PS1 Answers

Anna Voss

January 27, 2020

```
##Question 1
y <- c(105, 69, 86, 100, 82, 111, 104, 110, 87, 108, 87, 90, 94, 113, 11
#Use qnorm (n > 30)
z90 <- qnorm((1- 0.90)/2, lower.tail = FALSE)
n = length(y)
sample_mean <- mean(y)
sample_sd <- sd(y)
lower_90 <- sample_mean - (z90 * (sample_sd/sqrt(n)))
upper_90 <- sample_mean + (z90 * (sample_sd/sqrt(n)))
confint90 <- c(lower_90, upper_90)
confint90
# [94.13283,102.74717]

##Question 2
y <- c(105, 69, 86, 100, 82, 111, 104, 110, 87, 108, 87, 90, 94, 113, 11
##Data Normally distributed so can use 1 sample t-test
t.test(y, mu = 100)
#        One Sample t-test
# t = -0.59574, df = 24, p-value = 0.5569
# alternative hypothesis: true mean is not equal to 100
# 95 percent confidence interval:
#   93.03553 103.84447
# sample estimates:
#   mean of x
# 98.44

##Question 3
expenditure <- read.table("expenditure.txt", header=TRUE)

#a
library("tidyverse")
qplot(x = X1, y = Y, data = expenditure)
#X1 and Y show a strong positive correlation. On average, as per capita
qplot(x = X2, y = Y, data = expenditure)
#X2 and Y have a linear correlation. The number of resident per thousand
```

```
qplot(x = X3, y = Y, data = expenditure)
#X3 and Y have a postive correlation. On average, the number of people p

#b
ggplot(expenditure, aes(Region, Y)) +
geom_boxplot(aes(group=Region))
#Region 4 (West) has the highest per capita expenditure on public educat

#c
qplot(x = X1, y = Y, data = expenditure)
#X1 and Y show a strong positive correlation. On average, as per capita
qplot(x = X1, y = Y, data = expenditure, shape = Region)

ggplot(expenditure, aes(X1, Y)) +
geom_point(aes(shape = Region, color = Region)) +
scale_shape_identity() +
scale_color_gradient(low="blue", high="red")
```