

Lab 6 - Using Color Differentiation in Plots

Ariane Ducellier

10/17/2023

In this lab, we are going to visualize the relationship between loan amounts and home ownership for different credit grades (column name = grade).

Libraries

Load the necessary libraries.

```
library(tidyverse)

## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.3      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.0
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

Data

Load the data.

```
df <- read_csv("../data/LoanStats.csv")

## Rows: 42542 Columns: 145
## -- Column specification -----
## Delimiter: ","
## chr (29): id, term, int_rate, grade, sub_grade, emp_title, emp_length, home_...
## dbl (34): loan_amnt, funded_amnt, funded_amnt_inv, installment, annual_inc, ...
## lgl (82): member_id, url, initial_list_status, mths_since_last_major_derog, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
dfn <- df[, c("home_ownership", "loan_amnt", "grade")]
dfn <- na.omit(dfn)
dfn <- subset(dfn, !dfn$home_ownership %in% c("NONE"))
```

Plot

Complete the cell with the code to make the plot.