

**Q1:**

Data set:  $y \leftarrow c(105, 69, 86, 100, 82, 111, 104, 110, 87, 108, 87, 90, 94, 113, 112, 98, 80, 97, 95, 111, 114, 89, 95, 126, 98)$

**1.1**

<ul style="list-style-type: none"> <li>• <math>\bar{x}</math> is 98.44</li> <li>• <math>Z</math> is 1.645</li> <li>• <math>s</math> is 13.093</li> <li>• <math>n</math> is 25</li> </ul>	$1.654 = 98.44 \pm 90\% \frac{13.093}{\sqrt{25}}$
--	---

**1.2**

Level of significance:  $\alpha = 0.05$

If the teacher wants to know whether the student body's average IQ is more than the national average, she must test the opposite, vis-à-vis the null hypothesis.

Ergo, the  $H_0$  is: My students' have the same or lower IQ than the national average.  $H_0 : \mu \leq 100$

$$\frac{98.44 - 100}{13.093/\sqrt{25}} = -60$$

Degrees of Freedom:  $df = n - 1 \therefore 24 = 25 - 1$

Reject  $H_0$  if  $p\text{-value} \leq \alpha$

$t > -60 = p\text{-value is } .7215 > \alpha = 0.05$

$\therefore$

Do not reject  $H_0$

Q2:

## load libraries

set wd

## clear global .envir

```
library(knitr) library(rmarkdown)
```

```
include_graphics() # remove objects rm(list=ls()) # detach all libraries detachAllPackages <- function()
{ basic.packages <- c("package:stats", "package:graphics", "package:grDevices", "package:utils", "pack-
age:datasets", "package:methods", "package:base") package.list <- search()[ifelse(unlist(gregexpr("package:",
search()))==1, TRUE, FALSE)] package.list <- setdiff(package.list, basic.packages) if (length(package.list)>0)
for (package in package.list) detach(package, character.only=TRUE) } detachAllPackages()
```

## load libraries

```
pkgTest <- function(pkg){ new.pkg <- pkg[!(pkg %in% installed.packages()[, "Package"])] if (length(new.pkg))
install.packages(new.pkg, dependencies = TRUE) sapply(pkg, require, character.only = TRUE) }
```

here is where you load any necessary packages

ex: stringr

```
lapply(c("stringr"), pkgTest)
```

```
lapply(c(), pkgTest)
```

## set working directory

```
setwd("~/Documents/GitHub/QTM200Spring2021/problem_sets/PS1")
```

## Problem 1

```
y <- c(105, 69, 86, 100, 82, 111, 104, 110, 87, 108, 87, 90, 94, 113, 112, 98, 80, 97, 95, 111, 114, 89, 95, 126,
98)
```

## Problem 2

```
expenditure <- read.table("https://raw.githubusercontent.com/ASDS-TCD/StatsI_Fall2021/main/datas
ets/expenditure.txt", header=T) expenditure <- read.table("https://raw.githubusercontent.com/ASDS-
TCD/StatsI_Fall2021/main/datasets/expenditure.txt", header=T)
```

```
#Y - PC Exp. on H; X1 - PC Income; X2 - #ofRes. per 100k $Insecure; X3 - #ofPeopl per 1k in Urban
Areas str(expenditure) expenditure[2:5]
```

```
####Q2.1
```

```
#The higher income p.c., the more is spent on housing.
```

```
scatter.smooth(expenditure$Y, expenditure$X1)
```

```
#How much is spent on housing p.c. and amount of people financially insecure.
```

```
scatter.smooth(expenditure$Y, expenditure$X2)
```

```
#Money spent on housing p.c. and amount of people living in urban areas.
```

```

scatter.smooth(expenditure$Y, expenditure$X3)
#Income p.c. and financial insecurity
scatter.smooth(expenditure$X1, expenditure$X2)
#Income p.c. and people living in urban areas: higher the income, the less urban dwellers.
scatter.smooth(expenditure$X1, expenditure$X3)
#Correlation between financial insecurity and urban population.
scatter.smooth(expenditure$X2, expenditure$X3)

```

##I would be able to gather more information if I new how to label the dots by state abbreviation and region, denoted by 4 different colors.

####Q2.2

```

#Region 4 has highest p.c. expenditure on housing.
barplot(expenditure$Y, expenditure$Region, col="light green")

```

####Q2.3

#I can't figure out how to change the points.

```

scatter.smooth(expenditure$Y, expenditure$X1, expenditure$Region)

```







