程序代写代做 CS编程辅导

MTHM501 Formative Coursework 2

Visualisation

Question

Load the Animals.csv file from ELE. This consists of a datafile with three variables, "x", "y" and "z" and 3,026 observations (rows). "x" records weight in kilogrammes, "y" records height in metres and "z" records species. You should

1. Load the file into RStudio

2. Rename (use the "rename" function) the "x" variable to "Weight", the "y" variable to "Height" and the "z" variable "Species".

3. Rename the levels of "Species", a=Ostrich, b=Deer, c= Bear, d= Giant tortoise (use the "factor" function to reset the labels).

4. Use "mutate" to create a new variable called BMI (BMI is calculated as weight in kilogrammes divided by height in metres squared).

5. Create a scatter plot to explore the relationship between Height and Weight. What do you conclude about this relationship?

6. Create a scatter plot to explore the relationship between Height and Weight split by species (use facetting). What do you conclude about this relationship?

7. Use an appropriate graph to explore the BMI of the different species types.

Your report should consist of the annotated code you used and your plots and interpretations from parts 5,6 and 7.

Solution

```r
library(tidyverse)
library(ggplot2)
library(magrittr)
#Animals <- read_csv(file="//Users//mjk223//Dropbox//Working_with_Data//Week4//Animals.csv")
Animals <- read_csv(file="//Users//markkelson//Dropbox//Working_with_Data//2020-21//Week4//Animals.csv")

# Renaming variables
Animals <- rename(Animals,Weight=x,Height=y,Species=z)

#Relabelling the species variable
Animals <- mutate(Animals, Species = factor(Species,
                                            labels= c("Ostrich",
```

```
                                                              "Deer",
                                                              "Bear",
                                                              "Giant tortoise")))
# Simple scatter plot
ggplot(data=Animals,aes(x=Weight,y=Height)) + geom_point()+ geom_smooth(method="lm")
```
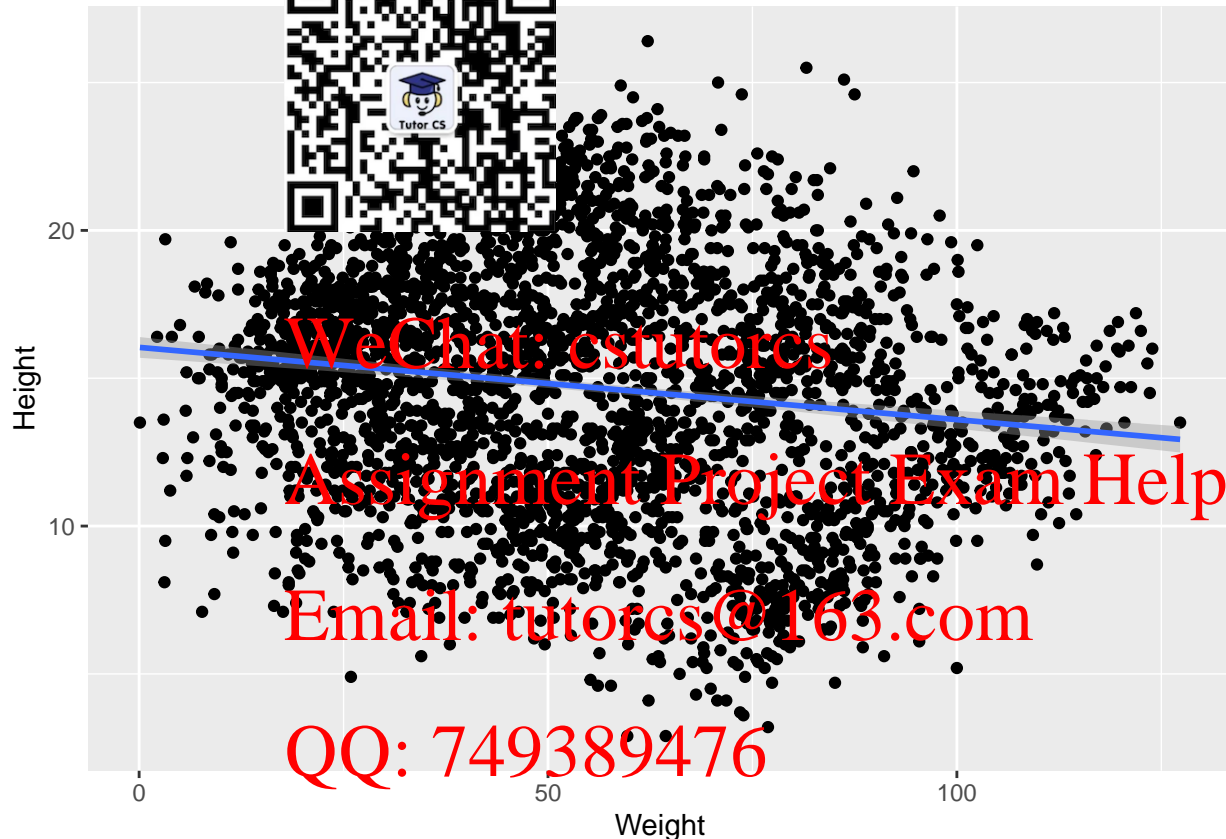

We would conclude there is a slight negative relationship between weight and height.

```
#This gives me the readPNG function
library(png)
#This gives me the getURL function
library(RCurl)
# this gives me the rasterGrob function
library(grid)
#I got these free png silhouettes from phylopic. This is completely optional

ostrichurl<-"https://images.phylopic.org/images/14ddbf4d-7749-4153-bbe1-8d0c6ffdf142/raster/435x512.png"
deerurl<-"https://images.phylopic.org/images/bb553480-e37f-4236-8c69-ce9fa8116b39/raster/396x512.png"
bearurl<-"https://images.phylopic.org/images/05f87521-20d4-4a05-8ac6-aa0bab7f1394/raster/512x355.png"
gianttortoiseurl<-"https://images.phylopic.org/images/67de5291-d93f-4b5b-9c5f-3a12de8daa6a/raster/512x30
#The location for the giant tortoise silhouette is this (it gets clipped above a bit)
#http://phylopic.org/assets/images/submissions/35c35a57-173f-424a-9f1a-8f4ccaddbae6.512.png


ostrich_logo <-  readPNG(getURLContent(ostrichurl),native=T)
deer_logo <-  readPNG(getURLContent(deerurl),native=T)
bear_logo <-  readPNG(getURLContent(bearurl),native=T)
```

```r
gianttortoise_logo <- readPNG(getURLContent(gianttortoiseurl), native=T)

#This creates my plot
my_plot <-
  ggplot(data=Animals,aes(x=Weight,y=Height,col=Species)) +
  geom_point(alpha=0. ...         ...ooth(method="lm",show.legend = FALSE) +
  facet_wrap(~Specie...
  #Removing the lege...
  theme(legend.posit...

# This function all...            ...to a facetted plot (it is optional- no marks awarded)
annotation_custom2 <...
function (grob, xmin...        , ymin = -Inf, ymax = Inf, data)
  { layer(data = data, stat = StatIdentity, position = PositionIdentity,
        geom = ggplot2:::GeomCustomAnn,
        inherit.aes = TRUE, params = list(grob = grob,
                                          xmin = xmin, xmax = xmax,
                                          ymin = ymin, ymax = ymax))}

a1 = annotation_custom2(rasterGrob(ostrich_logo, interpolate=TRUE),
                        xmin=80, xmax=120, ymin=3.75, ymax=20, data=Animals[1,])
a2 = annotation_custom2(rasterGrob(deer_logo, interpolate=TRUE),
                        xmin=80, xmax=120, ymin=3.75, ymax=20, data=Animals[932,])
a3 = annotation_custom2(rasterGrob(bear_logo, interpolate=TRUE),
                        xmin=0, xmax=50, ymin=15, ymax=25, data=Animals[1648,])
a4 = annotation_custom2(rasterGrob(gianttortoise_logo, interpolate=TRUE),
                        xmin=0, xmax=50, ymin=15, ymax=25, data=Animals[2459,])

my_plot + a1 + a2 + a3 + a4
```

We would conclude there is a strong positive relationship between weight and height.

```r
#Creating the BMI variable
Animals <- Animals %>% mutate(BMI = Weight/(Height^2))

# plotting on a log scale
ggplot(data=Animals, aes(x=Species, y=BMI, col=Species)) +
  geom_boxplot() + scale_y_continuous(trans='log2') +
  ylab("log BMI")
```

Mark scheme

- Loading the data - 1 mark

- Renaming the variables - 1 mark

- Renaming the factor levels of Species - 1 mark

- Creating BMI variable - 1 mark

- Creating the first scatter plot - 2 marks

- Interpreting the 1st scatter plot - 1 mark

- Creating the facetted scatter plot - 2 marks

- Interpreting the facetted scatter plot - 1 mark

- Using an appropriate summary of the BMI variable split by species (boxplot, violin plot etc...) - 1 mark

- Execution of plot - 2 marks

- Total: 13 marks