Effects of Question Type on E-Learning Performance

Hung Hong, Lane Harrison

December 12, 2019

Analysis of data collected from e-Learning Performance experiment run

References:
- https://github.com/mjskay/tidybayes
- @codementum

Libraries needed

```
library("jsonlite")
library(RCurl)
library(plyr)
library(tidyverse)
library(sjPlot)
library(sjmisc)
```

Grab the JSON file from firebase and convert into a list in R

```
tables <- read_json("https://elearning-20e46.firebaseio.com/.json", simplifyVector = TRUE)
# str(tables)
# tables$Session
```

Create the `trials` tibble

```
trials = tibble()
for (quiz in tables$quiz) {
  for (observation in quiz) {
      new_row <- as_tibble(observation)
      trials <- trials %>% bind_rows(new_row)
    # trials <- rbind(trials_df, data.frame(observation))
  }
}

# trials
```

Create the `surveys` tibble

```
surveys = tibble()
for (survey in tables$survey) {
    new_row <- as_tibble(survey)
    surveys <- surveys %>% bind_rows(new_row)
}

# surveys
```

Clean the data to make later function calls easier

Create `condition` column from `quiz` (to easily filter by condition)

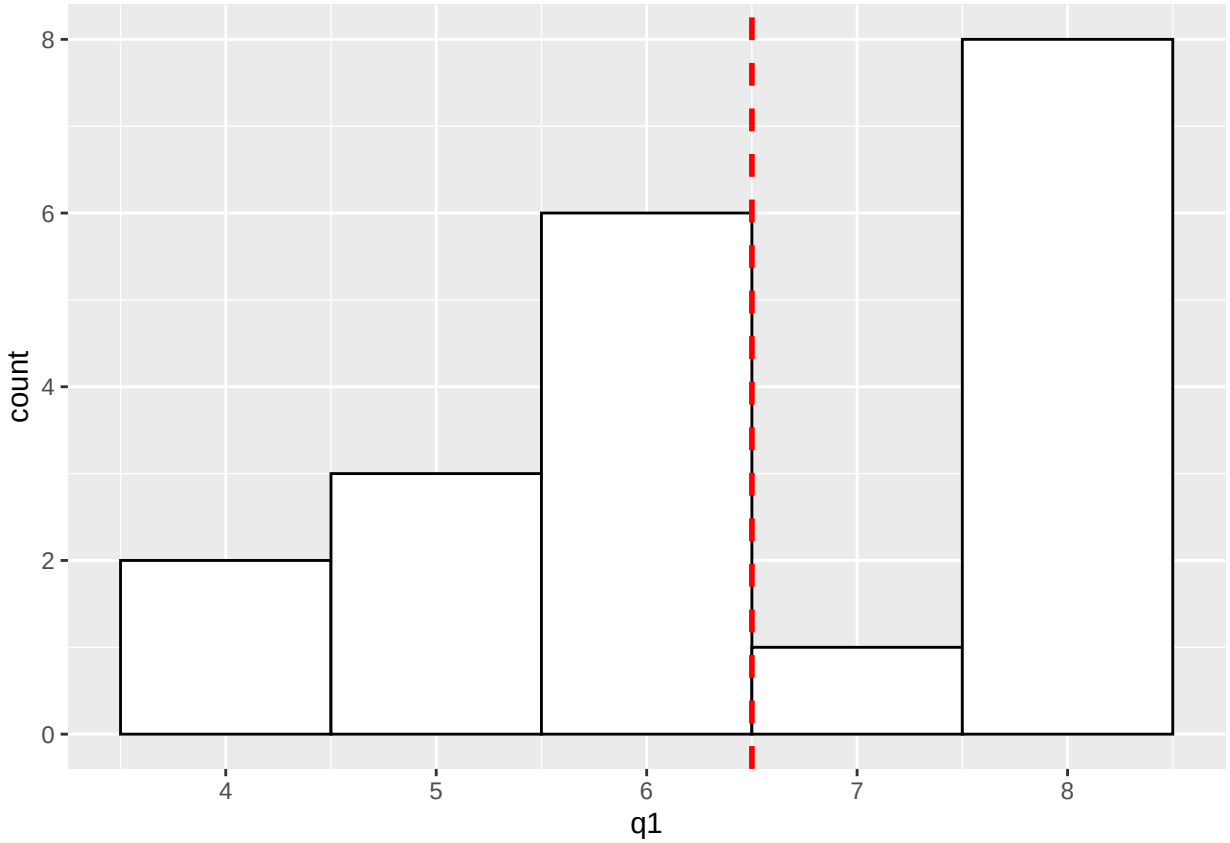
```
trials$condition <- NA
for (i in seq_along(trials$quiz)) {
  if (trials$quiz[i] == 1)
    condition <- 'Quiz 1'
  else if (trials$quiz[i] == 2)
    condition <- 'Quiz 2'
  else
    condition <- '???'

  trials$condition[i] <- condition
}
```

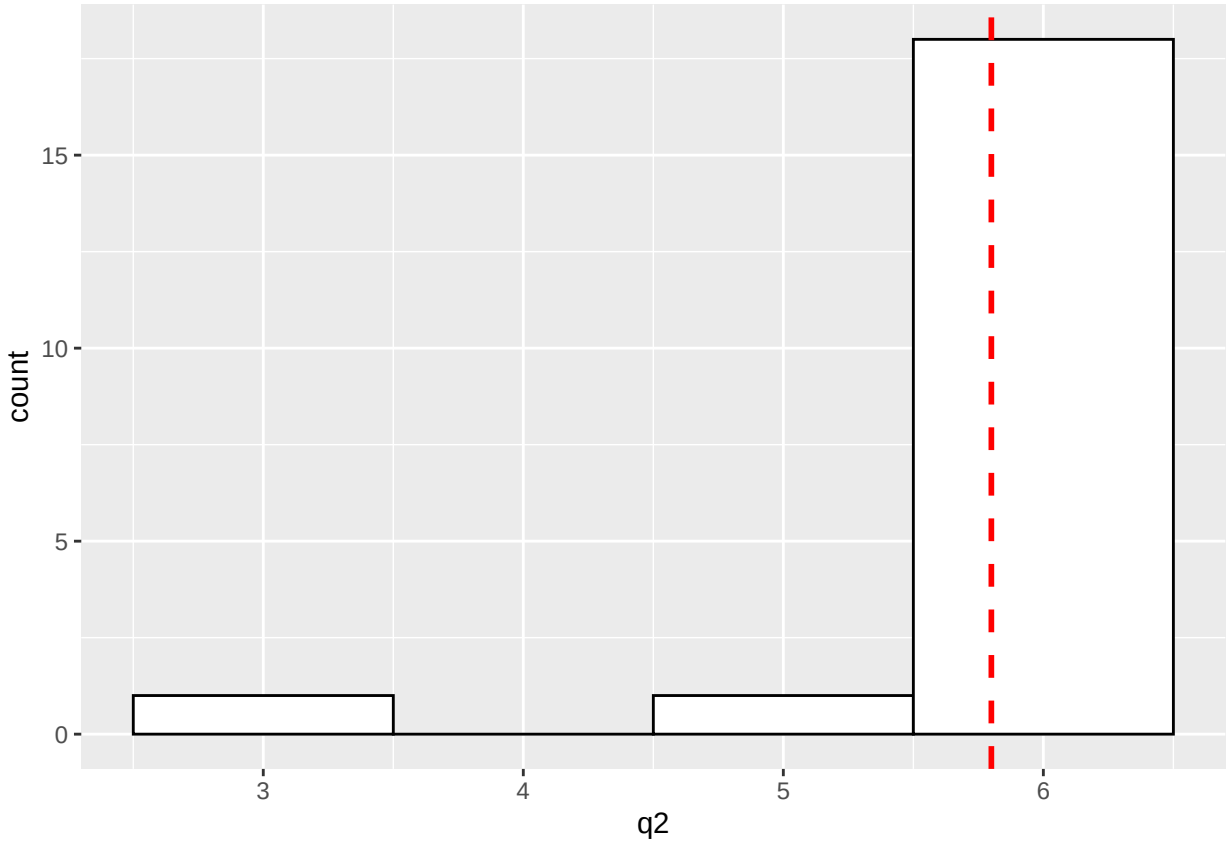
Statistical Analysis of Quiz Results

Histogram Plots of Quiz 1 Results

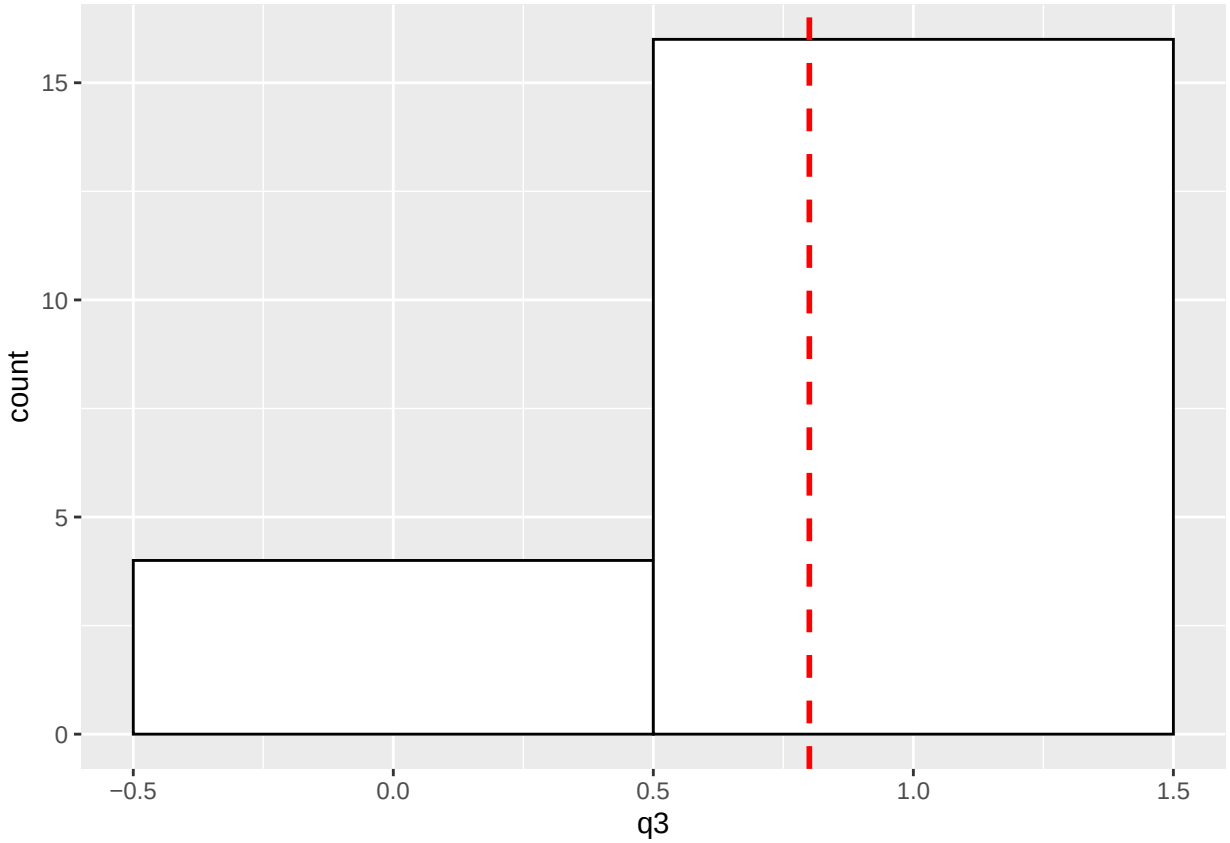
```
quiz1 <- trials %>%
  filter(condition == "Quiz 1")
quiz1 <- quiz1[,1:4]
ggplot(quiz1, aes(x=q1)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q1, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```



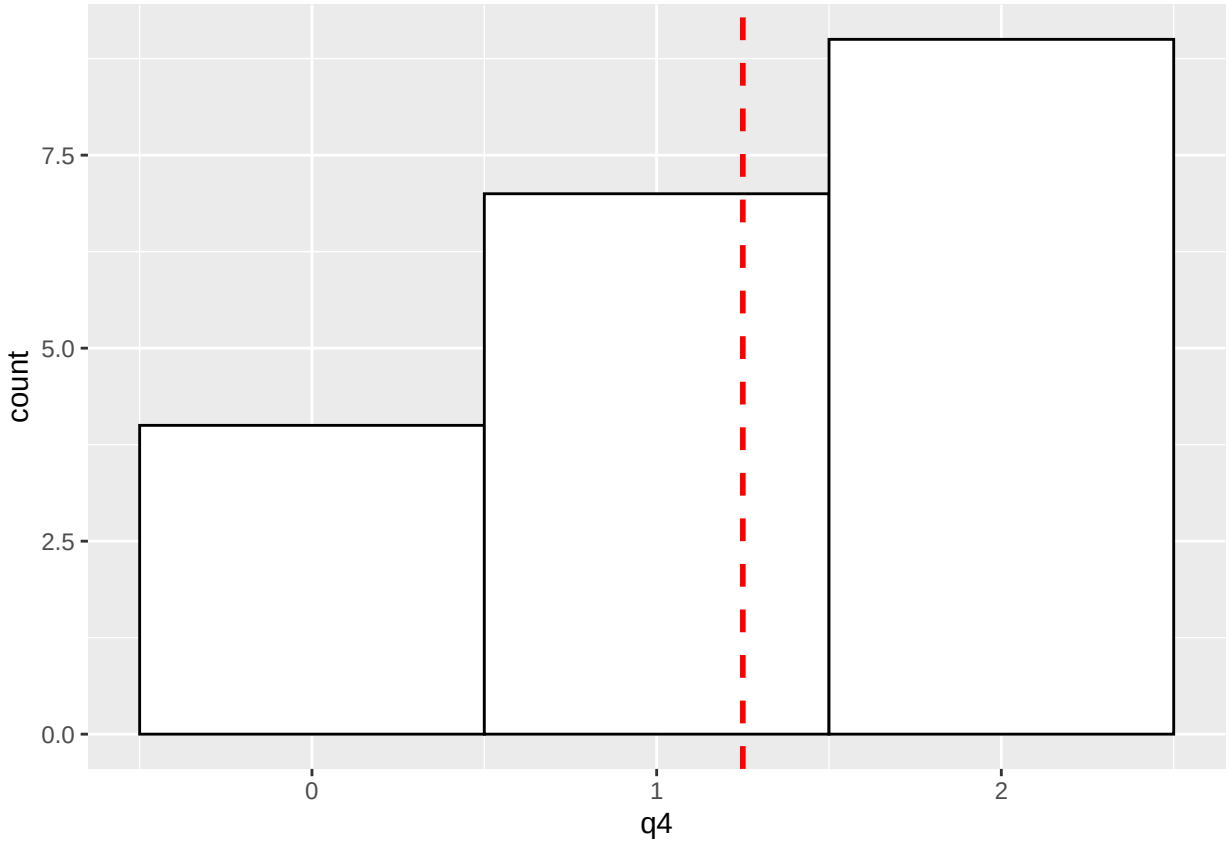

```
ggplot(quiz1, aes(x=q2)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q2, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```




```
ggplot(quiz1, aes(x=q3)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q3, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```




```
ggplot(quiz1, aes(x=q4)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q4, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```

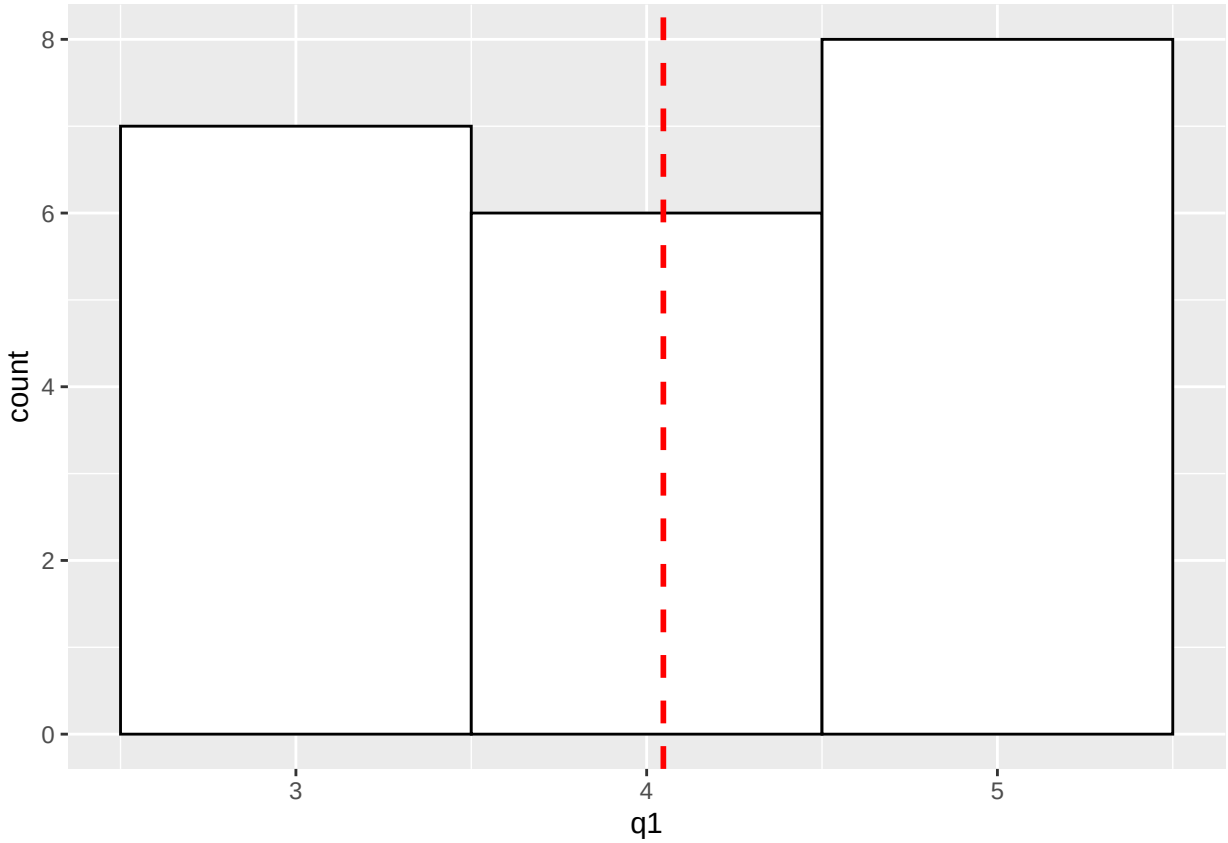



```
summary(quiz1)
```

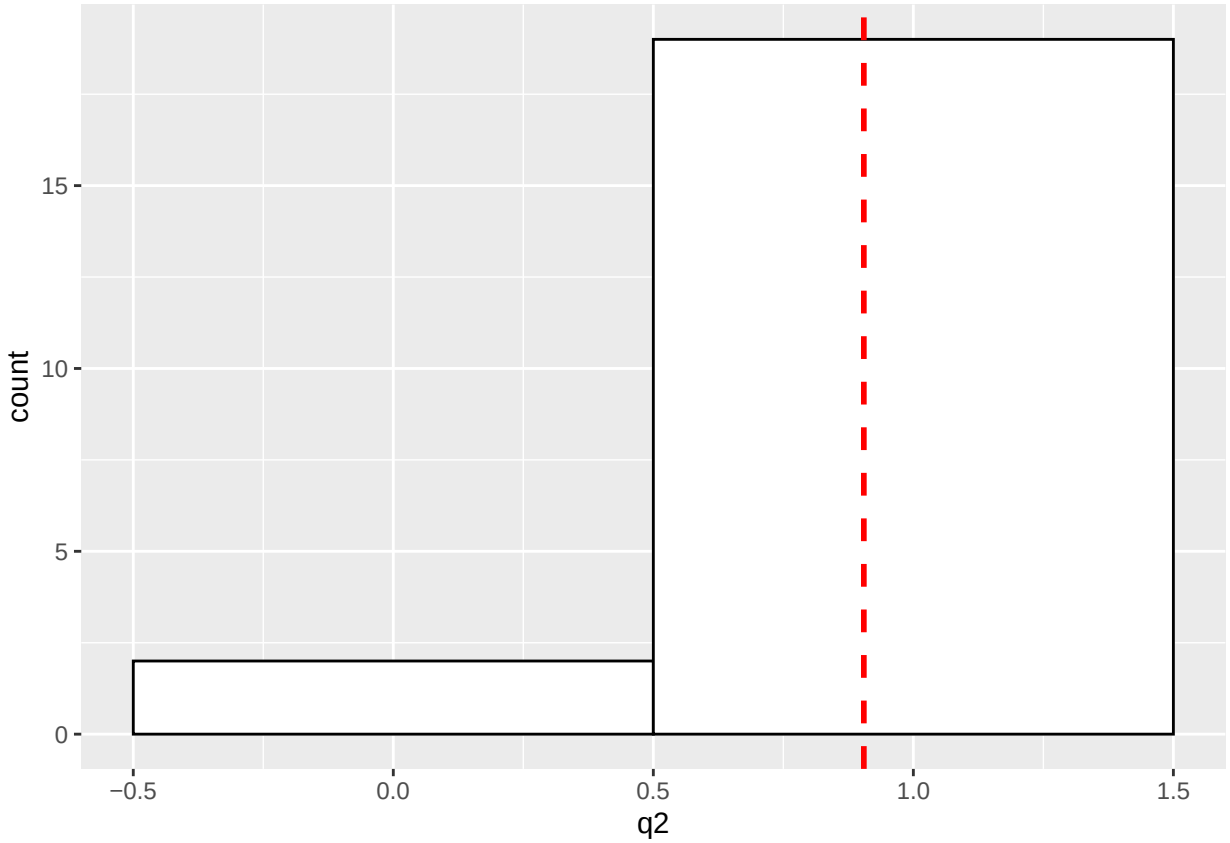
```
##        q1              q2              q3              q4
##  Min.   :4.00   Min.   :3.0   Min.   :0.0   Min.   :0.00
##  1st Qu.:5.75   1st Qu.:6.0   1st Qu.:1.0   1st Qu.:1.00
##  Median :6.00   Median :6.0   Median :1.0   Median :1.00
##  Mean   :6.50   Mean   :5.8   Mean   :0.8   Mean   :1.25
##  3rd Qu.:8.00   3rd Qu.:6.0   3rd Qu.:1.0   3rd Qu.:2.00
##  Max.   :8.00   Max.   :6.0   Max.   :1.0   Max.   :2.00
```

Histogram Plots of Quiz 2 Results

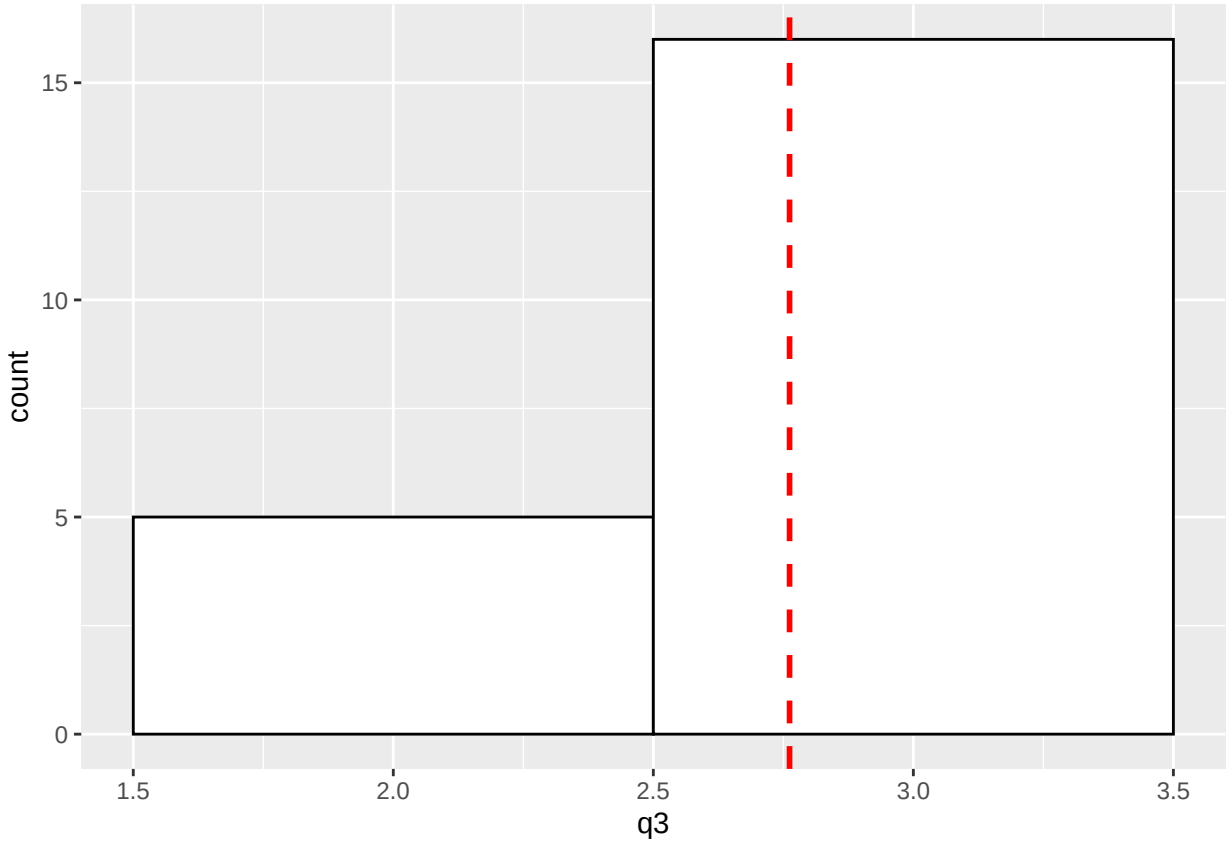
```
quiz2 <- trials %>%
  filter(condition == "Quiz 2")
quiz2 <- quiz2[,1:4]
ggplot(quiz2, aes(x=q1)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q1, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```



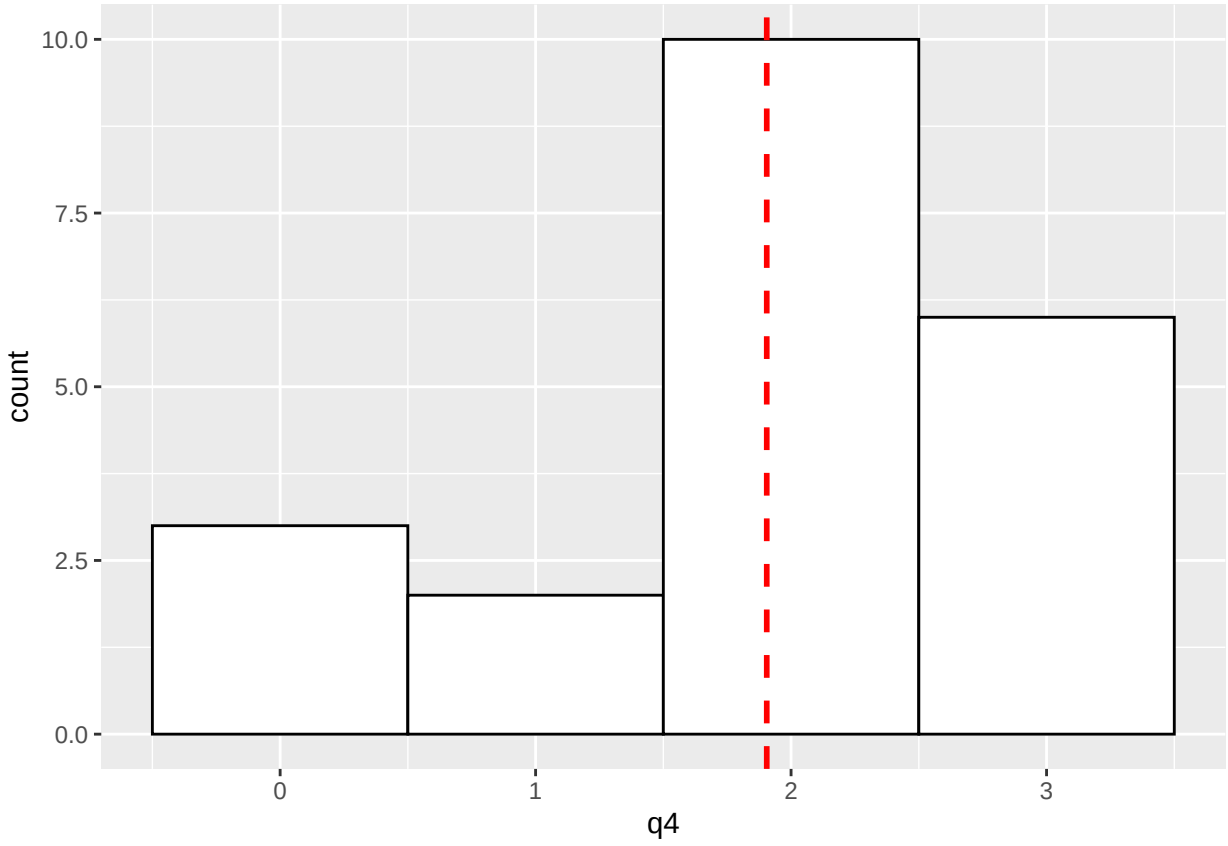

```
ggplot(quiz2, aes(x=q2)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q2, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```




```
ggplot(quiz2, aes(x=q3)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q3, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```




```
ggplot(quiz2, aes(x=q4)) + geom_histogram(binwidth=1, colour="black", fill="white") +
    geom_vline(aes(xintercept=mean(q4, na.rm=T)),   # Ignore NA values for mean
               color="red", linetype="dashed", size=1)
```




```
summary(quiz2)
```

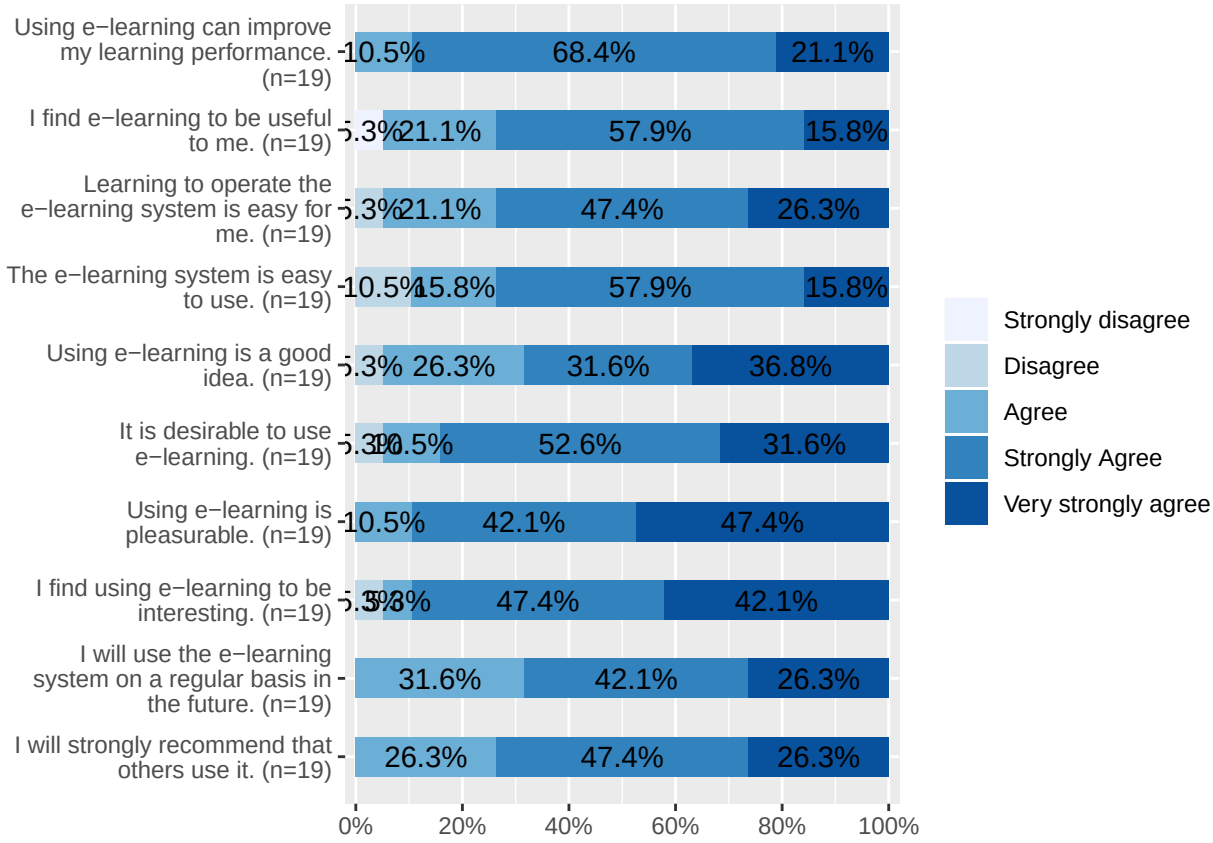
```
##        q1              q2              q3              q4
##  Min.   :3.000   Min.   :0.0000   Min.   :2.000   Min.   :0.000
##  1st Qu.:3.000   1st Qu.:1.0000   1st Qu.:3.000   1st Qu.:2.000
##  Median :4.000   Median :1.0000   Median :3.000   Median :2.000
##  Mean   :4.048   Mean   :0.9048   Mean   :2.762   Mean   :1.905
##  3rd Qu.:5.000   3rd Qu.:1.0000   3rd Qu.:3.000   3rd Qu.:3.000
##  Max.   :5.000   Max.   :1.0000   Max.   :3.000   Max.   :3.000
```

Due to the simple nature of the assignments, most participants perform well on the quizzes. It is noticeable that participants performed better in questions that contain answer as one of the options, such as Question 2 of Quiz 1 and Question 2, 3 of Quiz 3. We also notice a significantly lower performance for Question 1, 4 of Quiz 1 and Question 4 of Quiz 4, with the similarity of them being free response question with multiple answer fields.

Statistical Analysis of Survey Results

Stack Frequencies Plot of Opnions on E-Learning Platform

```
generalSurveys = surveys %>% select("q1", "q2", "q3", "q4", "q5", "q6", "q7", "q8", "q9", "q10")
surveyMat = data.matrix(generalSurveys)
generalSurveys <- as.data.frame(surveyMat)
rating <- list("Strongly disagree", "Disagree", "Agree", "Strongly Agree", "Very strongly agree")
questions <- list("Using e-learning can improve my learning performance.", "I find e-learning to be use
plot_stackfrq(generalSurveys, legend.labels = rating, axis.labels = rev(questions))
```



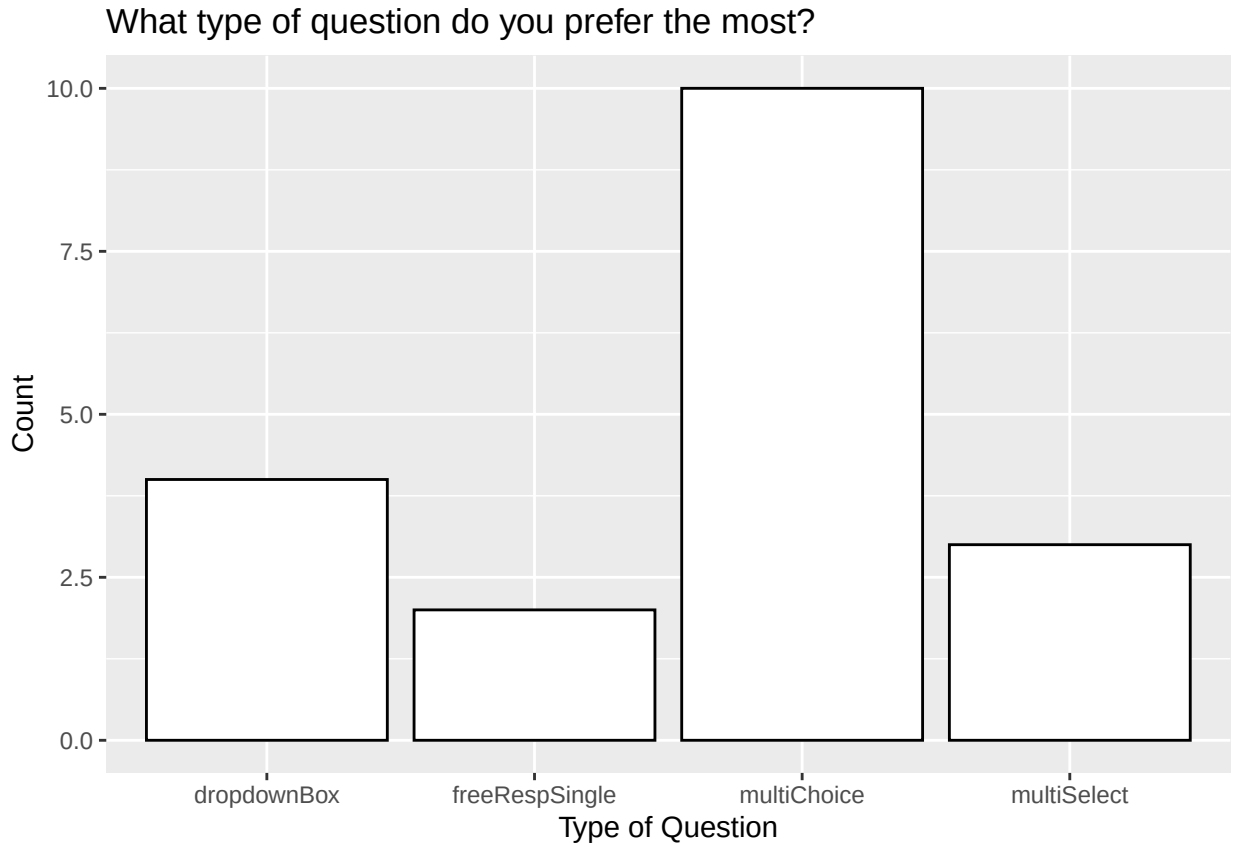
Despite being a Likert of 6, ranging from "Very strongly disagree" to "Very strongly agree", none of the participants opted for "Very strongly disagree" option, hence the stacked frequencies plot only shows 5 rating levels.

These questions are based from previous studies on e-learning continuance intention, concerning about perceived usefulness, perceived ease of use, attitude, perceived enjoyment, satisfaction and continuance intention. As seen from the plot, most participants show a positive attitude towards the e-learning system. Very few have some negative feedbacks regarding the ease of use of e-learning platform, which can also affect their level of enjoyment when using the platform. Despite that, all participants are willing to continue using the platform in the future and agree that the platform is effective in improving learning performance.

Stack Frequencies Plot of Opnions on Question Types in E-Learning Platform

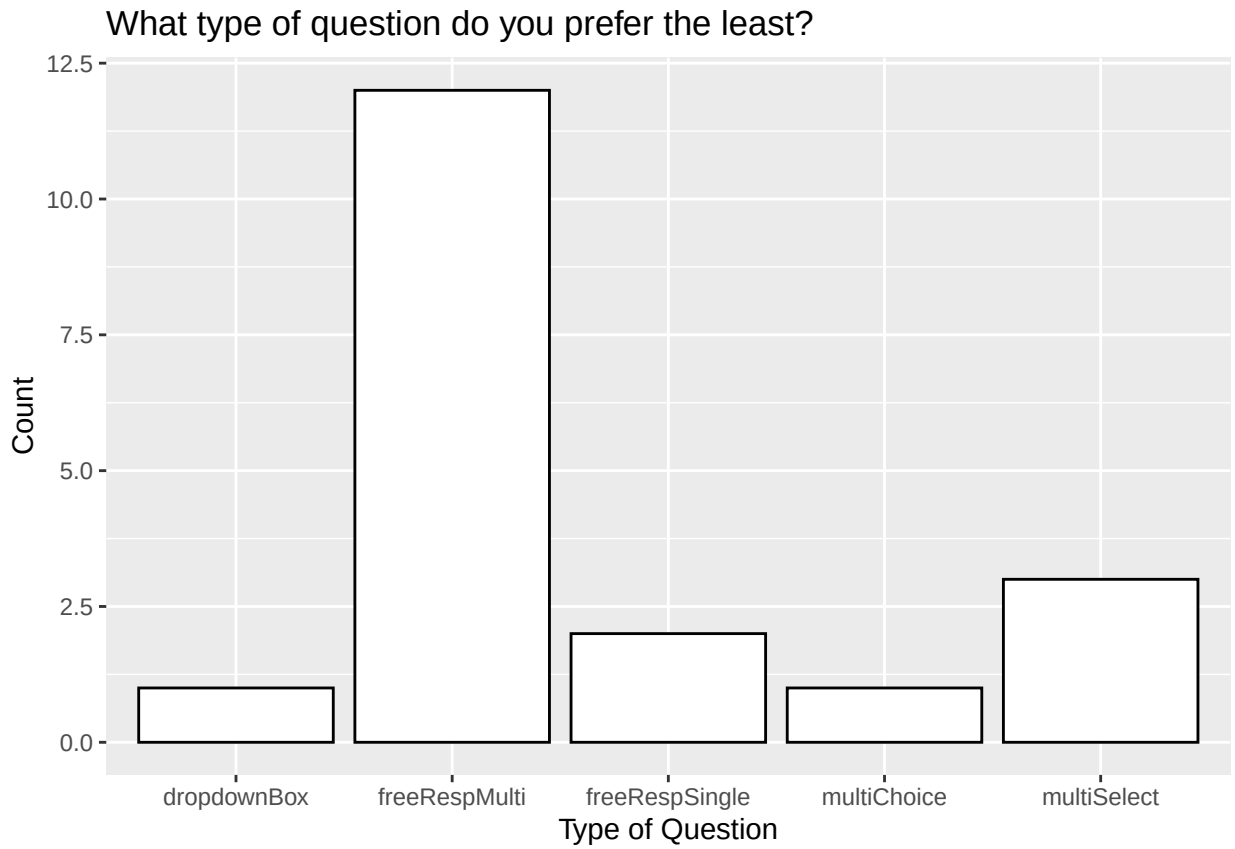
```
qnTypeSurveys = surveys %>% select("q11", "q12", "q13", "q14", "q15", "q16", "q17", "q18", "q19", "q20"
ggplot(qnTypeSurveys, aes(x=q11)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtit
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



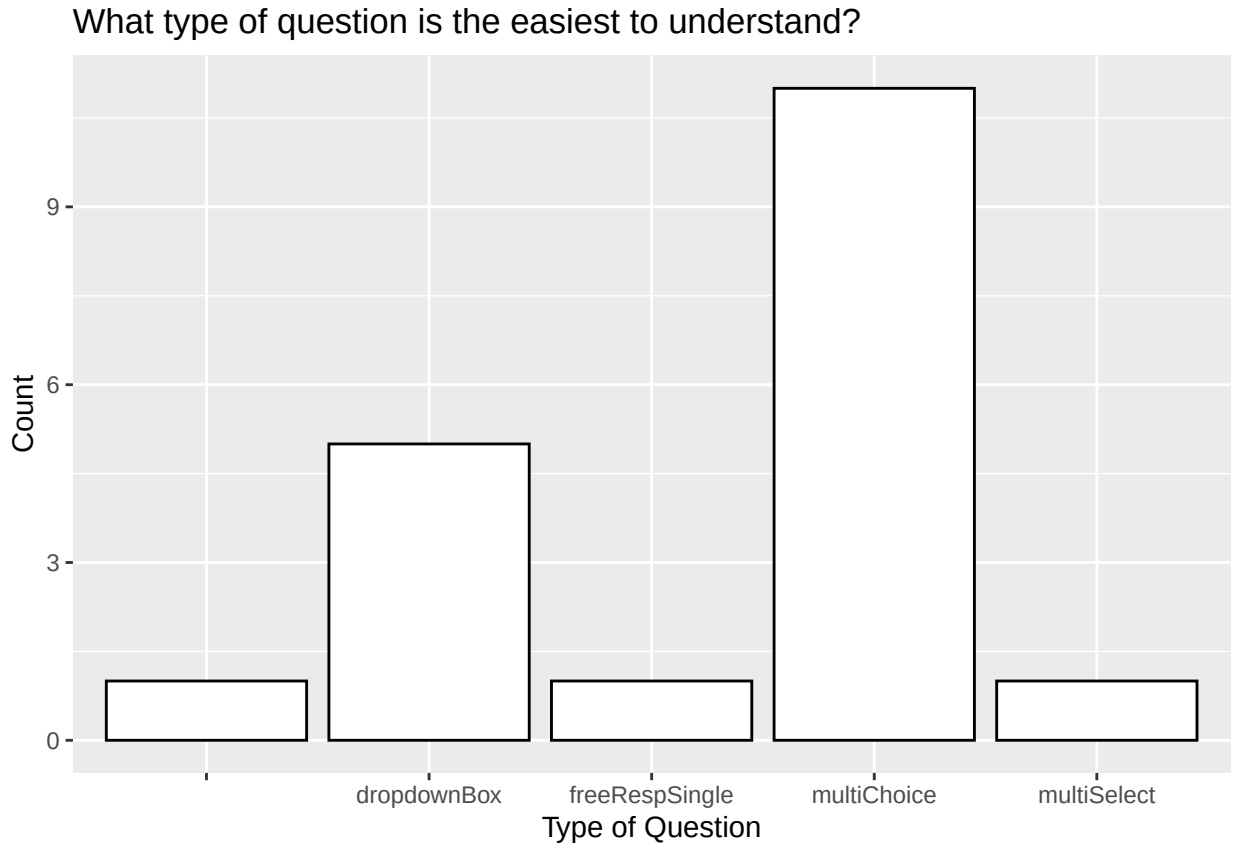

```
ggplot(qnTypeSurveys, aes(x=q12)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



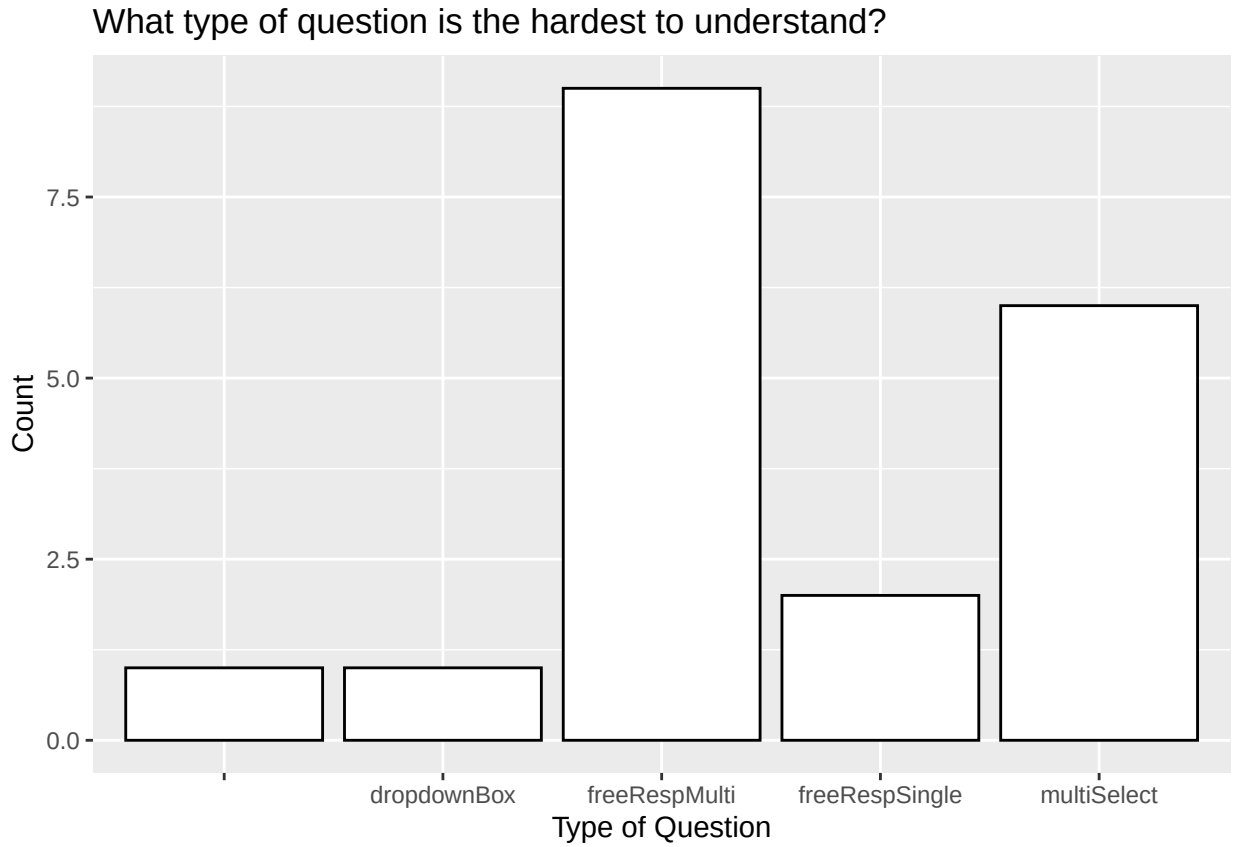

```
ggplot(qnTypeSurveys, aes(x=q13)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



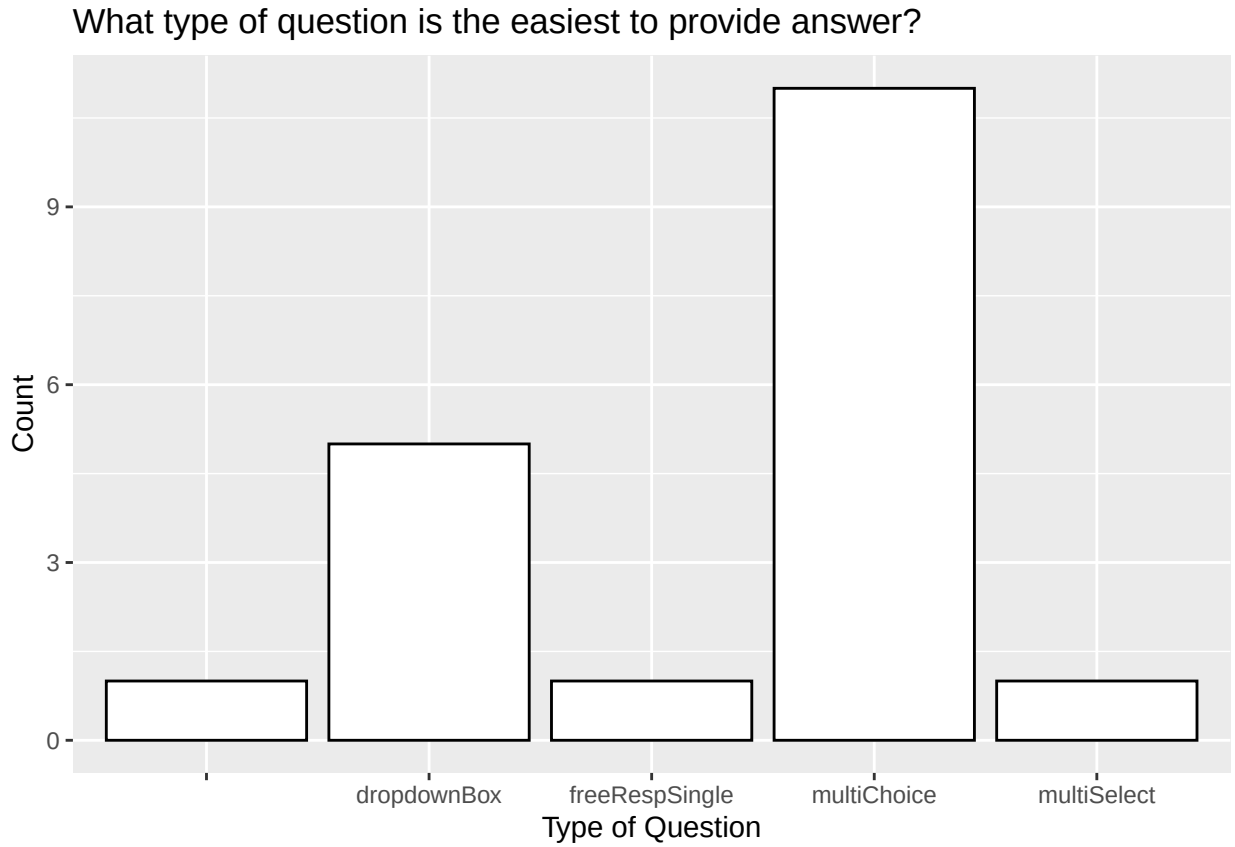

```
ggplot(qnTypeSurveys, aes(x=q14)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



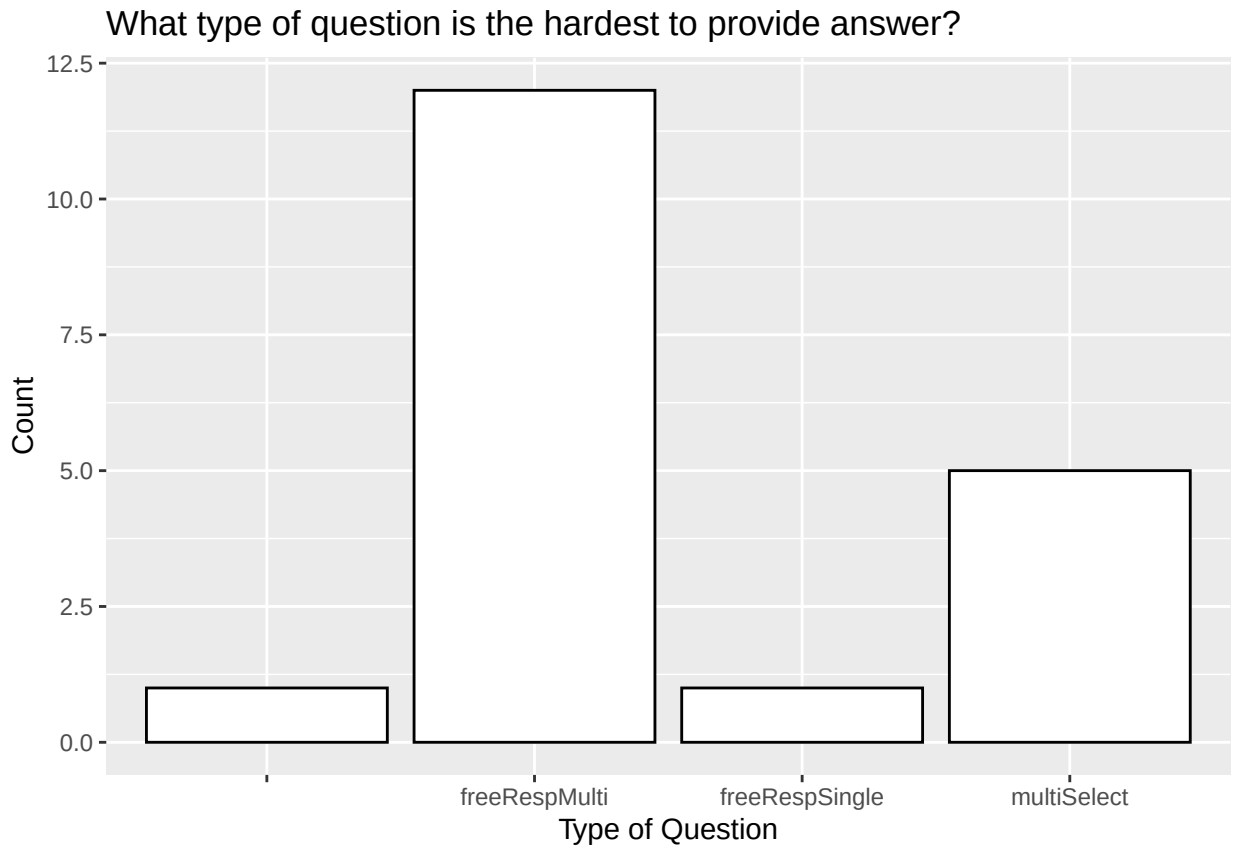

```
ggplot(qnTypeSurveys, aes(x=q15)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



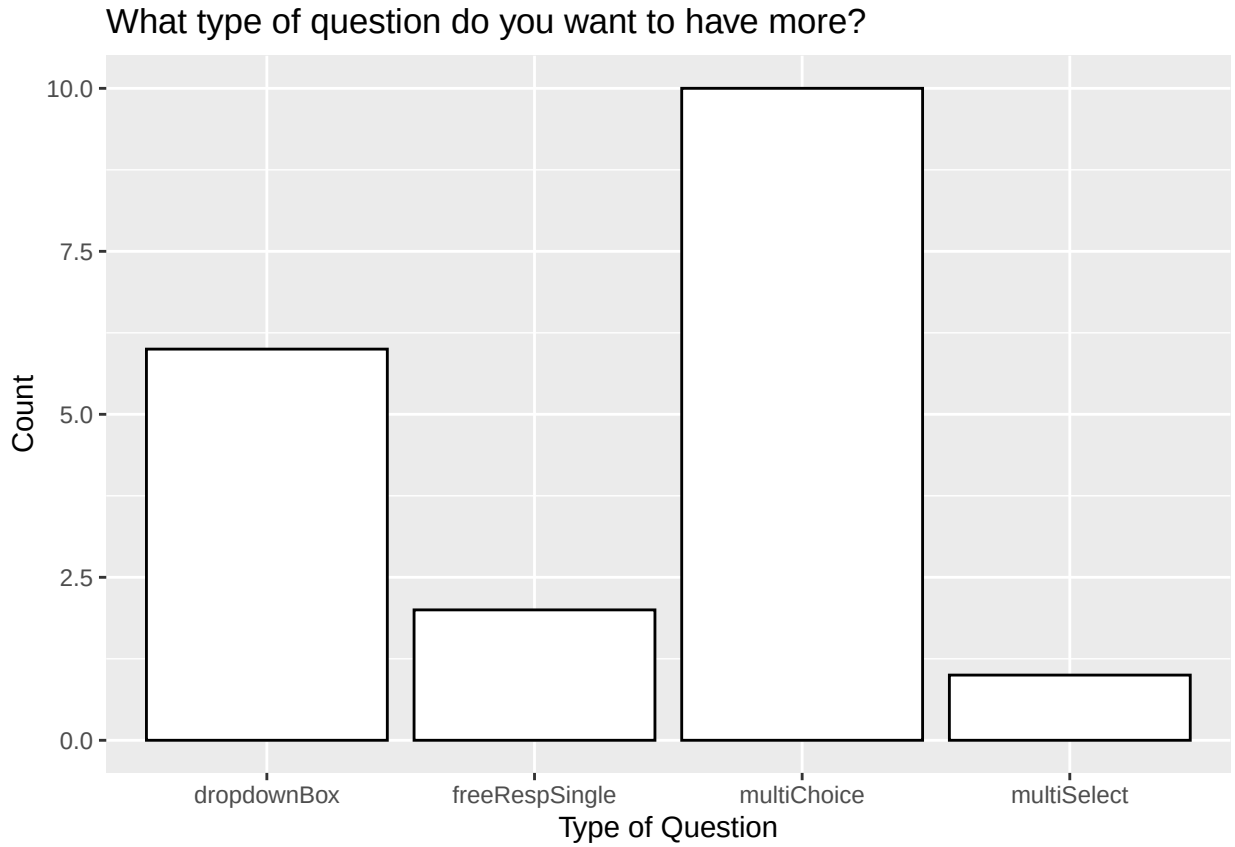

```
ggplot(qnTypeSurveys, aes(x=q16)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



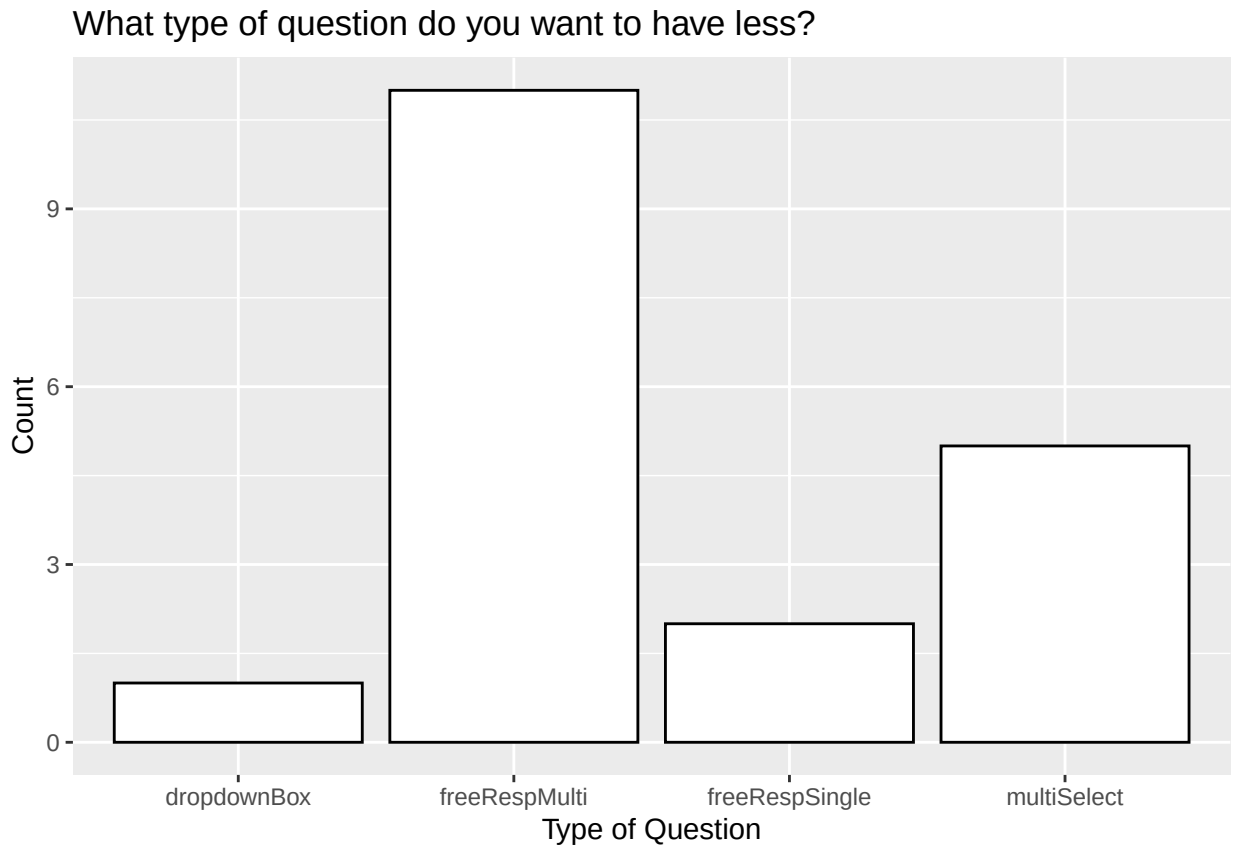

```
ggplot(qnTypeSurveys, aes(x=q17)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



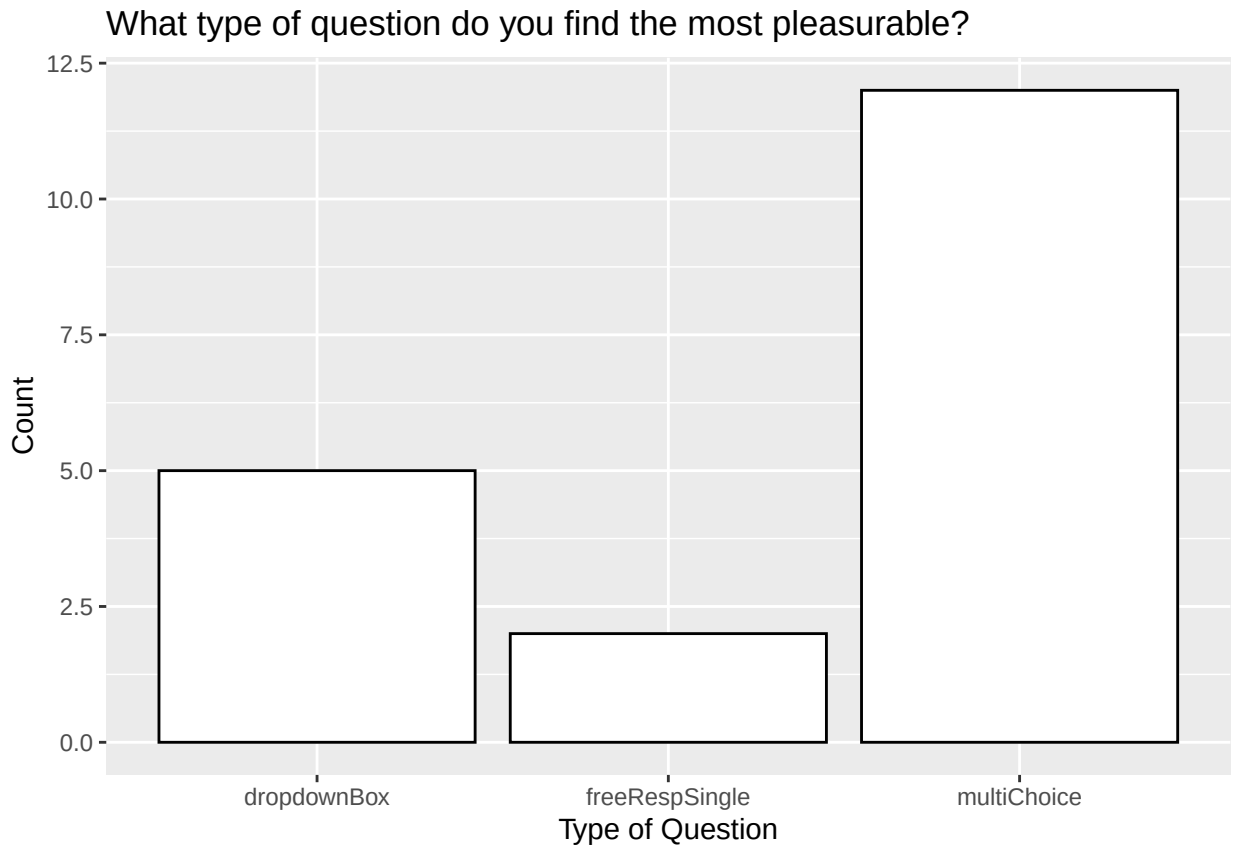

```
ggplot(qnTypeSurveys, aes(x=q18)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



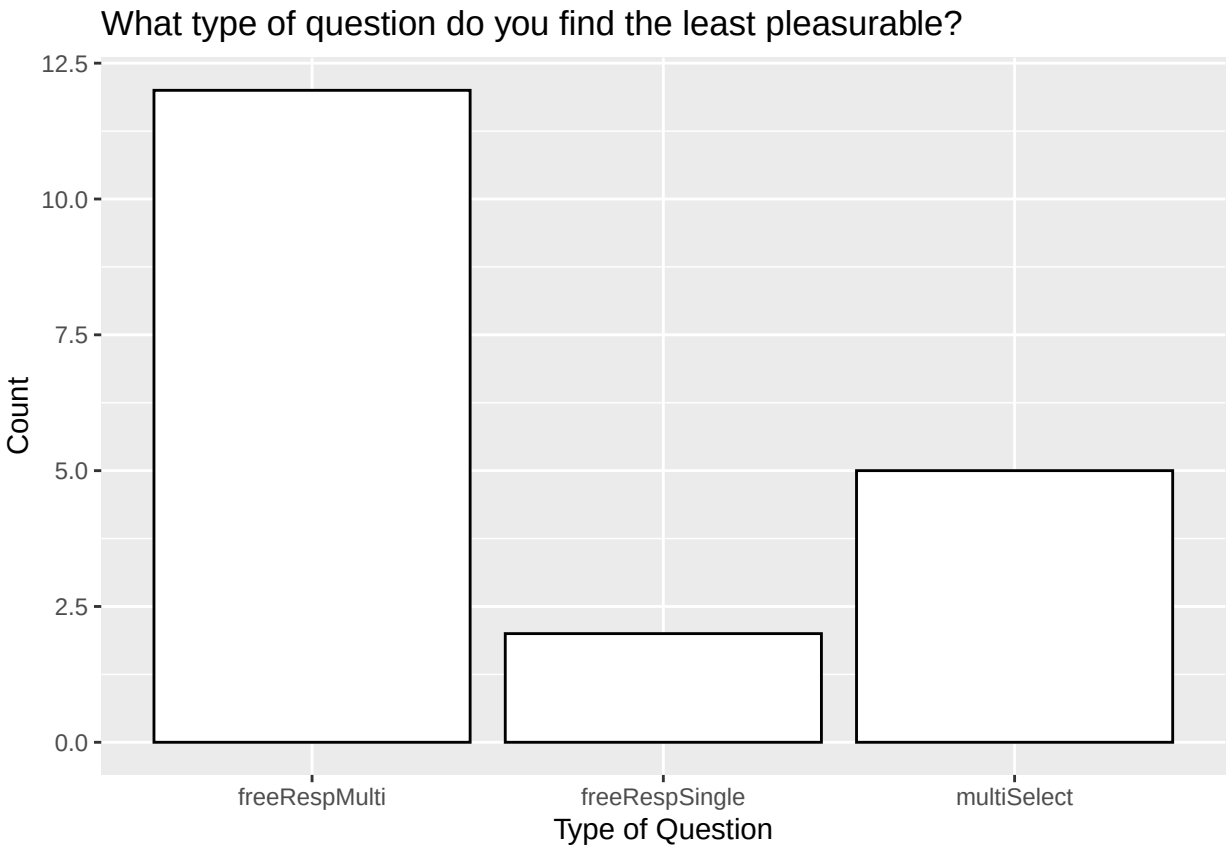

```
ggplot(qnTypeSurveys, aes(x=q19)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```




```
ggplot(qnTypeSurveys, aes(x=q20)) + geom_histogram(colour="black", fill="white",  stat="count") + ggtitl
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```

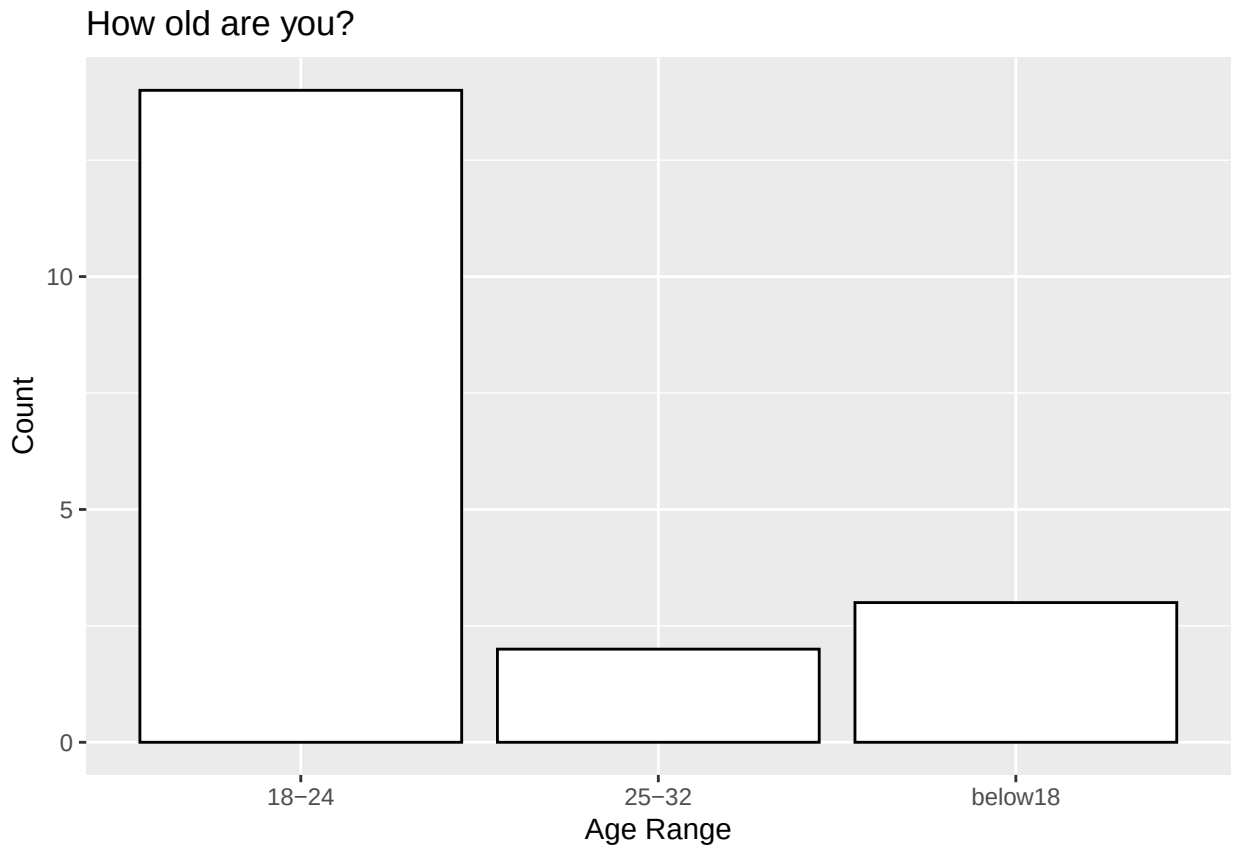


The next part of the survey focuses on the effectiveness of each type of question. Some of the plots have a "blank" bar, indicating that participants did not answer that particular question. These questions are built to concern about user's preference, with each category having both extremes to give a big picture of opinions towards question types. Looking at the plots, we can see that participants show positive feedbacks towards mainly Multiple Choice and Dropdown Box with Multiple Choice being the top choice. Meanwhile, Free Response with multiple answers and Multiple Select are less favored, with the former being the bottom choice in term of preference, easiest to understand and answer, pleasurability, and continuance intention. This results are expected, and they reflect well with the performance from the quizzes.

Demographic of Participants

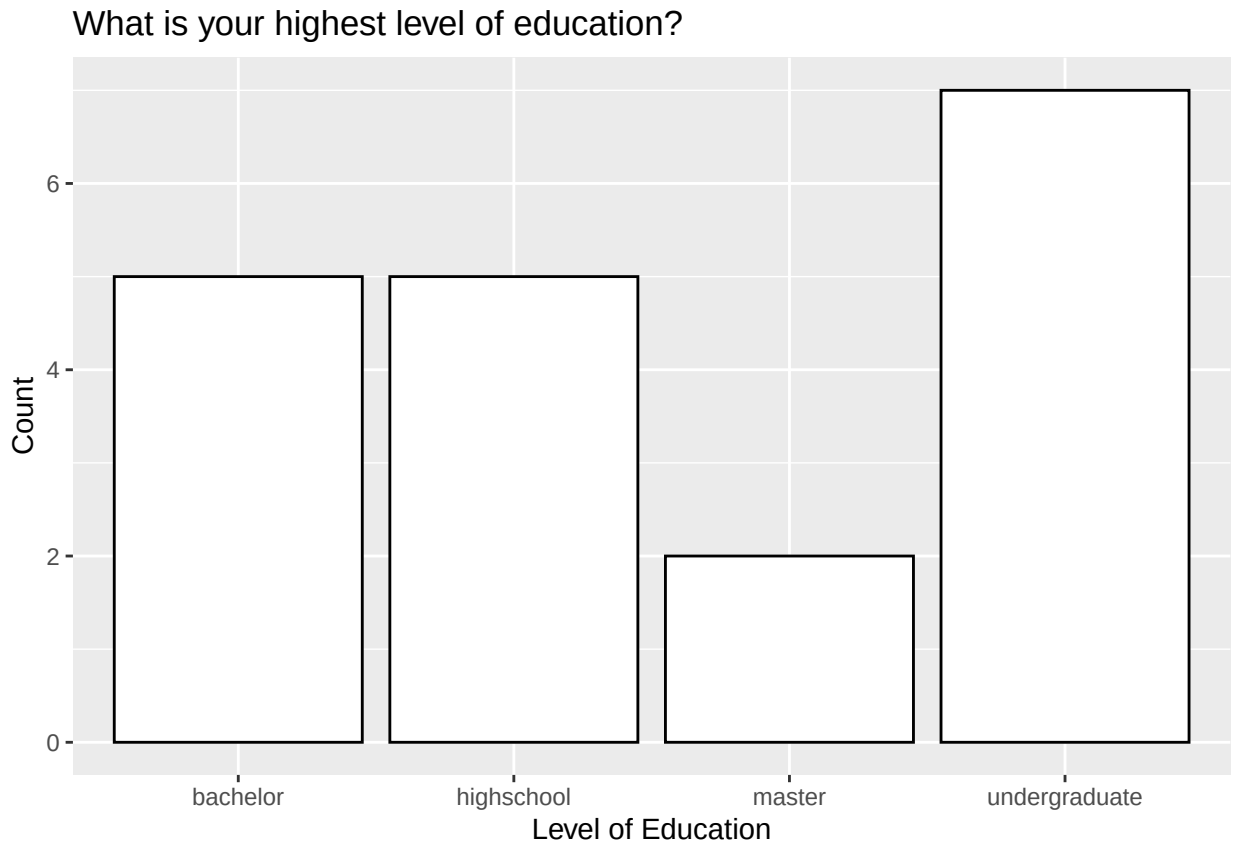
```
ggplot(surveys, aes(x=q21)) + geom_histogram(colour="black", fill="white",  stat="count", ) + ggtitle("
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



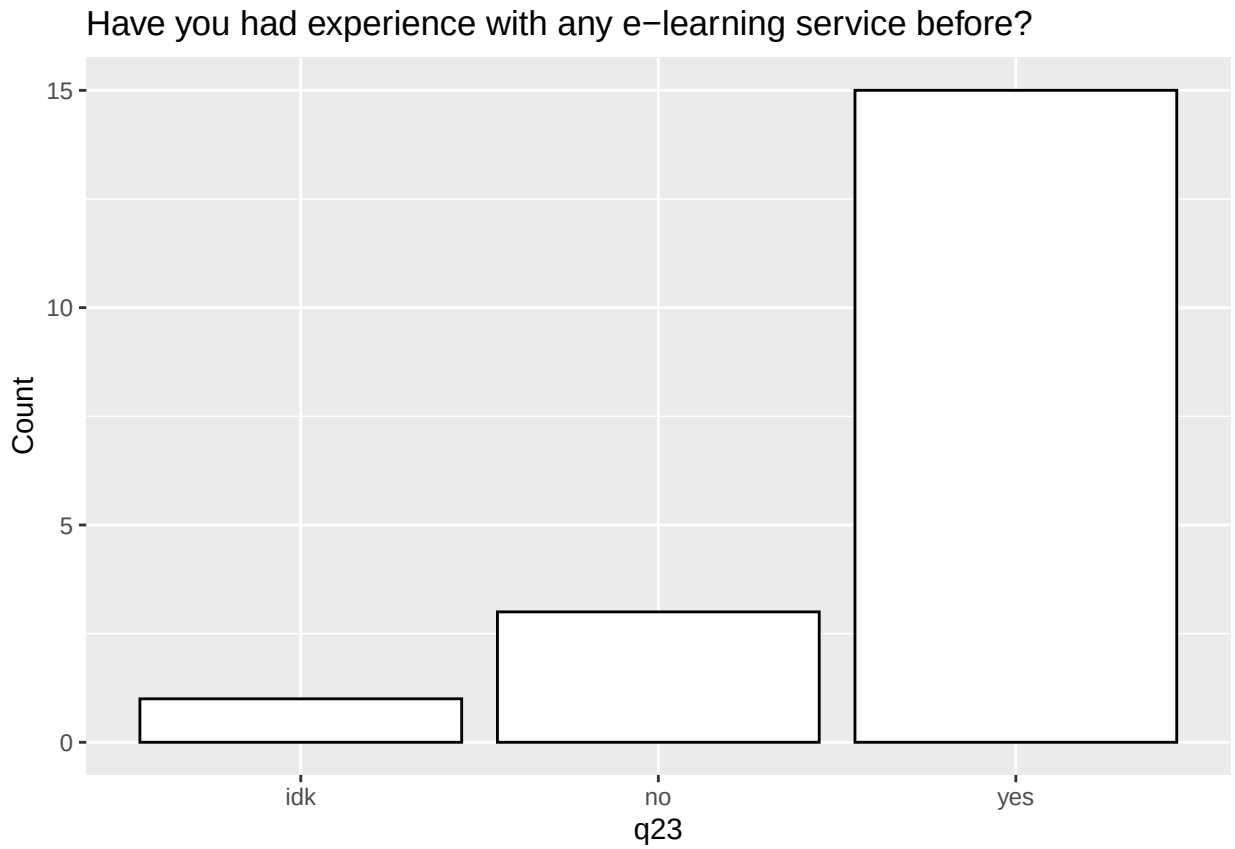

```
ggplot(surveys, aes(x=q22)) + geom_histogram(colour="black", fill="white",  stat="count", ) + ggtitle("
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



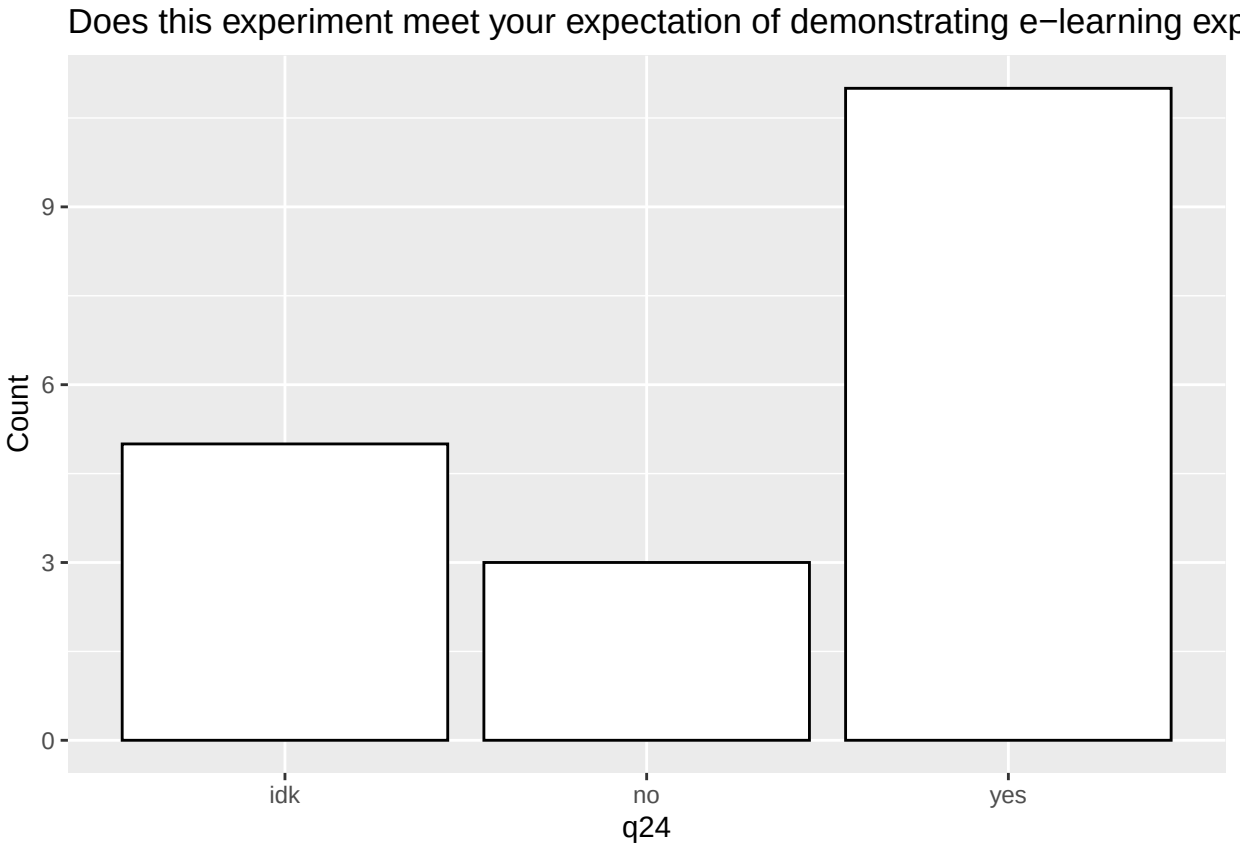

```
ggplot(surveys, aes(x=q23)) + geom_histogram(colour="black", fill="white",  stat="count", ) + ggtitle("
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```




```
ggplot(surveys, aes(x=q24)) + geom_histogram(colour="black", fill="white",  stat="count", ) + ggtitle("
```

```
## Warning: Ignoring unknown parameters: binwidth, bins, pad
```



Our participants are mainly high school and college students, with most being between 18 to 24 years old and are undergraduate students. This is intended, because most people are generally exposed to e-learning platform during their undergraduate years, with only a few continue to use the service as they enter higher educational level in the university. Some high school students nowadays are exposed to the e-learning platform early on, thanks to the advanced technology and the rise of popularity of e-learning platform in education.

Because there are different e-learning platform available nowadays, and our trial e-learning experiment cannot reflect exactly on the different formats of other e-learning platforms, some participants are unsure whether the experiment matches with their own e-learning experience.