

## R Resit Task

```
library(tidyverse)
library(ggplot2)
library(dplyr)

df <- read.csv("./DATA/loan_data.csv")

average_loan_amnt <- mean(df$loan_amnt, na.rm = TRUE)
average_loan_amnt

unique_purpose_count <- length(unique(df$purpose))
unique_purpose_count

df2 <- df %>% filter(home_ownership == 'RENT' | purpose != 'educational')

ggplot(df, aes(x = annual_inc)) +
  geom_histogram(binwidth = 5000, fill = 'skyblue', color = 'black' +
  labs(title = "Distribution of Annual Income", x = "Annual Income", y = "Frequency") +
  theme_minimal()

df3 = df %>%
  group_by(home_ownership) %>%
  summarise(std_dev_int_rate = sd(int_rate, na.rm = TRUE)) %>%
  arrange(std_dev_int_rate)
df3

df$issue_date = as.Date(df$issue_date)

defaults_over_time = df %>%
  mutate(issue_month = floor_date(issue_date, "month")) %>%
  group_by(issue_month) %>%
  summarise(default_count = sum(default = 1, na.rm = TRUE)) %>%
  arrange(issue_month)

ggplot(defaults_over_time, aes(x = issue_month, y = default_count))+
  geom_line(colour = "blue")+
  labs(title = "Evolution of Defaults Over Time", x = "Issue Month", y = "Number of
Defaults")+
  theme_minimal()
```