LLMs (Appendix)

Salman and Farhan

2023-07-14

Libraries and Loading Data

```
library(readr)
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(RColorBrewer)

data <- read_csv("C:/Users/Salman/Downloads/AI.csv")
```

```
## Rows: 231 Columns: 11
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (11): Timestamp, What is your age bracket?, What is your gender?, What i...
##
## 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.
```

Data Wranlging

```
colnames(data) <- c("time", "age", "gender", "education", "AI_freq",
                    "which_AI_used", "which_AI_prefer", "which_AI_helpful",
                    "bard_or_gpt", "gpt_or_bard", "room_improv"
                    )

data <- data %>%
  select(-1,-6)

data[, 1:9] <- lapply(data[, 1:9], factor)

summary(data)
```

```
##     age                    gender                 education        AI_freq
##  15-25:158   Female           : 63   Bachelor's        :140   Everytime:106
##  25-35: 48   Male             :167   Master's          : 35   Never    : 29
##  35-45: 25   Prefer Not to Say:  1   PhD               : 29   Often    : 57
##                                      Senior High School: 27   Rarely   : 39
##      which_AI_prefer     which_AI_helpful    bard_or_gpt     gpt_or_bard
##  Google BARD : 45    Google BARD : 47     Average  :150   Average  : 27
##  GPT 3/ GPT 4:186    GPT 3/ GPT 4:184     Excellent:  9   Excellent:120
##                                           Good     : 34   Good     : 48
##                                           Poor     : 38   Poor     : 36
##                      room_improv
##  Agree                     :177
##  Disagree                  : 19
##  Don't Know                : 16
##  Neither Agree nor Disagree: 19
```

```
str(data)
```

```
## tibble [231 x 9] (S3: tbl_df/tbl/data.frame)
##  $ age             : Factor w/ 3 levels "15-25","25-35",..: 2 3 2 2 3 2 1 1 2 1 ...
##  $ gender          : Factor w/ 3 levels "Female","Male",..: 2 2 2 2 2 1 2 2 2 2 ...
##  $ education       : Factor w/ 4 levels "Bachelor's","Master's",..: 3 3 2 3 3 3 1 1 1 1 ...
##  $ AI_freq         : Factor w/ 4 levels "Everytime","Never",..: 4 3 4 4 2 3 3 3 3 3 ...
##  $ which_AI_prefer : Factor w/ 2 levels "Google BARD",..: 2 2 2 2 2 2 2 2 2 2 ...
##  $ which_AI_helpful: Factor w/ 2 levels "Google BARD",..: 2 2 2 2 2 2 2 2 2 2 ...
##  $ bard_or_gpt     : Factor w/ 4 levels "Average","Excellent",..: 1 3 1 1 1 1 1 1 1 1 ...
##  $ gpt_or_bard     : Factor w/ 4 levels "Average","Excellent",..: 2 3 2 1 3 3 2 2 2 2 ...
##  $ room_improv     : Factor w/ 4 levels "Agree","Disagree",..: 1 1 1 1 3 1 1 1 1 1 ...
```

Exploratory Analysis

```
prop.table(table(data$age))
```

```
##
##     15-25     25-35     35-45
## 0.6839827 0.2077922 0.1082251
```

```r
prop.table(table(data$gender))
```

```
##
##            Female              Male Prefer Not to Say
##       0.272727273       0.722943723       0.004329004
```

```r
prop.table(table(data$education))
```

```
##
##         Bachelor's           Master's                PhD Senior High School
##          0.6060606          0.1515152          0.1255411          0.1168831
```

```r
prop.table(table(data$AI_freq))
```

```
##
## Everytime     Never     Often    Rarely
## 0.4588745 0.1255411 0.2467532 0.1688312
```

```r
prop.table(table(data$which_AI_prefer))
```

```
##
##  Google BARD GPT 3/ GPT 4
##    0.1948052    0.8051948
```

```r
prop.table(table(data$bard_or_gpt))
```

```
##
##    Average  Excellent       Good       Poor
## 0.64935065 0.03896104 0.14718615 0.16450216
```

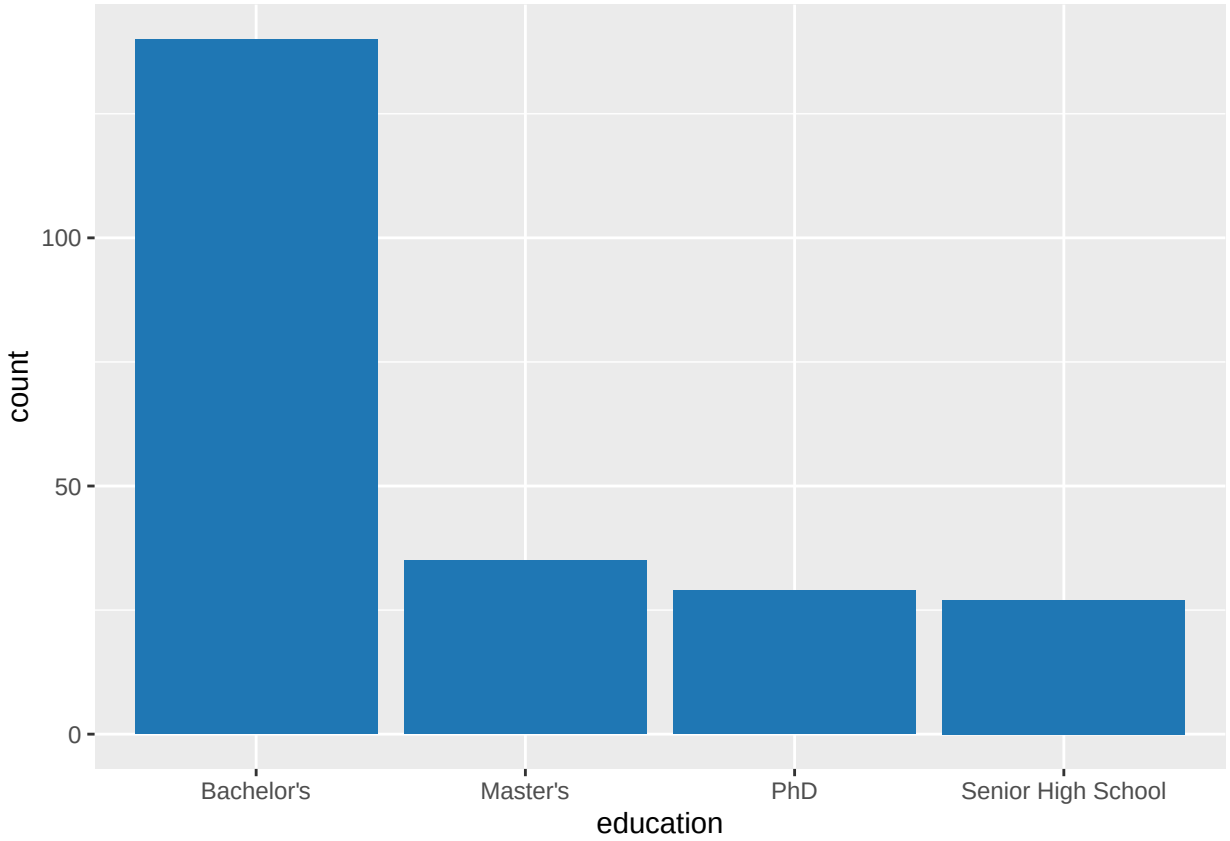
```r
prop.table(table(data$gpt_or_bard))
```

```
##
##   Average Excellent      Good      Poor
## 0.1168831 0.5194805 0.2077922 0.1558442
```

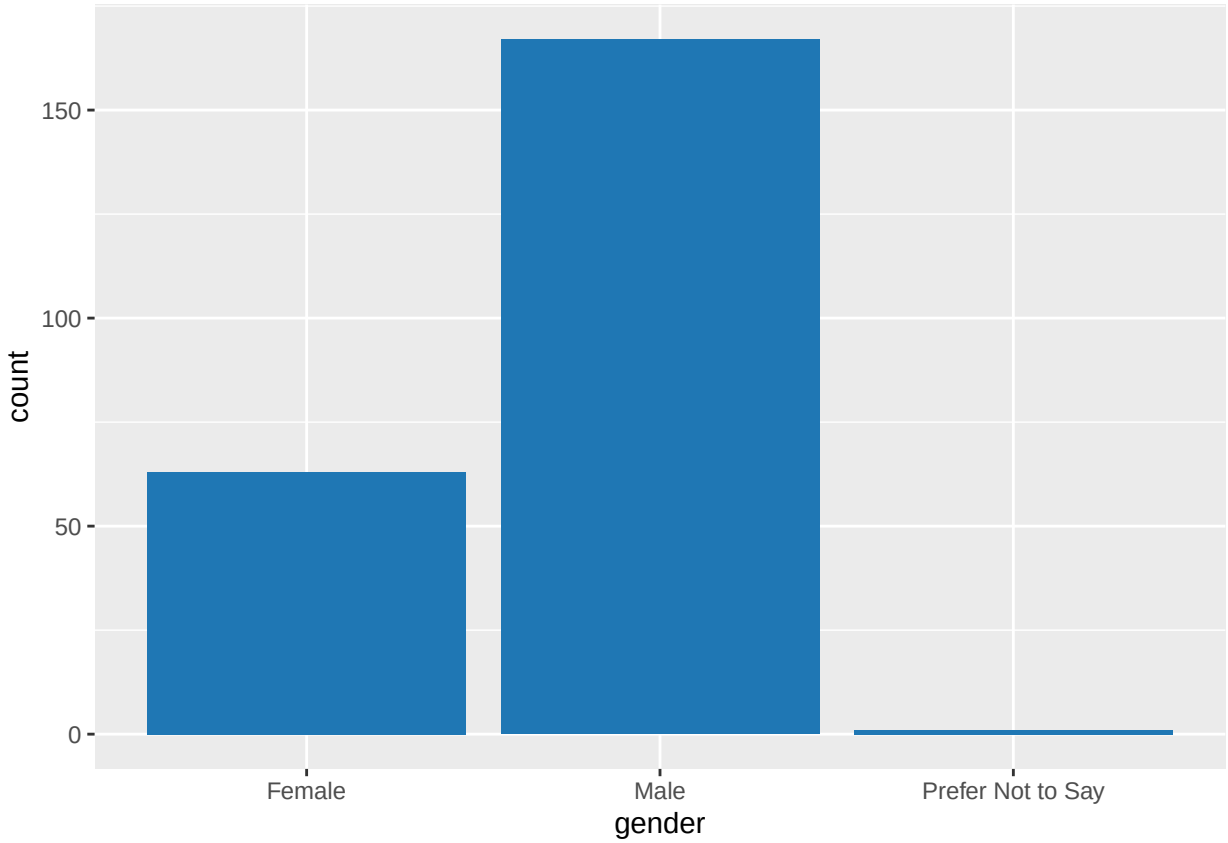
```r
prop.table(table(data$room_improv))
```

```
##
##                     Agree                  Disagree
##                0.76623377                0.08225108
##                Don't Know Neither Agree nor Disagree
##                0.06926407                0.08225108
```

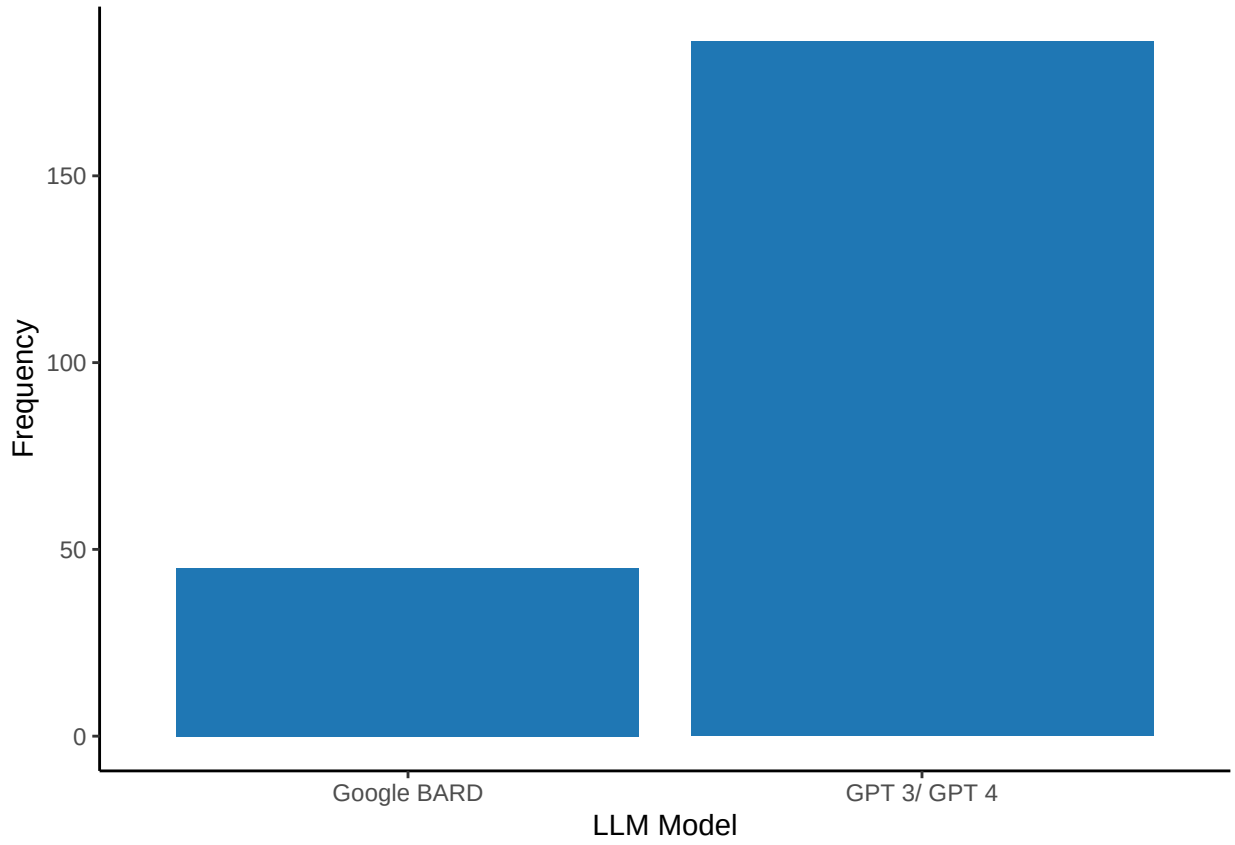
```r
ggplot(data, aes(education)) +
  geom_bar(fill = "#1F77B4")
```




```
ggplot(data, aes(gender)) +
  geom_bar(fill = "#1F77B4")
```

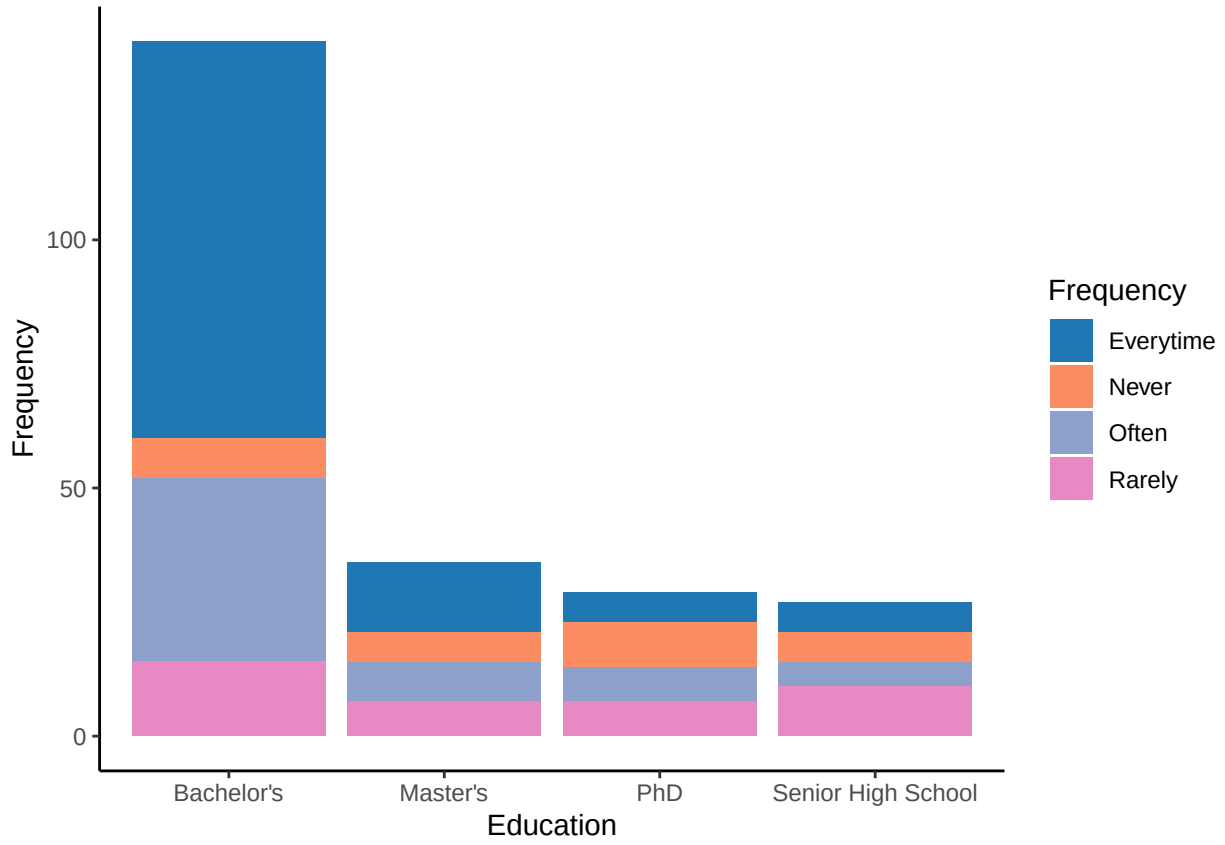



```
ggplot(data, aes(which_AI_prefer)) +
         geom_bar(fill = "#1F77B4") +
  labs(y = "Frequency", x = "LLM Model") +
  theme_classic()
```

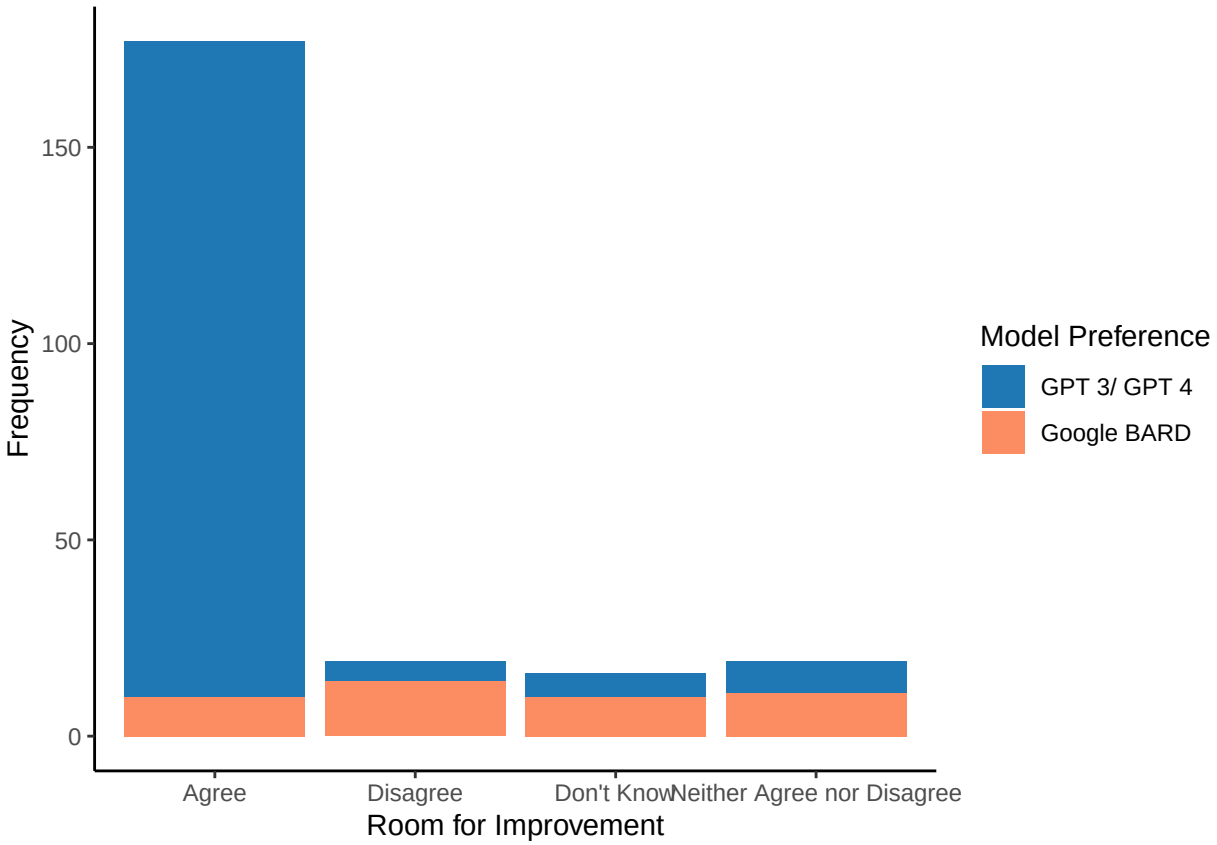



```
colors <- brewer.pal(4, "Set2")
colors[1] <- "#1F77B4"

ggplot(data, aes(education, fill = AI_freq)) +
  geom_bar() +
  theme_classic() +
  labs(x = "Education", y = "Frequency",
       fill = "Frequency") +
  scale_fill_manual(values = colors)
```




```
ggplot(data, aes(room_improv, fill = factor(which_AI_prefer,
                                            levels = c("GPT 3/ GPT 4",
                                                       "Google BARD"
                                                       )))) +
  geom_bar() +
  theme_classic() +
  labs(x = "Room for Improvement", y = "Frequency",
       fill = "Model Preference") +
  scale_fill_manual(values = colors)
```


Frequency
150
100
50
0
Agree
Disagree
Don't Know
Neither Agree nor Disagree
Room for Improvement
Model Preference
GPT 3/ GPT 4
Google BARD