M-KOPA
POWER TO OWN

M-KOPA
POWER TO OWN

FRAUD ANALYSE

By

ADEDOKUN ABDULMALIK ADEYEMI

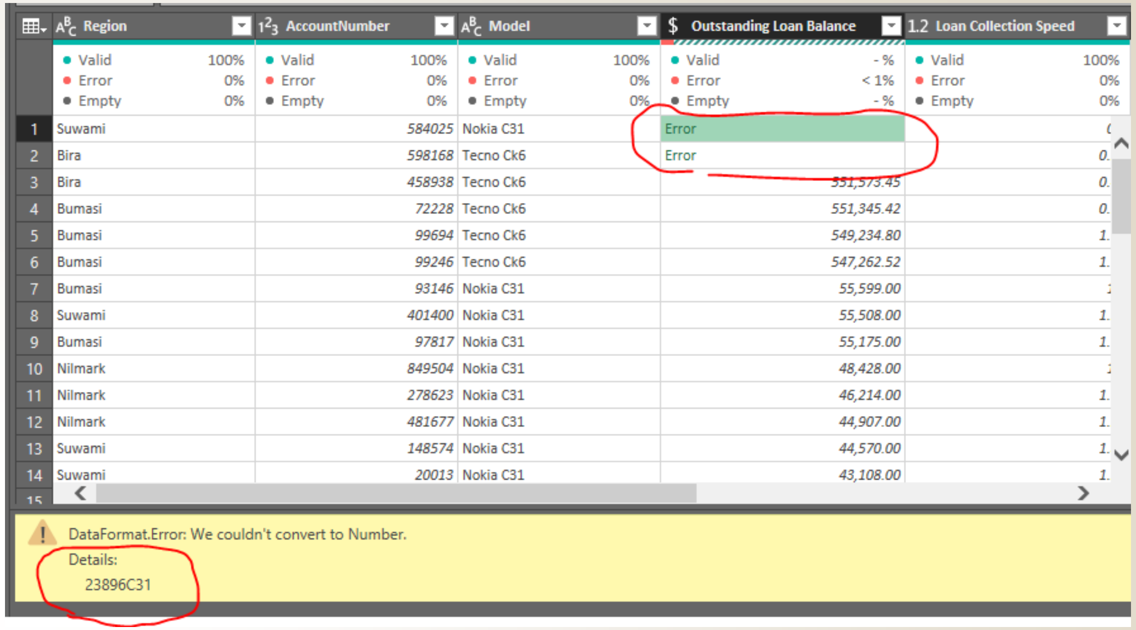
Table of Contents:

A. Data Cleaning and Preparation

Data cleaning was performed using Power Query. Several inconsistencies were identified:

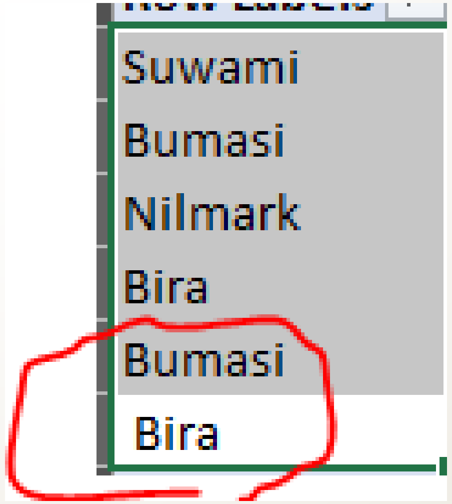
- Alphabetical characters within numerical values **(Column: Outstanding Loan Balance).**
- Misspelling of "investigated" as "investiagted"**(Column: Investigated).**
- Trailing spaces in "Bira" and "Bumasi" regions **(Column: Region).**

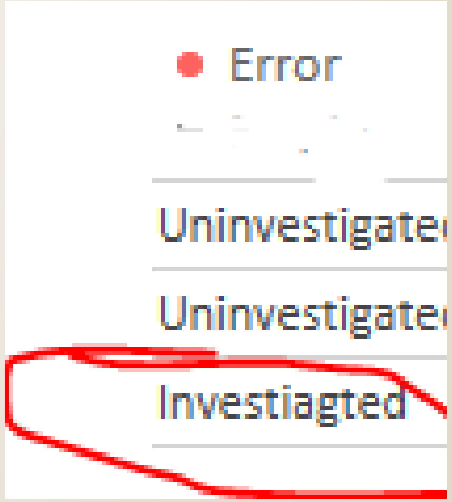
These errors were corrected to ensure data accuracy and reliability for subsequent analysis.



	Region	AccountNumber	Model	Outstanding Loan Balance	Loan Collection Speed
	Valid 100% / Error 0% / Empty 0%	Valid 100% / Error 0% / Empty 0%	Valid 100% / Error 0% / Empty 0%	Valid - % / Error < 1% / Empty - %	Valid 100% / Error 0% / Empty 0%
1	Suwami	584025	Nokia C31	Error	(
2	Bira	598168	Tecno Ck6	Error	0.
3	Bira	458938	Tecno Ck6	551,573.45	0.
4	Bumasi	72228	Tecno Ck6	551,345.42	0.
5	Bumasi	99694	Tecno Ck6	549,234.80	1.
6	Bumasi	99246	Tecno Ck6	547,262.52	1.
7	Bumasi	93146	Nokia C31	55,599.00	1
8	Suwami	401400	Nokia C31	55,508.00	1.
9	Bumasi	97817	Nokia C31	55,175.00	1.
10	Nilmark	849504	Nokia C31	48,428.00	2
11	Nilmark	278623	Nokia C31	46,214.00	1.
12	Nilmark	481677	Nokia C31	44,907.00	1.
13	Suwami	148574	Nokia C31	44,570.00	1.
14	Suwami	20013	Nokia C31	43,108.00	1.

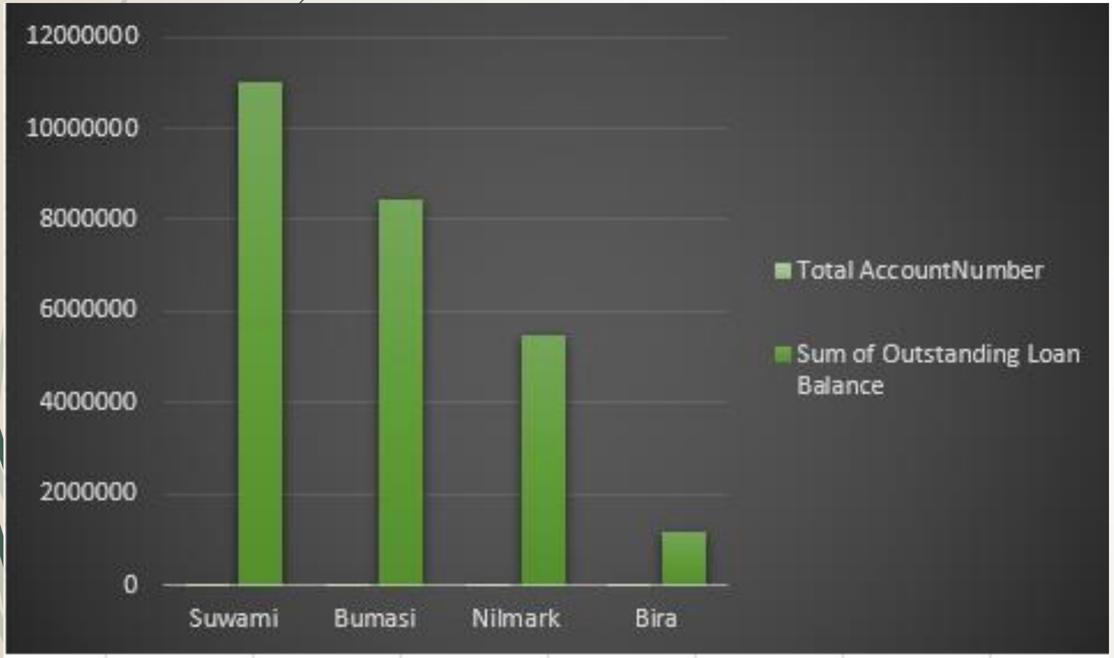
DataFormat.Error: We couldn't convert to Number.
Details:
23896C31





1. Which region has the highest Outstanding Loan Balance exposure

Row Labels	Total AccountNumber	Sum of Outstanding Loan Balance
Suwami	709	₦ 11,040,266.00
Bumasi	1016	₦ 8,455,168.00
Nilmark	352	₦ 5,496,709.00
Bira	266	₦ 1,182,567.00
Grand Total	**2343**	**₦ 26,174,710.00**



- Based on the provided table, we can determine the region with the highest Outstanding Loan Balance exposure by looking at the **"Sum of Outstanding Loan Balance"** column.
- **Conclusion:**

  Suwami has the highest Outstanding Loan Balance exposure with a total of **11,040,266.00**.
- **Evidence:**

  The table clearly indicates that Suwami has the largest sum of Outstanding Loan Balances compared to the other regions.
- **Reason:**
- The analyses suggests that Suwami might have a higher average loan balance per account compared to other regions.
- While Bumasi has more accounts (1016) compared to Suwami (709), the total outstanding loan balance for Suwami is significantly higher.

2. Ranking Regions by Potential Fraud Risk

Row Labels	Total AccountNumber	Sum of Outstanding Loan Balance
Suwami	709	₦ 11,040,266.00
Bumasi	1016	₦ 8,455,168.00
Nilmark	352	₦ 5,496,709.00
Bira	266	₦ 1,182,567.00
Grand Total	**2343**	**₦ 26,174,710.00**

Based on the data analyzed the potential risk factors can be ranked by regions as follows:

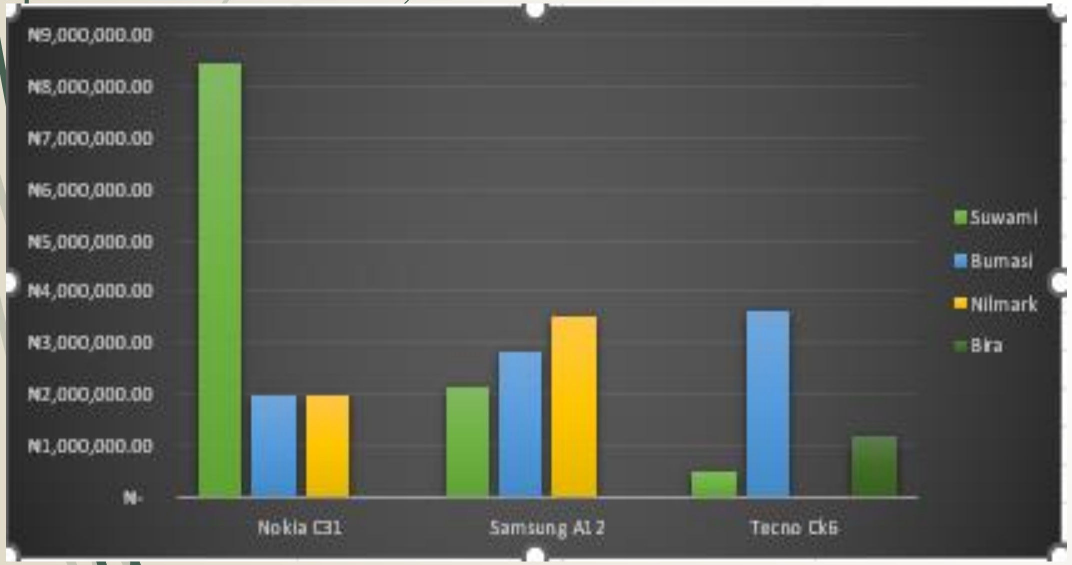
1. **Suwami:**
   - Highest total outstanding loan balance
   - Relatively high number of accounts (customer)
   - Potential for higher fraud risk due to higher average loan balance per account
2. **Bumasi**:
   - Second highest total outstanding loan balance
   - Highest number of accounts (customer)
   - Potential for higher fraud risk due to larger customer base
3. **Nilmark**:
   - Lower total outstanding loan balance compared to Suwami and Bumasi
   - Fewer accounts compared to the other two regions (i.e Suwami and Bumasi)
4. **Bira**:
   - Lowest total outstanding loan balance compared to all other region
   - Lowest number of accounts compared to all other region

Reasons for Ranking

1. Suwami ranks highest due to its combination of high total outstanding loan balance and a relatively large number of accounts. This suggests a potential higher risk profile(region).
2. Bumasi has the highest number of accounts, which could increase its vulnerability to fraud.
3. Nilmark and Bira have lower overall loan balances and fewer accounts, indicating a potentially lower fraud risk.

3. Using the data determine the most affected phone model, and region.

Sum of Outstanding Loan Balance	Region				
Phone Model	Suwami	Bumasi	Nilmark	Bira	Grand Total
Nokia C31	₦ 8,432,553.00	₦ 1,984,959.00	₦ 1,976,194.00		₦ 12,393,706.00
Samsung A12	₦ 2,125,184.00	₦ 2,833,330.00	₦ 3,520,515.00		₦ 8,479,029.00
Tecno Ck6	₦ 482,529.00	₦ 3,636,879.00		₦ 1,182,567.00	₦ 5,301,975.00
Grand Total	**₦ 11,040,266.00**	**₦ 8,455,168.00**	**₦ 5,496,709.00**	**₦ 1,182,567.00**	**₦ 26,174,710.00**



Most Affected Model:

- **Nokia C31** has the highest total outstanding loan balance at ₦12,393,706.00.

This indicates that Nokia C31 is the most affected model in terms of total outstanding loans.

Most Affected Region:

- Suwami has the highest total outstanding loan balance at ₦11,040,266.00.

This indicates that Suwami is the most affected region in terms of total outstanding loans.

Reasoning:

The model and region with the highest total outstanding loan balance are considered the most affected. This implies a higher number of loans or larger loan amounts outstanding in those specific categories.

4) Using the data show the most affected month by fraud.

Count of Investigated	Investigation Outcome			
Months	Confirmed Not Fraud	Confirmed to be Fraud	Not investigated	Grand Total
Jan	13	10	219	242
Feb	13	16	204	233
Mar	9	8	187	204
Apr	5	3	104	112
May	3	2	49	54
Jun			7	7
Jul	5	2	215	222
Aug	11	3	892	906
Sep	3	1	33	37
Oct	5	3	46	54
Nov	7	2	86	95
Dec	11	12	154	177
Grand Total	**85**	**62**	**2196**	**2343**

Based on the Analyses result the month with the highest number of **confirmed fraud cases** is **February** with **16** cases.

Therefore, **February** is the most affected month by fraud according to the given data.

- **Reason**:

February has the highest number of confirmed fraud cases. This suggests that there might have been a specific fraud pattern or vulnerability exploited during that month.

See next slide for further analyses

February Fraud Matrix

- Table 4.1

Region	AccountNumber	Model	Outstanding Loan Balance	Loan Collection Speed	Date of Sale	Investigated	Investiagtion outcome
Suwami	118510	Nokia C31	-1841	-0.08	17/02/2023	Investigated	Confirmed to be Fraud
Suwami	847632	Samsung	-387	-0.02	25/02/2023	Investigated	Confirmed to be Fraud
Suwami	400206	Samsung	235	0.01	27/02/2023	Investigated	Confirmed to be Fraud
Suwami	894039	Samsung	2442	0.16	20/02/2023	Investigated	Confirmed to be Fraud
Suwami	805265	Samsung	3587	0.25	27/02/2023	Investigated	Confirmed to be Fraud
Suwami	767638	Samsung	4026	0.27	22/02/2023	Investigated	Confirmed to be Fraud
Suwami	603338	Nokia C31	4167	0.19	12/02/2023	Investigated	Confirmed to be Fraud
Suwami	193386	Nokia C31	4226	0.21	24/02/2023	Investigated	Confirmed to be Fraud
Suwami	696486	Samsung	4419	0.28	10/02/2023	Investigated	Confirmed to be Fraud
Suwami	372900	Samsung	4463	0.29	12/02/2023	Investigated	Confirmed to be Fraud
Suwami	590985	Samsung	4570	0.28	04/02/2023	Investigated	Confirmed to be Fraud
Suwami	844775	Nokia C31	5213	0.25	16/02/2023	Investigated	Confirmed to be Fraud
Bumasi	17577	Nokia C31	5504	0.24	01/02/2023	Investigated	Confirmed to be Fraud
Suwami	452371	Nokia C31	5772	0.28	16/02/2023	Investigated	Confirmed to be Fraud
Bumasi	44981	Nokia C31	6019	0.29	18/02/2023	Investigated	Confirmed to be Fraud
Suwami	263183	Samsung	7163	0.48	19/02/2023	Investigated	Confirmed to be Fraud

Average: 0.21125 Numerical Count: 16 Min: -0.08 Max: 0.48 Sum: 3.38

- Table 4.2

Sum of Outstanding Loan Balance	Region		
Phone model	Bumasi	Suwami	Grand Total
Samsung A12		₦30,518.00	₦30,518.00
Nokia C31	₦ 11,523.00	₦17,537.00	₦29,060.00
Grand Total	**₦ 11,523.00**	**₦48,055.00**	**₦59,578.00**

- Analyses proof that their respective loan collection speed is below the overall average loan collection at **0.85 (table 4.1).**
- Also, Suwami Region has the highest loan outstanding balance which happened to be a provision for bad also solely on **Samsung A12 model.** Suwami might have a higher average loan balance per account.
- Furthermore, result shows that there are only **2 affected region (Bumasi and Suwami)** and **2 Phone model (Samsung A12 and Nokia C81).**
- As seen in the table 4.2 that **Bumasi** investigation outcome that was confirmed to be fraud was on both Phone Model compared to **Suwami** which was confirmed on one phone model only.

5) How could we potentially improve the fraud identification process.

Investiagation Status	Count of AccountNumber	Sum of Outstanding Loan Balance
Investigated	147	₦ 1,484,680.00
Uninvestigated	2196	₦ 24,690,030.00
Grand Total	**2343**	**₦ 26,174,710.00**

	Percentage (%) metrics
Investigated	**6%**
Uninvestigated	**94%**

From the analyses table above result show shows that there are huge number of accounts (customers) that as not been investigated totaling **2,196 uninvestigated acct** with the sum of their **outstanding balance of 24,690,030** compared to investigated account (customers) who are just **147 investigated acct** and total **outstanding balance of 1,484,680** which is **6% lower** compared to uninvestigated account **94% higher** which is venerable to potential fraud.

However, fraud identification process is as follows:

1. **Enhance Data Collection:**

- **Expand data points:** Collect additional information such as customer demographics, transaction history, device information, and location data.
- **Improve data quality:** Ensure data accuracy and consistency to enhance analysis effectiveness.

See next slide for further process

Fraud identification process.

2. **Implement Advanced Analytics:**

- **Machine learning:** Utilize machine learning algorithms to identify patterns and anomalies in the data, which can help predict fraudulent activities.
- **Statistical and Behavioral analytics:** using statistical methods to identify outliers and unusual trends also to analyze customer behavior to detect deviations from normal patterns (e.g. re-payment history, multiple payment method, etc.)

3. **Strengthen Fraud Prevention Measures:**

- **Identity verification:** Implement robust identity verification processes to prevent account takeover.
- **Fraud screening:** Utilize advanced fraud screening tools to detect suspicious activities in real-time.
- **Education and awareness:** Train employees to recognize fraud indicators and report suspicious behavior.

4. **Continuous Monitoring and Improvement:**

- **Regular reviews:** Conduct periodic reviews of fraud prevention and detection process to identify areas for improvement.

6) What operational improvements should we investigate to improve the fraud investigation process.

Improving the fraud investigation process involves streamlining workflows, optimizing resource allocation, and enhancing investigative capabilities. Here are some operational improvements to consider:

1. **Process Optimization:**

- **Standardized Investigation Protocols:** Establish clear and consistent procedures for fraud investigations to ensure efficiency and accuracy.
- **Case Management System:** Implement a centralized case management system to track investigations, share information, and monitor progress.
- **Prioritization Matrix:** Develop a system to prioritize cases based on potential loss, risk, and complexity.
- **Collaboration:** Foster collaboration between different departments involved in fraud investigation to enhance information sharing and knowledge transfer.

2. **Resource Allocation:**

- **Skill Development:** Invest in training staff (investigators) on fraud investigation techniques, forensic analysis, and emerging fraud trends.
- **Technology Adoption:** Leverage advanced tools and technologies to automate routine tasks and improve investigative efficiency.
- **Staffing Optimization:** Ensure adequate staffing levels to handle the investigation workload effectively.

3. **External Partnerships:**

- **Law Enforcement Collaboration:** Build strong relationships with law enforcement agencies to share information and coordinate investigations.
- **Industry Collaboration:** Collaborate with other organizations to share fraud intelligence and best practices.

7. Write an SQL query to replicate the results in DataSheet but only getting results for Suwami reg. Use the data on sheet named "Short schema"

```
SELECT
  ad.Region,
  ad.AccountNumber,
  pd.PhoneModel,
  pd.OutstandingLoanBalance AS "Outstanding Loan Balance",
  pd.CollectionSpeed AS "Loan Collection Speed",
  pd.Date AS "Date of Sale",
  ir.Investigated,
  ir.Investigation Outcome
FROM
  AccountsDetails AS ad
        INNER JOIN PaymentDetails AS pd ON ad.AccountNumber =
pd.AccountNumber
        LEFT JOIN InvestigationRecords AS ir ON a.AccountNumber =
ir.InvestigationRecords
WHERE
  region = 'Suwami';
```

- **Explanation**:

I use an **INNER JOIN** between ***AccountsDetails*** and ***PaymentsDetails*** to ensure that only accounts with corresponding payment information are included.
I also make use of **LEFT JOIN** between ***AccountsDetails*** and ***InvestigationRecords*** to include all accounts from ***AccountsDetails***, even if they don't have investigation records. The query accurately only retrieve the specified details for the '***Suwami***' region, including accounts without investigation records.

- **Observation Note:**

The provided SQL query references columns *DateOfSale, Investigated, and InvestigationOutcome* which do not exist in the specified table schema.
These columns are absent from the table structure, preventing their inclusion in the query results.

SECTION 2

1. Would you denote that investigated accounts perform similarly to non-investigated accounts when considering payment.

Investiagation Status	Count of AccountNumber	Sum of Outstanding Loan Balance
Investigated	147	₦ 1,484,680.00
Uninvestigated	2196	₦ 24,690,030.00
Grand Total	**2343**	**₦ 26,174,710.00**

- **Investigated accounts and non-investigated accounts exhibit significantly different payment performance.**

While investigated accounts represent only 6.3% of the total account number, they account for 5.66% of the total outstanding loan balance.

This suggests that uninvestigated accounts have a higher average outstanding loan balance compared to investigated accounts, indicating poorer payment performance.

2. Which region among the four would you consider to be the worst performing (or most vulnerable to fraud). And which Model in that region would you consider to be the worst model.

Region	Count of AccountNumber	Sum of Outstanding Loan Balance
Suwami	709	₦ 11,040,266.00
Bumasi	1016	₦ 8,455,168.00
Nilmark	352	₦ 5,496,709.00
Bira	266	₦ 1,182,567.00
Grand Total	**2343**	**₦ 26,174,710.00**

Phone Model	Sum of Outstanding Loan Balance
Nokia C31	₦ 12,393,706.00
Samsung A12	₦ 8,479,029.00
Tecno Ck6	₦ 5,301,975.00
Grand Total	**₦ 26,174,710.00**

Suwami has the highest total outstanding loan balance, indicating potential financial strain or higher risk exposure. This region might be considered the worst performing or most vulnerable to fraud based on this metric.

Nokia C31 has the highest outstanding loan balance, indicating potential financial strain or higher risk exposure. This model might be considered the worst performing or most vulnerable to fraud based on this metric.

3. Does the date a device was sold generally correlate to the balance a customer is left with. Explain (50 words)

Months	Count of AccountNumber	Average of Outstanding Loan Balance
Jan	242	₦ 16,033.55
Feb	233	₦ 14,086.84
Mar	204	₦ 11,711.48
Apr	112	₦ 10,387.19
May	54	₦ 7,557.46
Jun	7	₦ 5,556.86
Jul	222	₦ 4,579.27
Aug	906	₦ 5,023.39
Sep	37	₦ 15,335.92
Oct	54	₦ 24,170.94
Nov	95	₦ 46,002.37
Dec	177	₦ 18,091.41
Total	**2343**	**₦ 11,171.45**

The provided table outlines the relationship between months, the count of account numbers, and the average outstanding loan balance.

- **Key Observations:**
  - There's a significant fluctuation in the count of account numbers across different months.
  - August has the highest count of account numbers (906), followed by January (242) and February (233).
  - The average outstanding loan balance also varies significantly across months.
  - November has the highest average outstanding loan balance (46,002.37), followed by October (24,170.94) and September (15,335.92).
- **Potential Insights:**
  - Months with a higher count of account numbers don't necessarily correlate to higher average outstanding loan balances.
  - There seems to be a seasonal pattern in both account numbers and average outstanding loan balances.
  - Further analysis is required to understand the reasons behind these fluctuations.

4. The first stage when working with data is data cleaning. Explain the data cleaning process.

- **POWER QUERY**

Data cleaning involved handling missing values, inconsistencies, and outliers. This included removing irrelevant columns, formatting dates, and correcting data types.

Power Query's transformation steps were applied to streamline the process and ensure data accuracy for subsequent analysis. Sources and related content

5. The data was derived from three Tables -Tab tagged Short Schema
a) Write an SQL query to replicate the results in DataSheet but only getting results for Suwami reg

```
SELECT
  ad.Region,
  ad.AccountNumber,
  pd.PhoneModel,
  pd.OutstandingLoanBalance AS "Outstanding Loan Balance",
  pd.CollectionSpeed AS "Loan Collection Speed",
  pd.Date AS "Date of Sale",
  ir.Investigated,
  ir.Investigation Outcome
FROM
  AccountsDetails AS ad
      INNER JOIN PaymentDetails AS pd ON ad.AccountNumber =
pd.AccountNumber
      LEFT JOIN InvestigationRecords AS ir ON a.AccountNumber =
ir.InvestigationRecords
WHERE
  region = 'Suwami';
```

- **Explanation**:

I use an **INNER JOIN** between ***AccountsDetails*** and ***PaymentsDetails*** to ensure that only accounts with corresponding payment information are included. I also make use of **LEFT JOIN** between ***AccountsDetails*** and ***InvestigationRecords*** to include all accounts from ***AccountsDetails***, even if they don't have investigation records. The query accurately only retrieve the specified details for the '***Suwami***' region, including accounts without investigation records.

- **Observation Note:**

The provided SQL query references columns *DateOfSale, Investigated, and InvestigationOutcome* which do not exist in the specified table schema.

These columns are absent from the table structure, preventing their inclusion in the query results.

RECOMMENDATION SUMMARIZATION

- **Fraud Prevention and Detection**

1. Deepen fraud investigation.
2. Enhance fraud detection systems.
3. Train staff on fraud prevention.
4. Expand data for comprehensive analysis.
5. Continuously monitor for emerging threats.

- **Model and Regional Performance**

6. Investigate high outstanding loans for Nokia C31.
7. Review pricing and credit policies for Nokia C31.
8. Target marketing and loan restructuring for Nokia C31 users.
9. Conduct regional market analysis for Suwami.
10. Strengthen overall risk management and portfolio diversification.

THANK YOU

FRAUD ANALYSE

By

ADEDOKUN ABDULMALIK ADEYEMI