**Project Documentation**

**1. Understanding the Data**

converting raw files into structured Flat tables to analyze loan application data.

* Identifying the type of data available.
* Understanding column meanings.
* Defining data types (Integer, Text, Date, etc.).
* Analyzing relationships between different tables.

**2. Wireframing & Data Flow Design**

**Introduction**

Wireframing is the process of creating a visual blueprint for dashboards before development. It helps structure data presentation, define user interactions, and ensure clarity in design.

**Key Metrics (KPIs)**

* **Total Loan Applications**: Number of loan requests received.
* **Total Loan Approvals**: Applications that have been officially approved.
* **Total Loan Rejections**: Applications that were officially rejected.
* **Average Loan Processing TAT**: Time from application submission to approval (Line Chart).
* **Average Disbursement TAT**: Time from approval to fund disbursement (Clustered Bar Chart).
* **Branch-wise Performance**: Comparison of processing time across different branches (Bar Chart).
* **Center-wise Performance**: Analysis of processing time across different centers (Line Chart).
* **Geographic Performance**: Loan disbursement analysis by state (Map Chart).
* **Loan Processing TAT**: Time taken from submission to sanction (Bar Chart).
* **Disbursement TAT**: Time taken from sanction to disbursement (Bar Chart).
* **NEFT Processing TAT**: Time taken from disbursement to transaction (Line Chart).

**3. Data Structure**

A structured data flow system was implemented by converting raw data into an organized format.:

* **Primary Key**: LoanID, used to link multiple datasets.
* **Handled Datasets**:
  + Branch Info
  + Center Info
  + Client Data
  + Disbursement Data
  + Eligible/Not Eligible Data
  + High Ticket Loans
  + NEFT Transactions
  + Sanctions Data
* **Main Tables**:
  + **Client Table**: Contains all client-related details.
  + **Loan Table**: Includes branch, state, district, center name, loan status, dates (submission, sanction, disbursement), approval status, and amount.
  + **NEFT Table**: Stores transaction details such as transaction dates and amounts.
* **Data Integrity**: Used VLOOKUP for ensuring consistency with LoanID as a unique identifier.

**4. Data Cleaning & Processing (Apache Airflow)**

A **Python script** was created for data cleaning using **Apache Airflow DAGs**. Key operations:

* Dropping unnecessary columns.
* Handling missing values using **fillna()**.
* Cleaning string data by removing unwanted characters.

**Libraries Used:** Pandas, NumPy

**5. MSSQL Setup & Connection**

Steps to establish a connection between **MSSQL and Apache Airflow**:

* Installed MSSQL and created a **database schema**.
* Developed a **Python script with DAGs** in **VS Code**.
* Configured Airflow connection through the **Admin Panel**.
* Successfully loaded and verified data in MSSQL.

**6. Power BI Integration with MSSQL**

For **data visualization**, MSSQL was connected to **Power BI**:

* Installed MSSQL and configured **SQL Server Network Settings**.
* Enabled **TCP/IP** and restarted SQL Server.
* Established a connection using:
  + Server Name & Port
  + Database Name
  + Username & IP Address
  + Password
* Imported data into **Power BI**
* Implemented designed charts based on wireframes.

**Final Data Pipeline Integration:**

1. **Data Collection** → Load data into **Power BI**
2. **Data Cleaning** (Apache Airflow, Python)
3. **Data Transformation** → Convert into **Flat Tables**
4. **Load Data into MSSQL**
5. **Charts & Graphs** → Dashboard Design

**Dashboard Documentation**

## ****1. Overview****

The "Loan Processing TAT Analysis" dashboard provides insights into the turnaround time (TAT) for different stages of loan processing. It helps track loan submissions, sanctions, disbursements, and transaction completion times across multiple branches and states.

## ****2. Dashboard Components****

### ****2.1 Summary Metrics (Top Section)****

* **Total Loan Submitted:** Displays the total number of loan applications submitted ( 31K loans processed).
* **Average Sanction Time:** Average time taken for loan approval after submission.
* **Average Disbursement Time:** Average time taken to disburse a loan post-sanction.
* **Average Transaction Time:** Average time from loan disbursement to transaction completion.
* **Filters:** Dropdowns allow filtering by different stages (Submission, Sanction, Disbursement, Transaction).

### ****2 Performance Metrics (Branch-Level Analysis)****

**Fastest & Slowest Branch Loan Sanctions (Bar Charts)**

* **Top 5 Fastest Branches:** Lists branches with the quickest loan approval times.
* **Top 5 Slowest Branches:** Highlights branches with delays in loan approval.

**Disbursement and NEFT Performance**

* **Fastest Disbursement Branch:** Shows branches where loans are disbursed quickly.
* **Fastest NEFT TAT Branch:** Displays branches with the quickest NEFT transactions.
* **Slowest NEFT TAT Branch:** Highlights branches where NEFT processing is the slowest.

### ****3 Loan Processing Time Breakdown****

**Loan ID Sanction by TAT Category (Pie Chart)**

* Categorizes loan sanctions based on processing time:
  + **0-5 Days** (Majority, 98.15%)
  + **6-10 Days**
  + **11-20 Days**
  + **More than 20 Days**

**Loan Amount vs Processing Time (Bar Chart)**

* Compares the average time taken based on loan amount categories ( 0-50K, 50K-70K, etc.).

### ****4 State-Wise Loan Analysis****

***Total Active Loans by State (Map Visualization)***

* Displays the number of active loans geographically across states.

***Average Loan Sanction and Transaction by State***

* Compares sanction and transaction times across states like Uttar Pradesh, Odisha, Madhya Pradesh, Rajasthan, and Gujarat