Project Title: Loan Data Analysis Dashboard using Microsoft Excel
Prepared for: The Data Immersed (TDI)
Prepared by: Chinedu Uzorue

1. Project Scope

Objective: Develop an interactive loan data analysis dashboard to help TDI understand loan performance, manage risks, and improve customer satisfaction.

Deliverables:

1. An Excel-based interactive dashboard showcasing key metrics and trends.

2. Insights into loan disbursement, repayments, and customer demographics.

3. Actionable recommendations to optimize loan approval processes and profitability.

Key Focus Areas:

- Identifying default risk patterns.

- Analyzing borrower demographics and loan purposes.

- Optimizing interest rates and approval criteria.

- Visualizing loan trends across regions and over time.

2. Problem Statement

TDI Challenges:

- High loan default rates affecting profitability.

- Insufficient borrower risk profiling leading to suboptimal decisions.

- Need for better optimization of loan approval and interest rates.

The dashboard aims to address these challenges by providing actionable insights and facilitating data-driven decisions.

3. Excel File Overview

The Excel file contains two sheets:

1. **Dataset**: Contains the loan data for analysis.

2. **Data Dictionary**: Describes the columns in the dataset, providing details about their meaning.

Key Columns for Analysis:

1. **id**: Unique identifier for each loan.

2. **address_state**: Borrower's state.

3. **application_type**: Type of loan applicant (e.g., individual).

4. **emp_length**: Borrower's employment length.

5. **emp_title**: Borrower's job title.

6. **grade & sub_grade**: Credit grading.

7. **home_ownership**: Home ownership status.

8. **issue_date**: Date the loan was issued.

9. **last_credit_pull_date**: Most recent credit report pull date.

10. **term**: Loan term (e.g., 36 or 60 months).

11. **verification_status**: Income verification status.

12. **annual_income**: Borrower's annual income.

13. **dti**: Debt-to-Income ratio.

14. **installment**: Monthly installment amount.

15. **int_rate**: Interest rate.

16. **loan_amount**: Loan principal.

17. **total_acc**: Number of credit accounts the borrower has.

18. **total_payment**: Total payment made on the loan.

4. Key Tasks Identified

Data Cleaning:

- Fix date formats.

- Handle missing or inconsistent values in key columns.

Calculated Fields:

- Create metrics such as the Approval Rate or default rates.

Insights for Dashboard:

- Analyze loan default trends.

- Explore relationships between income, interest rates, and risk.

5. Data Cleaning and Transformation

The raw dataset required extensive preprocessing to ensure accuracy and usability for analysis.

Data Observations:

1. **Duplicates**: 16 duplicate rows identified for removal.

2. **Missing Values**:

   - **emp_title**: Missing values for job titles (1440 entries).

   - **annual_income, loan_amount, total_payment**: Few missing values in critical numeric columns.

3. **Data Types**:

- Date columns (e.g., issue_date, last_credit_pull_date) stored as strings need conversion to datetime format.

- Validation required for columns like id and member_id.

Cleaning Plan:

1. Remove duplicate rows.

2. Convert date columns to the correct format.

3. Handle missing values.

4. Ensure numeric columns like loan_amount, dti, and int_rate are consistent.

Steps Taken:

1. **Data Import and Format Standardization**:

   - First I loaded the raw data into Power Query in Excel for data cleaning and ensured consistent formatting for dates, numeric fields, and text.

   - Unified currency formats for monetary fields (e.g., loan amounts, repayments).

2. **Handling Missing Values**:

   - Replaced missing values in critical fields (e.g., credit scores) using median imputation.

   - Excluded records with incomplete loan purpose or region data.

3. **Duplicate Removal**:

   - Identified and removed duplicate entries based on unique identifiers (e.g., loan ID).

4. **New Column**:

   - Added new columns for Approval Status (Approved or Rejected)

5. **Categorical Data Standardization**:

   - Standardized state abbreviations for regional analysis.

6. Excel Dashboard Features

The interactive dashboard was designed to provide a seamless user experience, enabling TDI's management to explore data trends effectively.

Dashboard Layout:

1. **Overview Section**:

   - I used KPI card visuals (textbox in Excel) for Total loans disbursed and repayments.

   - I used a donut chart to visualize the Percentage of loans charged off (default rate).

   - To visualize the top loan purposes by value, I used a horizontal bar chart.

2. **Loan Trends Analysis**:

   - Monthly loan request trends was visualized using a line chart.

   - Seasonal spike identification.

3. **Borrower Demographics**:

   - Geographic distribution of loans.

   - Loan amounts segmented by borrower income and employment length – I used a histogram here.

4. **Loan Risk Insights**:

   - Default rates segmented by credit score, loan grade, and purpose.

   - Interactive slicers for filtering data by grade, region, and loan status.

Key Features:

1. **Interactive Filters**:

   - Slicers for loan purpose, region, loan grade, and status.

   - Drop-down menus for selecting specific time periods.

2. **Visual Elements**:

   - Bar charts for loan requests and approvals by region.

   - Line graphs for tracking repayment trends over time.

   - Pie charts for loan purposes.

Key Insights

The key insights from the dashboard are:

- Total sum of all loan disbursed is $436M while total repayments is $473M.
- 14% of all Loans are Charged off (Defaulted), this represents quite an amount and needs to be further investigated People take or request for loans most for Debt Consolidation Debt consolidation involves combining multiple debts into one by taking a new loan or credit card with a lower interest rate.
- Loans classified under Grade G have the highest approval rates This means that most of TDI's borrowers use the loans to service other loans, this is concerning.
- There has been a steady increase in the number of loan requests with December 2021 having the highest number of loan requests.
- Most of the borrowers are residents in California with 4,104 requests with New York coming at 2160 requests. Tennessee has the lowest loan requests at 4.
- Borrowers with more than 10 years in employment contributed the largest portion of loan requests.

Recommendations

Based on the insights from the dashboard, the follow are recommended:

1. **Address Loan Defaults (14% Charged Off)**:

   o   Introduce stricter creditworthiness checks, especially for borrowers seeking high-risk loans such as Debt Consolidation.

   o   Offer financial literacy programs to educate borrowers on managing consolidated debt effectively.

2. **Manage Dependence on Debt Consolidation Loans**:

   o   Diversify loan offerings to reduce reliance on debt consolidation as the primary loan purpose.

   o   Implement targeted marketing for other loan categories (e.g., business or education loans).

   o   Introduce financial advisory services to assist borrowers in resolving underlying financial challenges without over-reliance on loans.

3. **Optimize Loan Grading and Approval Policies**:

   o   Assess why Grade G loans have the highest approval rates and ensure grading criteria align with risk management goals.

   o   Adjust grading thresholds to improve balance between approval rates and borrower risk.

4. **Regional Expansion and Targeting**:

   o   Focus on underserved regions like Tennessee with campaigns to promote awareness and accessibility of loans.

   o   Continue leveraging high-demand areas like California and New York but evaluate regional repayment performance to tailor risk management strategies.

5. **Seasonal Demand Management**:

   o   Plan resources to manage seasonal spikes, such as the surge in loan requests in December 2021.

   o   Explore reasons for these trends and consider offering promotional interest rates or tailored products during peak periods.

Conclusion

The analysis highlights TDI's strong performance with $473M in repayments exceeding the $436M disbursed. However, the 14% default rate and over-reliance on debt consolidation loans indicate potential risks. Recommendations focus on reducing default rates, diversifying loan purposes, optimizing grading criteria, and expanding outreach to underserved regions. By implementing these strategies, TDI can enhance profitability, mitigate risks, and ensure sustainable growth in its loan portfolio.