Hypotheses testing

Question 1

A national research institute in Sri Lanka concludes that Sri Lankans watch television on average 25 hours per week. It seems likely that graduate students do not watch nearly this much television per week. To test this, the following data were gathered from a random sample of 50 graduate students.

```
tvhour <- c(24, 20, 29.3, 25.1, 30.6, 34.6, 30.0, 39.0, 33.7, 31.6,
       25.9, 34.4, 26.9, 23.0, 31.1, 29.3, 34.5, 35.1, 31.2, 33.2,
       30.2, 36.4, 37.5, 27.6, 24.6, 23.9, 27.0, 29.5, 30.1, 29.6,
       27.3, 31.2, 32.5, 25.7, 30.1, 24.2, 24.1, 26.4, 31.0, 20.7,
       33.5, 32.2, 34.7, 32.6, 33.5, 32.7, 25.6, 31.1, 32.9, 25.9)
```

i. State the appropriate null and alternative hypothesis.

ii. Perform an analysis to test the hypothesis in (i).

Question 2

The following data (in litres) which were selected randomly from a normally distributed population of values, represent measurements of a bottle content that is supposed to contain, on average 5L.

```
bottle.content <- c(5.1, 5.4, 5.3, 5.2, 5.5, 5.6, 5.4, 5.3, 5.4, 5.2,
                    5.8, 5.2, 5.2, 5.3, 5.1, 5.3, 5.4, 5.5, 5.5, 5.7)
```

Use these data and $\alpha = 0.01$ to test the hypothesis that the content average 5L.

Question 3

Use the following data to construct 80%, 90% and 95% confidence intervals to estimate μ.

```
 [1] 103.4 102.7  90.1  89.8  74.8 106.5 108.5  97.9 115.8 101.1  92.3 108.2
[13] 110.8  96.7  98.9 116.2 115.3  72.7  79.4 100.5
```

Plot all confidence intervals on a single graph plane.

Question 4

A mathematics teacher gives their students a calculus pretest on the first day of class, and a similar test towards the end of the course. The results are shown below.

before	after
72	65
57	91
71	67
64	66
55	85
60	60
67	97
65	72

before	after
84	54
64	65
56	70
60	75
65	75
69	70
75	72
67	78
74	75
81	90
80	85
71	76
68	65
70	83

Determine whether the students performed significantly better on the post test, using $\alpha = 0.05$

Question 5

Use `heights` from `alr3` package.

```
library(alr3)
data(heights)
head(heights) # First 6 rows of the data frame
```

```
  Mheight Dheight
1    59.7    55.1
2    58.2    56.5
3    60.6    56.0
4    60.7    56.8
5    61.8    56.0
6    55.5    57.9
```

Test if there is a statistically significant correlation between mothers' height and daughters' height.

Question 6

Use `penguins` dataset in `palmerpenguins`. You can use the following code to remove missing cases.

```
library(palmerpenguins)
library(tidyverse)
library(magrittr)
new_penguins <- penguins %>%
    filter(complete.cases(flipper_length_mm, body_mass_g, species, sex))
head(new_penguins)
```

```
# A tibble: 6 x 7
  species island bill_length_mm bill_depth_mm flipper_length_~ body_mass_g sex
  <fct>   <fct>           <dbl>         <dbl>            <int>       <int> <fct>
1 Adelie  Torge~           39.1          18.7              181        3750 male
```

```
2 Adelie  Torge~           39.5          17.4               186        3800 fema~
3 Adelie  Torge~           40.3          18                 195        3250 fema~
4 Adelie  Torge~           36.7          19.3               193        3450 fema~
5 Adelie  Torge~           39.3          20.6               190        3650 male
6 Adelie  Torge~           38.9          17.8               181        3625 fema~
```

a)

Test if there is statistically significant differences in

```
i. flipper length
ii. body mass
iii. bill_length
iv. bill_depth
```

between male penguins and female penguins.

b)

Test if there is statistically significant differences in

```
i. flipper length
ii. body mass
iii. bill_length
iv. bill_depth
```

between species types.