ASQ Final

# Fixing Data/ Load Libraries

library(readxl)  
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

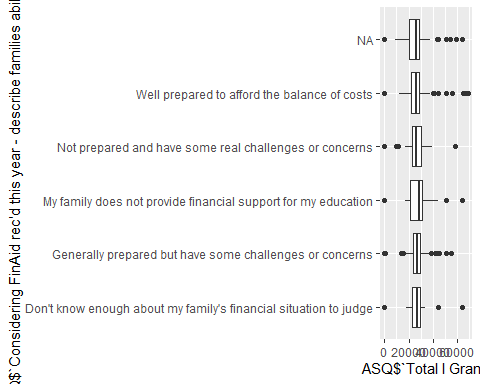
library(tidyr)  
library(writexl)  
library(ggplot2)  
  
ASQ <- read\_excel("ASQ - Completors - Fall 2021 20211021-092813.xlsx")  
  
ASQ$`Total I Grant` <- as.numeric(ASQ$`Total I Grant`)  
ASQ$`Merit Scholarship Amount` <- gsub("[\\$,]", "",ASQ$`Merit Scholarship Amount`)  
ASQ$`Merit Scholarship Amount` <- as.numeric(ASQ$`Merit Scholarship Amount`)  
ASQ$Admit <- as.numeric(ASQ$Admit)  
ASQ$Enrolled <- as.numeric(ASQ$Enrolled)  
ASQ$`Recalculated GPA` <- as.numeric(ASQ$`Recalculated GPA`)  
ASQ$`Which best describes your plans for fall 2021?` <- as.factor(ASQ$`Which best describes your plans for fall 2021?`)  
ASQ$`Merit Scholarship Amount` <- gsub("[\\$,]", "",ASQ$`Merit Scholarship Amount`)  
ASQ$Distance <- gsub("[,]" , "", ASQ$Distance)  
ASQ$Distance <- as.numeric(ASQ$Distance)

# Relationship between families ability to pay vs their grant

ggplot(ASQ, aes(ASQ$`Considering FinAid rec'd this year - describe families ability to pay`, ASQ$`Total I Grant` )) + geom\_boxplot() + coord\_flip()

## Warning: Use of `ASQ$`Considering FinAid rec'd this year - describe families  
## ability to pay`` is discouraged. Use `Considering FinAid rec'd this year -  
## describe families ability to pay` instead.

## Warning: Use of `ASQ$`Total I Grant`` is discouraged. Use `Total I Grant`  
## instead.

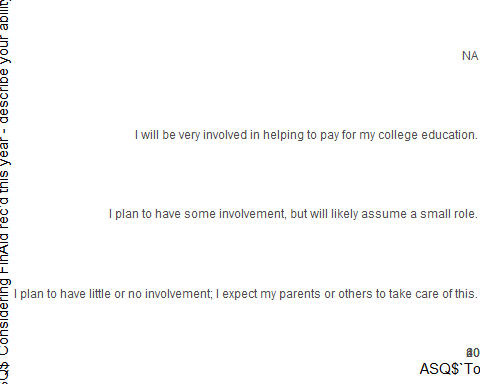


# Relationship between your ability to pay vs your grant

ggplot(ASQ, aes(ASQ$`Considering FinAid rec'd this year - describe your ability to pay`, ASQ$`Total I Grant` )) + geom\_boxplot() + coord\_flip()

## Warning: Use of `ASQ$`Considering FinAid rec'd this year - describe your ability  
## to pay`` is discouraged. Use `Considering FinAid rec'd this year - describe your  
## ability to pay` instead.

## Warning: Use of `ASQ$`Total I Grant`` is discouraged. Use `Total I Grant`  
## instead.



# Decision vs Enrollment

table(ASQ$`Which best describes your plans for fall 2021?`, ASQ$Enrolled)

##   
## 0 1  
## I do not plan to attend college this fall 128 0  
## I plan to attend a college other than Merrimack College 784 2  
## I plan to attend college this fall, but I still don't know where 39 3  
## I plan to attend Merrimack College 35 507

# Family Status vs Enrollment

table(ASQ$`Considering FinAid rec'd this year - describe families ability to pay`, ASQ$Enrolled)

##   
## 0 1  
## Don't know enough about my family's financial situation to judge 25 13  
## Generally prepared but have some challenges or concerns 196 137  
## My family does not provide financial support for my education 30 11  
## Not prepared and have some real challenges or concerns 51 46  
## Well prepared to afford the balance of costs 85 65

# Self status vs Enrollemnt

table(ASQ$`Considering FinAid rec'd this year - describe your ability to pay`,ASQ$Enrolled)

##   
## 0  
## I plan to have little or no involvement; I expect my parents or others to take care of this. 37  
## I plan to have some involvement, but will likely assume a small role. 182  
## I will be very involved in helping to pay for my college education. 166  
##   
## 1  
## I plan to have little or no involvement; I expect my parents or others to take care of this. 35  
## I plan to have some involvement, but will likely assume a small role. 117  
## I will be very involved in helping to pay for my college education. 117

# Zip Code analysis

ASQ %>%  
 group\_by(`Active Zip`) %>%  
 mutate(Frequancy = n()) %>%  
 filter(Frequancy > 3) %>%  
 summarise(Avg\_GPA = mean(`Recalculated GPA`),Avg\_Academic\_Rating = mean(`Academic Rating`),Admit\_Enrol = sum(Enrolled)/sum(Admit)\*100, Frequancy = n()) %>%  
 arrange(desc(Admit\_Enrol)) %>%  
 head(10)

## # A tibble: 10 x 5  
## `Active Zip` Avg\_GPA Avg\_Academic\_Rating Admit\_Enrol Frequancy  
## <chr> <dbl> <dbl> <dbl> <int>  
## 1 01701 3.24 3.2 100 5  
## 2 02050 3.10 3.38 87.5 8  
## 3 01940 2.85 4 80 5  
## 4 02043 2.57 3.8 80 5  
## 5 02056 3.2 3 80 5  
## 6 02052 3.21 3.25 75 4  
## 7 03842 3.42 2.75 75 4  
## 8 01876 3.61 3.08 69.2 13  
## 9 02330 3.75 2.4 60 5  
## 10 02339 3.81 2 60 5

# While there is interesting talking points towards why there are such high admit - enroll % per zipcode I feel there is not enough data to back up such findings.