# FML

Randheer Gonuguntla

2023-09-10

Using Credit Risk Dataset/ Source: https://www.kaggle.com/datasets/laotse/credit-risk-dataset

Importing Dataset

```
Credit_Dataset<-read.csv("C:\\Users\\sidda\\Downloads\\credit_risk_dataset.csv")
```

Drescriptive Statistics of the data

```
summary(Credit_Dataset)
```

```
##    person_age     person_income     person_home_ownership person_emp_length
##  Min.   : 20.00   Min.   :   4000   Length:32581          Min.   :  0.00
##  1st Qu.: 23.00   1st Qu.:  38500   Class :character      1st Qu.:  2.00
##  Median : 26.00   Median :  55000   Mode  :character      Median :  4.00
##  Mean   : 27.73   Mean   :  66075                         Mean   :  4.79
##  3rd Qu.: 30.00   3rd Qu.:  79200                         3rd Qu.:  7.00
##  Max.   :144.00   Max.   :6000000                         Max.   :123.00
##                                                           NA's   :895
##  loan_intent         loan_grade          loan_amnt     loan_int_rate
##  Length:32581       Length:32581       Min.   :  500   Min.   : 5.42
##  Class :character   Class :character   1st Qu.: 5000   1st Qu.: 7.90
##  Mode  :character   Mode  :character   Median : 8000   Median :10.99
##                                        Mean   : 9589   Mean   :11.01
##                                        3rd Qu.:12200   3rd Qu.:13.47
##                                        Max.   :35000   Max.   :23.22
##                                                        NA's   :3116
##   loan_status     loan_percent_income cb_person_default_on_file
##  Min.   :0.0000   Min.   :0.0000      Length:32581
##  1st Qu.:0.0000   1st Qu.:0.0900      Class :character
##  Median :0.0000   Median :0.1500      Mode  :character
##  Mean   :0.2182   Mean   :0.1702
##  3rd Qu.:0.0000   3rd Qu.:0.2300
##  Max.   :1.0000   Max.   :0.8300
##
##  cb_person_cred_hist_length
##  Min.   : 2.000
##  1st Qu.: 3.000
##  Median : 4.000
##  Mean   : 5.804
##  3rd Qu.: 8.000
##  Max.   :30.000
##
```

Transforming person income variable of the dataset

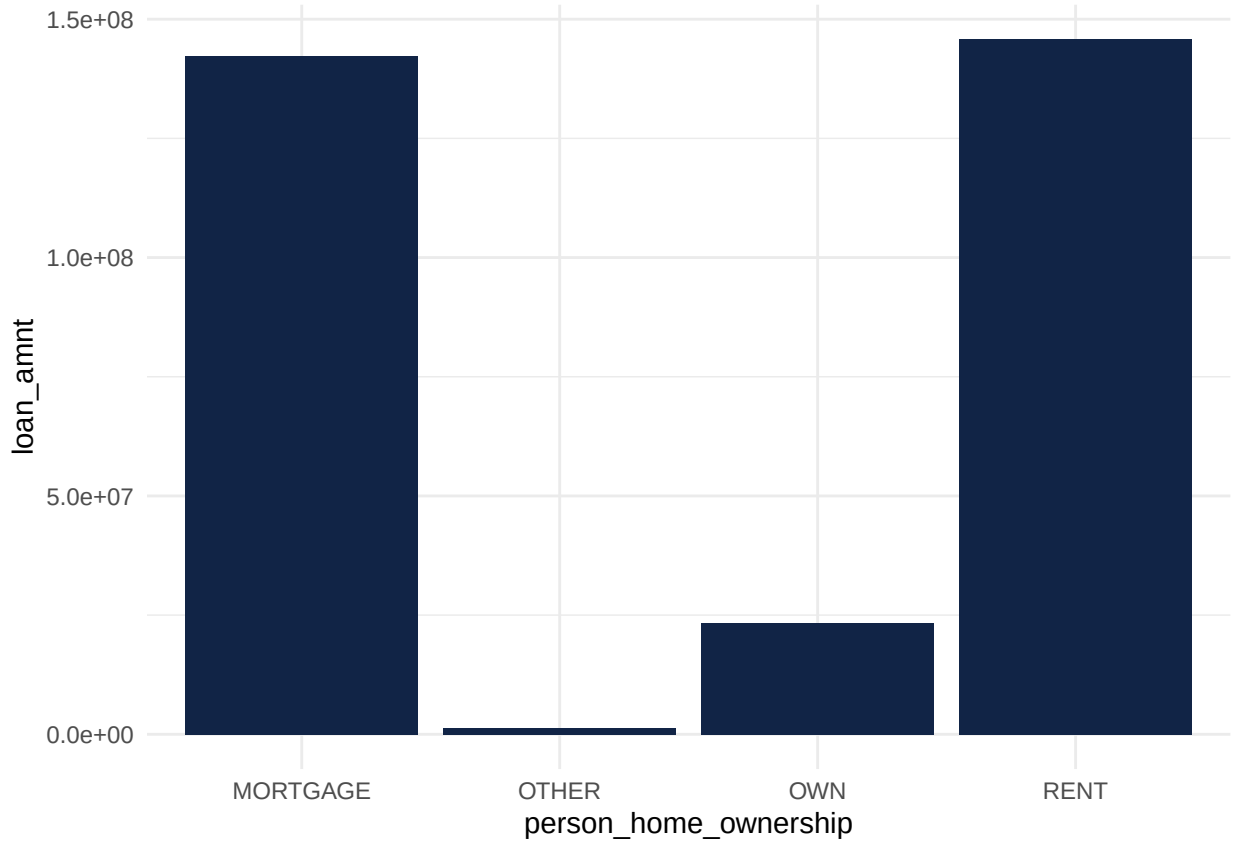
```
mean_income=mean(Credit_Dataset$person_income)
median_income=median(Credit_Dataset$person_income)
mode_income=mode(Credit_Dataset$person_income)
sd_income=sd(Credit_Dataset$person_income)


Transformed_income=(Credit_Dataset$person_income-median_income)/sd_income+ mean_income
```

Transforming loan amount

```
New_loanAmount= (Credit_Dataset$loan_amnt)/3
```

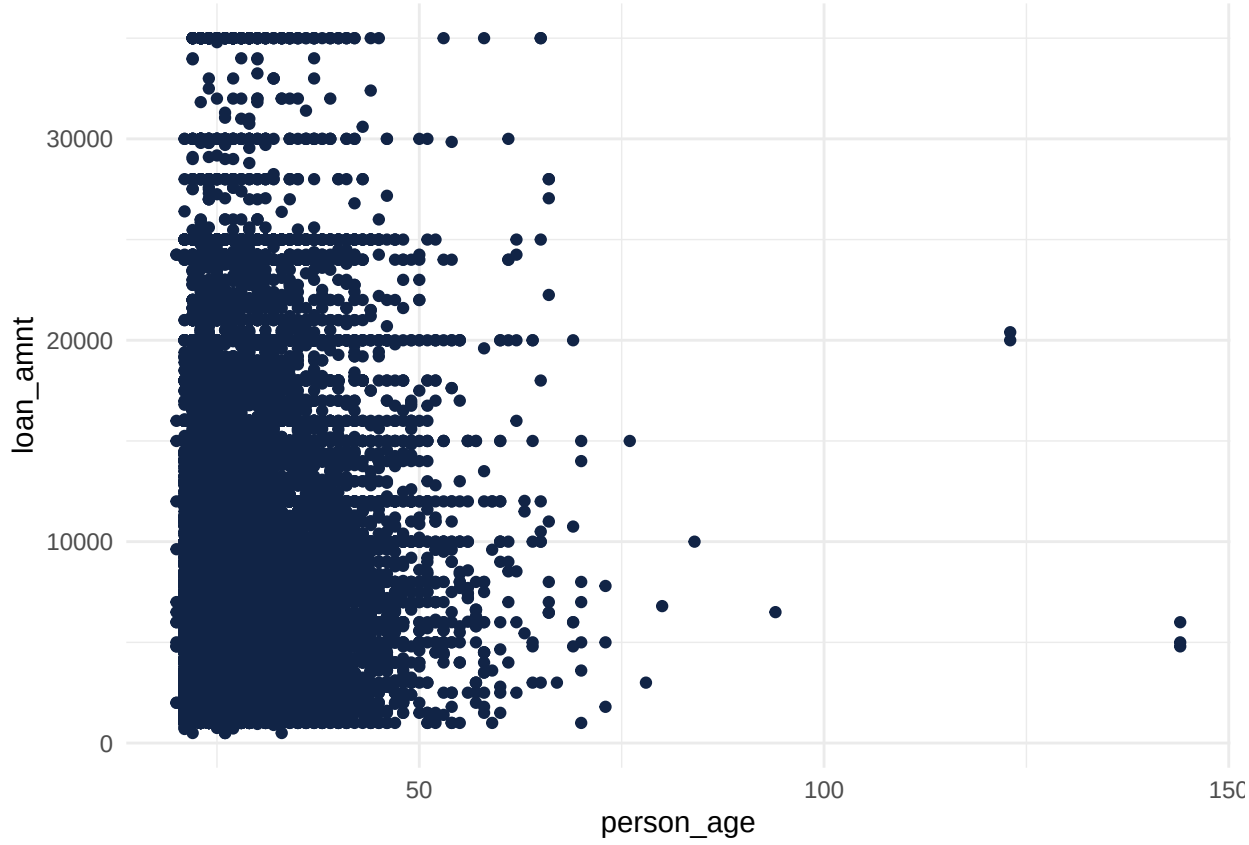
Plotting person home ownership vs loan amount

```
library(ggplot2)


library(ggplot2)

ggplot(Credit_Dataset) +
 aes(x = person_home_ownership, y = loan_amnt) +
 geom_col(fill = "#112446") +
 theme_minimal()
```



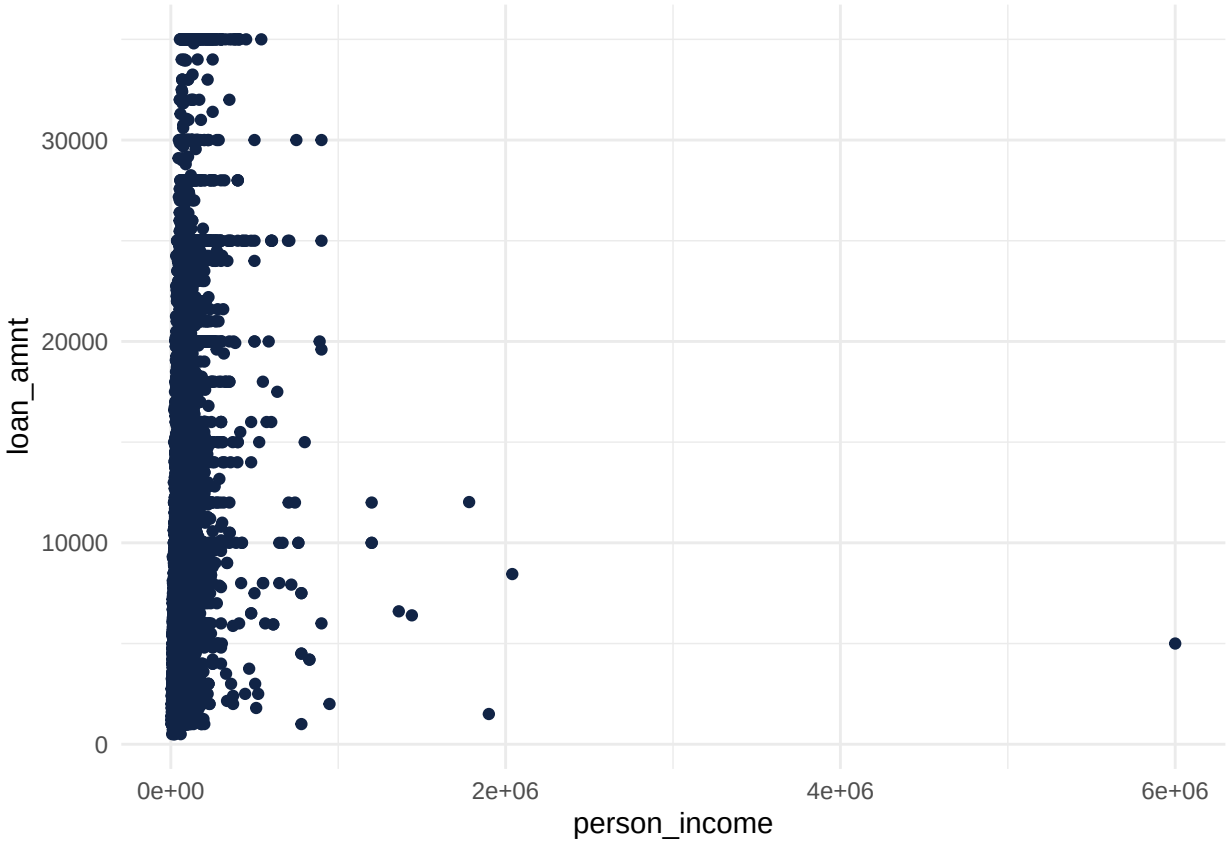

Scatter plot of Loan amount and person age

```
ggplot(Credit_Dataset) +
 aes(x = person_age, y = loan_amnt) +
 geom_point(shape = "circle", size = 1.5,
 colour = "#112446") +
 theme_minimal()
```




Scatter plot of personal income and loan amount

```
ggplot(Credit_Dataset) +
 aes(x = person_income, y = loan_amnt) +
 geom_point(shape = "circle", size = 1.5,
 colour = "#112446") +
 theme_minimal()
```


loan_amnt
30000
20000
10000
0
0e+00
2e+06
4e+06
6e+06
person_income