**SAMPLE PROJECT**

**LOAN DATA ANALYSIS**

**(CRM)**

**(GOTLA SHIVANI YADAV)**

**AIM -What is the aim of analysis**

1. **Customer segmentation**

*Demographic segmentation* based on their state, employment details and their home ownership.

*Behavioural segmentation* based on the loan purpose, term and credit grade Financial *segmentation* based on income, the loan amount taken and payed back, credit score assigned after a loan instalments payment starts, loan interest rates.

1. **Credit risk modelling**

This can be done with the payment dates and the customers delays towards the payments and the rate of defaulters based on their financial and employment status and also their credit scoring positions as per subgrade analysis to analyse the distribution of defaults across different sub grades to refine risk assessment criteria.

1. **Customer Lifetime Value (CLV)**

*CLV Calculation:* the net present value of expected future payments from each customer.

*Profitability Analysis*: Identify the most profitable customer segments and focus retention efforts on them.

1. **Loan performance analysis**

Defaulters rate can be analysed with the given loan status and also determine the loans which is good or bad based on customer payback.

1. **Customer analysis for retention and predictive future**

Identify customer who contribute towards company revenue generation and extending upsell opportunities to maximize profitability while minimizing default risk.

**FEATURES**

The data used for analysis is loan data from website called karggle.com. the data contains different components that are:

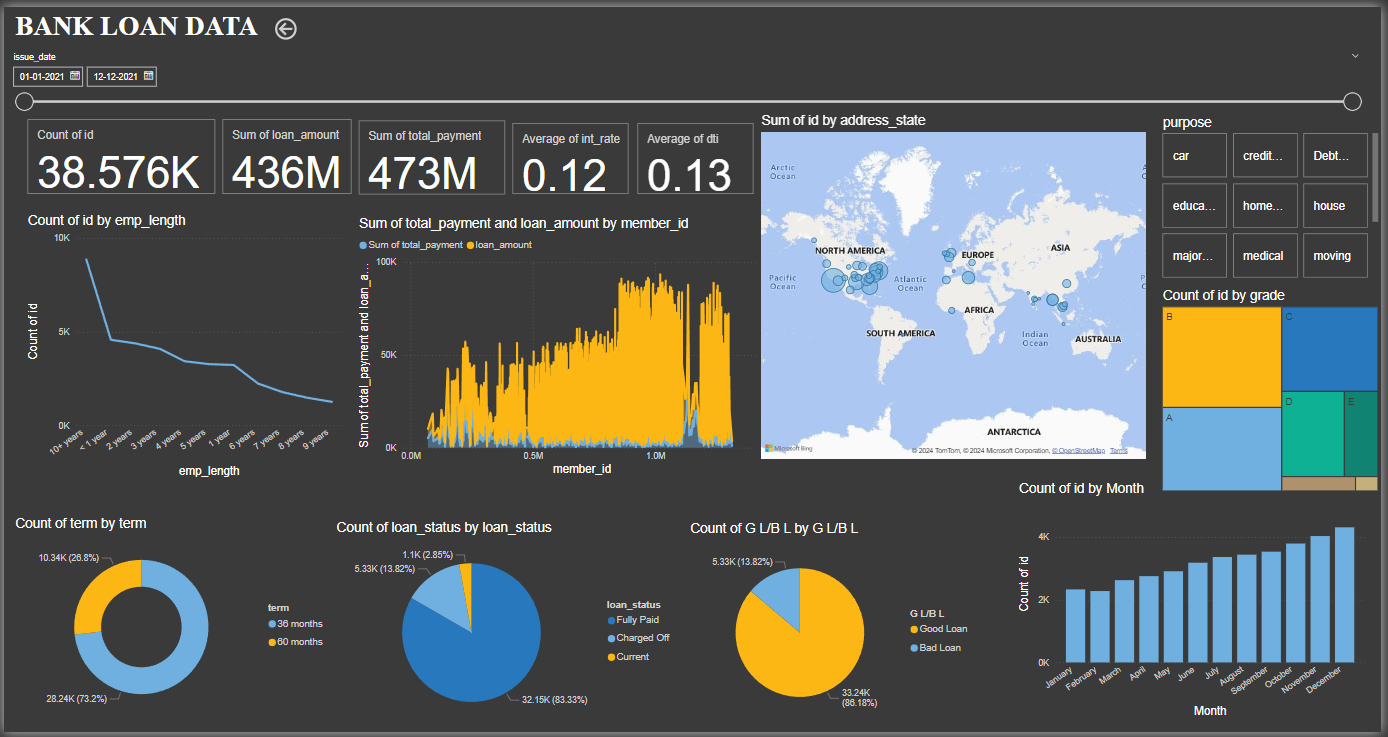
* **Id** - determines a unique identifier number for the loan application placed
* **Address state** - determines the state or country to which they belong to
* **Application type** - determines the loan application type whether it is an individual or group of customers taken
* **Emp-length** - determines the years on to how many years an employee is working in a particular organisation (1 year ,2years etc)
* **Emp-title** - determines the profession or an organisation a customer is employed
* **Grade** - determines different credit score category at the time on application based on their past credit payable pattern (e.g.: A determines highly eligible for loan and repayment score is high low risk, B refers a little bit less credit worthy than A can take a as best and b is better. So as)
* **Homeownership** - wheatear the customer is rental or owner of the house.
* **Issue date** -it determines the data at which the loan was issued to the customer
* **Last-credit-pull-date** - it is the date at which the new credit score was given based on the given loan repayment criteria
* **Last-payment-date** -it is the date at which the last payment was made
* **Loan status -** it determines whether the loan is fully paid(paid with the interest rate),charged off(not in apposition to pay so the lender stops attempting to collect the debt through regular means and may sell the debt to a collection agency) or current(all scheduled payments have been made on time, and there are no overdue amount).
* **G L/B L** - it determines whether the given loan is good or bad
* **Next-payment-date** - it is the date give to the customer for the next payment
* **Member -** it is a unique id that was given to the customer after the loan given
* **Purpose -** it determines for what purpose the loan is taken such as car, credit card, debit consolidation, wedding, medical etc.
* **Subgrade**- based on the credit payment pattern the customer are divided into sub categories of grade such as A-A1, A2, A3, A4, A5, B-B1, B2, B3, B4 and so on
* **Term**- determines the duration of loan taken that is 36 months or 60 months
* **Verification status**- it determines which status of loan whether it is verified or not
* **Annual income** -it determines the income earned by the employee annually
* **DTI –** Debt-to-Income (DTI) ratio is a key measure used by lenders to assess a borrower's ability to manage monthly payments and repay debts. It compares the borrower's total monthly debt payments to their gross monthly income
* **Instalments-** number of instalments given for payment
* **Int\_rate** -interest rate charged on the customer based on the credit score and the amount taken
* **Loan amount-** loan amount funded
* **Total\_acc-**total number of accounts that a customer holds on
* **Total payment –** total amount paid back

**METHODOLOGY**

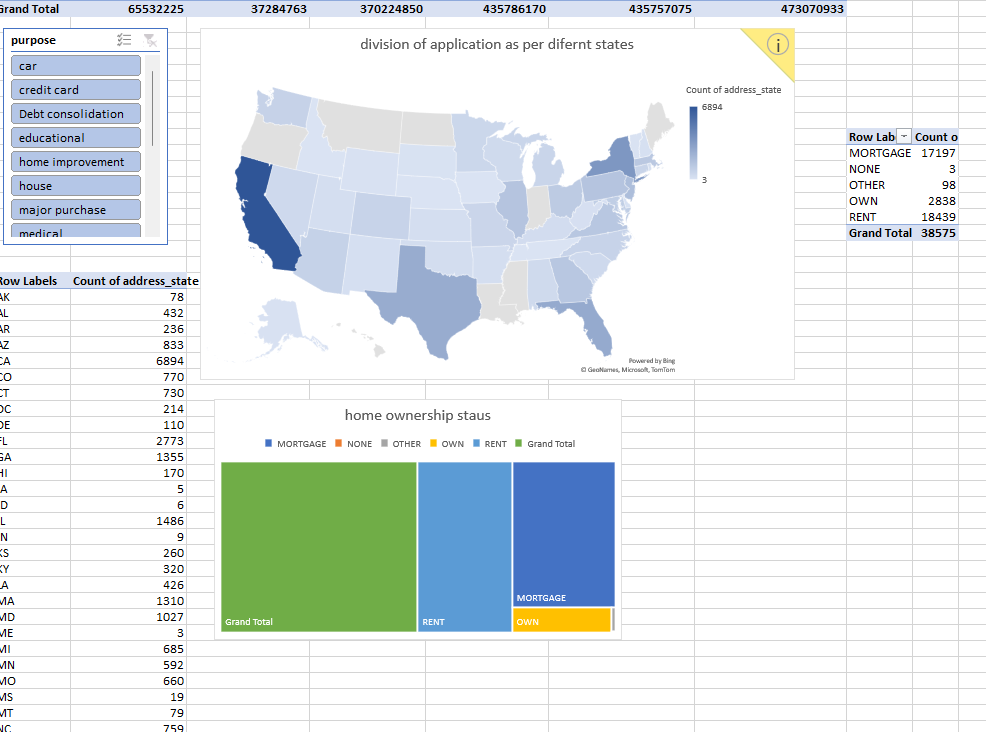
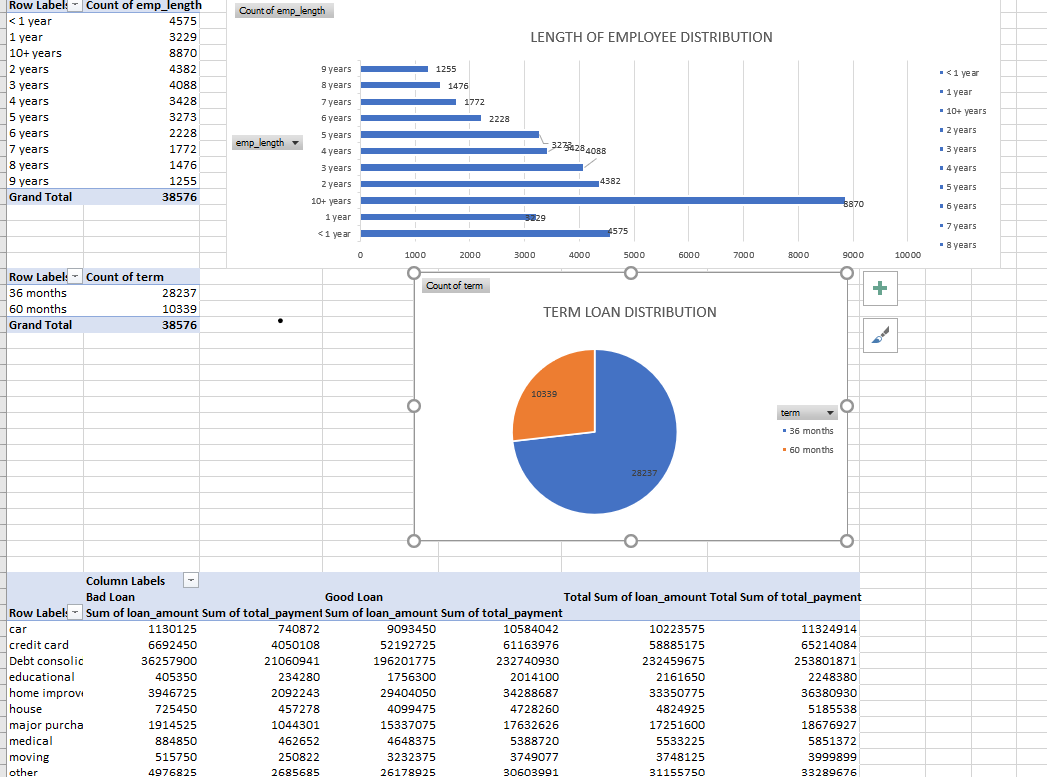
1. **Data collection** -loan data is collected through an online site called KARGGLE .COM regarding various details related to customers, loan amounts, interest etc.
2. **Descriptive analysis** – is done by using the transformed data in excel and analysis is done using formulas like – sum, average, count, count if to find amount funded, amount received, average interest rate and data rate, defaulters’ rate, how many loans give are considered as good and bad
3. **Visualization** -used different charts in excel to show visualizations like bar graphs, pie charts, box plotter and map and also used pivot tables and slicer to look values according to the criteria required
4. **Segmentation Analysis, Predictive Modeling Customer Lifetime Value (CLV) Analysis and other patterns analysis** – is done Using power bi which showed overview of customer analysis dashboard, loan amount analysis dashboard and summary dash board.

**OUTPUT**

SUMMARY DASHBAORD (SHEET 1 OF POWERB1)



EXCEL VISUVALIZATIONS

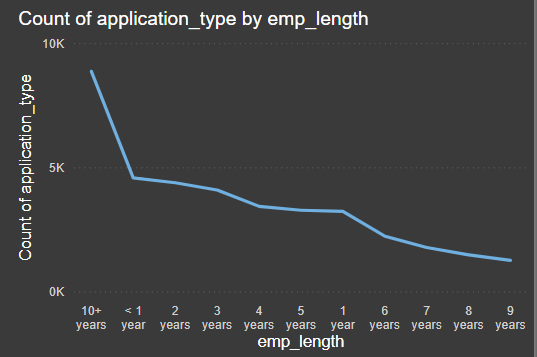


**RESULT ANALYSIS**

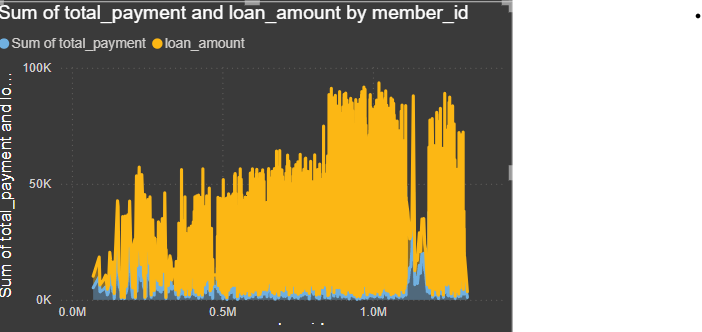
* THE dash board starts of with the heading of bank loan and the dates. these dates range starts from 1 Jan 2021 to 12 dec 2021 determining the applications starting date and the end date.
* Next is the total number of applications received -38.576k
* Total loan amount funded is 436M determining the bank's lending capacity and risk exposure.
* Total loan amount received is 473M. This metric can be used to identify high-value customers and potential churn risks.
* Average interest rate is almost 12% determining the bank's pricing strategy and competitiveness and also useful for identifying potential pricing optimization opportunities.
* Average DTI rate is almost 13% so Debt-to-Income (DTI) ratio is a key measure used by lenders to assess a borrower's ability to manage monthly payments and repay debts. It compares the borrower's total monthly debt payments to their gross monthly income.
* This determines various applications and from which geographical area people are prone to take loans and future marketing can be done to attract more customers in that regions.



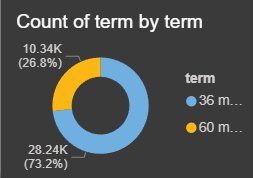
* Count of application as per their length of employment it is observed that people of term less than 5 years and more than 10 years in their professional life are prone to take loan so at the time of advertisement we can focus more on those group of customers and retain customers.



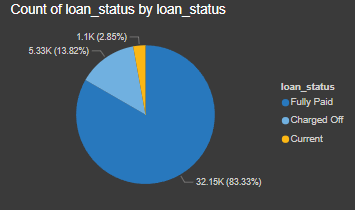
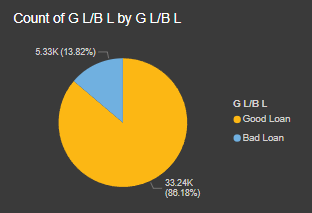
* Determining sum of total amount taken and paid back, the stands of both differ a lot this is because the data here is 1 year data and is the first year of loan application so its starting period the amount collected from the customer is very low expected to increase every year till the term ends and occupied with good number of customers for next 3 years as the loan terms are 36 months or 60 months.



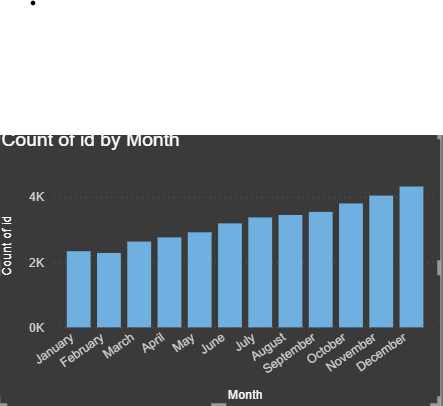
* It determines the count of term 36 months is more obtained by customer than term of 60 months, this identifies the customer behaviour towards the loan repayment faster and finish off rather than extending to more period.



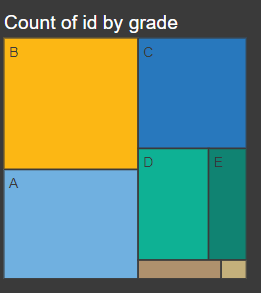
* Determining the default rate, it is observed that the charged off of loans is very less in count and termed as bad loan which is 13.82% of total this Iindicates potential credit risk and customer management issues. a good customer management can be done risk can be controlled and also these fully paid and current are considered under good loan criteria and charged of is considered as bad loan and there by upsells can be extended to the good loan customers.

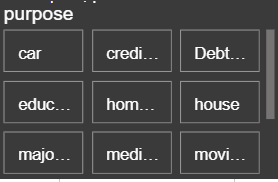
* This chat determines the stable growth in the number of customers every month so the economic growth that is increase in customer spending and borrowing, favorable interest rates, effective marketing campaigns by bank, improving customer experience and financial goals of consumers which drive them to increased borrowing. Shows loan disbursement trends over time. Can be used to identify seasonal patterns and customer behavior.



* This determines the credit score of customers like A is best, b is better, c is good so on as the alphabets goes on decreasing the credit scores decreases as per the plot determining more of a, b, c grade less risk criteria and payback is okay and can’t say loss but a little percentage of loss is there. Provides a granular view of loan risk profiles. Can be used for risk assessment and customer segmentation.



* A slicer was inserted so all the graphs will change according to the purpose. for example, if we click on car, we can get all the data related to car regarded how much loan was taken on car purchase purpose, how many applications, what is the term based on this can-do predictive analysis of customer in future.



**OVER ALL ANALYSIS**

* **Analysing future customer** based on report and data results like focusing on customers of north America as they prone to take more loans, customer whose employment length range is less than 5 years and more than 10 years so that we can advertise and **target those segments** to gain customers
* **Customer relation** should be well managed with the customers of good loan criteria so that the upsell extensions can be done for **customer life time value.**
* **Loan performance analysis** that is **defaulters’ rate** is analysed which is pretty less and can be controlled through good customer relationship management
* Implemented robust **credit scoring models** and early warning systems based on subgrade and repayment history. Continuously monitor interest rates to remain competitive and attract new customers.
* Shows **loan disbursement trends** over time and also **seasonal patterns** and **customer behavior towards loans** is analyzed.
* **Net present value** of expected future payments from each customer is analyzed.

Thank you