Regarding the monetary kpis and insights, I'm assuming the currency is in euros and the revenue is unrealized.

Task

A product manager wants to evaluate their lending product's profitability. You are provided with a table holding raw historical data for this product you would normally query from a transactional database. Please use the provided data set and your data analytics power to answer the following business questions in the form of a report or presentation.

1. What are some distinctive characteristics of this product?
2. What are some KPIs describing this product's performance and profitability?
3. What are some ways to make this product more profitable and what are the pros and cons behind them?
4. Are there any other useful data insights you would share with the product team management? → **PLEASE CHECK OUT THE INSIGHTS SECTION. ABOVE EACH VISUAL, THERE IS THE DESCRIPTION OF INSIGHT IT PROVIDES.**

Q1: What are some distinctive characteristics of this product?

1. We can **presume that this product has been launched** in the beginning of 2019.
2. The product has a moderate marketing campaign which creates demand for it, but not as significant as one would expect from a product.
3. The **most applied loan term** is 30 days followed by 90 days.
4. **Most of the requested loan amounts** are between 750 and 917 euros.
5. The product has a **consistent interest rate** across the timeline. There is not a single change, so we can say that this product has a fixed-rate interest rate which is usually found in mortgages, however, looking at the requested loan amounts, we can conclude that this is something else like **car loan**, **personal loan**, **student loan** or something else.
6. The **most issued loan amounts** are between 2380 and 2550 euros with 87.68% issuance rate.
7. The product is **mostly used** by people aged from 21 to 45.
8. Among the clients of the product, there is no correlation between their income and the loan issuance to them. Neither does client age nor loan amount.

Q2: What are some KPIs describing this product's performance and profitability?

1. The percentage of issued loans are not too good. When we look at the ***Percentage of successful and unsuccessful loan application results*** chart, we can see that we've only issued roughly ⅔ of all the loan applications. This in turn means less profit, which can be an issue in the long term.

2. ***Average requested loan term*** is the KPI which shows the average term requested for loan repayment. And it is roughly 43.5 days. So we can say that considering the most amount of requested loan term is 30 days, we should be expecting on average to be able to collect repayments in more than 30 days.
3. ***Average loan amount requested*** is the KPI which shows the average loan amount requested by the clients. In our case it's about 1085 euros.
4. ***Monthly total expected revenue with average interest rate*** visual shows us the unrealized revenue (for that reason I've named it as expected revenue) gained over months as well as the average interest rate per month.***Average Interest Rate*** as the name suggests, this KPI shows us the average interest rate we've imposed on the loan amounts.
5. ***Total Expected Revenue*** total unrealized revenue calculated by deducting the loan amount from loan due.
6. ***Average Interest Rate*** the average interest rate imposed on the loan amounts.
7. ***Total number of clients*** the total number of distinct clients.
8. ***Average success rate of loan application*** the metric shows on average what is the issuance rate of applied loan per client.
9. ***Average Delayed Repayment Time*** the average time usually it takes for a client to repay the loan after loan term.
10. ***Average Daily Fare Charged on Delayed Payments*** as there are fares on the delayed payments, this KPI shows on average how much do we charge the clients per day for each delayed day.
11. ***Total fare charged for delayed payments*** the total amount we've collected from clients as a fare on delayed payments.
12. ***Total Payments Received*** total payments made by clients.

Q3: What are some ways to make this product more profitable and what are the pros and cons behind them?

1. As a reference to 2nd point in Q1, I think there can be a better product marketing campaign to create more demand. Like decreased interest rate, buy now pay later or some more.
2. As a reference to 1st point in Q2, I think we should analyze the potential credit risk per client and choose a benchmark. Anyone who makes it above that benchmark should be granted the loan as this will increase our profitability. Make it too high, we lose clients and so profit, make it too less, we risk getting repayments back. So, there must be a proper analysis to set up a reliable benchmark.
3. As a reference to 3rd point, knowing the average requested loan amount, we can do our planning and budgeting as well as contingency budget according to that.
4. When we look at ***Percentage of loan amounts by bucket applied per loan term bucket*** visual, we can see that, for example, 50.22% of requested loan amounts which were requested with a loan term interval from 10 days to 15 days were roughly between

263 and 427 euros. This visual can be a reference point to build planning and estimation for potential loan application amounts and their correct classification.

5. As a reference to 4th point in Q2, we can say that the revenue has mostly been consistent, but in the last month it's decreased drastically. It **can't** be attributed to interest rate as we can see that it's been consistent, if not, the same. This amount is found by subtracting loan amount from loan due amount. So, we have to have a deep analysis to understand what are some other factors of this decrease. It may even be that a potential competitor has issued a similar product with better conditions for clients.
6. As a reference to 7th point in Q2, when considering that we have roughly 129000 loan applications and 83000 clients, on average each distinct client has made 1.5 loan applications. This means that there are returning clients which is another good indicator of profitability and correct customer success. However, it'll be good to also consider how are the returning customers are applying for 2nd time, as if they're making second application with increased liabilities and expenses, then it'll be a good choice to reconsider not issuing a loan to the given customer 2nd time.
7. As a reference to 10th point in Q2, we can see that the average daily fare charged on delayed payments is almost 70 euros. This can be a good action against delaying payments for clients, however, this can also demotivate a given customer to not pay the loan at all because the daily dare is huge and as they keep not paying back, they'll not want to pay anything at all. So we need to set a break-even point in here.
8. As a reference to 11th point in Q2, we can see that we've received roughly 5% of total payments from delayed payments fares. Keeping this consistent, we'll make more profits in case of more delayers.

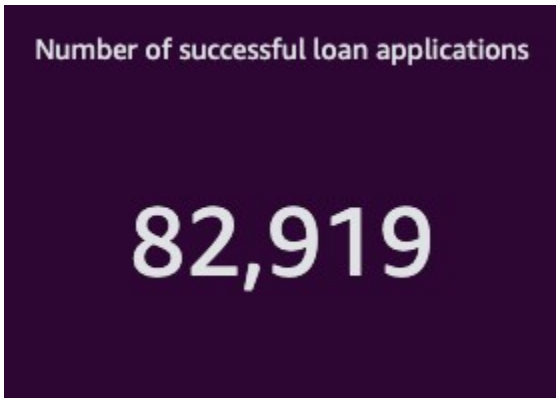
Interpretation of Insights:

Total Number of Loan Applications:

We start with this visual which shows the total number of loan applications regardless of its issuance result. Which means that regardless whether the loan was issued or not this number considers all the loan applications we have in our data set.

Number of successful loan applications:

This visual shows the number of loan applications which were issued successfully.

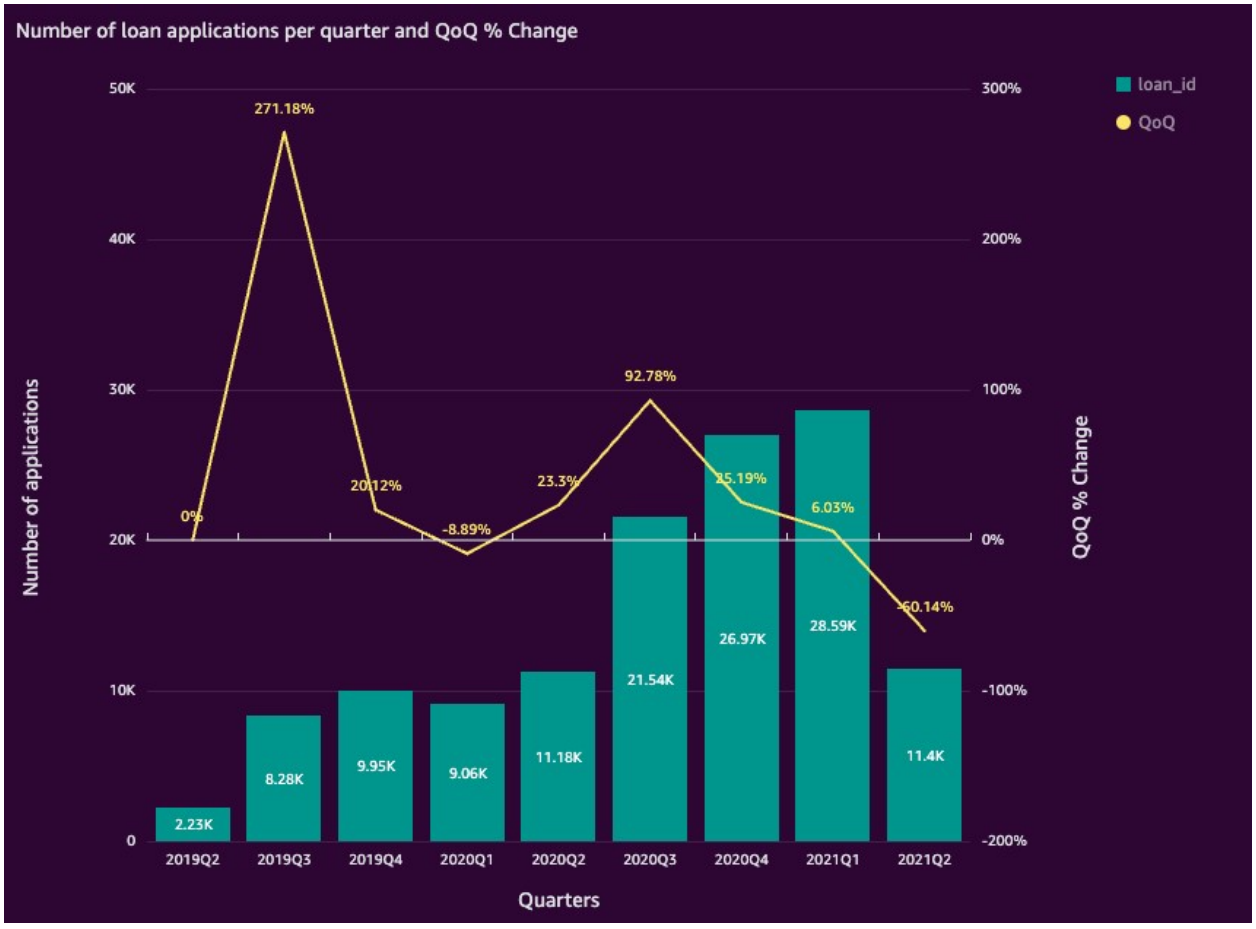


Number of unsuccessful loan applications:

This visual shows the number of loan applications which resulted as not being issued.

Number of loan applications per quarter and QoQ % Change:

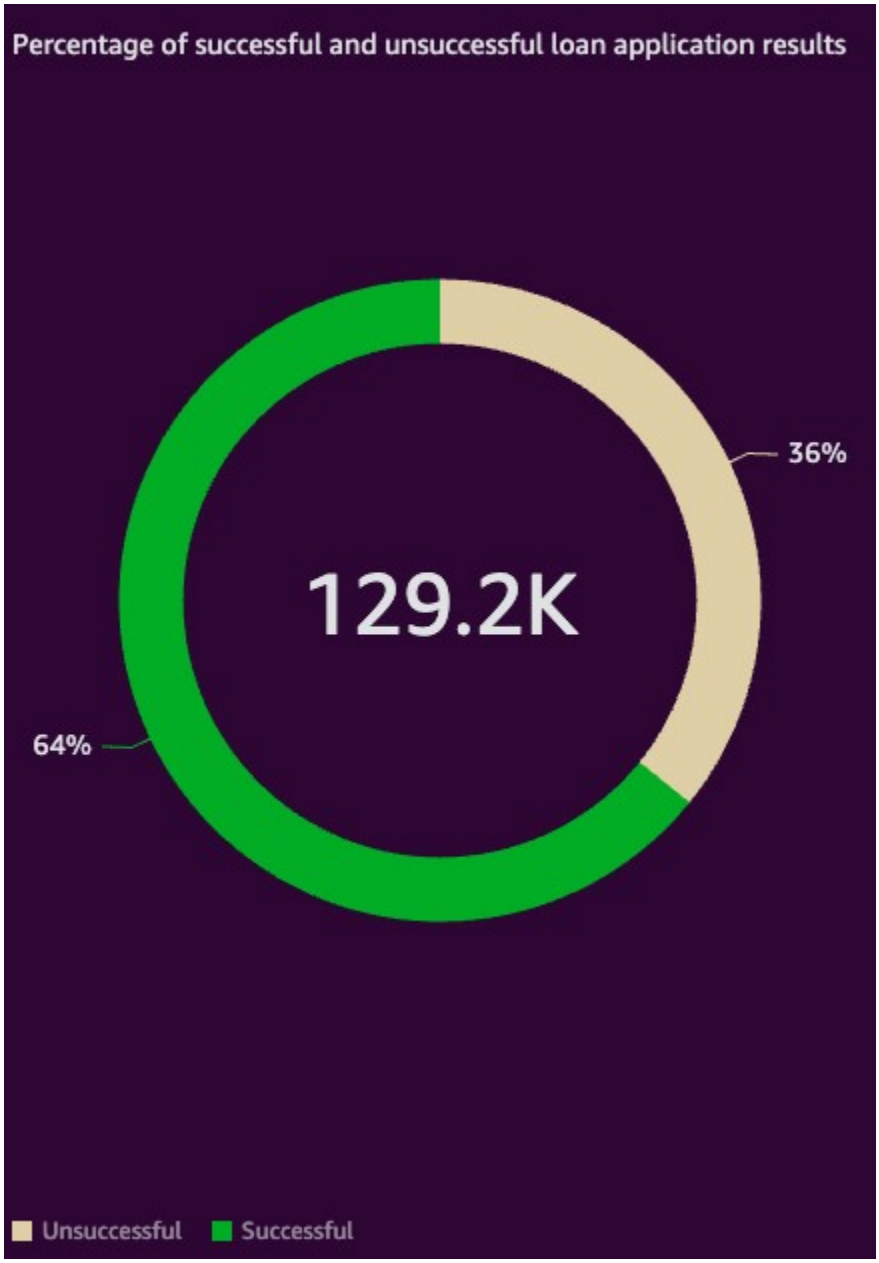
In this visual I have given the breakdown of the number of loan applications per quarter from 2019 quarter 2 till 2021 quarter 2 during these two years. From the chart, taking into account the loan creation date and number of loan applications, **we can presume that this product has been launched in the beginning of 2019.** For that reason, the QoQ % Change in 2019Q2 is quite significant. This can be attributed to the **potential good marketing campaign** as right after 3 months the product has been launched, we can see a quite a big demand by the number of loan applications. In general, we can see that the demand for the product has been quite consistent until the last quarter in the chart. This can be attributed to multiple factors, such as interest rate, however we already know that this decline has nothing to do with interest rate when we look at ***Monthly total expected revenue with average interest rate*** visual. So, this can be attributed to other factors, such as general economic condition, whether alike product has been launched by another company or so on.



Percentage of successful and unsuccessful loan application results:

In this visual we can see the total number of loan applications in the middle of the donut chart and we can see the successful and unsuccessful loan application results in the borders of donut

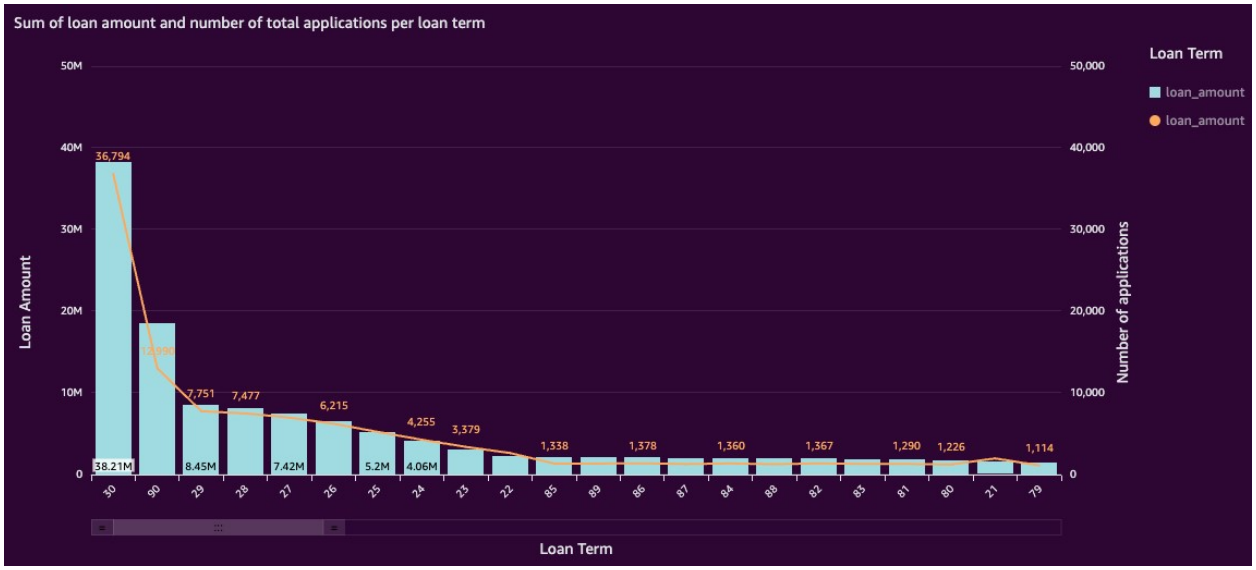
chart which means that we as a company have accepted 64% or roughly 2/3 of the total loan application numbers and we have decreased 36% or roughly 1/3 of them. This **can decrease our profitability**, as less loans issued, less profit will be made on these loans.



Sum of loan amount and number of total applications per loan term:

In this visual I wanted to describe the total loan amount that has been applied for per loan term and the number of applications for that loan term. We can see that the majority of applicants or clients have applied to 30 day loan terms and in fact this is linearly related with the total amount of loan application. Considering that the total value of applied loan amount is the most in 30 days loan term bar and following that comes 90 days then 29 days then 28 days and it goes in

a decreasing manner continuing on. So basically the people would like to pay the take a loan within 30 days the most.



Average requested loan term:

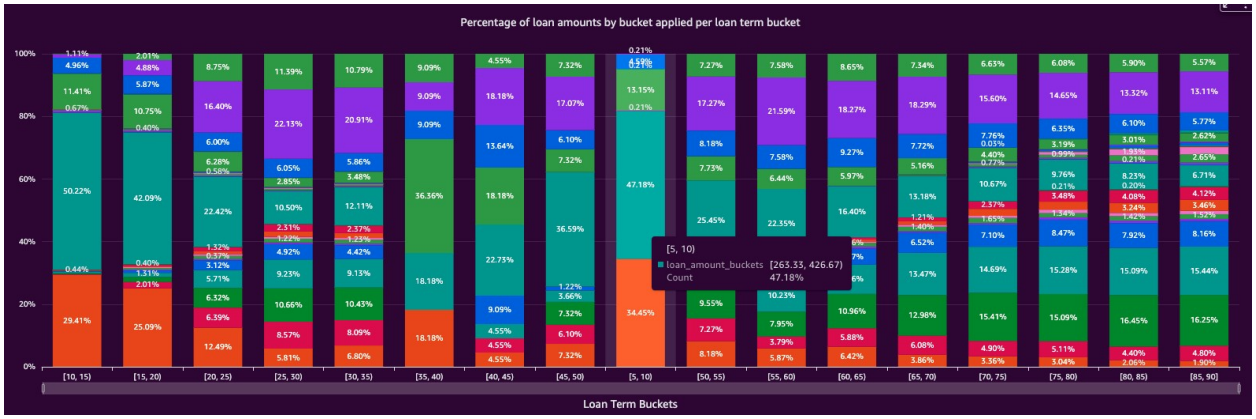
This visual shows the average days that have been requested by the applicants for repayment term and as we can see the average requested long-term days are roughly 43 days. This means that On average our applicants would like to pay their loans back within 44 days at most.

Average loan amount requested:

I think this visual is pretty self-explanatory that, the average loan amount requested was 1085.58 euros.

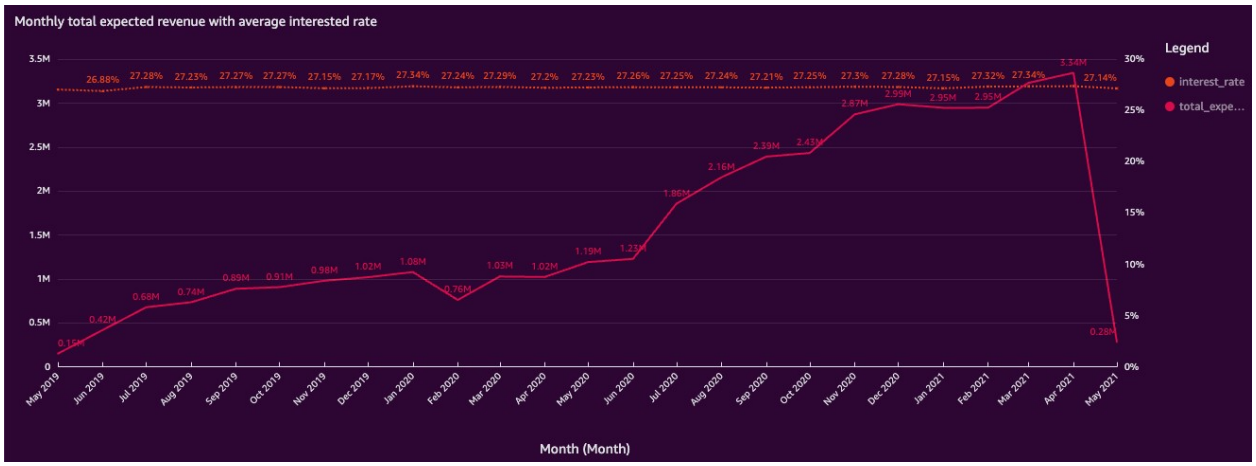
Percentage of loan amounts by bucket applied per loan term bucket:

This visual gives us one of the most important insights. When building this visual I put the loan-terms into buckets and also I put the loan amounts into buckets. So basically these buckets represent the total population of the data as if we were creating a histogram so they have their minimum values and maximum values intervals per bar. Now let's come to the point in which we understand what this visual actually shows to us. So this visual shows us in which loan term bucket the most loan amount bucket was requested by the applicants. As we can see from the attached screenshot that in the loan term bucket which consists of intervals of 5 days to 10 days the 47% of all loan amounts requested were among 263 and 426. In conclusion, this visual helps us to understand that usually when people request a certain amount, at what term do they request it so whether the loan amount being too big or too small depends on the loan term they have applied to.



Monthly total expected revenue with average interest rate:

This visual shows us the difference between loan due and loan amount requested, which is the margin our company would make that we would call revenue, per month along with the interest rate that was applied on those margins. The interest rate here has been calculated as the difference between loan due and loan amount divided by the loan amount. And then I have averaged those interest rates per month and as we can see the interest rate has been quite consistent with roughly 27% along all months. It is also quite useful to understand the total revenue decrease from April 2021 to May 2021 on the chart so that if someone gets suspicious that the reason why Revenue decreased was due to interest rate increase, this chart proves them wrong that this isn't the actual case and we need to look at other metrics or other parameters.



Total expected revenue:

This visual shows us the total expected Revenue that we can make within this time period given on the chart. The reason why I have called it the expected revenue is because not necessarily the loan due always is fully paid so whenever it is not fully paid we have to consider all those cases as well because on accounting level our company is using accruals accounting so we have to consider those revenues without actually those revenues are getting realized.

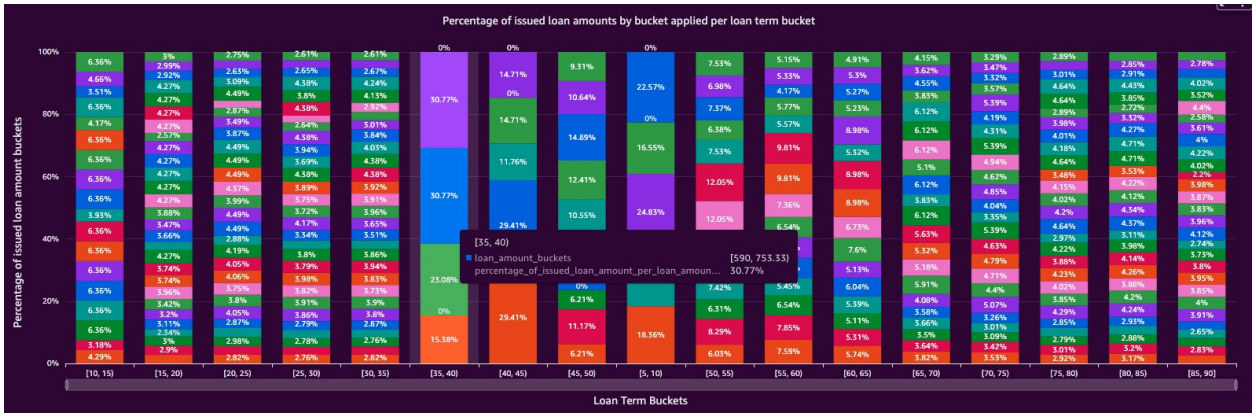
Average interest rate:

This visual shows us the average interest rate that we already have seen on the previous chart.

Percentage of issued loan amounts by bucket applied per loan term bucket:

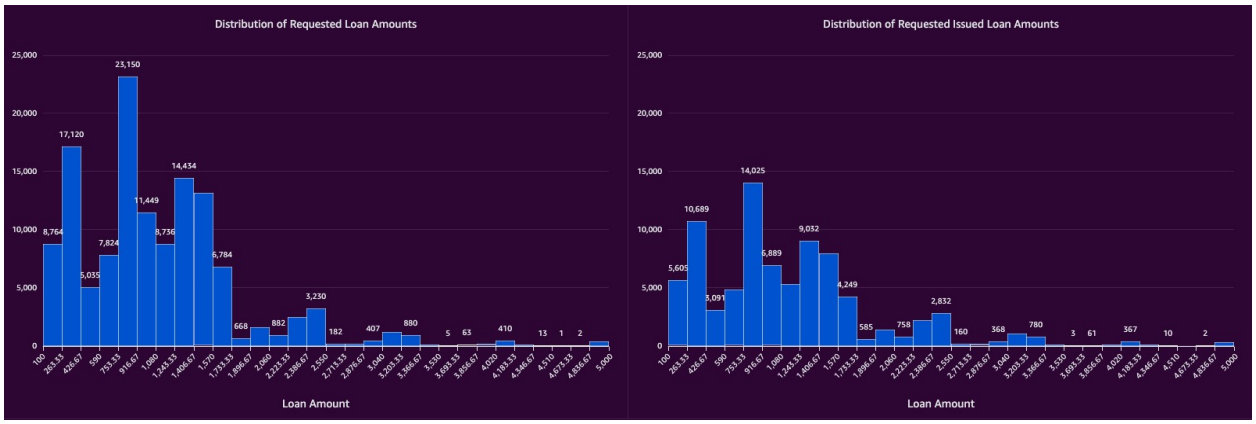
This visual shows us the percentage of applied loan amount intervals so-called loan amount buckets per term bucket. To put it easily let's assume that we are looking at the loan-term buckets where the term lies between 35 and 40 days interval and within this bar we can see the distribution of loan amount buckets that show the percentage of the range of loan amounts that were issued. When we look at the 35 and 40 day term interval and hover over the blue part of the bar we can see that the loan amount bucket which is the interval of 590 and 753 euros have been issued 30% of the time. So basically let's say that a client applies for a loan of 650 euros and they would like to repay it within 35 to 40 days, then we can generalize that out of 10 this

type of loan applications 3 will get issued. In the meantime we shouldn't forget that this chart only considers the issued loan amount and we do not consider the non issued loan amounts.



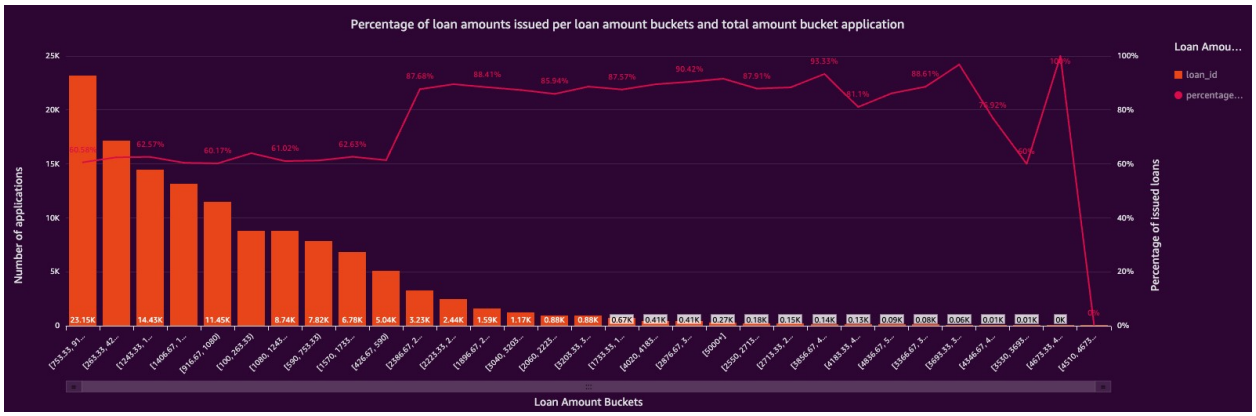
Distribution of Requested Loan amounts and issued loan amounts:

In these two charts standing next to each other on the left one we can see the distribution of requested loan amounts and on the right one we can see the distribution of requested amounts which were issued. So basically, the chart on the left shows us the frequency bars of the loan amount intervals and the right one shows the exact intervals however only the number of those loans within those intervals that were issued.



Percentage of loan amounts issued per loan amount buckets and total amount bucket application:

So this visual shows us the number of total applications per loan bucket and those numbers are shown in the bars. Following those bars we have a red line which has percentage numbers on it and those percentage numbers show the percentage of issued loan amounts out of all the applied loan amounts within the given loan amount interval.



Client Fact Summary Table

This visual is just a client fact summary table. It shows the client ID, client age, client income, client liabilities, client expenses, client debt, the number of applied loans by the given client, the number of issued loans for that client, the number of not issued loans for that client, and a column which is called success rate which shows the success of getting loan issued for the given client.

Client Fact Summary Table

Client	Client Age	Client Income	Client Liabilities	Client Expenses	Client Debt	Number of Applied Loans	Number of Issued Loans	Number of Not Issued Loans	Success Rate
663	23	3,652	923	33	0	14	13	1	92.31%
7,988	34	5,027	0	1,294	0	14	14	0	100%
109	30	2,042	589	1,033	0	13	11	2	81.82%
162	34	3,184	0	1,940	0	13	10	3	70%
1,555	23	3,312	0	123	0	13	12	1	91.67%
11,609	40	4,508	0	97	0	13	11	2	81.82%
1,163	28	2,825	0	477	50	12	11	1	90.91%
1,522	70	3,338	0	840	0	12	11	1	90.91%
2,612	34	4,728	592	367	0	12	10	2	80%
2,838	32	3,326	0	191	0	12	9	3	66.67%
2,896	34	4,275	0	251	0	12	11	1	90.91%
9,658	40	1,027	705	159	0	12	11	1	90.91%
15,565	31	3,445	227	181	0	12	11	1	90.91%
650	37	677	0	1,457	247	11	10	1	90%
966	45	4,751	2,588	92	0	11	10	1	90%
2,726	62	2,112	0	47	0	11	10	1	90%
3,276	29	1,664	0	1,507	0	11	9	2	77.78%
3,338	46	1,883	0	241	0	11	10	1	90%

Average client age:

The average age across all loan applicants.

Total number of clients:

Total number of distinct clients.

Average success rate of loan application:

The average success rate of getting loan issued after an application across all clients.

Client Payment Status Summary Table:

This table is quite an important table because it shows almost all the statistics related to client behavior when it comes to payments. It is to be considered that this table only accounts for the issued loans. The table contains client ID, loan ID which a given client has applied for. One

client can apply for multiple loans. Then we have loan created on column which is the date the loan is issued. Then we have the loan paid on date, which basically shows when a client made a repayment. Then we have loan term which is the number of days for repayment of the loan. Then we have loan due date column which is calculated by adding the loan term days on top of loan credit on date to understand when the loan is due in terms of date. Then we have loan amount which is the amount applied by the client and loan due which is the total amount of repayment expected by the client at the end of loan term. Then we have loan paid which is the amount of repayment made by the client. Then we have the remaining amount which is found by the deducting the loan paid from loan due and it shows us if the client has any outstanding payments or whether they have made an overpayment or if everything has gone well and they have 0 remaining amount. The remaining amount column has been conditionally formatted from lowest being green to highest being red. Then there comes repayment days passed. This column is a calculated field which shows the amount of days that have passed from the loan due date in case a client has made a delayed payment. It has also been conditionally formatted from lowest being green to highest being red. Then we have delayed payment factor which is only applicable to delayed payments and this is found by dividing the absolute value of remaining amount by the repayment days passed to understand how much a given client has incurred a delayed payment fee per delayed day. Then we have the unpaid loan overdue column and this column is a calculated field which is taking into account the maximum loan created on value and subtract all the long credit on values from this maximum value in case there are null values in loan paid on column. This helps us understand that in case there are some loans that have been taken from the company and they have not been paid up until the maximum date on loan credit on column and it let's us see how many days overdue are these loans. And finally we have the repayment status column which shows the repayment status of a given loan basically. Finally, we have the repayment status column which shows us the payment status of each given loan. This is calculated by taking into account the repayment days passed column and if the value in this column is equal to 0 then our values in the repayment status column are `paid on time`, if the repayment days passed is greater than 0 then we have `delayed payment`, and in any other case which obviously we can see from the chart that there are no values then we label them as `no payment`.

Client Payment Status Summary Table

Client	Loan Id	Loan Created On	Loan Paid On	Loan Term	Loan Due Date	Loan Amount	Loan Due	Loan Paid	Remaining Amount	Repayment Days Passed	Delayed Payment Factor	Unpaid Loan Overdue	Repayment Status
1	1	06-15-2019	08-23-2019	69	08-23-2019	900	1,183	1,183	0	0	0		Paid On Time
2	2	05-23-2019	08-21-2019	90	08-21-2019	700	935	935	0	0	0		Paid On Time
2	24,031	02-10-2020	null	27	03-08-2020	2,200	2,751		2,751		0	450 days	No Payment
6	6	05-29-2019	06-28-2019	30	06-28-2019	1,100	1,407	1,407	0	0	0		Paid On Time
7	7	06-02-2019	null	24	06-26-2019	1,500	1,834		1,834		0	703 days	No Payment
8	8	05-24-2019	06-23-2019	30	06-23-2019	200	256	256	0	0	0		Paid On Time
8	5,534	07-16-2019	08-15-2019	30	08-15-2019	2,400	3,068	3,068	0	0	0		Paid On Time
8	11,104	09-14-2019	11-29-2019	76	11-29-2019	2,300	3,038	3,038	0	0	0		Paid On Time
9	9	06-03-2019	09-01-2019	90	09-01-2019	1,500	2,005	2,005	0	0	0		Paid On Time
13	13	06-06-2019	07-06-2019	30	07-06-2019	1,000	1,279	1,279	0	0	0		Paid On Time
14	14	06-02-2019	06-23-2019	21	06-23-2019	1,200	1,433	1,433	0	0	0		Paid On Time
14	5,535	07-16-2019	null	78	10-02-2019	2,300	3,041		3,041		0	659 days	No Payment
16	16	05-24-2019	08-19-2019	86	08-18-2019	700	930	935.6	-5.6	1	5.6		Delayed Payment
16	11,105	09-14-2019	12-13-2019	90	12-13-2019	2,300	3,073	3,073	0	0	0		Paid On Time

Average Delayed Repayment Date:

This visual shows us on average how many days are payments delayed.

Average Daily Fare Charged on Delayed Payments:

This visual shows us the average fare charged on delayed payments. And how do we find these? As we mentioned before we have the delayed payment factor which is the amount charged per day for each repayment day passed and we average it across all the delayed payments.

Total fare charged for delayed payments:

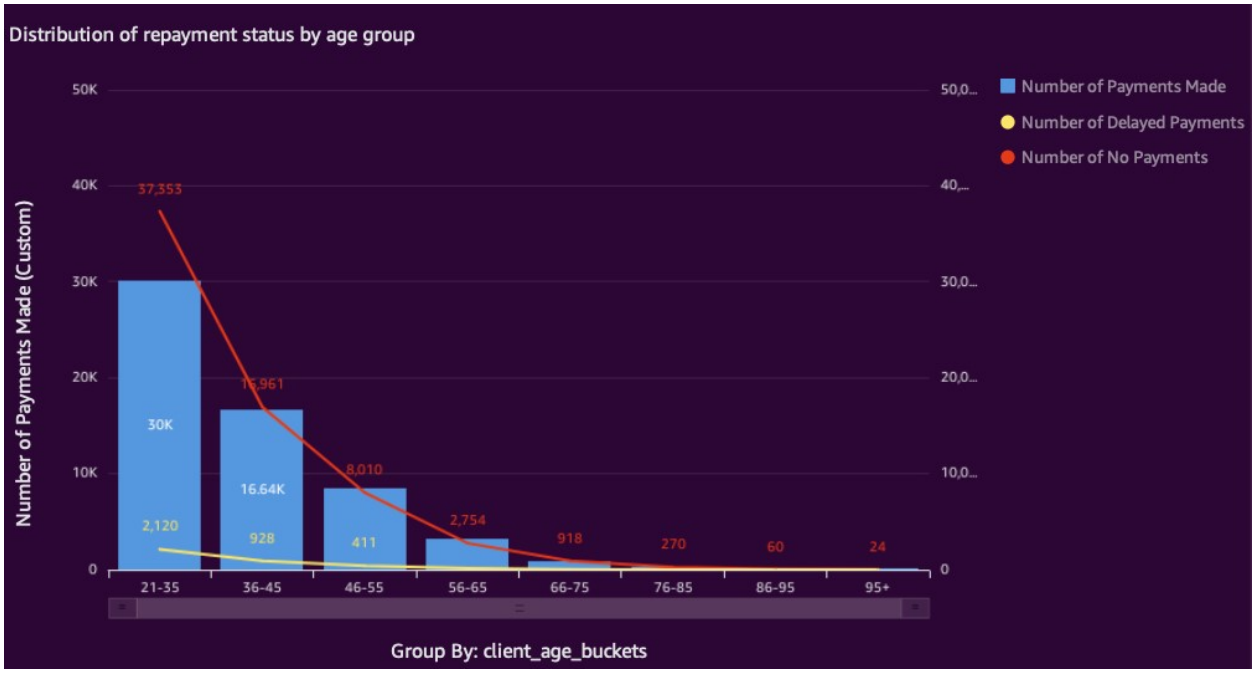
This visual shows us the total fare charge for delayed payments across all delayed payments.

Total payments received:

This visual shows us the total payments made by the clients.

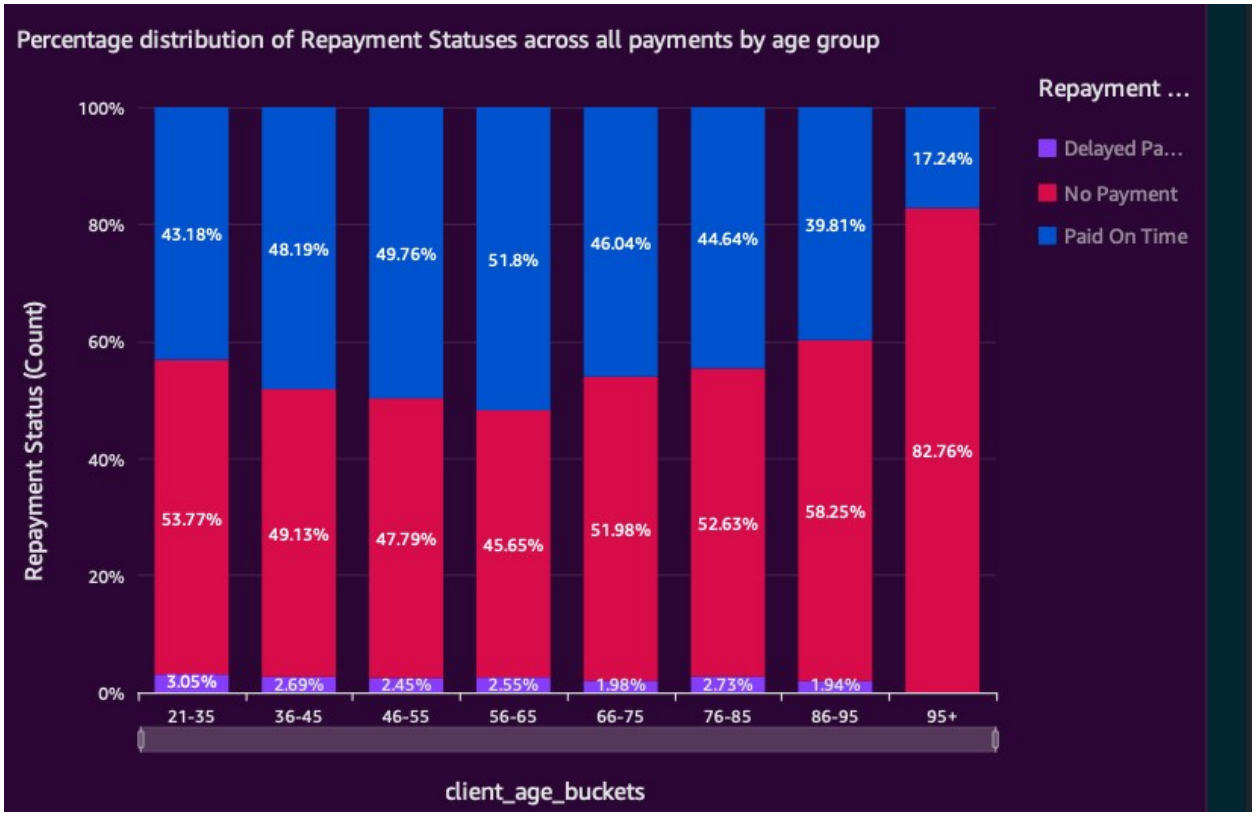
Distribution of repayment status by age group:

This visual shows us insights about the payment behavior of different age groups. The bars in the chart show us the number of successful payments made. The red line shows the number of no payments and the yellow line shows the number of delayed payments. From the chart, we can see that as most of our clients are in the 21-35 age group, it makes sense that proportionally the repayment status will be consistent with the number of clients per age group.



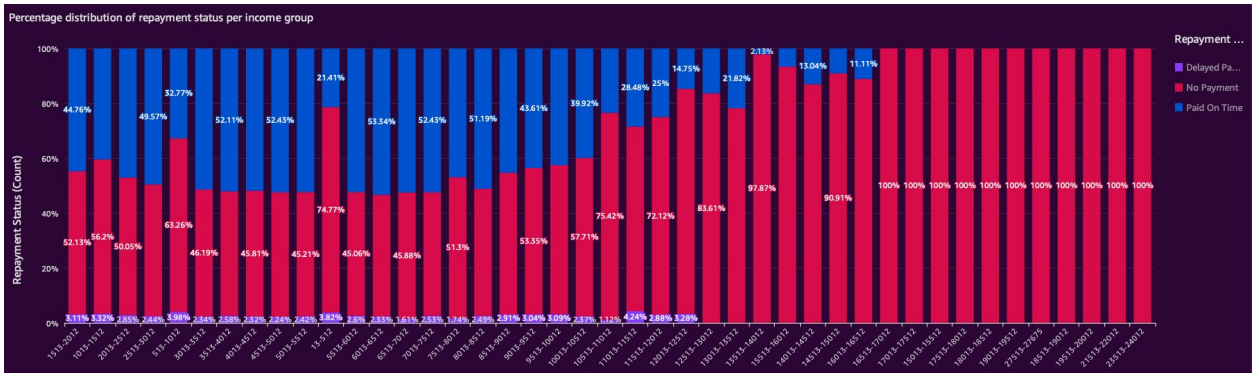
Percentage distribution of Repayment statuses across all payments by age group:

Following the above visual, in this visual we show the percentage of repayment statuses per age group total payment amount. The bars show the total number of issued loans per age group and the different stacks show the respective percentage of repayment status for these loans. For example, we can see from the chart that the most payments made on time across all age groups belong to people who are between the ages of 56 and 65. The age group which has made no payments at all mostly belong to the age group of 95+ aged people.



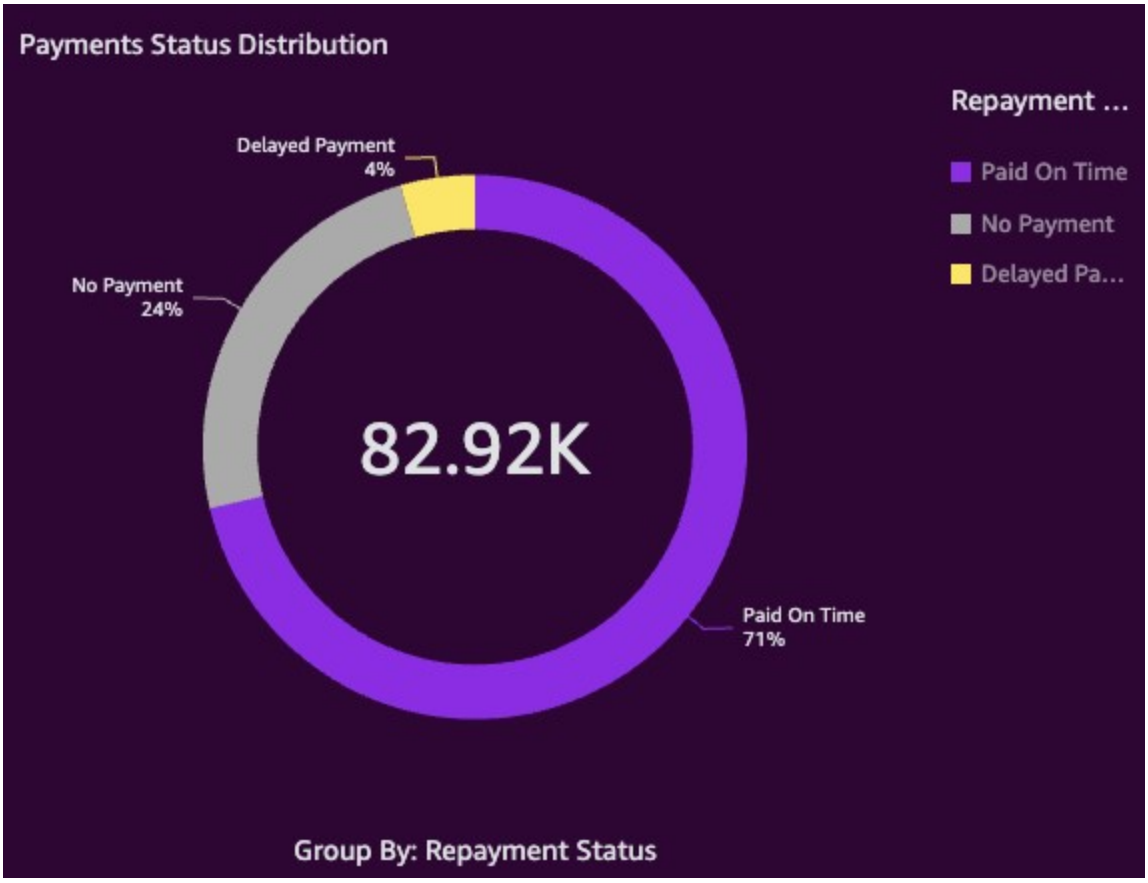
Percentage distribution of repayment status per income group:

In this visual we can see the same logic of repayment status distribution for income groups rather than age groups. For example, when looking at the chart, we can see that in general there is a tendency that the more the income of a client, they tend to not make payments on time. This can be attributed to the fact that usually the people who earn more wages are seniors or elderly and with this fact we can interpolate this insight on to the previous chart as well.



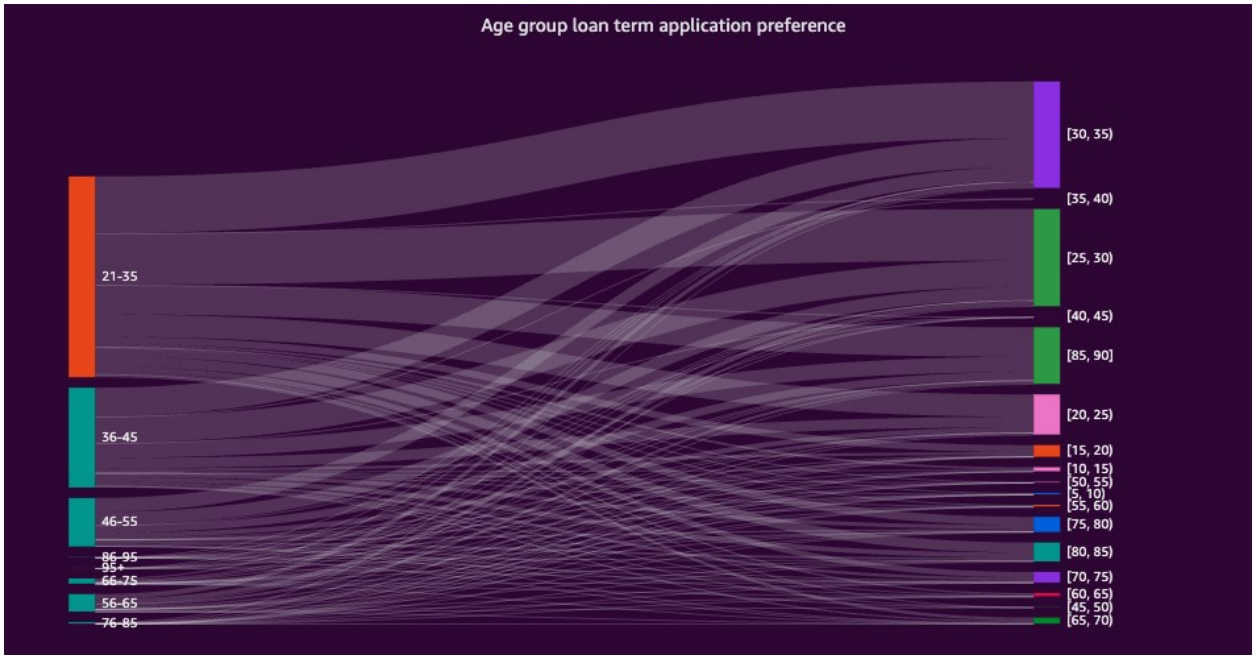
Payments status distribution:

This donut chart simply shows us the total number of issued loans in the middle and the respective payment status on those loans.



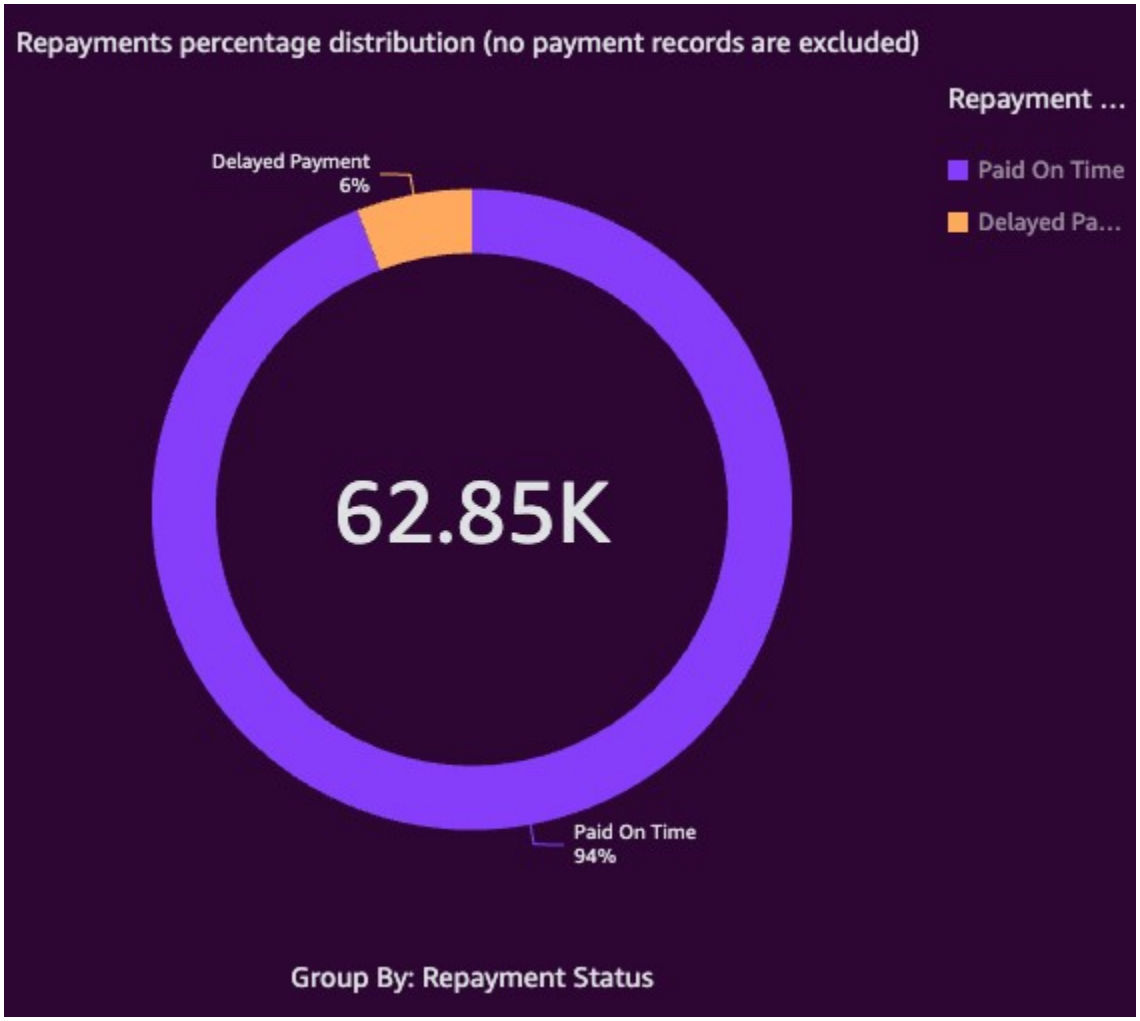
Age group loan term application preference:

This sankey diagram shows us the respective loan application term preferences per age group.



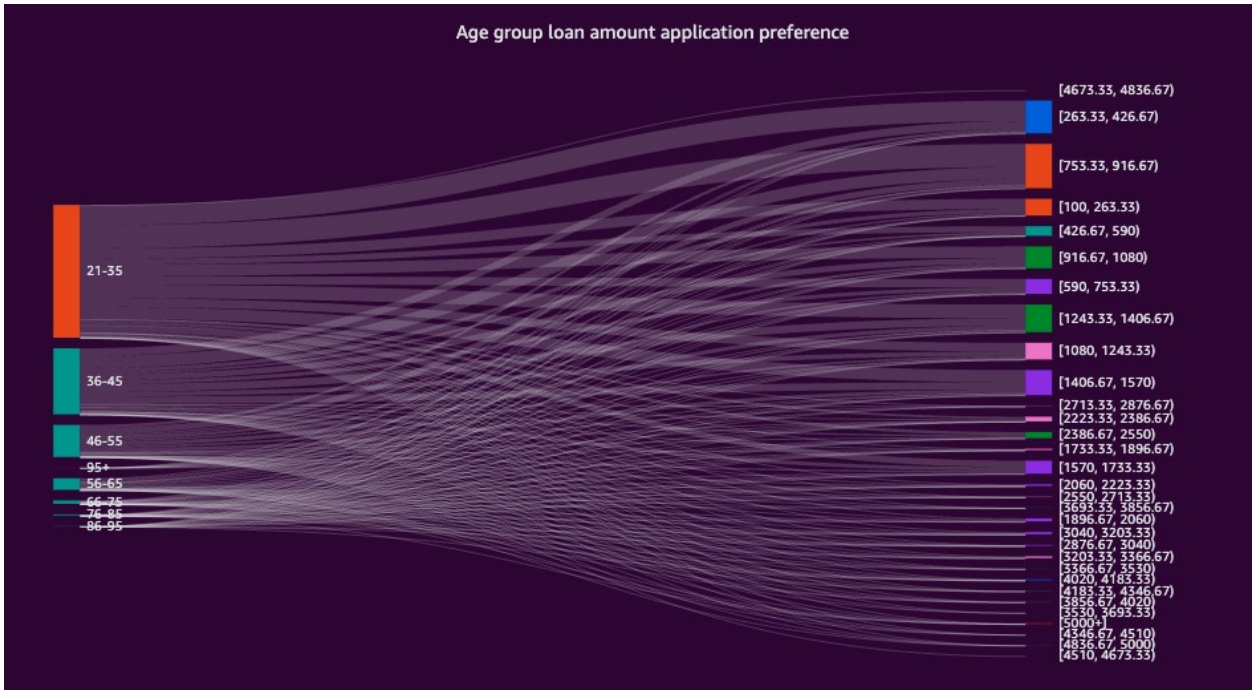
Repayments percentage distribution (no payments records are excluded):

This donut chart shows us the total number of payments made in the middle of the donut, and the percentage of payments status.



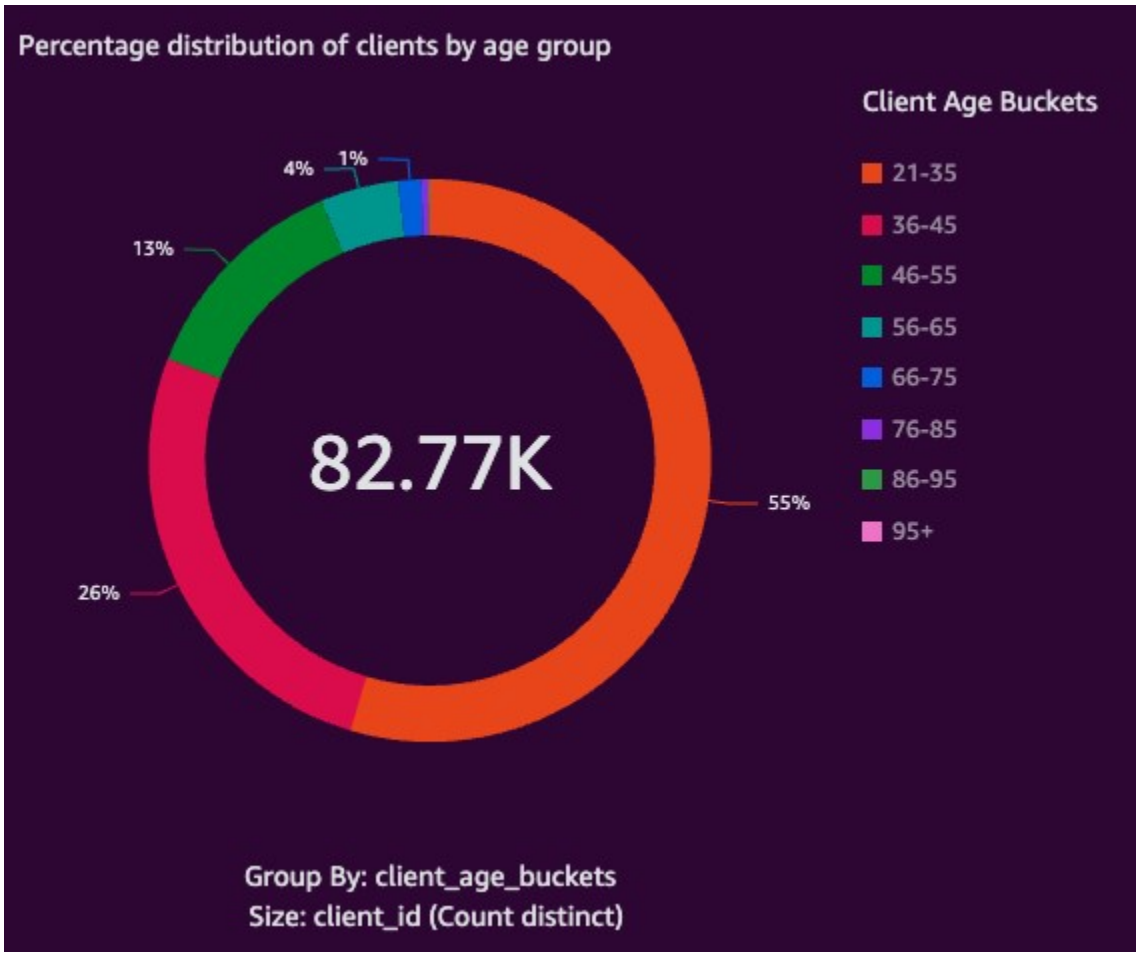
Age group loan amount application preference:

This sankey diagram shows us the loan amount preference of respective age group.



Percentage distribution of clients by age group:

This donut chart shows the total client base in the middle (distinct count if client ids) and their respective age group.



Correlations:

Here, I wanted to understand the relationship between some variables. For that reason I created some Point Biseral Correlations to understand:

- ## Correlation between Loan Amount and Loan Issuance:

  I wanted to understand if there is a relationship between loan amount and loan issuance. My hypothesis was that if the loan amount requested was higher, then the loan becomes less likely to be issued. But the correlation coefficient showed that there is almost no correlation between the two.

- ## Correlation between client age and issuance:

  I wanted to understand if there is a relationship between the client age and loan issuance. My hypothesis was that if the client's age is too young, then loan becomes less likely to be issued. But the correlation coefficient showed that there is almost no correlation between the two.

- ## Correlation between client age and loan payment:

  I wanted to understand if there is a relationship between the client age and loan payment. My hypothesis was that if the client's age is too young, then they become reluctant or delayed in terms of payments. But the correlation coefficient showed that there is almost no correlation between the two.

- ## Correlation between client income and issuance:

  I wanted to understand if there is a relationship between the client income and loan issuance. My hypothesis was that if the client's income is higher, then loan becomes more likely to be issued. But the correlation coefficient showed that there is very small correlation between them that it's not significant enough to conclude that opinion.

Correlation between Loan Amount and Loan Issuance	Correlation between client age and issuance
-0.12	0.03
Correlation between client age and loan payment	**Correlation between client income and issuance**
0.06	0.14