Data Analysis and Insights for Case Study Problem [Bank Data Warehouse]

Presentation by Kwabena Owusu Asante

# Agenda

- Case Study Overview
- The Consultant
- Process
- Data Analysis & Insights
- Summary

Case Study Overview

Case Study Problem Statement:

In the attached excel sheet (Bank Data Warehouse.xlsx), there are 5 tables as:
1. Customer – This contains the customer information.
2. Instrument – This contains the loans information.
3. Financial – This contains the transactions for the loans.
4. GL – This contains further information about the transactions.
5. Group GL – This contains the grouping information for the GL transactions.
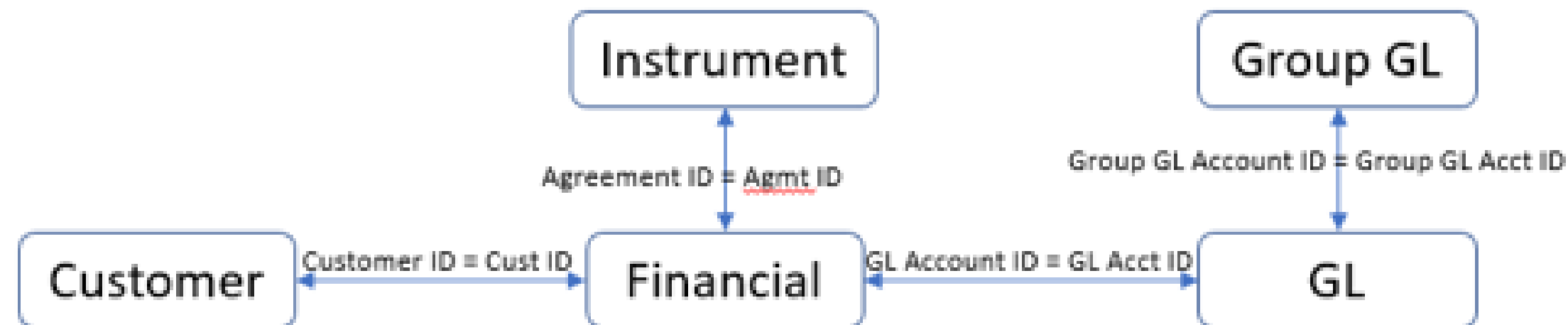
Case Study Overview

Case Study Tasks

What you need to do?

1. Read and profile the data. Explain your understanding about the data and provide insights.
2. Explain the assets and off balances for each customer sector category.
3. Explain the total amount for each sector category before and after adjustment. Hint: Adjustments have "ADJ" text in Source attribute.
4. Explain the amounts aggregated on countries. Show which country has maximum assets per NACE code and sector category.
5. What are the total assets for missing Customer Responsible Unit? Can you find any trend from customer or instrument perspective?
6. Considering the rating/scores as below priority order:
    a. "6+", "6-", "6" - Highest rating
    b. "0+", "0-", "0" – Lowest rating (Defaulted)
    c. "A+", "A-", "A" – Highest score
    d. "0+", "0-", "0" – Lowest score (Defaulted)
   e. U – Unassigned

Explain identified anomalies with performance of instruments.

7. Explain any other observation about the data.

Case Study Overview

The goal of this case study is to analyze and provide insights from a dataset containing five tables in the Bank Data Warehouse dataset. These tables include customer information, loan details, financial transactions, general ledger data, and grouping information for transactions. The analysis involves profiling the data, understanding asset distributions, adjustments, country-wise aggregations, missing data trends, and performance anomalies based on ratings and scores.
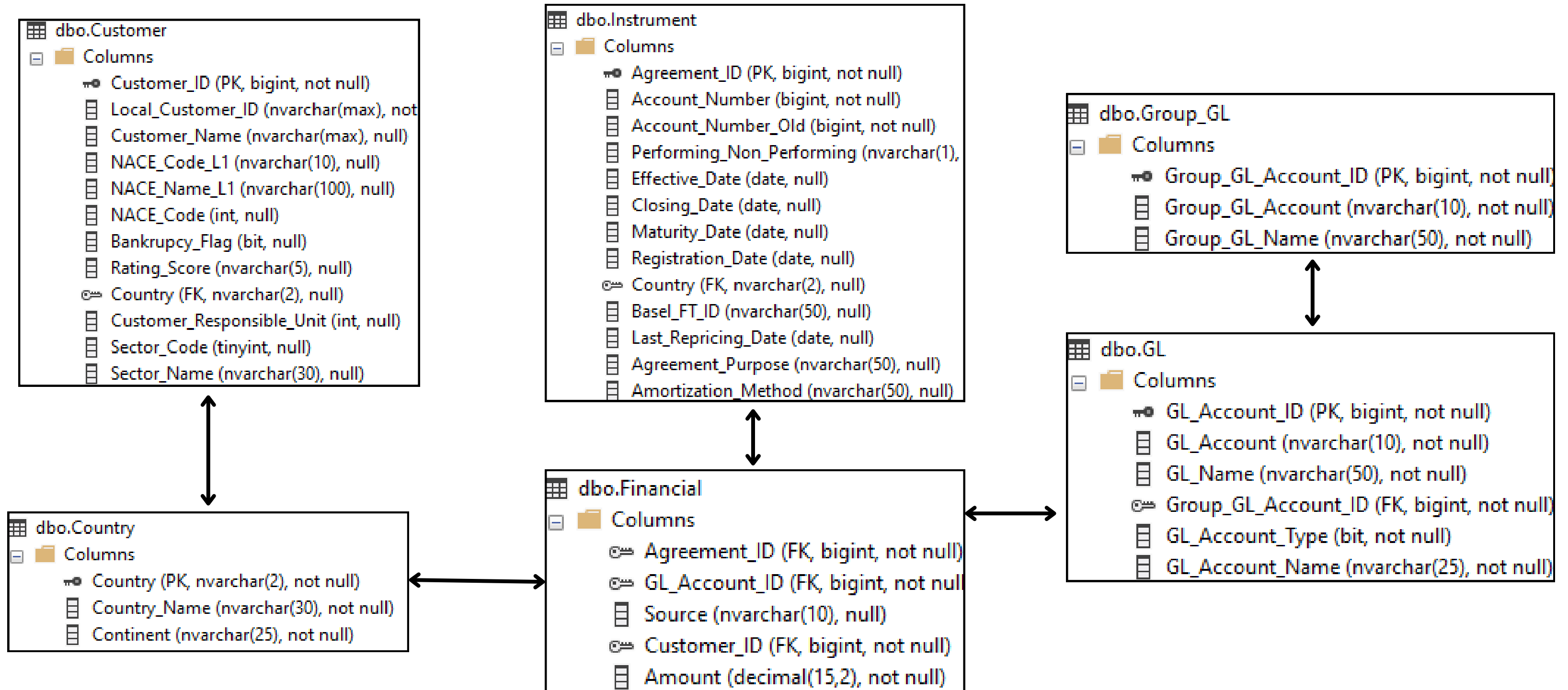
Data Model below :
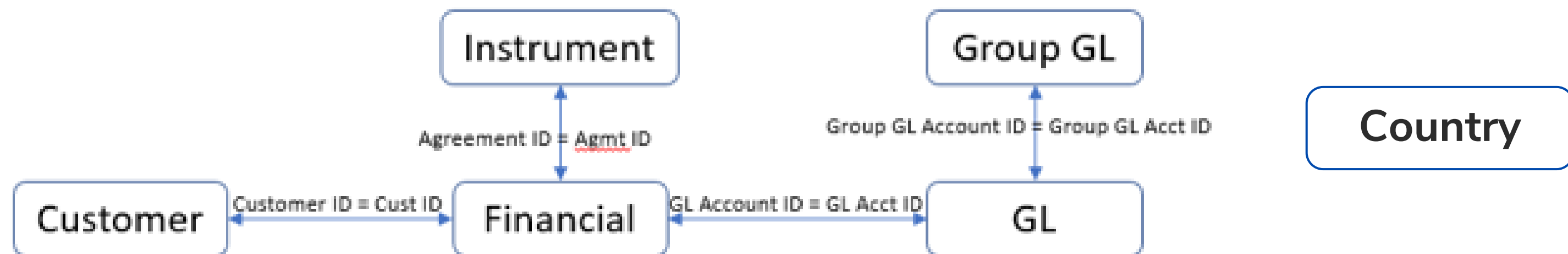
A new table **Country** was created to include country names and continent information using the ISO alpha-2 standard. This was done to derive country names and categorize them further into continents, enhancing the clarity of country and region-specific analyses.

Case Study Overview

Data Model with tables :

dbo.Customer
- Columns
  - Customer_ID (PK, bigint, not null)
  - Local_Customer_ID (nvarchar(max), not
  - Customer_Name (nvarchar(max), null)
  - NACE_Code_L1 (nvarchar(10), null)
  - NACE_Name_L1 (nvarchar(100), null)
  - NACE_Code (int, null)
  - Bankrupcy_Flag (bit, null)
  - Rating_Score (nvarchar(5), null)
  - Country (FK, nvarchar(2), null)
  - Customer_Responsible_Unit (int, null)
  - Sector_Code (tinyint, null)
  - Sector_Name (nvarchar(30), null)

dbo.Instrument
- Columns
  - Agreement_ID (PK, bigint, not null)
  - Account_Number (bigint, not null)
  - Account_Number_Old (bigint, not null)
  - Performing_Non_Performing (nvarchar(1),
  - Effective_Date (date, null)
  - Closing_Date (date, null)
  - Maturity_Date (date, null)
  - Registration_Date (date, null)
  - Country (FK, nvarchar(2), null)
  - Basel_FT_ID (nvarchar(50), null)
  - Last_Repricing_Date (date, null)
  - Agreement_Purpose (nvarchar(50), null)
  - Amortization_Method (nvarchar(50), null)

dbo.Group_GL
- Columns
  - Group_GL_Account_ID (PK, bigint, not null)
  - Group_GL_Account (nvarchar(10), not null)
  - Group_GL_Name (nvarchar(50), not null)

dbo.GL
- Columns
  - GL_Account_ID (PK, bigint, not null)
  - GL_Account (nvarchar(10), not null)
  - GL_Name (nvarchar(50), not null)
  - Group_GL_Account_ID (FK, bigint, not null)
  - GL_Account_Type (bit, not null)
  - GL_Account_Name (nvarchar(25), not null)

dbo.Country
- Columns
  - Country (PK, nvarchar(2), not null)
  - Country_Name (nvarchar(30), not null)
  - Continent (nvarchar(25), not null)

dbo.Financial
- Columns
  - Agreement_ID (FK, bigint, not null)
  - GL_Account_ID (FK, bigint, not null
  - Source (nvarchar(10), null)
  - Customer_ID (FK, bigint, not null)
  - Amount (decimal(15,2), not null)

The Consultant

- Professional with 12+ years of diverse background in Banking and Finance (Credit Risk & iT Support), Data Analysis and iT Consultancy
- Blend of technical, functional, client, consulting, and business management skills
- Proven ability to relate to all organizational levels
- Consistently add value through initiatives that increase revenue & enhance efficiency, and improve staff productivity
- Dedicated to leveraging expertise to deliver impactful solutions for business growth

Process

- Data Profiling and Understanding
- Table Creation for Data Model
- Addressing Case Study Tasks
- Data Analysis and Insights

Data Analysis & Insights [Customers]

SectorName	No Of Customers	%Percentage
Non financial corporations	5745	84.95%
Unclassified	471	6.96%
Households	292	4.32%
Other financial corporations	209	3.09%
General governments	24	0.35%
Credit institutions	21	0.31%
Central Banks	1	0.01%
Total	**6763**	**100.00%**

NACE_Name_L1	No Of Customers	%Percentage
REAL ESTATE ACTIVITIES	2113	31.24%
AGRICULTURE, FORESTRY AND FISHING	1336	19.75%
TRANSPORTATION AND STORAGE	528	7.81%
MANUFACTURING	493	7.29%
WHOLESALE AND RETAIL TRADE	387	5.72%
FINANCIAL AND INSURANCE ACTIVITIES	346	5.12%
CONSTRUCTION	308	4.55%
Total	**6763**	**100.00%**

Bankrupcy_Flag	No Of Customers
False	6305
	455
True	3
Total	**6763**

Priority Rating / Score	No Of Customers	%Percentage
⊞ **Unclassified**	**5913**	**87.43%**
⊟ **Lowest Rating (Defaulted)**	**355**	**5.25%**
PC0+	239	3.53%
PC0	92	1.36%
PC0-	24	0.35%
⊟ **Highest Rating**	**218**	**3.22%**
PC6-	95	1.40%
PC6+	66	0.98%
PC6	57	0.84%
⊟ **Unassigned**	**198**	**2.93%**
PPU	114	1.69%
PCU	84	1.24%
⊟ **Highest Score**	**67**	**0.99%**
PPA+	37	0.55%
PPA-	16	0.24%
PPA	14	0.21%
⊟ **Lowest Score (Defaulted)**	**12**	**0.18%**
PP0+	7	0.10%
PP0-	3	0.04%
PP0	2	0.03%
Total	**6763**	**100.00%**

Continent		No Of Customers	%Percentage
⊟ **Europe**		**5941**	**87.85%**
	Norway	2831	41.86%
	Denmark	2141	31.66%
	Finland	387	5.72%
	Sweden	361	5.34%
	Netherlands	30	0.44%
	United Kingdom	26	0.38%
	Poland	21	0.31%
	Germany	20	0.30%
	France	14	0.21%
	Belgium	13	0.19%
	Switzerland	12	0.18%
	Luxembourg	11	0.16%
	Russia	10	0.15%
	Jersey	8	0.12%
	Czech Republic	7	0.10%
	Faroe Islands	6	0.09%
	Italy	5	0.07%
	Malta	5	0.07%
	Slovakia	5	0.07%
	Estonia	4	0.06%
	Ireland	3	0.04%
Total		**6763**	**100.00%**

Continent	No Of Customers	%Percentage
Europe	5941	87.85%
	455	6.73%
Africa	118	1.74%
North America	98	1.45%
Asia	90	1.33%
Marshall Islands	45	0.67%
South America	16	0.24%
Total	**6763**	**100.00%**

Continent	Customer_Responsible_Units	%Percentage
Europe	296	95.79%
North America	33	10.68%
Asia	24	7.77%
Africa	9	2.91%
South America	9	2.91%
Marshall Islands	4	1.29%
	1	0.32%
Total	**309**	**100.00%**

Customer_Responsible_Unit	No_Of_Customers	%Percentage
	470	6.95%
1050002484	282	4.17%
1040221871	273	4.04%
1040331010	268	3.96%
1050002482	261	3.86%
Total	**6763**	**100.00%**

Data Analysis & Insights [Loans]

Agreement_Purpose	No of loans	%GT Count of Agreement_ID
FINANCING	10695	45.63%
GENERAL CORPORATE PURPOSES	5640	24.06%
REFINANCE	2206	9.41%
TERM FACILITY	1627	6.94%
SHIP FINANCE	1179	5.03%
REAL ESTATE	785	3.35%
REVOLVING FACILITY	723	3.08%
	277	1.18%
ACQUISITION	225	0.96%
ASSET PURCHASED	17	0.07%
UNCOMMITTED FACILITY	17	0.07%
CAPITAL EXPENDITURES	16	0.07%
WORKING CAPITAL	10	0.04%
TERM GUARANTEE FACILITY	8	0.03%
NOT SPECIFIED	5	0.02%
COMMERCIAL PAPER BACK-UP	4	0.02%
REVOLVING GUARANTEE FACILITY	3	0.01%
SWINGLINE FACILITY	3	0.01%
Total	**23441**	**100.00%**

SectorName_New	No of loans	%GT Count of Agreement_ID
Non financial corporations	20162	86.01%
Unclassified	1594	6.80%
Households	856	3.65%
Other financial corporations	708	3.02%
General governments	99	0.42%
Credit institutions	70	0.30%
Central Banks	1	0.00%
Total	**23441**	**100.00%**

Sector_Name	Assets	Off-balance-sheet items	Total
Central Banks	-457,756.50		**-457,756.50**
Households	-413,853,666.60	-11,516,032.62	**-425,369,699.22**
General governments	-1,003,070,078.63	-867,417,981.77	**-1,870,488,060.40**
Other financial corporations	-3,107,686,912.27	-1,427,111,834.77	**-4,534,798,747.04**
	-3,480,890,963.14	-2,874,168,375.78	**-6,355,059,338.92**
Non financial corporations	-41,152,739,476.88	-22,332,378,500.49	**-63,485,117,977.37**
Credit institutions	-4,278,054,820.97	-62,011,185,208.10	**-66,289,240,029.07**
Total	**-53,436,753,674.99**	**-89,523,777,933.53**	**-142,960,531,608.52**

Performing_Non_Performing	Count of Agreement_ID	%Percentage
Y	23396	99.81%
N	45	0.19%
Total	**23441**	**100.00%**

Amortization_Method	Count of Agreement_ID	%Percentage
IRREGULAR (CURRENCY LOANS)	9263	50.39%
AMORTIZATION FREE (BULLET)	7831	29.47%
	6231	19.32%
SERIAL AMORTIZATION	116	0.81%
Total	**23441**	**100.00%**

Assets

Sector_Name
- Central Banks
- Households
- General gove...
- Other financi...
- (Blank)
- Credit institut...
- Non financial...

0bn (0%)
-3.48bn (6.51%)
-41.15bn (77.01%)

Off-Balance Sheet Items

Sector_Name
- Households
- General gover...
- Other financia...
- (Blank)
- Non financial ...
- Credit instituti...

-0.01bn (0.01%)
-22.33bn (24.95%)
-62.01bn (69.27%)
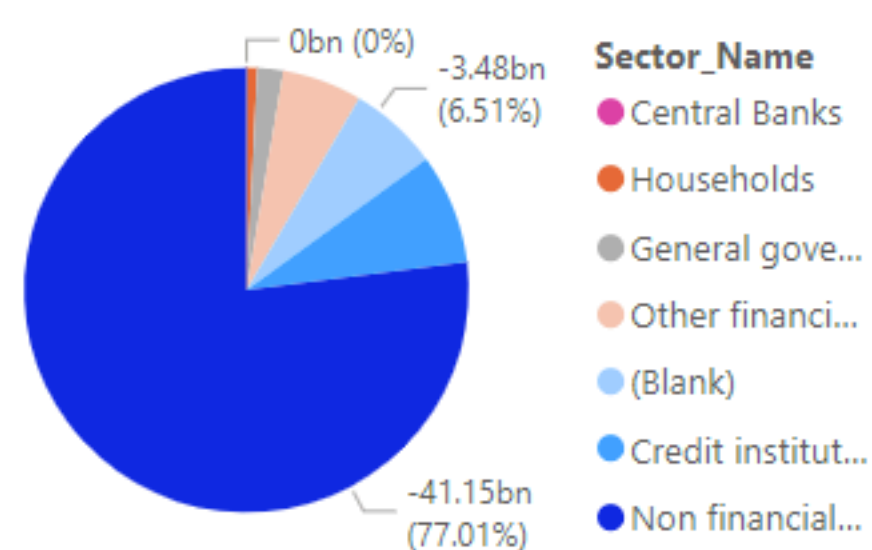
Summary

- Strong customer presence in Europe.
- Significant lending in real estate and agriculture sectors.
- Stable and reliable customer base: over 90% of loans are performing, and there are low bankruptcy rates among customers.
- Need for improved data management: the presence of NULL values in datasets suggests data quality issues, which are essential to address for robust risk management strategies.
- Invest in data analytics expertise to extract valuable insights, improve financial reporting, and enhance regulatory compliance.

Thank You

Presentation by Kwabena Owusu Asante