---

title: "Lab 2 - feedback solutions"

author: "Ines"

date: "2022/23"

output:

html\_document: default

pdf\_document: default

---

```{r setupE}

knitr::opts\_chunk$set(echo = TRUE)

```

Loading libraries. Note that tidyverse includes ggplot2, which is the library we need tor plotting data.

```{r include=FALSE}

library(tidyverse)

```

### Histogram

Histogram with the pressure1 dataset, displaying the pressure.

First load the dataset, then produce the histogram.

Note: if you have loaded the data using the "Import Dataset" facility the R variable containing the dataset is likely to be called "pressure1", so use this name below instead of "pressure1Data.

```{r}

m <- c("CMM004", "CMM007", "CMM020", "CMM022", "CMM024", "CMM028", "CMM012","CMM513")

mn <- c("SPE", "Intranet Systems Dev", "Data Vis and Analysis", "Databases", "OOP", "ITI and Administration", "Project Investigation", "MSc Project")

cr <- c(15,15,15,15,15,15, 15, 45)

whatToStudy <- data.frame(module = m, moduleName=mn,

credits = cr, stringsAsFactors=T)

```

If we rather have a tibble we can use the tibble function - note that the vectors m, mn and cr were created above.

```{r}

whatToStudyT <- tibble(module = m, moduleName=mn, credits = cr)

whatToStudyT <- whatToStudyT |> mutate(across(where(is.character),as\_factor))

```

Accessing the 2nd column by name

```{r}

whatToStudy$moduleName

```

Accessing the 2nd column by index.

```{r}

whatToStudy[,2]

```

Accessing the 4th row.

```{r}

whatToStudy[4,]

```

Accessing the 3rd element in the moduleName column.

```{r}

whatToStudy$moduleName[3]

```

Accessing the 3rd element in the 2nd column.

```{r}

whatToStudy[3,2]

```

Getting the dimensions

```{r}

dim(whatToStudy)

```

Getting the number of rows

```{r}

nrow(whatToStudy)

```

Getting the number of columns

```{r}

ncol(whatToStudy)

```

```{r}

summary(whatToStudy)

```

## Exercise 1

Accessing the 2nd column by index.

```{r}

whatToStudyT[,2]

```

Accessing the 4th row.

```{r}

whatToStudyT[4,]

```

Accessing the 3rd element in the moduleName column.

```{r}

whatToStudyT$moduleName[3]

```

Accessing the 3rd element in the 2nd column.

```{r}

whatToStudyT[3,2]

```

Getting the dimensions

```{r}

dim(whatToStudyT)

```

Getting the number of rows

```{r}

nrow(whatToStudyT)

```

Getting the number of columns

```{r}

ncol(whatToStudyT)

```

```{r}

summary(whatToStudyT)

```

## Loading the pressureTemp dataset.

Loading the data onto a data frame

```{r}

pressureTemp <- read.csv("pressureTemp.csv", header=T, stringsAsFactors =T)

summary(pressureTemp)

```

Alternatively, load the data onto a tibble.

```{r}

pressureTempTib <- read\_csv("pressureTemp.csv" )

pressureTempTib <- pressureTempTib |>

mutate(across(where(is.character),as\_factor))

```

# Plots

## Histogram for pressure

```{r}

p <- ggplot(data = pressureTemp, aes(x=Pressure))

p <- p + geom\_histogram()

p

```

Colouring it blue.

```{r}

p <- ggplot(data = pressureTemp, aes(x=Pressure))

p <- p + geom\_histogram(fill="blue")

p

```

Setting the bin width to 4.

```{r}

p <- ggplot(data = pressureTemp, aes(x=Pressure))

p <- p + geom\_histogram(binwidth=8, fill="blue")

p

```

Specifying that 10 bins are required (no binwidth in this case!).

```{r}

p <- ggplot(data = pressureTemp, aes(x=Pressure))

p <- p + geom\_histogram(bins=10, fill="blue")

p

```

## Exercise 2

```{r}

p <- ggplot(data = pressureTemp, aes(x=Temperature))

p <- p + geom\_histogram()

p

```

## Exercise 3

```{r}

p <- ggplot(data = pressureTemp, aes(x=Temperature))

p <- p + geom\_histogram(fill="orange")

p

```

## Exercise 4

```{r}

p <- ggplot(data = pressureTemp, aes(x=Temperature))

p <- p + geom\_histogram(bins=7, fill="lightblue")

p

```

## Exercise 5

```{r}

p <- ggplot(data = pressureTemp, aes(x=Temperature))

p <- p + geom\_histogram(binwidth=2, fill="violet")

p

```

## Exercise 6

```{r}

p <- ggplot(data = pressureTemp, aes(x=Readby))

p <- p + geom\_bar( fill="red")

p

```

## Exercise 7

```{r}

p <- ggplot(data = pressureTemp, aes(x=Readby))

p <- p + geom\_bar( fill="brown")

p

```

## Reading the consumption dataset

```{r}

consumption <- read.csv("consumption.csv", header=T, stringsAsFactors = T)

```

## Column plot for Energy in year X2035.

```{r}

p <- ggplot(data = consumption, aes(x=Energy, y=X2035))

p <- p + geom\_col(fill="powderblue")

p

```

## Exercise 8

```{r}

p <- ggplot(consumption, aes(x=Energy, y=X2035))

p <- p + geom\_col(fill="pink")

p

```

## Exercise 9

Loading the data file.

```{r}

consumptionTransposed <- read.csv("consumptionTransposed.csv",

header=T, stringsAsFactors = T)

```

## Exercise 10

```{r}

p <- ggplot(consumptionTransposed, aes(x=Year, y=Oil))

p <- p + geom\_col(fill="pink")

p

```

## Exercise 11

```{r}

consumptionTransposed$Year <- as.factor(consumptionTransposed$Year)

p <- ggplot(consumptionTransposed, aes(x=Year, y=Oil))

p <- p + geom\_col(fill="pink")

p

```

The year values are taken as numeric, so the space between columns varies.

## Exercise 12

```{r}

consumptionTransposed$Year <- as.factor(consumptionTransposed$Year)

```

## Exercise 13

```{r}

p <- ggplot(consumptionTransposed, aes(x=Year, y=Coal))

p <- p + geom\_col(fill="red")

p

```

## Boxplots

```{r}

p <- ggplot(data= pressureTemp, aes(x=Readby, y=Temperature))

p <- p+ geom\_boxplot()

p

```

## Exercise 14

```{r}

p <- ggplot(pressureTemp, aes(x=Readby, y=Temperature))

p <- p+ geom\_boxplot(fill="red")

p

```

## Exercise 15

```{r}

p <- ggplot(pressureTemp, aes(x=Readby, y=Temperature))

p <- p+ geom\_boxplot(fill=c("pink", "blue", "green"))

p

```

## Exercise 16

```{r}

p <- ggplot(pressureTemp, aes(x=Readby, y=Temperature))

p <- p+ geom\_boxplot(fill=terrain.colors(3))

p

```

```{r}

p <- ggplot(pressureTemp, aes(x=Readby, y=Temperature))

p <- p+ geom\_boxplot(fill=topo.colors(3))

p

```

## Exercise 17

```{r}

p <- ggplot(pressureTemp, aes(x=Readby, y=Pressure))

p <- p+ geom\_boxplot(fill=topo.colors(3))

p

```