Name: **Yuchao Zhao**

UmichID: **76821781**

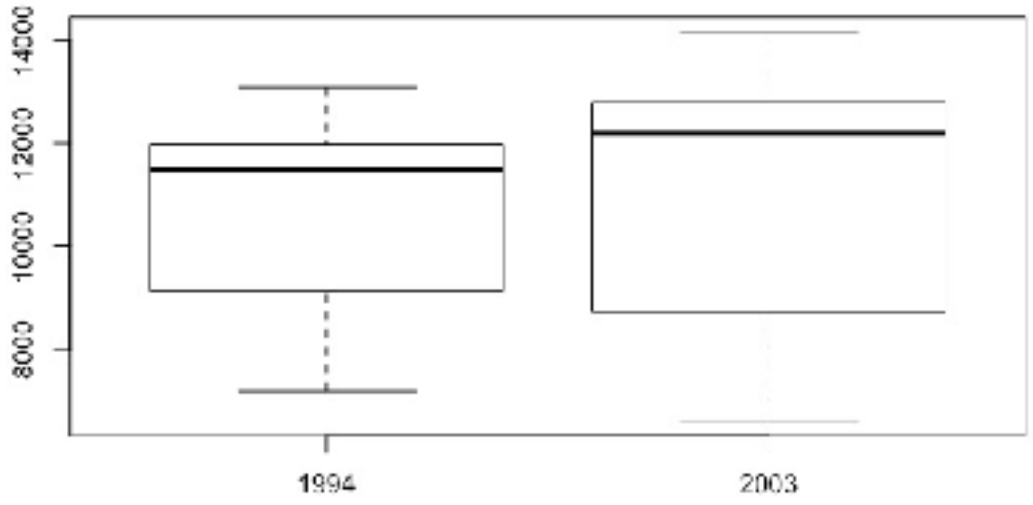
What dataset are you working with: **US_births_1994_2003**

List 3 questions that you can ask with your dataset.
Q1: Is there any difference between the average amount of daily new baby between 1994 and 2003?
Q2: In 1994, is there any differences of the average amount of daily new baby among different month? In another way, is some month will have higher amount of new born baby in 1994?
Q3:

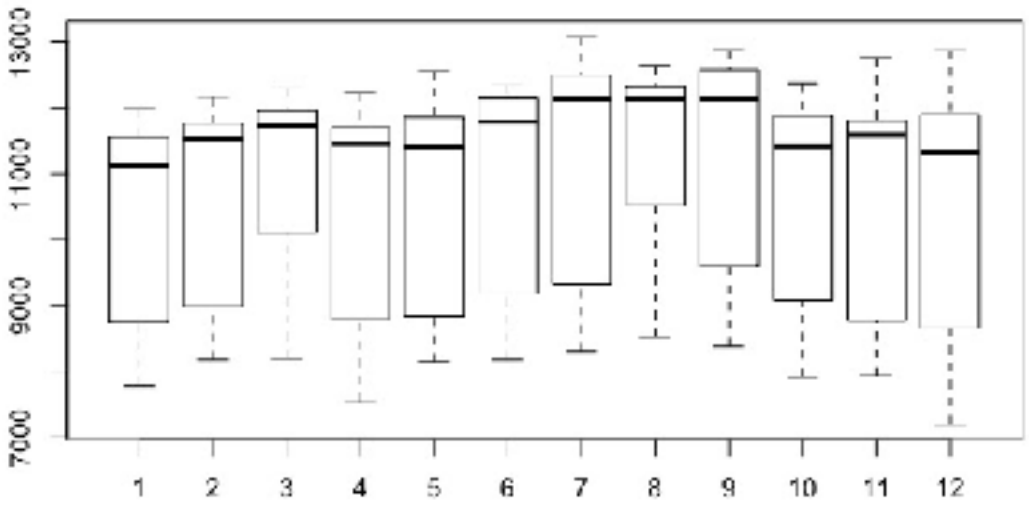
List the associated null hypothesis for each question:
Q1: The average amount of daily new born baby in 1994 is not significantly different with 2003.
Q2: The average amount of daily new born baby in 1994 is not significantly different among each months.
Q3:

What statistical test(s) will you use to answer each of the questions:
Q1: Welch two sample T-test
Q2: Anova
Q3:

Make a visual plot showing the relationship that you will analyze statistically (e.g. boxplot for t-test or ANOVA; scatterplot for regression; table for chi-square).
Q1:



Q2:



Q3:

Do your data meet the assumptions required for the statistical test you want to run? Please state the assumptions you examined and whether or not your data meet those assumptions:

Q1: data are continuous? (meet)
samples are randomly selected? (meet)
observations are independent? (meet)
values are nearly normal or sample size is large enough? (meet the sample size)
equal variance? (not meet by run code var.test)

Q2: The populations of interest must be normally distributed? (not meet by run qqnorm, but I can still assume normality since the sample size is large enough)
samples are independent? (meet)
each much have same variance? (obviously not meet by checking the box plot)

Q3:

Run the statistical test! Put your results here:

Q1: p-value = 0.007743
Q2: p-value = 0.476
Q3:

Interpret your results!

Q1: True difference in means is not equal to 0. Which means the amount of daily new born baby in 1994 is significantly different with 2003.

Q2: We cannot reject the null hypotheses. Which means there is no significant differences of the true average amount of daily new baby among each month in 1994.

Q3: