# Homework 4

Michael Gillis

1. Does the amount of emails sent to high school students by a college increase as the college observed gets closer to the student?

2. The data I will be using was collected from Sep 5th 2019 to Sep 5th 2021 by flagging each college email in my personal inbox and then recording which college the email came from.
   a. Link to the csv file I will be using: EmailStats.csv
   b. Link to R file that I used to compile summary stats/charts: EmailPlots.R
   c. Link to my original excel file: CollegeEmails.xlsx

3. The variables that will help me answer my question are the count of emails by each individual school, and the proximity of the school to my hometown. Other explanatory variables to keep in mind are average GPA, acceptance rate, enrollment size, and endowment.
   a. The main explanatory variable is proximity in miles, which is numeric.
   b. The main response variable is count per school, which is also numeric.

4. During my final two years of high school (roughly when the results of my Junior year PSAT results were published), I began receiving an influx of emails from many different colleges across the United States. I began to track these emails coming into my inbox, just by filling out a table with each school name (and adding a new column if the email came from a new school), and recording the date as a

"tally" of sorts. Later, as I continued to collect the "count" data, I realized I should record some other variables other than just the count that could be found easily with some light research, like proximity of the school from my hometown (miles), or average GPA (4.0 scale). I had a suspicion that I was getting many more emails from schools in the Boston area, or even just in New England, as opposed to schools that were further away, thus leading to my question: does count increase as proximity decreases?

5.

| Variable	Variable Type	Standard Deviation	Median (Q1, Q3)
Count	Response	33.46996	42 (18, 65)
Proximity	Main Explanatory	743.8066	312 (137.5, 921.5)
Average GPA	Other Explanatory	0.3265764	3.65 (3.448, 3.815)
Acceptance Rate	Other Explanatory	0.2600796	0.669 (0.365, 0.779)
Enrollment Size	Other Explanatory	12472.55	5578 (2700, 10840)
Endowment	Other Explanatory	6743.314	474.5 (197.5, 1390.0)

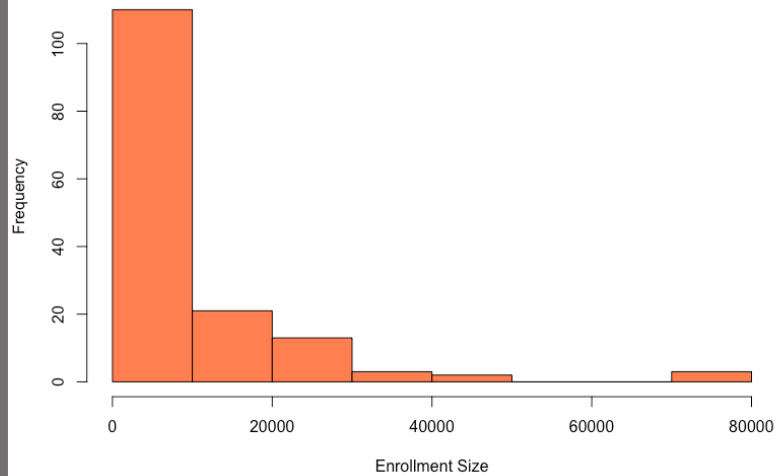## 6. Charts of the response variable (Email Count Per Day)
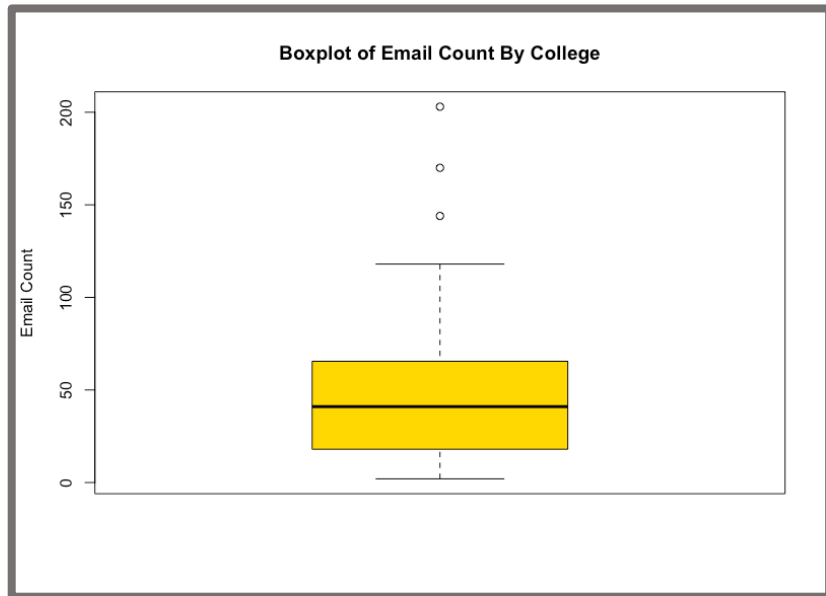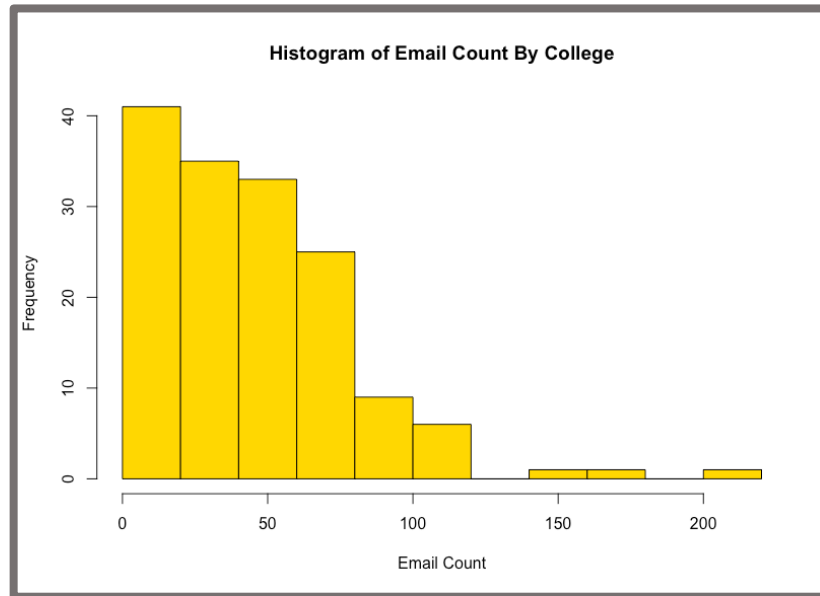
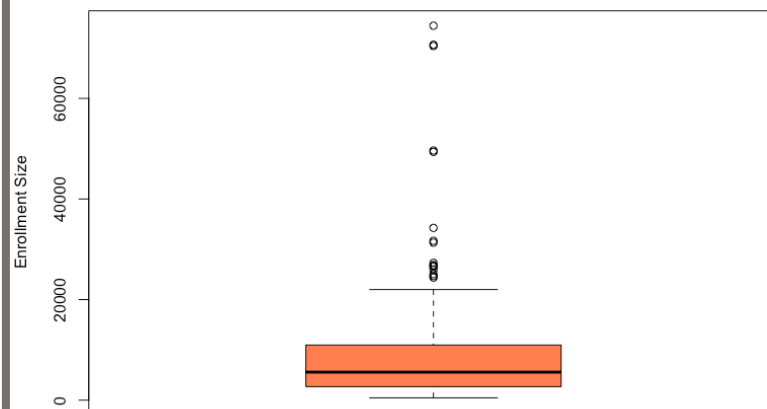## 7. Charts for the main explanatory variable (Proximity).
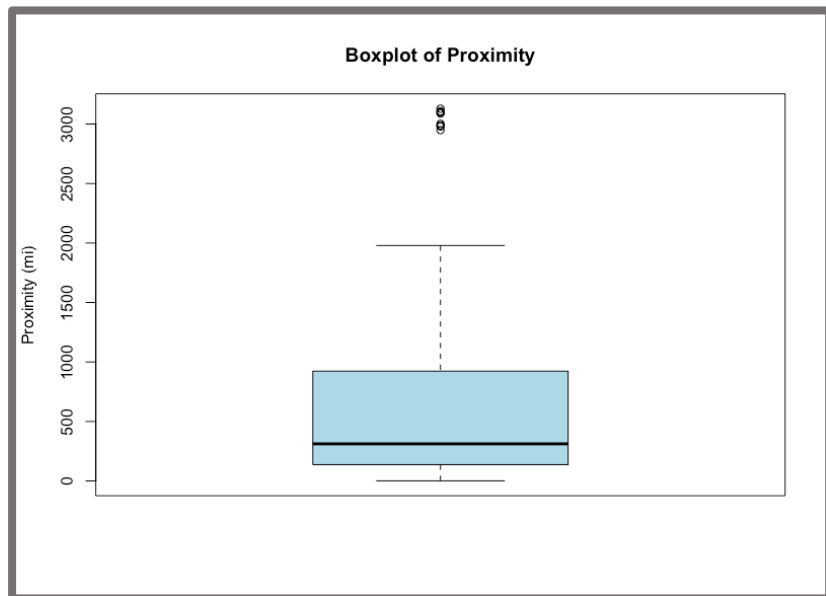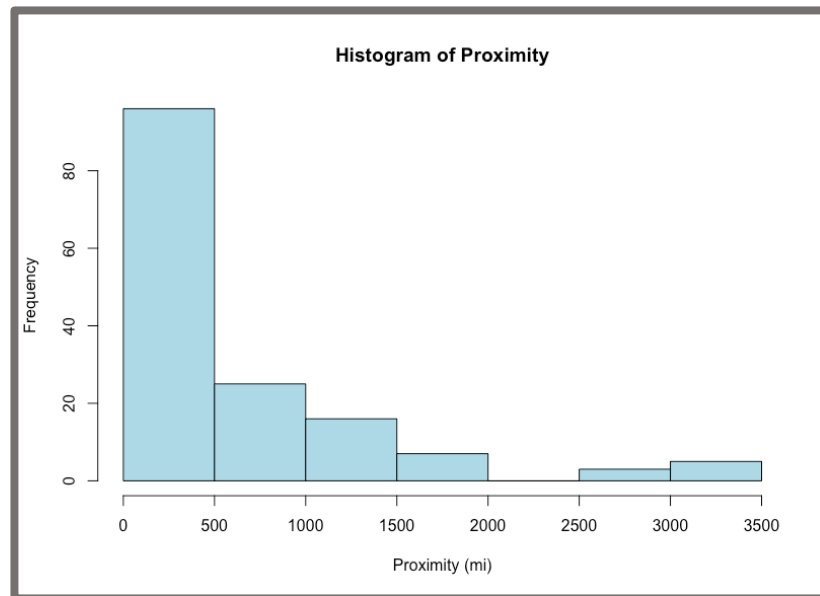
8. Charts for other explanatory variables (GPA, Accepthence Rate, Enrollment Size, Endowment).
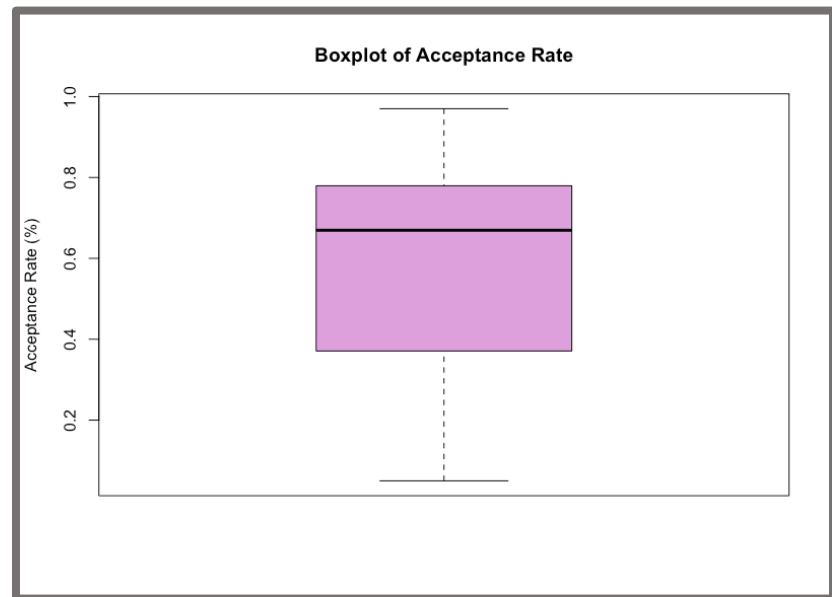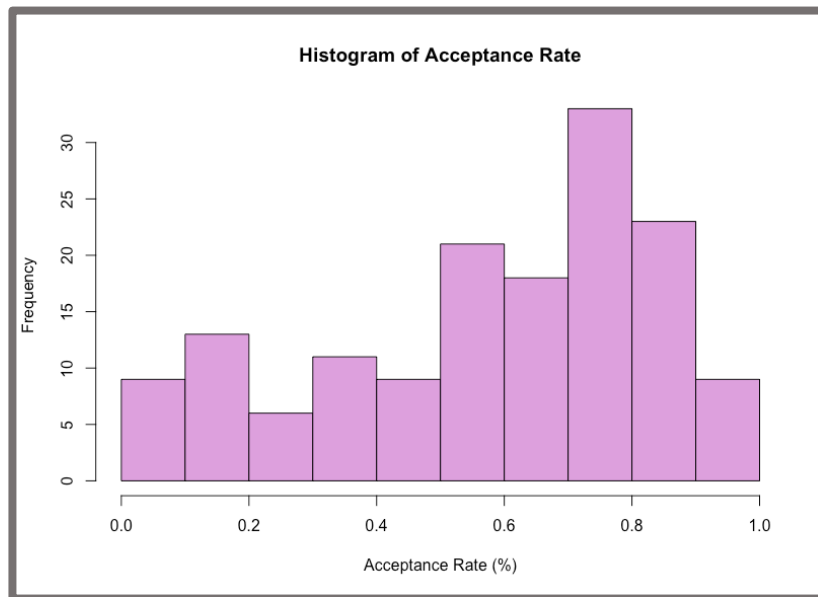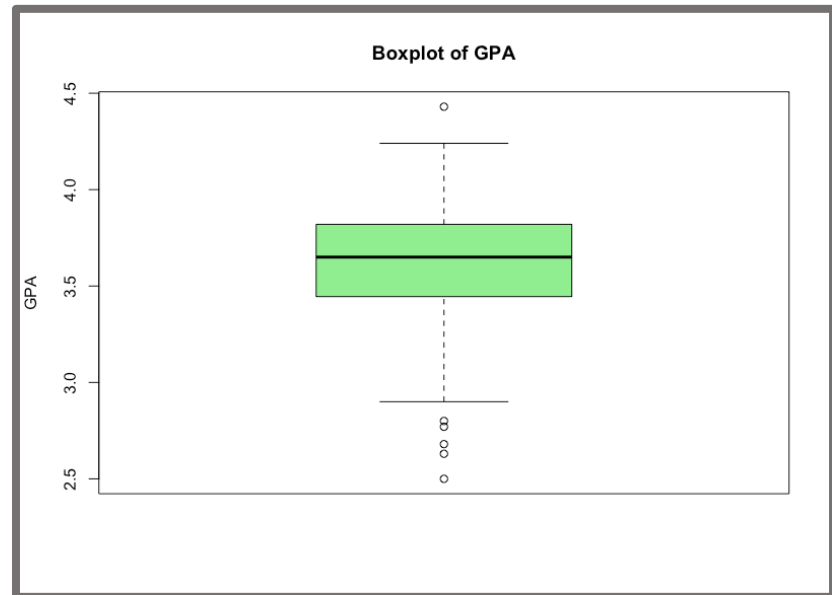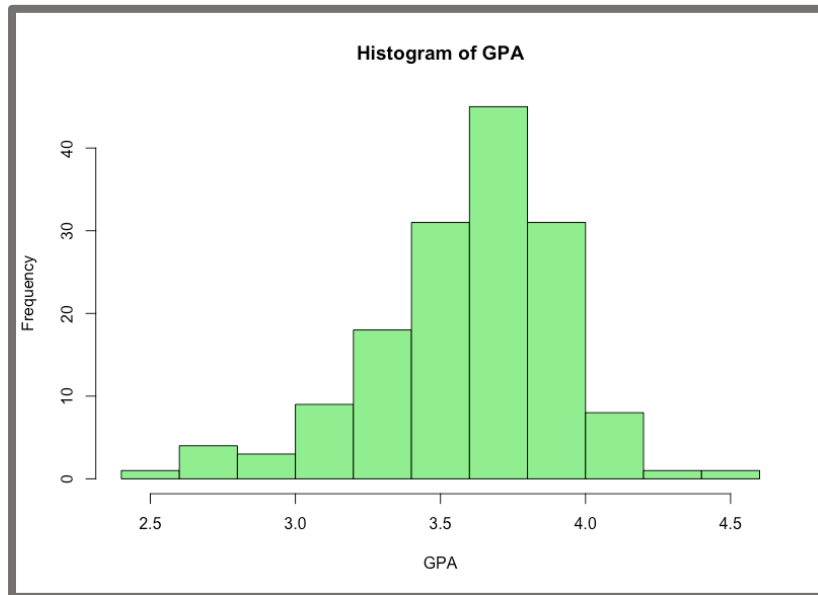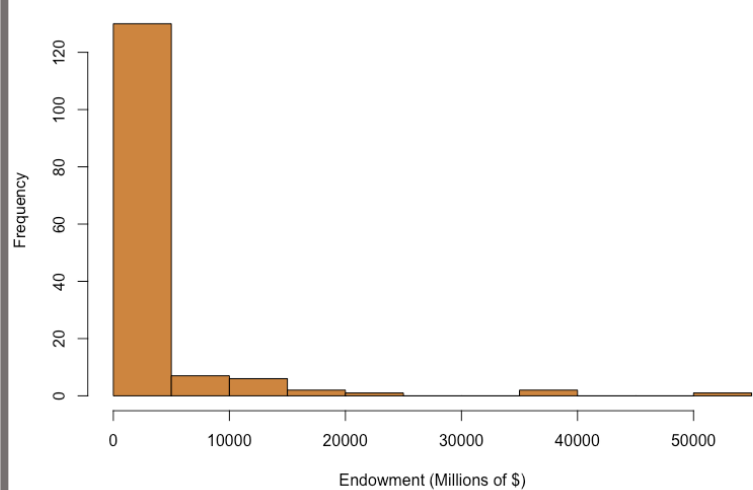
Histogram of Enrollment Size

Frequency

Enrollment Size

Boxplot of Enrollment Size

Enrollment Size

Histogram of Endowment

Frequency

Endowment (Millions of $)

Boxplot of Endowment

Endowment (Millions of $)