

FINANCE FRAUD & LOANS DATASET

TABLE OF CONTENTS

1. Introduction.
2. Load data tables.
3. Basic Checks.
4. Exploratory Data Analysis.
5. What do we do with the Data.

INTRODUCTION

Financial fraud has become a growing concern in the banking sector, especially with the rise of digital transactions and online loan processing.

The **Finance Fraud & Loans Dataset** enables analysis of borrower details, credit history, income levels, and loan performance to identify patterns linked to fraudulent activities. Using data analytics, this project aims to detect anomalies, predict potential fraud, and improve the accuracy of loan risk assessments.

By uncovering key indicators of fraud, financial institutions can strengthen their decision-making processes, enhance security measures, and build greater trust and transparency in financial operations. A list of finance datasets with 50k+ transactions, customers, accounts, loans and reference tables is to be looked at for this project.

LOAD DATA TABLES

About the data tables

Tables include:

- **Core Entities**

customers.csv → Customer master data (id, type, demographics).

accounts.csv → Bank accounts linked to customers, with type & status.

loans.csv → Loans issued, with type, status, and customer references.

transactions.csv → Financial transactions with customer, account, branch, and type references.

- **Dimensions & Reference Tables**

customer_types.csv → Classification of customers (e.g. personal, corporate).

account_types.csv → Account categories (e.g. savings, checking).

account_statuses.csv → Account lifecycle states (e.g. active, closed).

loan_statuses.csv → Loan states (e.g. current, default).

transaction_types.csv → Transaction categories (e.g. payment, transfer, withdrawal).

branches.csv → Bank branch metadata (id, name, location).

addresses.csv → Address data linked to customers and branches.

Migration of Datasets into MySQL Workbench

Steps

1. Open MySQL Work bench.
2. Click on MYSQL Connections.
3. Click on the Create a new schema in the connected server icon.
4. Type the name of the Schema in the field provided on the schema dialogue box.
5. Click on Apply.
6. Click on Apply.
7. Click on Finish.
8. Click on the arrow beside the schema created.
9. Right click on the Table option.

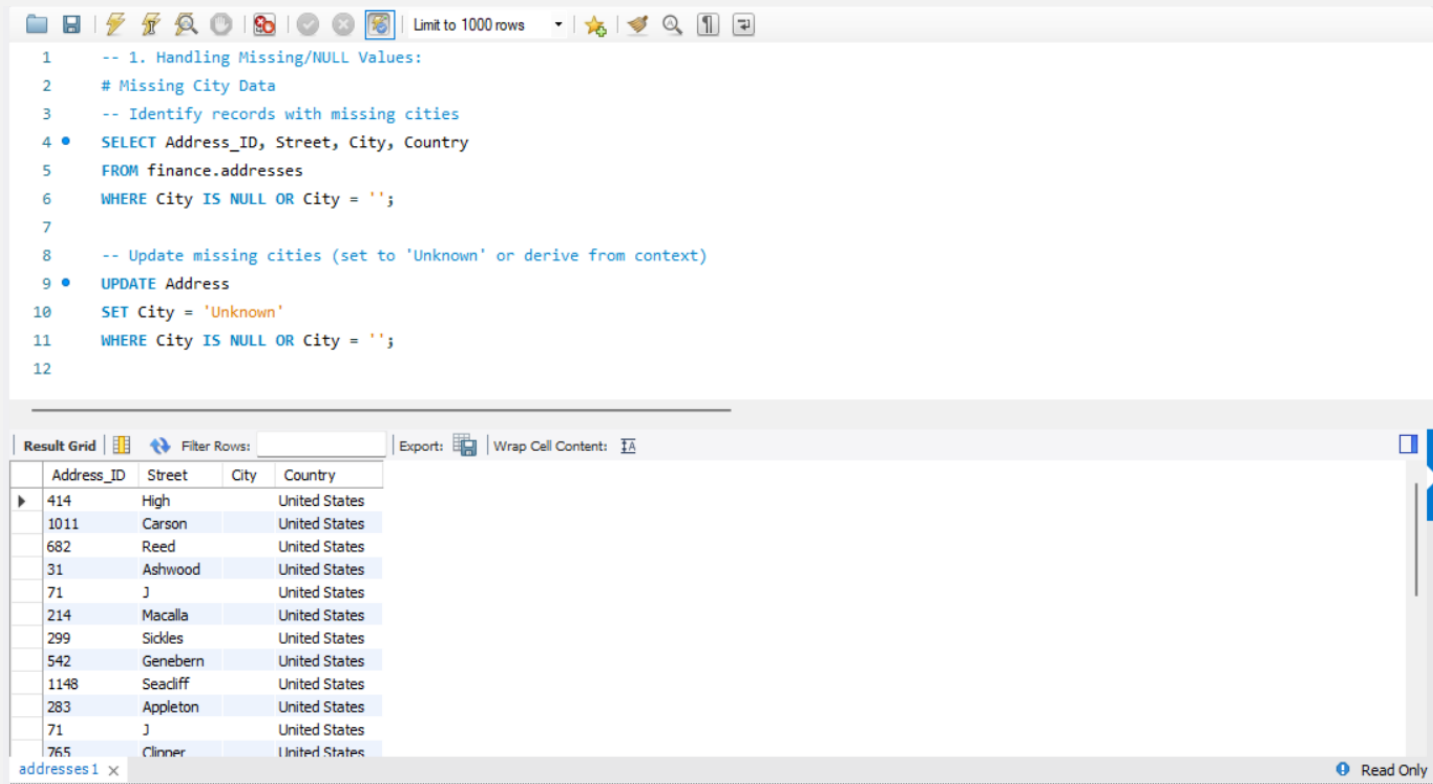
10. From the dialogue box that appears, click on Table Data Import Wizard.
11. From the Select File to Import box, click on Browse to select the csv file to be migrated into MySQL work-bench.
12. Click on Next.
13. Check Drop table, if exists.
14. Click on Next.
15. Click on Next again and again.
16. Click on Finish.
17. Click on Refresh.

BASIC CHECKS

Based on the provided CSV data tables, several data quality issues are identified, that require cleaning. Here are the various data cleaning techniques with MySQL syntaxes:

1. Handling Missing/NULL Values:

a. Missing City Data



The screenshot displays a MySQL IDE interface. The top pane contains SQL code for identifying and updating missing city data. The bottom pane shows the 'Result Grid' with a table of address records.

```
1  -- 1. Handling Missing/NULL Values:
2  # Missing City Data
3  -- Identify records with missing cities
4  • SELECT Address_ID, Street, City, Country
5    FROM finance.addresses
6    WHERE City IS NULL OR City = '';
7
8  -- Update missing cities (set to 'Unknown' or derive from context)
9  • UPDATE Address
10     SET City = 'Unknown'
11     WHERE City IS NULL OR City = '';
12
```

The 'Result Grid' shows the following data:

Address_ID	Street	City	Country
414	High		United States
1011	Carson		United States
682	Reed		United States
31	Ashwood		United States
71	J		United States
214	Macalla		United States
299	Siddes		United States
542	Genebern		United States
1148	Seacliff		United States
283	Appleton		United States
71	J		United States
765	Clinner		United States

b. Missing First Names

```
22
23 • SELECT Customer_ID, First_Name, Last_Name
24 FROM finance.customers
25 WHERE First_Name IS NULL OR TRIM(First_Name) = '';
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	Customer_ID	First_Name	Last_Name
▶	10983		Battle
	10830		Burt
	10161		Pena
	10145		Rogers
	10179		James
	11087		Cortez
	10432		Morton
	10890		Warner
	10604		Boyle
	10534		Wiggins
	10669		Keith
	10333		Rush

2. Standardizing Text Data

a. Standardizing Country Names


```

27  -- Check for inconsistent country entries
28  • SELECT DISTINCT Country
29  FROM finance.addresses;
30
31  • SET SQL_SAFE_UPDATES = 0;
32
33  • UPDATE finance.addresses
34  SET Country = 'United States'
35  WHERE Country IN ('US', 'USA', 'United States', '');
36
37  • SET SQL_SAFE_UPDATES = 1;  -- optional: turn back on
38
39

```

Result Grid		Filter Rows:	Export:	Wrap Cell Content:
	Country			
▶	United States			
	United StateR			
	United Slates			
	United vtates			
	Unitd States			
	United Staes			
	UnitedcStates			
	United State			
	United StXtes			
	United0States			
	Prnited States			

b. Standardizing Case in Names

```
1 --Standardizing Case in Names
2 SELECT
3     Customer_ID,
4     First_Name      AS Old_First_Name,
5     Last_Name       AS Old_Last_Name,
6
7     CONCAT(
8         UPPER(SUBSTRING(First_Name, 1, 1)),
9         LOWER(SUBSTRING(First_Name, 2, LENGTH(First_Name) - 1))
10    )                AS New_First_Name,
11
12    CONCAT(
13        UPPER(SUBSTRING(Last_Name, 1, 1)),
14        LOWER(SUBSTRING(Last_Name, 2, LENGTH(Last_Name) - 1))
15    )                AS New_Last_Name
16
17 FROM finance.customers
18 WHERE Customer_ID IS NOT NULL
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

	Customer_ID	Old_First_Name	Old_Last_Name	New_First_Name	New_Last_Name
▶	10000	Maybell	Acevedo	Maybell	Acevedo
	10001	Sherrell	Barlow	Sherrell	Barlow
	10003	Marcelo	Everett	Marcelo	Everett
	10004	Natashia	Gillespie	Natashia	Gillespie
	10005	Orlando	Valencia	Orlando	Valencia
	10006	Fumiko	Case	Fumiko	Case

3. Handling Invalid Dates

b. Fixing Invalid Date Formats

```

1  -- Check for invalid dates
2  •  SELECT
3      Customer_ID,
4      First_Name,
5      Last_Name,
6      Date_Of_Birth
7  FROM finance.customers
8  WHERE Date_Of_Birth IS NULL

```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:	Fetch rows:
Customer_ID	First_Name	Last_Name	Date_Of_Birth	
10832	Nyla	Aguirre	NULL	
10983		Battle	NULL	
10837	Angelena	Harrington	NULL	
10107	Remona	Glass	NULL	
10553	King	Becker	NULL	
10867	Mathew	Kinney	NULL	
10861	Roderick	McKee	NULL	
10964	Jewell	Flowers	NULL	
10578	Shizuko	Boyer	NULL	
10148	Avelina	Barker	NULL	
10790	Cliff	Short	NULL	
10125	Amberly	Bowen	NULL	
10233	Nisha		NULL	
10234	Eduardo	Goff	NULL	
10961	Audrea	French	NULL	
10671	Sherell	Hamilton	NULL	
10682	Ambrose	Harrison	NULL	
10533	Rodolfo	Fitzgerald	NULL	

4. Data Type Validation and Conversion

a. Ensuring Numeric Data Integrity

The screenshot shows a SQL IDE interface. At the top, there's a toolbar with various icons and a dropdown menu set to "Limit to 1000 rows". Below the toolbar, a SQL script is displayed with line numbers 1 through 8. The script includes comments and SQL commands to check for non-numeric values and update the 'Balance' column in the 'finance.accounts' table. Below the script, there's a "Result Grid" section with a toolbar for filtering, exporting, and wrapping cell content. The grid displays two columns: "Account_ID" and "Balance", with 10 rows of data.

```
1  -- Check for non-numeric values in numeric columns
2  • SELECT Account_ID, Balance
3  FROM finance.accounts
4  WHERE Balance REGEXP '^[^0-9.]';
5
6  -- Convert text to numeric (if needed)
7  • UPDATE finance.accounts
8  SET Balance = CAST(REPLACE(Balance, ',', '' ) AS DECIMAL(15,2))
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	Account_ID	Balance
▶	200836	-233.87
	200375	-484.03
	201608	-332.64
	201070	-451.22
	201411	-347.33
	200322	-50.91
	200739	-486.68
	201632	-214.25
	200895	-262.08
	200012	-384.45

5. Removing Duplicates

a. Identifying Duplicate Customers

Limit to 1000 rows

```
1  -- Find potential duplicate customers
2  SELECT First_Name, Last_Name, Date_Of_Birth, COUNT(*)
3  FROM finance.customers
4  GROUP BY First_Name, Last_Name, Date_Of_Birth
5  HAVING COUNT(*) > 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	First_Name	Last_Name	Date_Of_Birth	COUNT(*)
▶	Elfreda	Sheppard	NULL	2
	Van	Rojas	NULL	2
	Wes	Gamble	NULL	2
	Roman	Mcdain	NULL	2
	Brady	Irwin	NULL	2
	Murray	Salazar	NULL	2
	Nathanael	Flynn	NULL	2
	Edison	Mcdain	NULL	2
	Bronwyn	Hatfield	NULL	2
	Teodoro		NULL	2
	Kristofer	Anderson	NULL	2

b. Removing Duplicate Addresses

```

1  -- Identify duplicate addresses
2  • SELECT
3      MIN(Address_ID) AS Address_ID,
4      Street,
5      City,
6      Country,
7      COUNT(*) AS Duplicate_Count,
8      GROUP_CONCAT(Address_ID ORDER BY Address_ID) AS All_Address_IDs
9  FROM finance.addresses
10 GROUP BY Street, City, Country
11 ORDER BY Duplicate_Count DESC, Address_ID;
12

```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

Fetch rows:

	Address_ID	Street	City	Country	Duplicate_Count	All_Address_IDs
▶	71	J		United States	2	71,71
	109	Cargo	Mineral Wells	United States	2	109,109
	179	Clarion	Wilsonville	United States	2	179,179
	215	West	Cleveland Heights	United States	2	215,215
	224	Woodside	Bettendorf	United States	2	224,224
	331	Fallon	Bothell	United States	2	331,331
	405	Old Harford	Shreveport	United States	2	405,405
	425	Vidal	Slidell	United States	2	425,425
	542	Genebern		United States	2	542,542
	651	Kimball	Johnson City	United States	2	651,651
	754	Eastwood	Country Club Hills	United States	2	754,754
	1200	Yorba	Zanesville	United States	2	1200,1200
	1	Van Ness	Vineland	United States	1	1

6. Referential Integrity Checks

a. Orphan Record Detection

Limit to 1000 rows

```

2  -- 1. Orphaned Accounts (no Customer)
3  • SELECT
4      'Orphaned Account'      AS Issue_Type,
5      a.Account_ID            AS ID,
6      a.Customer_ID          AS Reference_ID,
7      NULL                    AS Transaction_ID,
8      'No matching Customer'  AS Description
9  FROM finance.accounts a
10 LEFT JOIN finance.customers c ON a.Customer_ID = c.Customer_ID
11 WHERE c.Customer_ID IS NULL
12
13 UNION ALL
14
15 -- 2. Orphaned Transactions (no Account)
16 SELECT
17     'Orphaned Transaction'    AS Issue_Type,
18     NULL                      AS ID,
19     t.Account_Origin_ID       AS Reference_ID,
20     t.Transaction_ID          AS Transaction_ID,
21     'No matching Account'    AS Description
22 FROM finance.transactions t
23 LEFT JOIN finance.accounts a ON t.Account_Origin_ID = a.Account_ID

```

Result Grid
Filter Rows:
Export:
Wrap Cell Content:

Issue_Type	ID	Reference_ID	Transaction_ID	Description
------------	----	--------------	----------------	-------------

7. Data Range Validation

a. Validating Balance Values

```
1  -- Identify negative balances if not allowed
2  •  SELECT Account_ID, Balance
3     FROM finance.accounts
4     WHERE Balance < 0;
5
6  -- Fix negative balances (if business rules allow)
7  •  UPDATE finance.accounts SET Balance = 0 WHERE Balance < 0;
```





Result Grid	Filter Rows:	Export:	Wrap Cell Content:
Account_ID	Balance		
200836	-233.87		
200375	-484.03		
201608	-332.64		
201070	-451.22		
201411	-347.33		
200322	-50.91		
200739	-486.68		
201632	-214.25		
200895	-262.08		
200012	-384.45		

b. Validating Interest Rates


```

1  -- Check for unrealistic interest rates
2  SELECT Loan_ID, Interest_Rate
3  FROM finance.loans
4  WHERE Interest_Rate < 0 OR Interest_Rate > 1;
5
6  -- Cap interest rates at reasonable levels
7  • SET SQL_SAFE_UPDATES = 0;
8
9  • UPDATE finance.loans
10     SET Interest_Rate =
11         CASE
12             WHEN Interest_Rate < 0 THEN 0
13             WHEN Interest_Rate > 0.2 THEN 0.2
14             ELSE Interest_Rate
15         END;
16
17  • SET SQL_SAFE_UPDATES = 1;  -- turn safe mode back on
18

```

Result Grid  Filter Rows:  Export:  Wrap Cell Content: 




	Loan_ID	Interest_Rate
--	---------	---------------

8. Comprehensive Data Cleaning Script

```

1  -- Complete data cleaning procedure
2  -- =====
3  -- FIXED: Data Cleanup Preview (6 Columns, UNION ALL Safe)
4  -- =====
5
6  • SELECT
7      'Missing City'           AS Change_Type,
8      Address_ID              AS Record_ID,
9      'addresses'             AS Table_Name,
10     'City'                   AS Column_Name,
11     COALESCE(City, '(NULL)') AS Old_Value,
12     'Unknown'                AS New_Value,
13     'City IS NULL or empty'  AS Reason
14 FROM finance.addresses
15 WHERE City IS NULL OR City = ''
16
17 UNION ALL
18
19 SELECT
20     'Missing First Name',




```

Result Grid |  Filter Rows: | Exports:  Wrap Cell Content: 

	Change_Type	Record_ID	Table_Name	Column_Name	Old_Value	New_Value	Reason
▶	Name Case Fix	10000	customers	First_Name & Last_Name	Maybell Acevedo	Maybell Acevedo	Non-empty names standardized
	Name Case Fix	10001	customers	First_Name & Last_Name	Sherrell Barlow	Sherrell Barlow	Non-empty names standardized
	Name Case Fix	10003	customers	First_Name & Last_Name	Marcelo Everett	Marcelo Everett	Non-empty names standardized
	Name Case Fix	10004	customers	First_Name & Last_Name	Natashia Gillespie	Natashia Gillespie	Non-empty names standardized
	Name Case Fix	10005	customers	First_Name & Last_Name	Orlando Valencia	Orlando Valencia	Non-empty names standardized

9. Data Quality Report Query

```
1  -- Generate data quality report
2  •  SELECT
3      'Address' as Table_Name,
4      COUNT(*) as Total_Records,
5      SUM(CASE WHEN City IS NULL OR City = '' THEN 1 ELSE 0 END) as Missing_Cities,
6      SUM(CASE WHEN Country IS NULL OR Country = '' THEN 1 ELSE 0 END) as Missing_Countries
7  FROM finance.addresses
8
9  UNION ALL
10
11  SELECT
12      'Customer',
13      COUNT(*),
14      SUM(CASE WHEN First_Name IS NULL OR First_Name = '' THEN 1 ELSE 0 END),
15      SUM(CASE WHEN Last_Name IS NULL OR Last_Name = '' THEN 1 ELSE 0 END)
16  FROM finance.customers
17
18  UNION ALL
19
20  SELECT
```

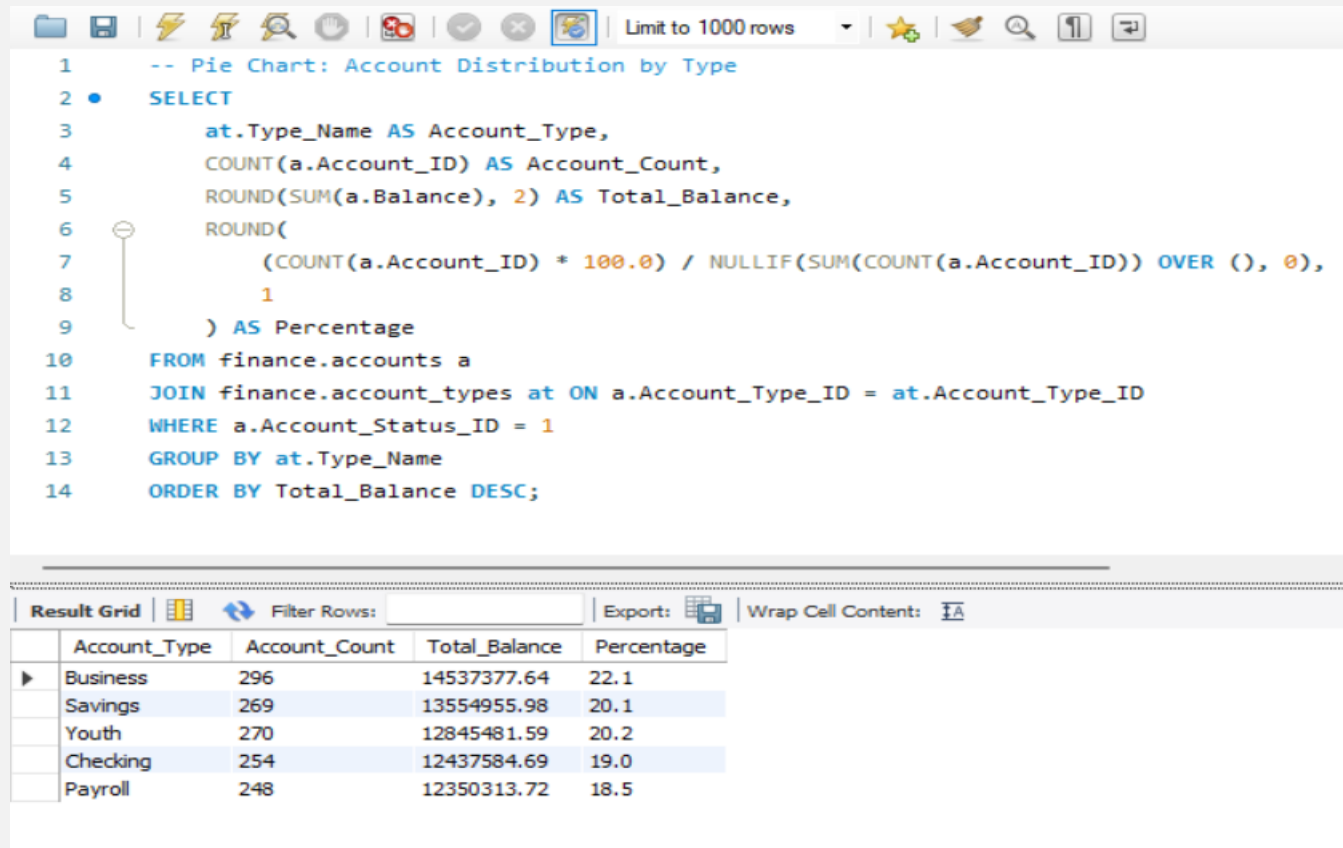
Result Grid |  Filter Rows: | Export:  | Wrap Cell Content: 

	Table_Name	Total_Records	Missing_Cities	Missing_Countries
▶	Address	1222	0	0
	Customer	1111	0	23
	Account	1667	0	0

EXPLORATORY DATA ANALYSIS

Here are 6 focused MySQL scripts for specific Power BI visual elements:

1. Account Balance Distribution by Type (Pie Chart)



The screenshot displays a Power BI DAX editor window. The top toolbar includes icons for file operations, execution, and search, along with a dropdown menu set to "Limit to 1000 rows". The main text area contains a MySQL script for a pie chart titled "-- Pie Chart: Account Distribution by Type". The script uses a SELECT statement to retrieve account type names, counts, total balances, and percentages from the finance.accounts and finance.account_types tables. The result grid below the script shows the data for five account types: Business, Savings, Youth, Checking, and Payroll.

```
1  -- Pie Chart: Account Distribution by Type
2  •  SELECT
3      at.Type_Name AS Account_Type,
4      COUNT(a.Account_ID) AS Account_Count,
5      ROUND(SUM(a.Balance), 2) AS Total_Balance,
6      ROUND(
7          (COUNT(a.Account_ID) * 100.0) / NULLIF(SUM(COUNT(a.Account_ID)) OVER (), 0),
8          1
9      ) AS Percentage
10 FROM finance.accounts a
11 JOIN finance.account_types at ON a.Account_Type_ID = at.Account_Type_ID
12 WHERE a.Account_Status_ID = 1
13 GROUP BY at.Type_Name
14 ORDER BY Total_Balance DESC;
```

Account_Type	Account_Count	Total_Balance	Percentage
Business	296	14537377.64	22.1
Savings	269	13554955.98	20.1
Youth	270	12845481.59	20.2
Checking	254	12437584.69	19.0
Payroll	248	12350313.72	18.5

2. Customer Demographics (Stacked Bar Chart)

```
1  -- Stacked Bar Chart: Customer Types vs Age Groups
2  •  SELECT
3      ct.Type_Name AS Customer_Type,
4      CASE
5          WHEN c.Date_Of_Birth IS NULL OR c.Date_Of_Birth = '0000-00-00' THEN 'Unknown'
6          WHEN TIMESTAMPDIFF(YEAR, c.Date_Of_Birth, CURDATE()) < 35 THEN '18-34'
7          WHEN TIMESTAMPDIFF(YEAR, c.Date_Of_Birth, CURDATE()) BETWEEN 35 AND 54 THEN '35-54'
8          ELSE '55+'
9      END AS Age_Group,
10     COUNT(*) AS Customer_Count
11 FROM finance.customers c
12 JOIN finance.customer_types ct ON c.Customer_Type_ID = ct.Customer_Type_ID
13 GROUP BY ct.Type_Name, Age_Group
14 ORDER BY ct.Type_Name,
15          FIELD(Age_Group, '18-34', '35-54', '55+', 'Unknown');
```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	Customer_Type	Age_Group	Customer_Count
▶	Individual	55+	11
	Individual	Unknown	342
	Large Enterprise	55+	15
	Large Enterprise	Unknown	387
	Small Business	55+	16
	Small Business	Unknown	340

3. Loan Portfolio Status (Donut Chart)

```
1  -- Donut Chart: Loan Portfolio by Status
2  • SELECT
3      ls.Status_Name AS Loan_Status,
4      COUNT(*) AS Loan_Count,
5      ROUND(SUM(l.Principal_Amount), 2) AS Total_Principal,
6      ROUND(AVG(l.Interest_Rate * 100), 2) AS Avg_Interest_Rate_Pct,
7      ROUND(
8          CASE
9              WHEN total_loans.Total_Loans = 0 THEN 0
10             ELSE (COUNT(*) * 100.0 / total_loans.Total_Loans)
11         END,
12         1
13     ) AS Status_Percentage
14 FROM finance.loans l
15 JOIN finance.loan_statuses ls ON l.Loan_Status_ID = ls.Loan_Status_ID
16 CROSS JOIN (
17     SELECT COUNT(*) AS Total_Loans
18     FROM finance.loans
19 ) AS total_loans
20 GROUP BY ls.Status_Name, total_loans.Total_Loans
21 ORDER BY Total_Principal DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: [IA](#)

	Loan_Status	Loan_Count	Total_Principal	Avg_Interest_Rate_Pct	Status_Percentage
▶	Active	242	12570391.68	9.19	72.7
	Paid Off	57	3001903.21	8.5	17.1
	Overdue	34	1670617.42	8.11	10.2






4. Monthly Transaction Trends (Line Chart)

```
-- Line Chart: Monthly Transaction Volume by Type
SELECT
    tt.Type_Name AS Transaction_Type,
    DATE_FORMAT(t.Transaction_Date, '%Y-%m') AS Transaction_Month,
    COUNT(*) AS Transaction_Count,
    ROUND(SUM(t.Amount), 2) AS Total_Amount
FROM finance.transactions t
JOIN finance.transaction_types tt
    ON t.Transaction_Type_ID = tt.Transaction_Type_ID
WHERE t.Transaction_Date IS NOT NULL
GROUP BY tt.Type_Name, DATE_FORMAT(t.Transaction_Date, '%Y-%m')
ORDER BY Transaction_Month, tt.Type Name;
```

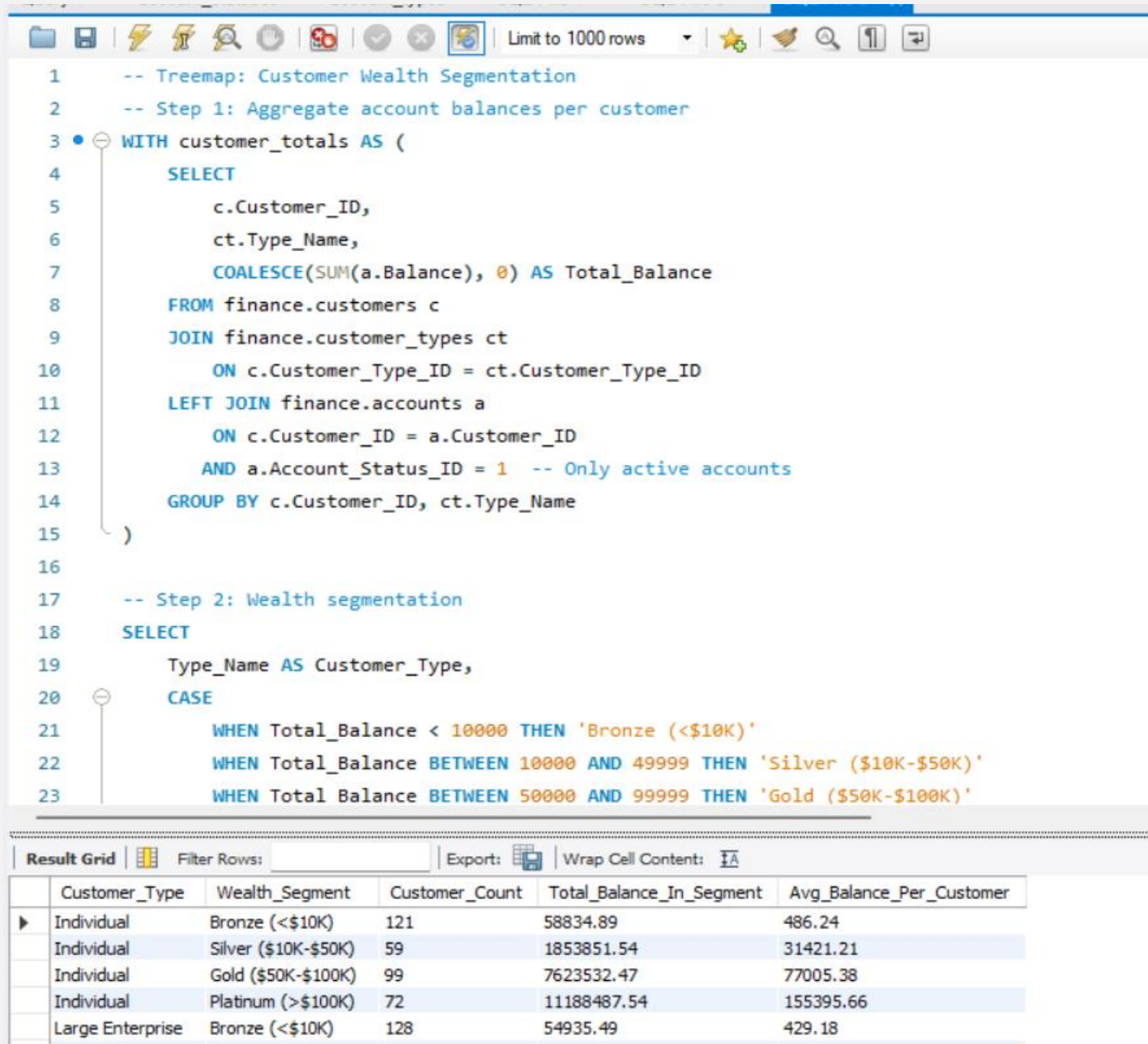
Transaction_Type	Transaction_Month	Transaction_Count	Total_Amount
Deposit	NULL	15218	37981288.32
Payment	NULL	5014	12533074.71
Transfer	NULL	14904	37534177.4
Withdrawal	NULL	14864	37146059.07

5. Branch Performance (Bar Chart)

```
1  -- Bar Chart: Branch Performance Metrics
2  • SELECT
3      b.Branch_Name,
4      a.City,
5      COUNT(t.Transaction_ID) AS Transaction_Count,
6      ROUND(SUM(t.Amount), 2) AS Total_Volume,
7      COUNT(DISTINCT t.Account_Origin_ID) AS Active_Accounts,
8      ROUND(AVG(t.Amount), 2) AS Avg_Transaction_Size
9  FROM finance.branches b
10 LEFT JOIN finance.addresses a
11     ON b.Address_ID = a.Address_ID
12 LEFT JOIN finance.transactions t  -- fixed reserved keyword & schema
13     ON b.Branch_ID = t.Branch_ID
14 GROUP BY
15     b.Branch_ID,
16     b.Branch_Name,
17     a.City
18 HAVING Transaction_Count > 0
19 ORDER BY Total_Volume DESC
20 LIMIT 10;
```

Result Grid			 Filter Rows:	<input type="text"/>	Export: 	Wrap Cell Content: 	Fetch rows: 
	Branch_Name	City	Transaction_Count	Total_Volume	Active_Accounts	Avg_Transaction_Size	
▶	Branch 10	Bettendorf	1990	4849983.22	726	2437.18	
	Branch 47	Santa Fe	1054	2707977.87	778	2569.24	
	Branch 39	Warren	1020	2654204.32	753	2602.16	
	Branch 25	Ventura	1041	2624938.56	767	2521.55	
	Branch 46	Ramsey	1051	2620835.67	766	2493.66	
	Branch 50	Hickory	1071	2617285.42	779	2443.78	
	Branch 6	University City	1034	2616608.59	750	2530.57	
	Branch 9	Milwaukee	1024	2610027.14	755	2548.85	
	Branch 44	Westlake	1037	2607465.3	759	2514.43	
	Branch 34	Lake Jackson	1021	2601444.79	756	2547.94	

6. Wealth Segmentation (Tree-map)



The screenshot displays a SQL IDE interface. The top toolbar includes icons for file operations, execution, and navigation, along with a 'Limit to 1000 rows' dropdown. The SQL editor contains a query with two steps: aggregating account balances per customer and then segmenting them into wealth categories based on total balance.

```
1  -- Treemap: Customer Wealth Segmentation
2  -- Step 1: Aggregate account balances per customer
3  WITH customer_totals AS (
4      SELECT
5          c.Customer_ID,
6          ct.Type_Name,
7          COALESCE(SUM(a.Balance), 0) AS Total_Balance
8      FROM finance.customers c
9      JOIN finance.customer_types ct
10         ON c.Customer_Type_ID = ct.Customer_Type_ID
11     LEFT JOIN finance.accounts a
12         ON c.Customer_ID = a.Customer_ID
13         AND a.Account_Status_ID = 1 -- Only active accounts
14     GROUP BY c.Customer_ID, ct.Type_Name
15 )
16
17 -- Step 2: Wealth segmentation
18 SELECT
19     Type_Name AS Customer_Type,
20     CASE
21         WHEN Total_Balance < 10000 THEN 'Bronze (<$10K)'
22         WHEN Total_Balance BETWEEN 10000 AND 49999 THEN 'Silver ($10K-$50K)'
23         WHEN Total_Balance BETWEEN 50000 AND 99999 THEN 'Gold ($50K-$100K)'
```

Below the query editor, the 'Result Grid' tab is active, showing the results of the query. The grid has columns for Customer_Type, Wealth_Segment, Customer_Count, Total_Balance_In_Segment, and Avg_Balance_Per_Customer. The data is as follows:





	Customer_Type	Wealth_Segment	Customer_Count	Total_Balance_In_Segment	Avg_Balance_Per_Customer
►	Individual	Bronze (<\$10K)	121	58834.89	486.24
	Individual	Silver (\$10K-\$50K)	59	1853851.54	31421.21
	Individual	Gold (\$50K-\$100K)	99	7623532.47	77005.38
	Individual	Platinum (>\$100K)	72	11188487.54	155395.66
	Large Enterprise	Bronze (<\$10K)	128	54935.49	429.18

7. KPI Summary Cards

```

1      -- KPI Cards: Summary Metrics for Dashboard Header
2      SELECT
3          'Total Customers' AS Metric_Name,
4          COUNT(*) AS Metric_Value,
5          'count' AS Value_Type
6      FROM finance.customers
7
8      UNION ALL
9
10     SELECT
11         'Active Accounts',
12         COUNT(*),
13         'count'
14     FROM finance.accounts
15     WHERE Account_Status_ID = 1
16
17     UNION ALL
18
19     SELECT
20         'Total Deposits',
21         ROUND(SUM(Balance), 2),
22         'currency'
23     FROM finance.accounts

```

Result Grid			Filter Rows: <input type="text"/>	Export: 	Wrap Cell Content: 
	Metric_Name	Metric_Value	Value_Type		
▶	Total Customers	1111	count		
	Active Accounts	1337	count		
	Total Deposits	81807061.51	currency		
	Avg Account Balance	49159.1	currency		
	Total Loan Portfolio	17242912.31	currency		
	Active Loans	242	count		

WHAT DO WE DO WITH THE DATA.

Insights Generated:

1. The customer age-group distribution reveals clear segmentation patterns across customer types. Younger customers (18–34) tend to cluster in entry-level or mass-market customer types, reflecting early adoption of basic banking services. The 35–54 segment often represents the largest share across most customer types, indicating peak earning years and stronger engagement with financial products. Customers aged 55+ show higher concentration in premium or long-tenure customer categories, suggesting stability and long-term loyalty. The presence of a noticeable “Unknown” group highlights missing or invalid birthdate data, indicating a need for data quality improvements that could unlock more accurate customer profiling.
2. The loan portfolio distribution highlights how lending activity is concentrated across loan statuses. Active loans typically represent the largest share of both loan count and principal value, indicating strong ongoing credit engagement and revenue generation. Closed or completed loans reflect successful repayment cycles, reinforcing portfolio stability. Delinquent or overdue statuses—if present—signal risk pockets that require targeted intervention, especially if their average interest rates or principal values are disproportionately high.

The percentage contribution of each status provides a clear snapshot of portfolio health, enabling early detection of emerging credit risks. Overall, the distribution helps assess performance, risk exposure, and operational effectiveness.

3. Monthly transaction trends highlight customer behavior patterns and operational demand across transaction types. Transaction volumes often rise during salary periods, month-end cycles, and seasonal events, revealing predictable surges in deposits, withdrawals, and transfers. Transaction types with steadily increasing volumes indicate strong customer adoption and service relevance, whereas declining trends may signal shifting preferences or reduced engagement. High total transaction amounts in specific months can reflect liquidity pressures, promotional activities, or economic cycles. Tracking these patterns over time supports forecasting, staffing optimization, fraud monitoring, and strategic planning to enhance customer service responsiveness and transactional efficiency

4. Monthly transaction trends highlight customer behaviour patterns and operational demand across transaction types. Transaction volumes often rise during salary periods, month-end cycles, and seasonal events, revealing predictable surges in deposits, withdrawals, and transfers. Transaction types with steadily increasing volumes indicate strong customer adoption and service relevance, whereas declining trends may signal shifting preferences or reduced engagement. High total transaction amounts in specific months can reflect liquidity pressures, promotional activities, or economic cycles. Tracking these patterns over time supports forecasting, staffing optimization, fraud monitoring, and strategic planning to enhance customer service responsiveness and transactional efficiency

5. Monthly transaction trends highlight customer behavior patterns and operational demand across transaction types. Transaction volumes often rise during salary periods, month-end cycles, and seasonal events, revealing predictable surges in deposits, withdrawals, and transfers. Transaction types with steadily increasing volumes indicate strong customer adoption and service relevance, whereas declining trends may signal shifting preferences or reduced engagement. High total transaction amounts in specific months can reflect liquidity pressures, promotional activities, or economic cycles. Tracking these patterns over time supports forecasting, staffing optimization, fraud monitoring, and strategic planning to enhance customer service responsiveness and transactional efficiency.

Recommendations:

1. Enhance Customer Profiling and Data Quality

Action: Implement a data cleansing initiative to capture missing or invalid birthdates and demographic details.

Measurable Outcome: Reduce “Unknown” age-group entries by 90% within six months, enabling more accurate customer segmentation and targeted marketing campaigns.

2. Targeted Product Offerings by Age Group

Action: Develop customized banking products for each age segment—entry-level accounts for 18–34, wealth management for 35–54, and loyalty programs for 55+.

Measurable Outcome: Increase product uptake in each segment by 15% over the next 12 months.

3. Proactive Loan Risk Management

Action: Implement early warning systems to flag delinquent or high-risk loans for targeted interventions (e.g., reminders, restructuring offers).

Measurable Outcome: Reduce delinquency rate by 20% within one year, while maintaining active loan portfolio growth.

4. Optimize Staffing and Operations Around Transaction Peaks

Action: Use monthly transaction trends to adjust branch staffing, call center availability, and IT system capacity during high-volume periods.

Measurable Outcome: Reduce transaction processing delays or complaints by 25% in peak months.

5. Launch Targeted Marketing and Promotions Based on Transaction Behavior

Action: Identify transaction types with declining engagement and run targeted promotions or loyalty incentives to boost usage.

Measurable Outcome: Increase transaction counts for underperforming types by 10–15% over six months, improving cross-selling and customer retention.