Data exploration

Including packages

```
library(data.table)
```

```
## Warning: package 'data.table' was built under R version 4.2.3
```

```
library(magrittr)
```

```
## Warning: package 'magrittr' was built under R version 4.2.3
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.2.2
```

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✓ tibble  3.1.8      ✓ dplyr   1.0.10
## ✓ tidyr   1.3.0      ✓ stringr 1.5.0
## ✓ readr   2.1.3      ✓ forcats 1.0.0
## ✓ purrr   1.0.1
```

```
## Warning: package 'tidyr' was built under R version 4.2.2
```

```
## Warning: package 'readr' was built under R version 4.2.2
```

```
## Warning: package 'purrr' was built under R version 4.2.2
```

```
## Warning: package 'dplyr' was built under R version 4.2.2
```

```
## Warning: package 'stringr' was built under R version 4.2.2
```

```
## Warning: package 'forcats' was built under R version 4.2.2
```

```
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::between()   masks data.table::between()
## ✖ tidyr::extract()   masks magrittr::extract()
## ✖ dplyr::filter()    masks stats::filter()
## ✖ dplyr::first()     masks data.table::first()
## ✖ dplyr::lag()       masks stats::lag()
## ✖ dplyr::last()      masks data.table::last()
## ✖ purrr::set_names() masks magrittr::set_names()
## ✖ purrr::transpose() masks data.table::transpose()
```

```
library(patchwork)
```

```
## Warning: package 'patchwork' was built under R version 4.2.3
```

Fetch data

Read CSV File

```
survey_df <- read.csv("from_yao/SimulatedUsers-Final_August4.csv")
# survey_df
```

Convert to `data.table`

```
dt <- data.table(survey_df)
survey_dt <- dt[-(1:2)]
# summary(survey_dt)
```

Data Wrangling

StartDate, EndDate and RecordedDate are dates

```
survey_dt$StartDate <- as.Date(survey_dt$StartDate)
```

Demographic Questions

```
survey_dt$D1 <- as.integer(survey_dt$D1)
survey_dt$D2 <- as.factor(survey_dt$D2)
survey_dt$D3 <- as.factor(survey_dt$D3)
```

Single choice predictions - Birds

```
bird_options <- c("Crested Auklet", "Least Auklet", "Parakeet Auklet", "Rhinoceros Auklet")
numbercol_to_birdnames <- function(s) { factor(s, levels = c("1", "2", "3", "4"), labels = bird_options) }
survey_dt <- survey_dt %>% mutate(across(c(Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, Q10, Q11, Q12, Q13, Q14, Q15, Q16, Q17, Q18,
Q19, Q20), numbercol_to_birdnames))
```

Single Choice predictions - Street signs

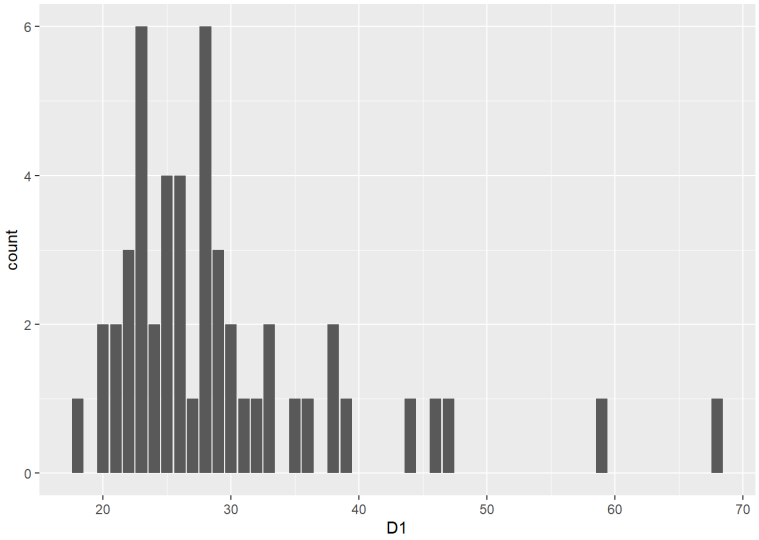
```
sign_options <- c("Left-bend", "Right-bend", "Ahead-or-left", "Ahead-or-right")
numbercol_to_signnames <- function(s) { factor(s, levels = c("1", "2", "3", "4"), labels = sign_options) }
survey_dt <- survey_dt %>% mutate(across(c(Q1.1, Q2.1, Q3.1, Q4.1, Q5.1, Q6.1, Q7.1, Q8.1, Q9.1, Q10.1, Q11.1, Q12.1, Q13.1,
Q14.1, Q15.1, Q16.1, Q17.1, Q18.1, Q19.1, Q20.1), numbercol_to_signnames))
```

Visualizations

Demographics

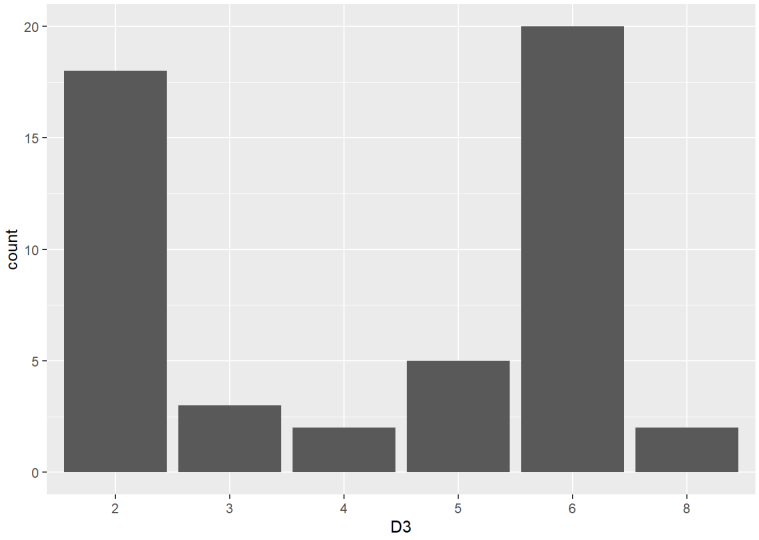
Ages

```
ggplot(survey_dt, aes(x = D1)) + geom_bar()
```



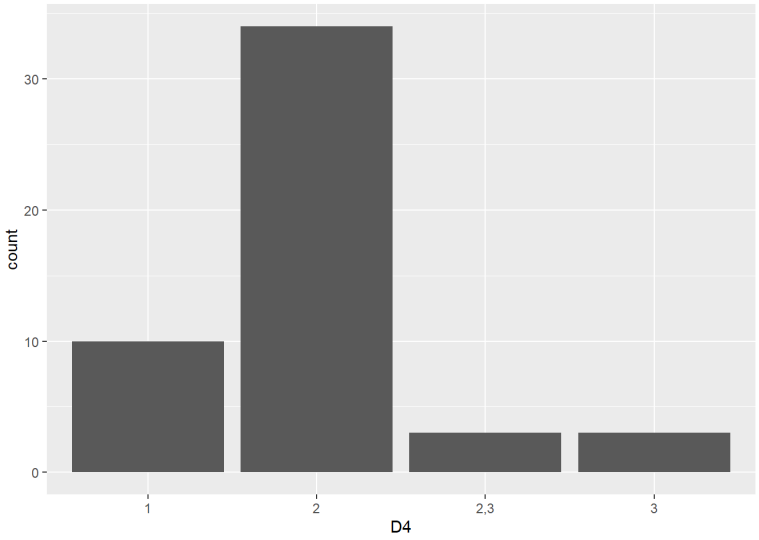
Employment

```
ggplot(survey_dt, aes(x = D3)) + geom_bar()
```



Prior Experience

```
ggplot(survey_dt, aes(x = D4)) + geom_bar()
```



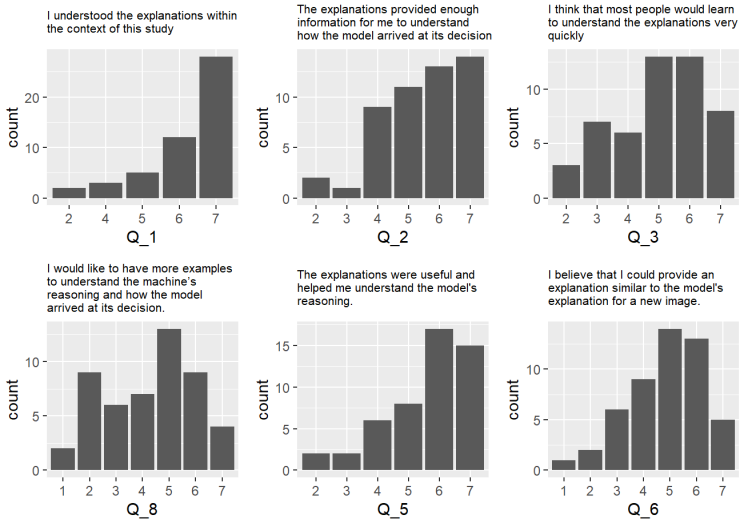
Bird Species

Guessing Predictions

```
# TODO
```

Subjective Understanding

```
sub1_p <- ggplot(survey_dt, aes(x = Q_1)) + geom_bar() + ggtitle("I understood the explanations within\nthe context of this
study") + theme(plot.title = element_text(size=8))
sub2_p <-ggplot(survey_dt, aes(x = Q_2)) + geom_bar() + ggtitle("The explanations provided enough\ninformation for me to und
erstand\nhow the model arrived at its decision") + theme(plot.title = element_text(size=8))
sub3_p <-ggplot(survey_dt, aes(x = Q_3)) + geom_bar() + ggtitle("I think that most people would learn\nto understand the exp
lanations very\nquickly") + theme(plot.title = element_text(size=8))
sub4_p <-ggplot(survey_dt, aes(x = Q_8)) + geom_bar() + ggtitle("I would like to have more examples\nto understand the machi
ne's\nreasoning and how the model\narrived at its decision.") + theme(plot.title = element_text(size=8))
sub5_p <-ggplot(survey_dt, aes(x = Q_5)) + geom_bar() + ggtitle("The explanations were useful and\nhelped me understand the
model's\nreasoning.") + theme(plot.title = element_text(size=8))
sub6_p <-ggplot(survey_dt, aes(x = Q_6)) + geom_bar() + ggtitle("I believe that I could provide an\nexplanation similar to t
he model's\nexplanation for a new image.") + theme(plot.title = element_text(size=8))
sub1_p + sub2_p + sub3_p + sub4_p + sub5_p + sub6_p + plot_layout(ncol = 3)
```



Street Signs

Guessing Predictions

```
# TODO
```

Subjective Understanding

```
sub1.1_p <- ggplot(survey_dt, aes(x = Q_1.1)) + geom_bar() + ggtitle("I understood the explanations within\nthe context of t
his study") + theme(plot.title = element_text(size=8))
sub2.1_p <-ggplot(survey_dt, aes(x = Q_2.1)) + geom_bar() + ggtitle("The explanations provided enough\ninformation for me to
understand\nhow the model arrived at its decision") + theme(plot.title = element_text(size=8))
sub3.1_p <-ggplot(survey_dt, aes(x = Q_3.1)) + geom_bar() + ggtitle("I think that most people would learn\nto understand the
explanations very\nquickly") + theme(plot.title = element_text(size=8))
sub4.1_p <-ggplot(survey_dt, aes(x = Q_8.1)) + geom_bar() + ggtitle("I would like to have more examples\nto understand the m
achine's\nreasoning and how the model\narrived at its decision.") + theme(plot.title = element_text(size=8))
sub5.1_p <-ggplot(survey_dt, aes(x = Q_5.1)) + geom_bar() + ggtitle("The explanations were useful and\nhelped me understand
the model's\nreasoning.") + theme(plot.title = element_text(size=8))
sub6.1_p <-ggplot(survey_dt, aes(x = Q_6.1)) + geom_bar() + ggtitle("I believe that I could provide an\nexplanation similar
to the model's\nexplanation for a new image.") + theme(plot.title = element_text(size=8))
sub1.1_p + sub2.1_p + sub3.1_p + sub4.1_p + sub5.1_p + sub6.1_p + plot_layout(ncol = 3)
```