacked Bar Chart 8](#_Toc183092360)

[7.5 Geographical Filled map 9](#_Toc183092361)

[8. Conclusion 9](#_Toc183092362)

# Problem Statement

**Chip vs. Swipe Performance: Do chip-based transactions result in fewer errors or discrepancies compared to swipe transactions?**

## 1.1 Why is it important to solve this problem.

* **Reduced Fraud Risk:** Chip-based transactions are generally more secure than swipe transactions due to their dynamic authentication process.
* **Faster Transactions:** Chip transactions can often be processed more quickly than swipe transactions, especially in high-volume environments.
* **Improved Cash Flow:** Efficient transaction processing can lead to faster payment processing and improved cash flow.
* **Risk Mitigation:** By analyzing performance metrics, businesses can identify potential compliance gaps and take corrective action.

## 1.2 Insights that can be gained from solving the problem

**By comparing chip and swipe transaction data, businesses can gain valuable insights into:**

* **Error Rates:** Identifying which transaction method has a higher error rate can help pinpoint areas for improvement in training, equipment, or processes.
* **Transaction Times:** Analyzing transaction times can help optimize checkout processes and reduce customer wait times.
* **Fraud Trends:** Monitoring fraud rates for each transaction method can help identify emerging threats and implement preventive measures.
* **Customer Preferences:** Understanding customer preferences for different payment methods can inform future strategies and investments.

# Data Requirement

* To effectively analyze chip vs. swipe performance, we need a dataset that includes the following information:

## 2.1 Transaction Data

* **Transaction ID:** A unique identifier for each transaction.
* **Date and Time:** The timestamp of the transaction.
* **Card Type:** Whether the card used was chip-based or swipe-based.
* **Transaction Amount:** The monetary value of the transaction.
* **Transaction Status:** The outcome of the transaction (e.g., approved, declined, error).
* **Merchant Category Code (MCC):** The type of merchant where the transaction occurred.
* **Merchant Location:** The geographic location of the merchant.

## 2.2 Error Data:

* **Error Type:** The specific type of error that occurred (e.g., card declined, system error, authorization failure).
* **Error Description:** A detailed description of the error.
* **Error Resolution Time:** The time taken to resolve the error.

## Possible Data Sources:

* **Point-of-Sale (POS) Systems:**

1. **Internal POS Systems:** If your business has its own POS system, it can provide valuable transaction data.
2. **Third-Party POS Providers:** Companies like Square, Clover, and Toast can provide aggregated data on transaction performance.

* **Payment Processors:**

1. **Payment Gateways:** Processors like Stripe, PayPal, and Adyen can provide detailed transaction data, including information on card type and error rates.
2. **Card Networks:** Visa, Mastercard, and American Express can provide aggregated data on card performance and industry trends.

* **Financial Institutions:**

1. **Banks and Credit Unions:** Banks and credit unions can provide data on card usage, transaction volume, and error rates.

By leveraging these data sources and carefully analyzing the relevant metrics, businesses can gain valuable insights into the performance of chip and swipe transactions, ultimately improving customer experience and reducing operational costs.

# Data Collection and Data Validation

* Given the sensitive nature of transaction data, which is typically not publicly accessible, we relied on a reputable online dataset source like Kaggle.
* This platform provides access to a variety of datasets, including those related to credit card transactions.
* We utilized a specific dataset focused on North American card usage, encompassing both online and in-store (swipe) transactions.
* The dataset further granularizes the data to the state and city levels, offering a detailed view of transaction patterns.

The dataset contains the following columns:

1. **id**: Transaction ID.
2. **date**: Date and time of the transaction.
3. **client\_id**: Unique identifier for the client.
4. **card\_id**: Identifier for the card used.
5. **amount**: Transaction amount (negative values might represent refunds or errors).
6. **use\_chip**: Transaction method (e.g., Swipe Transaction).
7. **merchant\_id**: Identifier for the merchant.
8. **merchant\_city**: City of the merchant.
9. **merchant\_state**: State of the merchant.
10. **zip**: Zip code of the merchant.
11. **mcc**: Merchant category code.
12. **errors**: Indicates issues (if any) with the transaction.

A screenshot of a computer

Description automatically generated

# Data Cleaning

* Data cleaning is a critical step in data analysis.
* It involves identifying and correcting errors, inconsistencies, and missing values.
* By cleaning data, we ensure data accuracy, reliability, and consistency.
* This process enhances the quality of insights derived from data analysis and improves the overall decision-making process

## 4.1 Steps to clean:

1. Remove duplicates.
2. Handle missing or invalid data in critical fields.
3. Normalize column names (if necessary).
4. Ensure proper data types (e.g., date, numeric).
5. Verify validity of monetary amounts (e.g., no unrealistic values).
6. Retain only relevant columns if specified.

The dataset has been cleaned with the following steps completed:

1. **Removed duplicates.**
2. **Handled missing values** in critical columns (id, date, client\_id, card\_id, amount, merchant\_id).
3. **Normalized column names** to lowercase with underscores.
4. **Ensured proper data types** for date and amount.
5. **Filtered monetary values** to include only realistic amounts (e.g., -1,000,000 to 1,000,000).

A screen shot of a table

Description automatically generated

# Tools used for Visualization of the problem - Power BI

## Power BI

* Power BI is a powerful business intelligence tool developed by Microsoft.
* It helped us transform raw data into visually appealing and interactive insights.
* With Power BI, we can create stunning reports, dashboards, and visualizations that make complex data easy to understand and interpret.

## 5.2 Why visualization is effective

* **Clear Comparison:** The bar chart directly compares the performance of chip and swipe transactions, making it easy to identify any differences in error rates.
* **Visual Impact:** The use of color and different bar heights enhances the visual appeal and understanding of the data.
* **Actionable Insights:** Visualization can help identify areas where improvements are needed, such as optimizing chip reader infrastructure or enhancing fraud detection systems.

## Charts implemented in our dashboard

* **Bar Graph** – to get a quantitative comparison in terms of total amount transaction between the transaction type
* **Pie Chart** – Percentage distribution of errors in the total transaction amount.
* **Funnel Chart** – To identify which error occurred the most.
* **Stacked Bar Chart** – To identify the occurrence of errors under both the transaction types.
* **Geographical Filled map** – To understand the distribution of swipe transactions and online transactions amongst the different states.

# Dashboard

A screenshot of a computer

Description automatically generated

# Description of the Dashboard

## Bar Graph

A screenshot of a graph

Description automatically generated

* The chart displays the total amount of transactions by the method used, either "Swipe Transaction" or "Online Transaction."
* **Swipe Transaction:** This method has a significantly higher total amount, represented by the tall purple bar.
* **Online Transaction:** This method has a smaller total amount, shown by the shorter red bar.
* Overall, the chart indicates that most of the transactions were performed using the swipe method, while online transactions contributed a smaller portion to the total amount

## Pie Chart

A screenshot of a graph

Description automatically generated

* The pie chart displays the sum of amounts by transaction type.
* The majority of transactions, **accounting for 84.98%, are "Swipe Transactions**," while **"Online** **Transactions" make up 15.02%.**

## Funnel Chart

A screenshot of a computer

Description automatically generated

* This funnel chart visualizes the breakdown of "use\_chip" events based on different error categories.
* It shows a significant drop-off at the **"Insufficient Balance" stage, indicating that this is the primary reason for failed transactions.**
* The other error categories have a minimal impact on the overall count.

## Stacked Bar Chart

A screenshot of a graph

Description automatically generated

* **Swipe Transactions:** Most of the amount is attributed to "9864," with a small portion due to "Bad Card Number" and "Bad CVV" errors.
* **Online Transactions:** Like Swipe Transactions, "9864" dominates the amount, with a slightly larger proportion attributed to errors compared to Swipe Transactions.
* The chart indicates that while both **Swipe and Online transactions have a significant portion of the amount associated with errors, Online transactions seem to have a slightly higher error rate.**

## Geographical Filled map

A map of the world

Description automatically generated

* The map shows the distribution of online and swipe transactions across various regions of the world.
* There are clusters of both online and swipe transactions in North America, Europe, and Asia.
* There seems to be a higher concentration of swipe transactions in North America compared to online transactions.
* Europe appears to have a more balanced distribution between online and swipe transactions.
* Asia shows a significant number of both online and swipe transactions.

# Conclusion

* **Global Distribution of Transactions:** Both online and swipe transactions are prevalent in North America, Europe, and Asia. However, North America shows a stronger preference for swipe transactions, while Europe exhibits a more balanced distribution between the two methods.
* **Error Analysis:** Most transaction amounts, both for swipe and online transactions, are associated with the "9864" category. However, online transactions seem to have a slightly higher error rate compared to swipe transactions, particularly in terms of insufficient balance.
* **Transaction Volume:** Swipe transactions significantly outnumber online transactions, accounting for over 84% of the total transaction volume.
* **Transaction Amount:** Swipe transactions also dominate in terms of total transaction amount, contributing a much larger portion compared to online transactions.
* **Primary Reason for Failed Transactions:** Insufficient balance is the primary cause of transaction failures, as indicated by the significant drop-off at this stage in the funnel chart. Other error categories have a minimal impact on the overall transaction count.

**THANK YOU**