---

title: "Lab 4

- feedback solutions"

author: "Ines Arana"

date: "2022/23"

output: html\_document

---

```{r setup, include=FALSE}

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

```

Setting the working directory to the folder where this file is.

```{r, warning=F}

rm(list=ls())

```

Loading library (package) tidyverse

```{r, , warning = F}

library(tidyverse)

```

Loading all the data.

```{r}

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

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

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

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

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

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

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

```

## Using a text geom

Instead of points text labels are used in this plot.

```{r}

p <- ggplot(importsExportsAmerica, aes(crude.imports,crude.exports, label=country))

p <- p+geom\_text()

p

```

## Bar plots

### Stacking

```{r}

p <- ggplot(projectTeam, aes(site, fill = experience))

p <- p + geom\_bar(position = "stack")

p

```

### Dodging

```{r}

p <- ggplot(projectTeam, aes(site, fill = experience))

p <- p + geom\_bar(position = "dodge")

p

```

### Filling

```{r}

p <- ggplot(projectTeam, aes(site, fill=experience))

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

p

```

## Identity

Note that in this case some bars are hidden behind other bars so this is not very useful.

```{r}

p <- ggplot(projectTeam, aes(site, fill = experience))

p <- p + geom\_bar(position = "identity")

p

```

## Adjusting axis

The following plot has no axis adjustments.

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE, colour=Source))

p <- p+geom\_point()

p

```

### Adjusting the x axis

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE, colour=Source))

p <- p + xlim(2010,2035)

p <- p+geom\_point()

p

```

## Adding titles, subtitles, x and y axis labels and levend label

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE, colour=Source))

p <- p+geom\_point()

p<- p+ labs(x="Year",

y= "Millions of tons of oil equivalent",

title ="Energy consumption by source",

subtitle = "Oil is excluded",

colour = "Energy source")

p

```

## Changing colours

```{r}

colours <-c(coal="black", gas="yellow2", renewables="green")

p <- ggplot(consumptionNoOil, aes(Year, MTOE, colour=Source))

p <- p +geom\_point()

p<-p+labs(x= "Year",

y= "Millions of tons of oil equivalent",

title = "Energy consumed",

colour="Energy source")

p<- p + scale\_colour\_manual(values=colours)

p

```

## Adding lines

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE, colour=Source))

p <- p+geom\_point()

p <- p + geom\_vline(aes(xintercept = Year ), linetype = "longdash", colour="green")

p

```

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE, colour=Source))

p <- p+geom\_point()

p <- p + geom\_vline(aes(xintercept = 2015), linetype = "longdash", colour="green")

p

```

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE))

p <- p+geom\_point()

p <- p + geom\_hline(aes(yintercept = 100), linetype = "dotdash", colour="green")

p

```

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE, colour=Source))

p <- p+geom\_point()

p <- p + geom\_hline(aes(yintercept = 100), linetype = "dotted", colour="darkgreen")

p <- p + geom\_vline(aes(xintercept = 2015), linetype = "twodash", colour="blue")

p

```

```{r}

p <- ggplot(consumptionNoOil, aes(Year,MTOE, colour=Source))

p <- p+geom\_point()

p <- p + geom\_hline(aes(yintercept = 100), linetype = "dashed", colour="green")

p <- p + geom\_vline(aes(xintercept = 2015), linetype = "dashed", colour="blue")

p

```

## Pie chart

```{r}

pieData <- projectTeam |> count(site, name="Number")

names(pieData) <- c("site","numberStaff")

## order data according to the site (important for placing labels later on)

pieData <- arrange(pieData,desc(site))

# create new column with position for label

pieData <- pieData |> mutate(positionLabel = cumsum(numberStaff) - 0.5\* numberStaff)

# creating plot

p <- ggplot(pieData, aes(x="", y= numberStaff, fill = site))

p <- p + geom\_bar(width = 1, stat="identity")

p <- p + coord\_polar("y", start=0)

# Changing the colour scheme

p <- p + scale\_fill\_manual(values = c("red","blue", "lightblue"))

# removing outer "ring"

p <- p + theme\_void()

p

```

# Exercises

## Exercise 1

```{r}

p <- ggplot(ProvenOilReservesZerosT, aes(MTBarrels, Denmark))

p <- p + geom\_point()

p <- p + labs(x="Year", y="Million tons oil equivalent" , title="Proven Oil Reserves in Denmark")

p

```

## Exercise 2

This produces far too many vertical lines!

```{r}

p <- ggplot(ProvenOilReservesZerosT, aes(MTBarrels, Denmark))

p <- p + geom\_point()

p <- p + labs(x="Year", y="Million tons oil equivalent" , title="Proven Oil Reserves in Denmark")

p <- p + geom\_vline(aes(xintercept = MTBarrels),linetype = "longdash")

p

```

Exercise 2 solution - with only a couple of vertical lines!

```{r}

p <- ggplot(ProvenOilReservesZerosT, aes(MTBarrels, Denmark))

p <- p + geom\_point()

p <- p + labs(x="Year", y="Million tons oil equivalent" , title="Proven Oil Reserves in Denmark")

# lines put at 1990 and 2005. These are just examples.

p <- p + geom\_vline(aes(xintercept = 1990))

p <- p + geom\_vline(aes(xintercept = 2005))

p

```

## Exercise 3

```{r}

p <- ggplot(ProvenOilReservesZerosT, aes(MTBarrels, Italy))

p <- p + geom\_line(colour="orange", size=2)

p <- p + labs(x="Year", y="Million tons oil equivalent" , title="Proven Oil Reserves in Italy")

p <- p + geom\_hline(aes(yintercept = 0.9))

p

```

## Exercise 4

In the solution below note that using colour to emphasize size may not be the best option.

```{r}

### horizontal blue line added at 5.

p <- ggplot(ProvenOilReservesZerosT, aes(MTBarrels, United.Kingdom, colour=United.Kingdom))

p <- p + geom\_point()

p <- p + labs(x="Year", y="Million tons oil equivalent" , title="Proven Oil Reserves in the UK", colour="Size of Reserves")

### horizontal blue line added at 5.

p <- p + geom\_hline(aes(yintercept = 5), colour="blue")

p

```

## Exercise 5

```{r}

p <- ggplot(ProvenOilReserveWEurope, aes(Year, MT.Barrels, group=Country))

p <- p + geom\_line(aes(colour=factor(Country)))

p <- p + labs(x="Year", y="Million tons barrels" , title="Proven Oil Reserves in Western Europe", colour= "Country")

p

```

## Exercise 6

```{r}

p <- ggplot(channels, aes(length, area, size=n))

p <- p + geom\_point()

p <- p + labs(x="length", y="area" , title="Perception - length vs. area", size= "Size of stimuli")

p

```

## Exercise 7

```{r}

p <- ggplot(channels, aes(length, area, size=n))

p <- p + geom\_point()

p <- p + labs(x="length", y="area" , title="Perception - length vs. area", size= "Size of stimuli")

p <- p + ylim(0,5)

p

```

## Exercise 8

### stacking

```{r}

p <- ggplot(universities, aes(status, fill = school))

p <- p + geom\_bar(position = "stack")

p

```

#### Dodging

```{r}

p <- ggplot(universities, aes(status, fill = school))

p <- p + geom\_bar(position = "dodge")

p

```

#### Filling

```{r}

p <- ggplot(universities, aes(status, fill = school))

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

p

```

#### Identity

This does not make any sense as some data is "hidden" behind other data.

```{r}

p <- ggplot(universities, aes(status, fill = school))

p <- p + geom\_bar(position = "identity")

p

```

## Exercise 9

```{r}

pieData <- projectTeam |> count(site, name="Number")

universities2 <- universities |> filter(status %in% c("student", "both"))

# Counting the number of rows per school.

#pieData <- count(universities2$school)

names(pieData) <- c("school","numberStudents")

#Oordering data according to the school name (important for placing labels later on)

pieData <- arrange(pieData,desc(school))

# Creating a new column called positionLabel with position for label

pieData <- mutate(pieData, positionLabel = cumsum(pieData$numberStudents) - 0.5\*pieData$numberStudents)

# creating pie chart

p <- ggplot(pieData, aes(x="", y= numberStudents, fill = school))

p <- p + geom\_bar(width = 1, stat="identity")

p <- p + coord\_polar("y", start=0)

# Adding text labels

p <- p + geom\_text(aes(y = positionLabel, label = numberStudents))

# Changing the colour scheme

p <- p + scale\_fill\_manual(values = c("wheat", "lightblue", "cornflowerblue", "blue1"))

# removing outer "ring"

p <- p + theme\_void()

p

```