**Credit Default Risk Analytics Project Documentation**

**Executive Summary**

This document details the creation of a Power BI dashboard aimed at deeply understanding credit default risk within our customer accounts. The project's main goals were to pinpoint what causes defaults, keep track of the health of our credit portfolio, and build data-driven plans to better manage risk. Through a careful process involving SQL for data setup, Excel for initial checks, and Power BI for the dashboard, we found important patterns in customer demographics and behaviour that are linked to defaulting. The dashboard offers a live look at risk trends and customer groups, ending with clear suggestions to lower future default rates and make our credit operations work better.

**Project Objective**

To design a dynamic system for reporting key performance numbers that will:

* Keep an eye on how customers pay back their loans.
* Find customers who are at high risk based on how late they are with payments and how likely they are to default.
* Analyse how billing and payment amounts change over a six-month period.
* Sort customer behaviour based on details like their age, gender, marital status, and education.
* Help decision-makers spot potential credit risks and understand the overall health of our credit accounts.

**Problem Statement (For Reporting & Business Analysis)**

To effectively reduce and understand what truly drives credit default risk and its financial effects, this project aims to look closely at customer repayment habits and how they use their credit, using credit card data. The goal is to build a strong dashboard for reporting key numbers. This dashboard will find the main reasons for default risk, payment delays, and trends in bills and payments across different customer groups (like age, gender, education, and marital status). Ultimately, this will give us clear, practical information to help us better assess risk and make improved decisions in managing credit.

**1. Project Overview**

**1.1. Solution Overview: The Credit Default Risk Analytics Dashboard**

The main outcome of this project is an interactive Power BI dashboard. This dashboard acts as a central tool for analysis, showing important performance numbers (KPIs), past trends, and detailed breakdowns of default rates across different customer groups. It lets users actively filter and look into the data, turning raw information into useful knowledge.

**2. Data & Methodology**

**2.1. Data Source(s)**

Our analysis used a dataset that contained 30,000 customer records. This data included various details about customers, their past payment history, bill amounts, payment amounts, and credit limit information.

**2.2. Data Acquisition & Pre-processing (SQL)**

The first part of the project involved preparing the data carefully to make sure it was good quality and ready for analysis.

* **Extraction:** We pulled the original data from its source using specific SQL commands.
* **Cleaning & Transformation:** We did a lot of cleaning, changing, and organizing of the raw data directly in SQL. This made sure the information was consistent and didn't have common data problems.
* **Output:** The result was a clean, changed, and ready-to-use Cleaned\_Data table in the SQL database. This table was the main starting point for the next steps in our analysis.

**2.3. Feature Engineering & Derivations**

To make our analysis richer and find deeper insights, we created several important new pieces of information (features) and added them to the Cleaned\_Data table during the SQL phase. These new details are essential for sorting customers and understanding their risk:

* repayment\_category: This sorted customer payment habits (like Full Payer, Partial Payer, Zero Payer).
* utilization\_category: This grouped how much credit customers used (like Low, Mid, High, Over Utilizer).
* max\_delay\_segment: This identified the worst payment delay a customer had experienced.
* delay\_trend\_0\_2, delay\_trend\_2\_3: These measured how a customer's recent payment habits were changing (getting better, staying the same, or getting worse) over 0-2 and 2-3 month periods.
* cleaned\_sex, education\_segment, marriage: These were cleaned and sorted demographic details.
* utilization\_ratio, repayment\_ratio: These were calculated to give more detailed financial insights.
* **Monthly Mapping:** We thought of the PAY\_X, BILL\_AMT\_X, and PAY\_AMT\_X columns as representing specific months (April to September). This formed the basis for looking at trends over time.

**2.4. Data Profiling & Initial Insights (Excel Pivot Tables)**

After preparing the data in SQL, we brought the Cleaned\_Data into Excel for an initial, detailed look.

* **Comparing Information:** We used Excel Pivot Tables a lot to compare different customer traits with whether they defaulted (default\_payment\_next\_month).
* **Using Percentages:** We carefully applied percentages (like % of Row Total, % of Column Total) to make sure we could accurately compare different groups, no matter their size. We also kept a close watch on the number of people in each sample to confirm our findings were reliable.
* **Early Discoveries:** This step helped us quickly spot the first key groups of customers, understand their habits (payment, credit use), and see how past payment delays affected their risk of defaulting. This gave us a solid foundation before we started building the dashboard.

**3. Dashboard Development & Key Metrics**

**3.1. Power BI Data Model**

The Cleaned\_Data table, which now included all the new information we created, was brought directly from SQL into Power BI Desktop. We then set up the data model, making sure the connections between tables were correct (if we created any extra tables in Power BI, like one for monthly dates) to allow for accurate calculations and filtering.

**3.2. Key Performance Indicators (KPIs) & Definitions**

The dashboard clearly shows the following important performance numbers (KPIs), which are calculated using DAX in Power BI. These give a quick overview of how healthy our credit accounts are:

|  |  |  |
| --- | --- | --- |
| **KPI** | **Definition** | **Power BI Visual** |
| **Total Customers** | The total count of unique customer accounts in the dataset. | Card |
| **Overall Portfolio Default Rate** | The overall percentage of customers who defaulted in the next month (default\_payment\_next\_month = 1) across the entire portfolio. | Card |
| **Average Portfolio Credit Utilization** | The weighted average utilization, calculated as (Sum of BILL\_AMT1 for all active accounts / Sum of LIMIT\_BAL for all active accounts) \* 100. Accounts are considered 'active' if their LIMIT\_BAL is greater than 0. | Card |
| **Total Loan Amount** | The sum of credit limits (LIMIT\_BAL) extended to all customers in the portfolio, representing the total credit exposure. | Card |
| **Monthly Default Rate** | The percentage of customers who defaulted in a specific month (defined as PAY\_X >= 2 for that month), relative to the total eligible accounts for that month. This number is vital for looking at trends. | Line Chart |
| **Monthly Bill & Pay Amounts** | The total bill amounts (BILL\_AMT\_X) and total payment amounts (PAY\_AMT\_X) for each respective month. This shows how spending and repayment volumes change over time. | Line Chart |
| **Default Rate by Demographic** | The default rate broken down by various demographic categories (Gender, Age Group, Education Segment, Marriage Status). | Dynamic Bar Chart (Horizontal) |
| **Default Rate by Behavioral Factor** | The default rate broken down by various behavioral categories (Max Delay Segment, Delay Trend, Utilization Category, Repayment Category). | Dynamic Column Chart |
| **Customer Distribution** | The proportion of total customers in each category for Marriage Status and Age Group. | Pie/Donut Charts |

**3.3. Dashboard Structure & Visual Components**

The dashboard is designed with a clear, easy-to-understand layout to help the user follow the analysis:

* **Header:** A clear title and branding elements at the top.
* **Top KPIs:** Four distinct card visuals that give an immediate summary of key numbers for our credit accounts.
* **Left Sidebar (Global Filters):** A dedicated panel on the left side holding interactive filters for Gender, Marriage Status, Age Group, Education Segment, and Credit Limit (a filter that lets you pick a range). These filters let you narrow down all the information on the dashboard.
* **Main Content Area (Right Side):** This area is organized into logical sections:
  + **Trend Analysis:** Shows line charts for Monthly Default Rate and Monthly Bill & Pay Amounts, displaying how performance changes over time.
  + **Demographic Insights:** Includes a filter that changes what you see (a Field Parameter) for picking demographic categories. This filter controls a horizontal bar chart that shows default rates for the chosen group. This section also has pie/donut charts showing how customers are spread out by Marriage Status and Age Group.
  + **Behavioral Insights:** Similar to demographics, a filter that changes what you see (Field Parameter) lets you pick behavioral factors. This controls a column chart that shows default rates based on Max Delay Segment, Delay Trend, Utilization Category, and Repayment Category.
  + **Key Findings & Recommendations:** A dedicated text box at the bottom-right that briefly summarizes the most important discoveries and suggests practical steps.

**3.4. Interactivity & User Experience**

The dashboard is highly interactive, using Power BI's features to make it easy and engaging to use:

* **Dynamic Filters:** These allow users to switch between different ways of looking at the data without making the dashboard cluttered.
* **Cross-Filtering:** Clicking on any part of a chart (like a bar or a slice) automatically filters all other related charts on the dashboard. This lets users explore the data in detail.
* **Clear Look:** Consistent colors, fonts, and chart styles make sure everything is easy to read and looks professional.

**4. Key Findings & Strategic Recommendations**

Based on our thorough look at what causes credit defaults, we found the following main points. These lead to practical suggestions for better managing risk:

**4.1. Overall Portfolio Health**

* The overall default rate is **22.12%**. This tells us the current total risk level of our credit accounts.
* Average credit utilization is **30.58%**. This suggests a moderate level of credit usage across the portfolio.
* The total amount of credit we've given out is **5.02 billion**. This highlights how much money is potentially at risk.

**4.2. Trend Insights**

* **Monthly Default Rate:** The monthly default rate dropped from April (10.43%) to May (14.70%), then showed a consistent increase until August (9.89%), before dropping again in September (10.26%). This indicates fluctuating risk levels over the six-month period.
* **Bill & Pay Amounts:** The total bill amount trend shows a consistent decrease from April to September. In contrast, the total payment amount graph appears relatively stable, without significant spikes or degradation, remaining quite a straight line. This suggests a potential for growing outstanding debt as bill amounts decline while payments remain flat.

**4.3. Demographic Risk Profiles**

* **Higher Default Rates:** Our analysis consistently shows higher default rates among certain groups: **Males (24.17%), High School educated customers (25.16%), those in the 50+ Age Group (who consistently have the highest rate), and Married customers (23.47%).**
* **Customer Distribution:** Even though some groups have higher default rates, their size within our total customer base [e.g., 'varied a lot; for example, the 30-39 age group is the biggest part of our customers, even though their default rate is only moderate'].

**4.4. Behavioural Risk Drivers**

* **Delay Severity is Key:** The most important factor predicting default is how severe a payment delay is. The risk of defaulting goes up sharply from Minor Delay (~25%) to Severe Delay (~46%).
* **Worsening Trends are Early Warnings:** Customers whose payment habits are getting worse ('Worsening Delay Trend' or Positive) show a clear increase in default risk. Even accounts that seem to be improving ('Negative Trend') can still have a surprisingly high default risk (66-75%), suggesting that past serious payment problems can have a lasting effect.
* **Credit Use Patterns:** Using a lot of credit or going over the limit are strong signs of financial difficulty and higher default risk, especially for customers in the 50+ Age Group and those with Severe Delays. We also noticed that even customers who use very little credit aren't completely safe from minor delays.

**4.5. Strategic Recommendations**

Based on the comprehensive analysis of credit default drivers within the provided dataset, the following actionable recommendations are proposed. These strategies aim to enhance risk management and potentially reduce future default rates, using data-driven insights to inform targeted initiatives.

* **Improve Credit Checks and Onboarding for High-Risk Demographic Groups:**
  + **Insight:** Our analysis showed that Males have a higher default rate (24.17%) than Females (20.78%). We also found higher risk among High School educated customers (25.16% default rate), the 50+ Age Group (consistently the highest default rate), and Married customers (23.47% default rate).
  + **Specific Actions:**
    - **Stricter Credit Assessment:** For new applicants in these high-risk groups (Males, High School educated, 50+ age group, Married), we suggest using tougher credit scoring rules. This might mean requiring higher minimum credit scores or asking for more financial documents during the application process.
    - **Adjusted Credit Limits:** For new approvals in these groups, we should consider starting with lower credit limits. This helps manage early risk and potential exposure effectively.
    - **Tailored Welcome Messages:** We should create special welcome materials. These materials should focus on responsible credit use, the importance of paying on time, and the dangers of using too much credit. These should be specifically designed for and sent to these higher-risk demographic groups.
* **Set Up a Tiered, Proactive Collections and Support Plan Based on Payment Delays:**
  + **Insight:** Our analysis clearly showed that how severe a payment delay is (MAX\_Delay\_Segment) is a key sign of future default (No Delay: ~11.71%; Minor Delay: ~24.99%; Severe Delay: ~46.30%). Changes in payment habits were also critical: payments getting worse ('Positive Trend') greatly increase risk, while stable payments ('No Change') show the lowest default rates. It's important to note that even when payments improve ('Negative Trend'), the default risk can still be surprisingly high (66-75%), suggesting that past serious payment problems have a lasting effect.
  + **Specific Actions:**
    - **Early Action for Worsening Trends:** If customers' payment habits are getting worse (e.g., +1, +2 delay), we should immediately send automated, personalized messages (like SMS reminders or email alerts) to stop the problem from getting worse.
    - **Targeted Reminders for Minor Delays:** For customers with Minor Delays (especially those who often pay in full), we recommend a softer, automated reminder strategy (e.g., SMS, email) to encourage full payment. This helps avoid more aggressive collection methods.
    - **Intensive Collections and Help for Severe Delays:** Customers with Severe Delays, especially those making only partial payments (93.27%) or no payments (Zero Payers), should be our top priority for strong collection efforts. For these high-risk accounts, it's crucial to offer Hardship Programs, Account Restructuring, or Payment Relief options (like temporary payment breaks, lower interest rates, longer loan terms, or re-aging accounts under strict rules) to prevent complete losses and help customers recover if possible.
* **Improve Credit Limit Management and Watch Credit Use:**
  + **Insight:** Customers who use a lot of credit or go over their limit (High Utilizers and Over Utilizers), especially in the 50+ Age Group and those with Severe Delays, were found to be key risk indicators. We also saw that even customers who use very little credit can still have Minor Delays, and their defaulting needs specific attention.
  + **Specific Actions:**
    - **Proactive Credit Use Alerts:** We suggest setting up automatic alerts for customers whose credit use gets close to their limit (e.g., 80% or 90%). These alerts should offer financial advice or suggest ways to manage spending before they max out.
    - **Changing Credit Limits:** For customers who use a lot of credit and also show worsening delay trends or make partial/no payments, we should consider temporarily freezing or lowering their credit limit. This helps stop them from getting into more debt and reduces our risk.
    - **Investigate Defaults in Low Utilizers:** For customers who use little credit but still default or have Minor Delays, we should look deeper into why this is happening. This might involve talking to them to understand if it's due to sudden income problems, fraud, or other reasons not clear from just looking at the credit use.
* **Create Specific Financial Wellness and Support Programs for Different Groups:**
  + **Insight:** We found that specific groups, including the 50+ Age Group, High School educated individuals, and Males, consistently carried higher risk. We also saw different patterns for Single and Married customers regarding how severe their delays were.
  + **Specific Actions:**
    - **Tailored Financial Education:** We recommend creating simple financial education materials that focus on managing credit responsibly, budgeting, and what happens if payments are missed. This information should be easy to get for High School educated customers and the 50+ age group.
    - **Age-Specific Support:** For the 50+ age group, we suggest looking further into their unique financial challenges (like changes in retirement income or medical bills) through surveys or discussions. We could then create special support programs or products, possibly including specific Hardship Programs designed just for this group.
    - **Targeted Counselling:** We recommend offering debt management advice or financial planning help to Single customers with Severe Delays and Married individuals who are struggling with Partial Payments and Minor Delays.
* **Set Up Strong Monitoring and Regular Policy Reviews:**
  + **Insight:** To know if any changes we make to our plans or policies are working, we need to constantly measure and adapt.
  + **Specific Actions:**
    - **Dashboard Use:** We recommend using the dynamic KPI reporting dashboard to constantly watch the identified risk groups and payment patterns in real-time. This dashboard should be the main tool for keeping an eye on risk.
    - **Regular Policy Review:** It's crucial to regularly review our credit policies, rules for approving credit, and collection plans (for example, every three months). These reviews should use what we learn from the dashboard and how well our suggestions are working. We also suggest adding Vintage Analysis to see how groups of customers perform over a long time and how policy changes or economic shifts affect default rates.

**4.6. Illustrative Customer-Specific Intervention Strategies**

To further show how these suggestions can be put into practice for individual customers, here are some example customer profiles and the specific plans for them:

* **Profile: Customer A (High Risk, Potentially Recoverable)**
  + **Characteristics:** Male, 55 years old (50+ Age Group), University educated, Married, MAX\_Delay\_Segment: Severe Delay (e.g., 3-4 months), Repayment Category: Partial Payer, Utilization Category: High Utilizer, Delay Trend: No Change.
  + **Recommended Strategy:** For this customer, we should have a dedicated collections agent call them to understand their financial difficulties. The plan is to offer to restructure their account with a reduced monthly payment and a longer loan term. This aims to give them a manageable way to pay, prevent them from defaulting completely, and help them get back on track.
* **Profile: Customer B (Extreme Risk, Last Resort Intervention)**
  + **Characteristics:** Female, 30 years old, High School educated, Single, MAX\_Delay\_Segment: Severe Delay (e.g., 7+ months), Repayment Category: Zero Payer, Utilization Category: Over Utilizer, Delay Trend: Strong Worsening (+3, +4).
  + **Recommended Strategy:** This customer is at extremely high risk of defaulting very soon. The plan is to contact them urgently to offer a Hardship Program that might include a temporary lower interest rate or an offer to settle their debt for a lower amount. This is a last resort to get some money back and avoid a total loss. If they don't respond, we should also consider legal action.
* **Profile: Customer C (Early Warning, Proactive Support)**
  + **Characteristics:** Male, 25 years old, Graduate School educated, Single, MAX\_Delay\_Segment: Minor Delay (e.g., 1 month), Repayment Category: Zero Payer, Utilization Category: Low Utilizer, Delay Trend: Positive Trend (e.g., from -1 to 0, or 0 to 1).
  + **Recommended Strategy:** Even though this customer uses little credit, their Zero Payer status and worsening trend are important early signs. The plan is to send an immediate, personalized email or text message about the missed payment. This should be followed by a gentle phone call to understand why they missed the payment and offer basic tips on managing money or a simple payment reminder. The goal is to stop the problem from getting worse.

**5. Technical Validation & Quality Assurance**

To ensure the accuracy and reliability of all metrics and findings in the Power BI dashboard, a rigorous cross-verification process was conducted using direct SQL queries on the source Cleaned\_Data table.

**5.1. KPI Cross-Verification (SQL Queries & Rationale)**

Each KPI calculated in Power BI was independently validated against SQL aggregations. This involved:

* **Direct Aggregations:** For core KPIs (Total Customers, Overall Default Rate, Total Loan Amount), straightforward SUM() and COUNT(DISTINCT) queries were used.
* **Simulating Power BI Transformations:** For KPIs relying on Power BI's internal transformations (e.g., unpivoting PAY\_X columns for monthly trends), SQL queries were constructed using UNION ALL (or UNPIVOT where applicable) to replicate the Power BI data structure and calculation logic. This ensured that the SQL output precisely matched the Power BI visuals.

**5.2. Data Consistency Checks**

Throughout the project, consistency checks were performed to ensure that data types, definitions, and relationships were maintained across SQL, Excel, and Power BI environments.

**6. Sources**

* **Dataset:**[**https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset/data**](https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset/data)
* **SQL querries:**[**C:\Users\ASUS\OneDrive\Desktop\New folder\My Projects\Default Of Credit Card Clients Dataset\CREDIT CARD SQL QUERRIES.docx**](file:///C:\Users\ASUS\OneDrive\Desktop\New%20folder\My%20Projects\Default%20Of%20Credit%20Card%20Clients%20Dataset\CREDIT%20CARD%20SQL%20QUERRIES.docx)
* **Dashboard:**[**C:\Users\ASUS\OneDrive\Desktop\New folder\My Projects\Default Of Credit Card Clients Dataset\Credit\_card\_dashboards.pbix**](file:///C:\Users\ASUS\OneDrive\Desktop\New%20folder\My%20Projects\Default%20Of%20Credit%20Card%20Clients%20Dataset\Credit_card_dashboards.pbix)