HW4_jayapats

Swetha Jayapathy

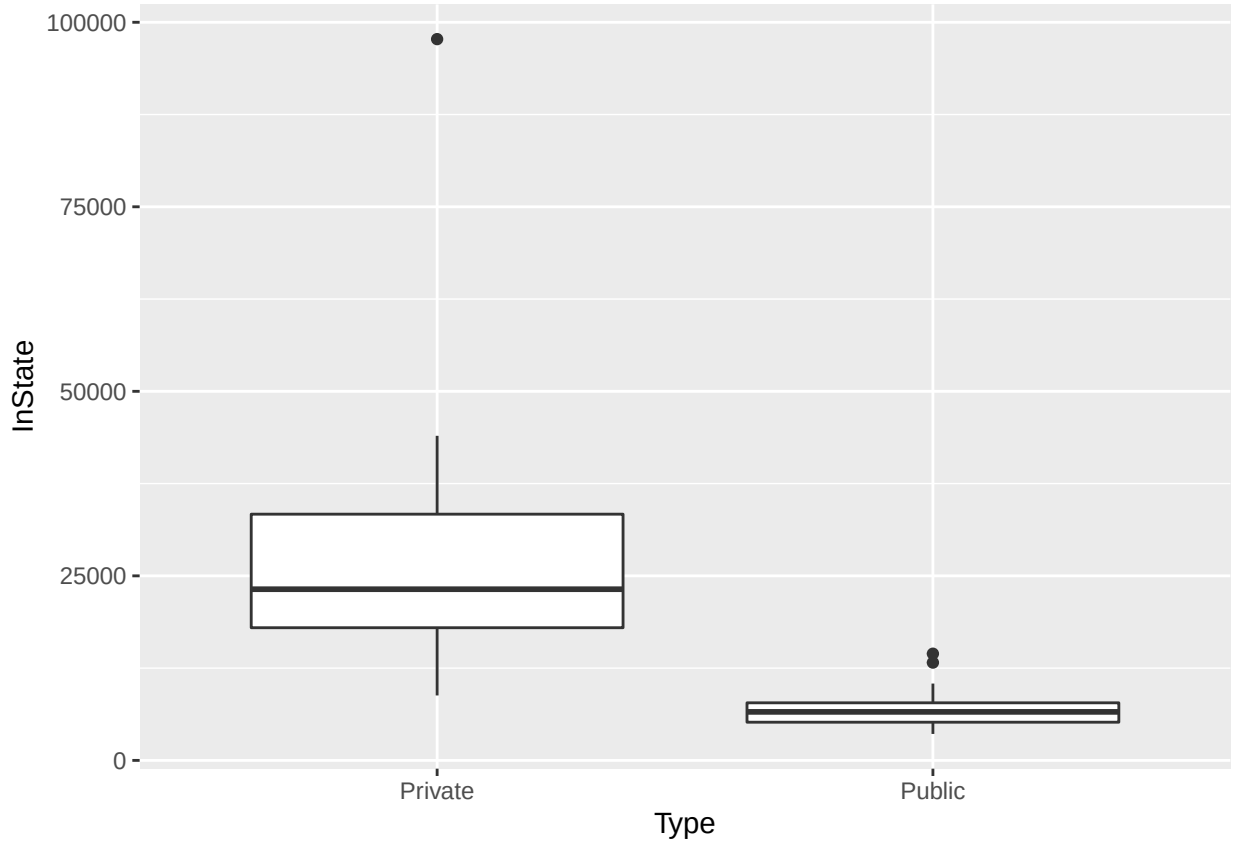
10/27/2020

R Markdown

```
library(Sleuth3)
library(ggplot2)
```

Question 1a

```
#Question 1a
qplot(Type, InState, data=ex0332, geom="boxplot")
```



Question 1b The plot shows that the equal variance assumption is violated as we can see that the InState tuition fees for Private colleges have a bigger spread(box is taller) compared to the InState Public colleges. The sizes of the boxes seems to bee too different.

Question 1c

μ1 - Mean InState tuition at private universities.

μ2 - Mean InState tuition at public universities.

Null Hypothesis (H0): μ1 - μ2 = 0 There is no difference in the mean InState tuition fees for the Private Universities and that of Public Universities.

Alternate Hypothesis (HA): μ1 - μ2 > 0 The mean InState tuition fees for the Private colleges is higher than that of Public colleges.

Question 1d

```
t.test(InState~Type,data=ex0332,alternative="greater")
```

Question 1e

Statistical Conclusion: There is a significant evidence that the mean InState tuition fees is higher for Private Universities than the Public Universities with a one sided Welch's test p value < 0.0001

Question 1f

Statistical Conclusion: We estimate increase in mean InState tuition fees between Private Universities and Public Universities in the range $13951.51 and $28563.45 with 95% Confidence level.

Question 2a

Consider the following:

Y = Survival time of random participant in Control group.

Y + Shift_parameter = Survival time of same participant in therapy group.

Where the shift parameter represents the difference between the two groups.

Null Hypothesis: The Shift parameter is zero, there is no effect of therapy on the life expectancy.

Alternate Hypothesis: The Shift parameter is non-Zero, there is an effect of therapy on the life expectancy.

Question 2b

```
wilcox.test(Survival~Group, data=ex0431, exact=FALSE, correct=FALSE,conf.int=TRUE, alternative="less")
```

Question 2c

There is no evidence that the survival time of breast cancer patients on therapy is higher than those without the therapy which is indicated by the p-value of 0.1308 (one-sided Wilcoxon rank sum test).

Question 3

```
HW2Dat <- read.csv("D:/D drive contents/Fall 2020/Stats 511/Class/HW/HW2/NavDat.csv")
diffs <- with(HW2Dat, DT-M)
l <- length(diffs)
g <- length(which(diffs>0))
```

```
binom.test(g, l, alternative="two.sided")
```

```
##
##  Exact binomial test
##
```

```
## data:  g and l
## number of successes = 28, number of trials = 36, p-value = 0.001193
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.6084824 0.8988495
## sample estimates:
## probability of success
##              0.7777778
```