Credit risk analysis using SQL

The simulated credit bureau dataset will be used to identify potential business opportunities for a client.

Queries will be created to answer the following questions:

1. What is the main motivation for credit acquisition?
2. What is the default rate for the loan with the highest demand?
3. Which loans have the lowest default rate?
4. Is there an ideal age group that can be targeted for loans?
5. What is the average loan amount for age groups with 30% or below utilisation?

Data pre-processing

```
proc sql;
create table credit.newbureau as
        select id,cb_person_cred_hist_length,
        case
                when cb_person_default_on_file = "Y" then  1
                when cb_person_default_on_file = "N" then  0
                end as default
        from credit.bureau;
quit;
```

Perform inner join of credit data and newbureau data using sql:

```
proc sql;
create table credit.mergecb as
        select *
        from credit.credit as c inner join credit.newbureau as b
        on c.id = b.id ;
quit;
```

1. What is the main motivation for credit acquisition?

```
Title 'Main reason for credit acquisition';
proc sql ;
        select loan_intent ,count(loan_intent) as count
        from credit.mergecb
        group by loan_intent
        order by count desc;
quit;
```

Educational loans are the main motivation for credit acquisition with 6453 educational loans acquired. This accounts for 19.8% of all acquired loans.

2. What is the default rate for the loan with highest demand?

```
Title "Default for loan_intent: Education";
proc sql ;
        select default as Default,count(loan_intent) as Number_of_Defaults
        from credit.mergecb
        where loan_intent = "EDUCATION"
        group by default ;
quit;
```

Only 1102 of the educational loans were categorised as default. This only accounts for 17.07% of the total loans that were categorised as default.

3. Which loans have the lowest default rate?

```
Title "Default level as per loan_intent";
proc sql ;
        select loan_intent,sum(default =1) as Default, sum(default =0) as NotDefault, calculated
Default/(calculated NotDefault+calculated Default) *100 as Percentage_Defaulted
        from credit.mergecb
        group by loan_intent
        order by Percentage_Defaulted asc;
quit;
```

Educational loans have the lowest default rate of only 17.07%

4. Which age group utilises credit best

```
Title "Analysis of age groups credit utilisation ";
proc sql ;
select
case
        when 0<=person_age<=20 then "other"
        when 20<=person_age<30 then "20-29"
        when 30<=person_age<40 then "30-39"
        when 40<=person_age<50 then "40-49"
        when 50<=person_age<60 then "50-59"
        when 60<=person_age<=65 then "60-65"
        when person_age>=66 then "other"
        end as Age_group, count(loan_percent_income) as Utilisation,
        sum(default = 1) as Defaulted, sum(default = 1)/count(loan_percent_income) *100 as
Percentage_Defaulted,
        round(avg(cb_person_cred_hist_length)) as Average_Credit_Years
from credit.mergecb
where loan_percent_income <=0.3
group by Age_group
;
quit;
```

Age group 50-59 has the best utilisation of credit (credit utilisation of 30% or less). It has the lowest default rate of 15.56% as compared to the other age groups and it has an average active credit history of 21 years, which is one of the highest.

5. What is the average loan amount for the age groups with 30% or below utilisation?

```
Title "Average loan amount per age groups ";
proc sql ;
select
case
```

```
        when 0<=person_age<=20 then "other"
        when 20<=person_age<30 then "20-29"
        when 30<=person_age<40 then "30-39"
        when 40<=person_age<50 then "40-49"
        when 50<=person_age<60 then "50-59"
        when 60<=person_age<=65 then "60-65"
        when person_age>=66 then "other"
        end as Age_group, avg(loan_amnt) as Average_credit
from credit.mergecb
where loan_percent_income <=0.3
group by Age_group
;
quit;
```

Age_group	Average_credit
20-29	8591.833
30-39	9175.463
40-49	9376.457
50-59	8548.547
60-65	11247.83
other	8790