

BANK LOAN ANALYTICS: POWER BI DASHBOARD

INTRODUCTION

PROJECT OVERVIEW

- **Introduction to the Project:** This project aims to develop an interactive loan performance dashboard using Power BI to analyse and visualize loan sales data for a bank.
- **Description of the Dashboard:** The dashboard will provide insights into loan sales, regional performance, KPIs and sales trends over time.
- **Importance of Data Visualization:** Utilizing Power BI for data visualization enables stakeholders to gain actionable insights from the loan sales data, aiding in strategic decision making and performance monitoring.
- **Tools and Technologies Used:** The project utilizes Power BI and the most important Power BI components such as *Power Query*, *DAX*, *Model view*, etc. for dashboard creation.
- **Timeline and Scope:** The project will cover loan sales data for the years 2014-17, focusing on key metrics and trends.

GOALS AND OBJECTIVES

- **Analytical Goals:** The primary objective is to analyse loan sales performance across loan categories, regions, branch types and customer segments to identify growth opportunities and areas for improvement.
- **Key Metrics and KPIs:** Key metrics such as total loan sales, loan disbursement trend, Region wise target, and branch wise disbursement volumes will be visualized to track performance against targets.
- **Dashboard Features:** The dashboard allows sales managers to drill down into specific regions and loan categories, view quarterly, monthly and annual sales trend, and identify top performing regions and branches.
- **Data Exploration Goals:** Exploring historical sales data will help uncover seasonal trends, customer borrowing behaviour, and the effectiveness of marketing campaigns.
- **Stakeholder Engagement:** Engaging with sales teams and executives to understand their data needs ensures that the dashboard aligns with their strategic objectives and supports decision-making processes.

DATA SOURCING & INTEGRATION

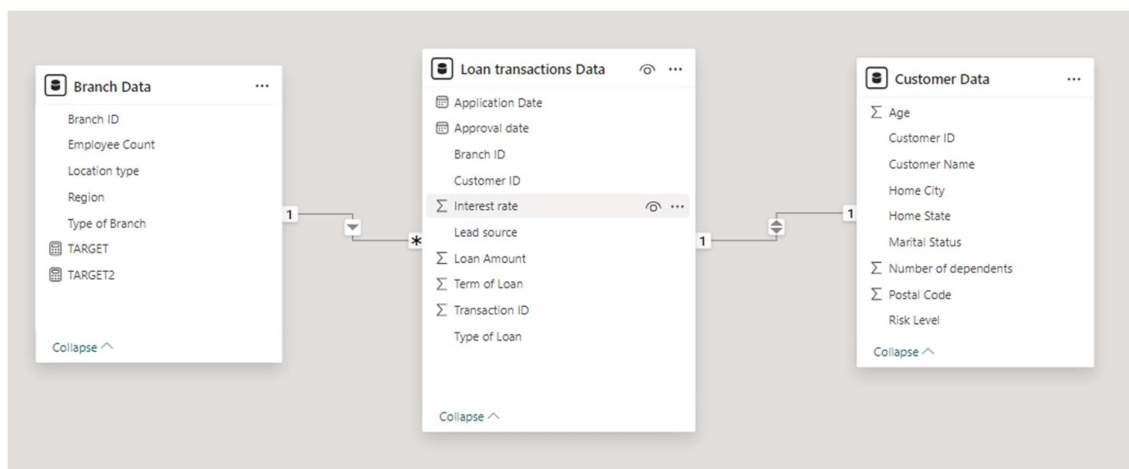
SOURCE DATASETS UTILIZED

- **Source Format:** The source datasets are in the format of *Excel Worksheets* and *CSV* files.
- **Dataset Details:** Since this project analyses the performance of a bank in terms of loan disbursement, the source datasets include the following

- *Loan Data*: Contains loan types and loan transaction details.
- *Branch Data*: Contains the branch details of the bank.
- *Customer Data*: Contains details about the customers of the bank.

LOADING AND INTEGRATION

- **Loading the Data from Sources**: The datasets are now loaded into Power BI one by one and the different data sources are now integrated together using Power BI's *Model View*.
- **Integrating the Source Datasets**: The *cardinality* and *modality* among the datasets are also established in this step. In this case, the loan transactions data has a *many to one* relationship with branch data and *one to one* relationship with customer data. This is due to the fact that loan transactions data has *foreign keys* Customer ID and Branch ID which are *primary keys* in branch data and customer data. A screenshot of the integrated data model is shown below.



DATA CLEANING AND PREPROCESSING

DATA TRANSFORMATION

- **Use First Row as Headers**: Using *Power Query*, efficient data transformation can be performed in Power BI. The first step in data transformation involves using *First Row as Headers*. This ensures that proper column name is assigned to the columns rather than Power BI using its own column names. Here, the Customer Data table had to be transformed using this technique.
- **Assigning Relevant Data Types**: The second step in data transformation involves type conversion. Some of the columns will not have relevant data types assigned by default. For them, we have to manually go ahead and convert to the necessary datatypes. In this case, the Approval data column was in text format by default. We had to manually convert it to Date datatype.

DATA CLEANING

- **Removing Null Values:** Null values in dataset can skew analysis results and affect data visualization. Power Query provides tools to filter out or replace null values. In this case, the null values were removed since the dataset is huge and doesn't affect it in a big way.
- **Dealing with Errors:** Errors in data, such as #N/A or #VALUE, can disrupt data processing. Power Query offers functions to handle errors gracefully, such as replacing errors with default values or custom error messages. Here, errors were observed in certain columns and they were replaced with the text 'NA'.
- **Removing Duplicate Values:** Duplicate values in dataset can lead to inaccuracies in analysis and visualization. Power Query offers features to identify and remove duplicate rows. Few duplicate entries were observed in a table and they were removed.

The screenshot shows the Power Query Editor interface. On the left, the 'Queries' pane lists 'Customer Data', 'Branch Data', and 'Loan transactions Data'. The main area displays a table with columns: Customer ID, Customer Name, Marital Status, Age, and Number of dependents. The formula bar shows the query definition: `= Table.SelectRows(#"Changed Type", each not List.IsEmpty(List.RemoveMatchingItems(Record.FieldValues`. On the right, the 'Query Settings' pane is open, showing the 'APPLIED STEPS' list. The steps listed are 'Source', 'Promoted Headers', 'Changed Type', and 'Removed Blank Rows'. The 'Removed Blank Rows' step is highlighted with a red box.

	Customer ID	Customer Name	Marital Status	Age	Number of dependents
1	CG-12520	Claire Gute	Married	67	
2	DV-13045	Darrin Van Huff	Married	31	
3	SO-20335	Sean O'Donnell	Unmarried	65	
4	BH-11710	Brosina Hoffman	Married	20	
5	AA-10480	Andrew Allen	Unmarried	50	
6	IM-15070	Irene Maddox	Married	66	
7	HP-14815	Harold Pawlan	Married	20	
8	PK-19075	Pete Kriz	Married	46	
9	AG-10270	Alejandro Grove	Married	18	

EDA AND VISUALIZATION

ANALYSIS

- **Loan Disbursement Trend:** Analysing the trend of loan disbursement by the bank over a period of time. Annual, Quarterly and Monthly trend are also identified to find the peak year, quarter and month in which the most loan amount is disbursed.
- **Customer Segmentation:** The dashboard helps in region wise customer segmentation based on marital status. This helps in identifying which type of customer to focus in each region.
- **Geographical Analysis:** Volume of loan disbursed for each region is tracked. Region wise KPIs are also calculated to track if desired target is achieved. Top 5 Cities and States that received the most loan amount is also identified.
- **Branch & Location Analysis:** Identifying which branch type disburses the most loan amount every year and also the location that disburses maximum loan.
- **Risk Analysis:** Analysing the volume of loan disbursed based on risk level of the loans.

BANK PERFORMANCE DASHBOARD



TOTAL CUSTOMERS

793



TOTAL CITIES

252



TOTAL STATES

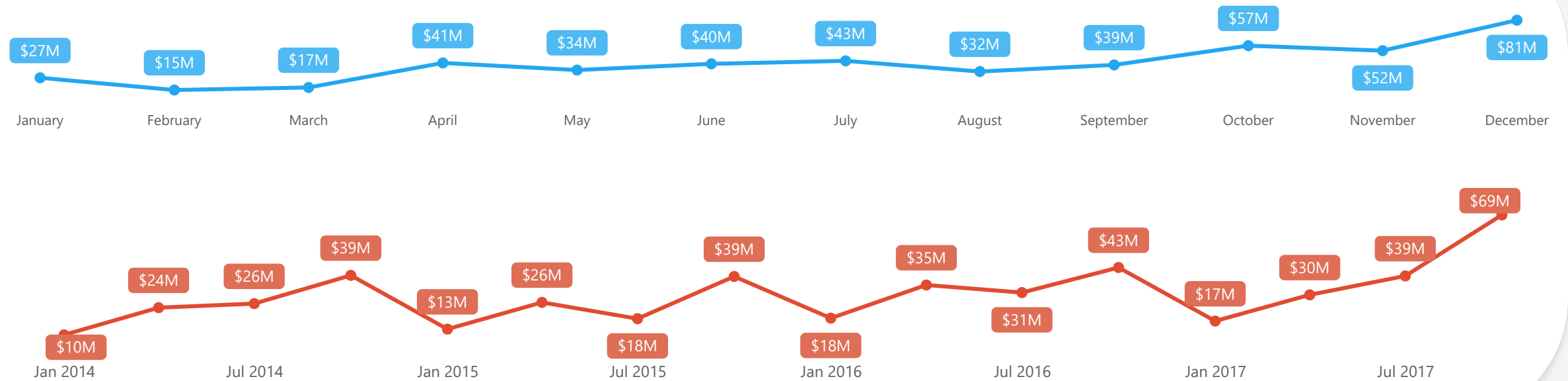
41



TOTAL BRANCHES

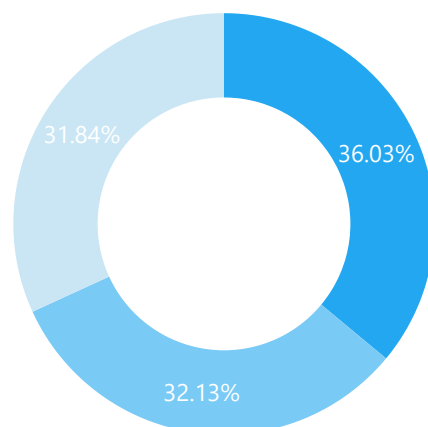
625

LOAN DISBURSMENT TREND



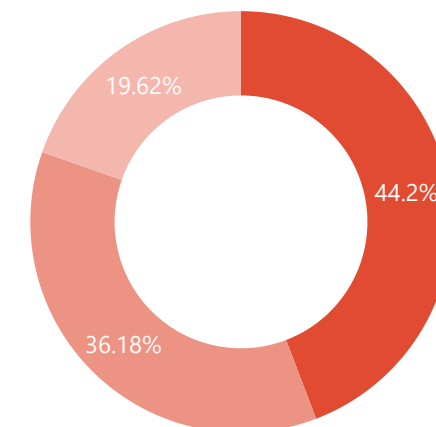
LOAN DISBURSMENT BY LOCATION

Urban Metro Rural



LOAN DISBURSMENT BY BRANCH TYPE

Traditional Digi-Smart Traditional-upgraded



REGIONWISE KPIs

\$30.72M✓

CENTRAL TARGET: \$30.00M (+2.41%)

\$51.32M!

EAST TARGET: \$55.00M (-6.69%)

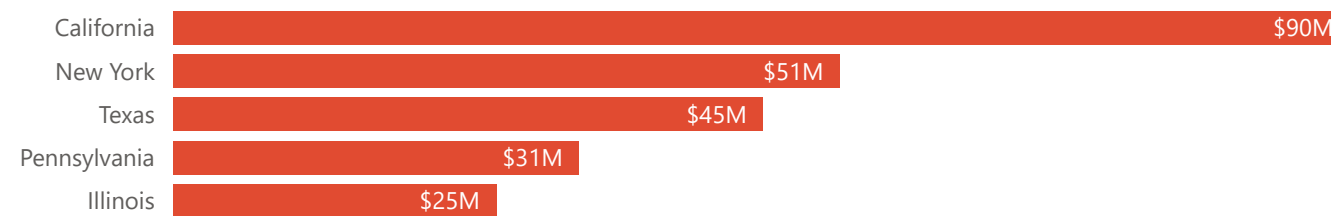
\$57.07M✓

WEST TARGET: \$55.00M (+3.77%)

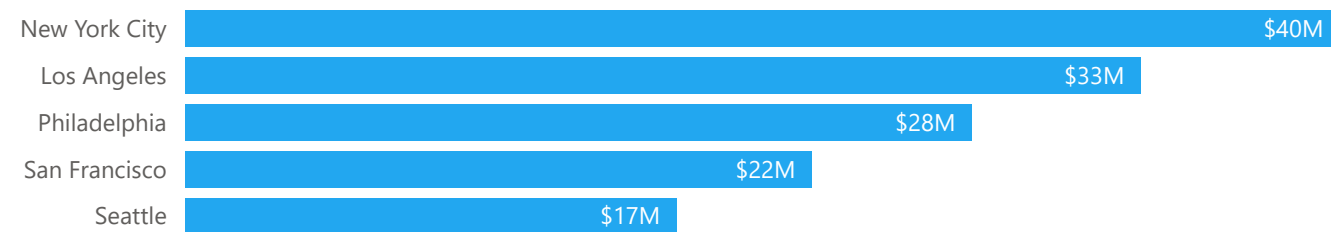
\$15.53M!

SOUTH TARGET: \$30.00M (-48.22%)

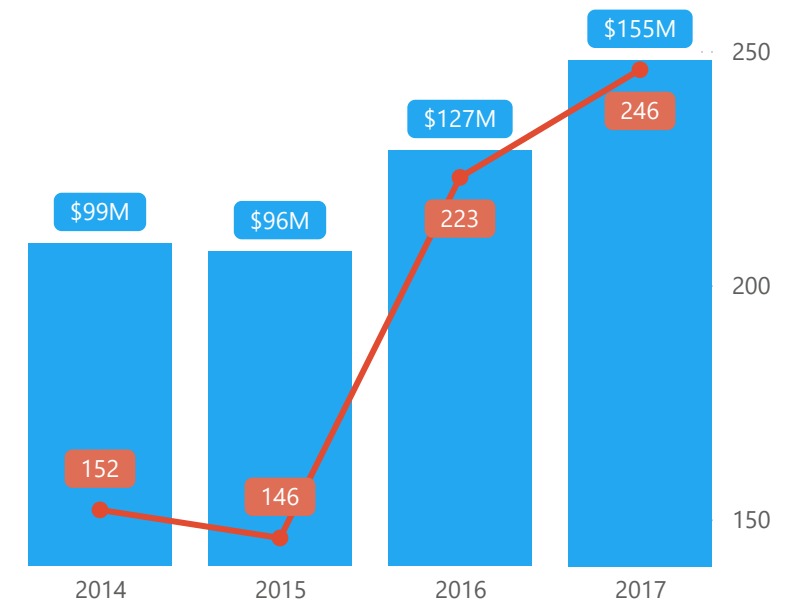
TOP 5 STATES THAT RECEIVED THE MOST LOAN AMOUNT



TOP 5 CITIES THAT RECEIVED THE MOST LOAN AMOUNT

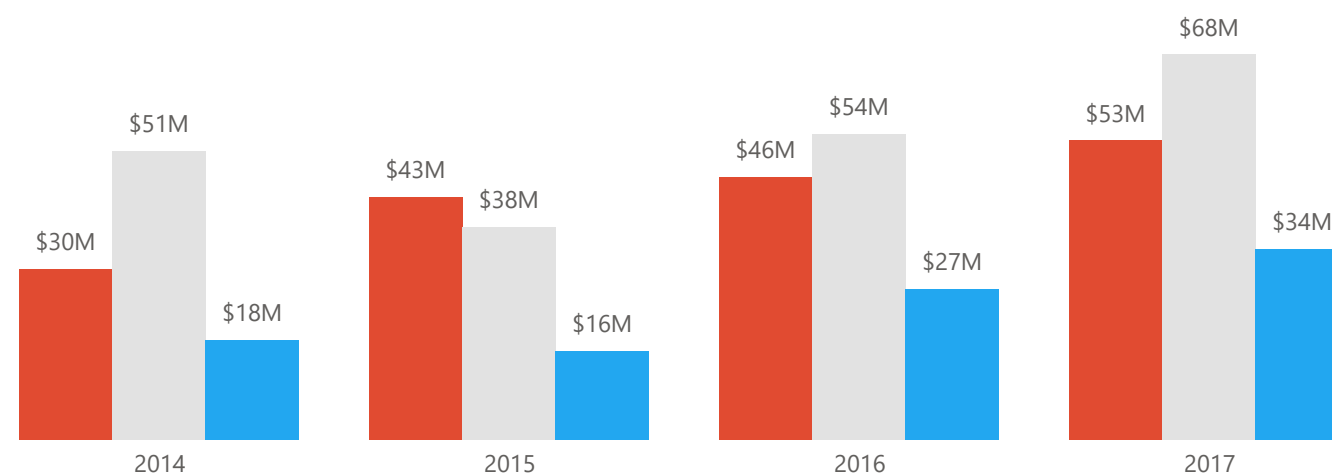


ANNUAL LOAN AMOUNT & NUMBER OF LOANS



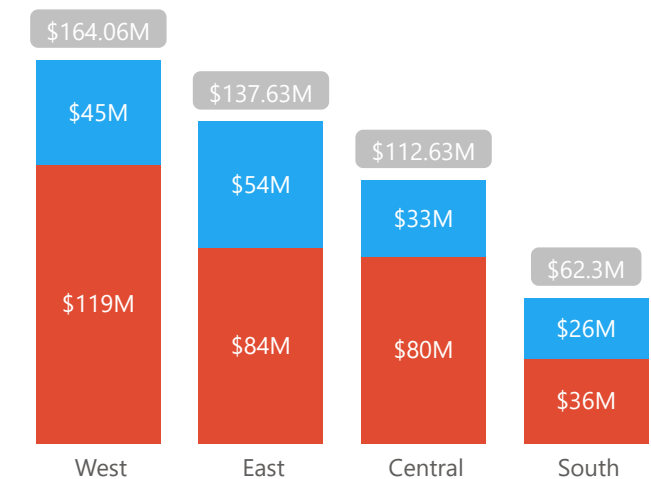
ANNUAL LOAN DISBURSMENT BY DIFFERENT BRANCH TYPES

● Digi-Smart ● Traditional ● Traditional-upgraded



REGIONWISE LOAN AMOUNT

● Married ● Unmarried



RESULTS AND CONCLUSION

FINDINGS AND KEY INSIGHTS

- **Loan Seasonality:** The highest loan amounts are typically disbursed during the *final quarter* of the year, particularly in *October, November, and December*.
- **Customer Preference for Traditional Bank Branches:** A significant *44%* of the total loan amount disbursed is attributed to customers who prefer availing loans from traditional bank branches.
- **Top Regions for Loan Volume:** *New York City* stands out as the top city, having received a remarkable loan volume totalling *\$40 million*. Additionally, *California* emerges as the leading state, securing the highest loan volume amounting to *\$90 million*.
- **Marital Status Influence on Loan Amounts:** Regardless of the region, *married* customers tend to avail the highest loan amounts compared to their unmarried counterparts.
- **Regional Loan Distribution and Performance:** The *West* region emerges as the top recipient of loans, with both the loan target and disbursement being successfully met in the West and Central regions. However, the East and South regions lag behind in meeting the loan target.
- **Equal Loan Distribution Across Locations:** Loan amounts are *distributed almost evenly* across different location types, including urban, rural, and metro areas, reflecting a balanced outreach strategy.

LIMITATIONS

- **Historical Data Availability:** Although the dataset includes data from 2014, a broader historical dataset could offer deeper insights and more accurate trend analysis.
- **Performance Challenges with Large Datasets:** When working with extensive datasets in Power BI, there's a possibility of encountering performance issues, such as prolonged loading times or reduced responsiveness. This may particularly occur with complex queries or calculations, impacting the user experience.
- **Complex Data Transformation:** While Power Query provides a diverse set of data transformation capabilities, intricate transformations may necessitate advanced scripting or custom functions beyond the scope of standard Power BI functionalities.
- **Customization Limitations in Visualizations:** Despite the range of built-in visualizations in Power BI, achieving highly specialized or customized visualizations may pose challenges due to limited customization options. For instance, certain advanced chart types, such as histograms, are not available by default.

CONCLUSION

In conclusion, the Power BI data analysis project has provided valuable insights into the bank's loan disbursement patterns, customer preferences, and regional performance. Through meticulous analysis and visualization of loan data, we have uncovered key trends and insights that can inform strategic decision-making and drive business growth.

The project has highlighted the seasonality of loan disbursements, with the final quarter of the year witnessing the highest loan amounts. Additionally, it has identified the strong preference among customers for traditional bank branches, with significant loan amounts attributed to this channel.

Furthermore, the analysis has shed light on the top regions for loan volume, with New York City and California emerging as key hubs for loan disbursement. The influence of marital status on loan amounts and the regional distribution of loans have also been examined, providing valuable insights into customer behaviour and market dynamics.

Despite these valuable insights, the project has also identified limitations such as historical data availability and performance challenges with large datasets in Power BI. However, these limitations have been mitigated through careful analysis and data transformation techniques.