

The difference in time women and men spend on various activities

210193570

02/05/2022

Contents

Data Origins	1
Research Questions	2
Data Preparation and Visualisations	2
The amount of time spent on various activities	3
The difference in the time spent on various activities	8
Summary	9
Brief interpretation of graphs	9
Brief thoughts	9
What else can be done	9

#loading libraries

```
library(tidyverse)
library(RColorBrewer)
library(here)
```

Data Origins

The data set used for my visualisation was retrieved from Kaggle and posted by user Felipe Chapa. It is an OECD data set which contains the amount of time women and men from different countries spend on various activities during the day. It includes variables such as *country*, *description (of the activity)*, *sex*, *age* and *minutes per day*.

#load data set

```
timeuse_raw <- read_csv("~/PSY6422project/PSY6422project/data/timeuse_raw.csv")
```

#show first few rows of raw data

```
head(timeuse_raw)
```

```
## # A tibble: 6 x 7
```

	Country_Code	Country	DESC	Description	Sex	Age	Minutes_per_day
	<chr>	<chr>	<chr>	<chr>	<chr>	<chr>	<dbl>
## 1	AUS	Australia	UPW	Unpaid work	Women	15-64	311
## 2	AUS	Australia	UPW	Unpaid work	Men	15-64	172.
## 3	AUT	Austria	UPW	Unpaid work	Women	15-64	269.
## 4	AUT	Austria	UPW	Unpaid work	Men	15-64	135.
## 5	BEL	Belgium	UPW	Unpaid work	Women	15-64	237.
## 6	BEL	Belgium	UPW	Unpaid work	Men	15-64	144.

#show last few rows of raw data

```
tail(timeuse_raw)
```

```
## # A tibble: 6 x 7
```

	Country_Code	Country	DESC	Description	Sex	Age	Minutes_per_day
--	--------------	---------	------	-------------	-----	-----	-----------------

```
##   <chr>          <chr>      <chr> <chr>          <chr> <chr>          <dbl>
## 1 LTU           Lithuania OTH   Other          Women 15-64          8.69
## 2 LTU           Lithuania OTH   Other          Men    15-64          7.61
## 3 LUX           Luxembourg PAW   Paid work or study Women 15-64          239.
## 4 LUX           Luxembourg PAW   Paid work or study Men    15-64          330.
## 5 LTU           Lithuania PAW   Paid work or study Women 15-64          279.
## 6 LTU           Lithuania PAW   Paid work or study Men    15-64          354.
```

```
#show summary of raw data
summary(timeuse_raw)
```

```
## Country_Code      Country      DESC      Description
## Length:329        Length:329      Length:329      Length:329
## Class :character   Class :character   Class :character   Class :character
## Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
## Sex              Age              Minutes_per_day
## Length:329        Length:329        Min.   : 3.26
## Class :character   Class :character   1st Qu.:135.28
## Mode  :character   Mode  :character   Median :266.20
##                      Mean    :288.04
##                      3rd Qu.:351.94
##                      Max.    :760.89
```

Research Questions

Main research question: **What is the difference in the time women and men spend on different activities during the day?** I will also visualise the amount of time women and men spend on various activities to find out whether there is a clear pattern in the data or if there are some inconsistencies between countries.

Data Preparation and Visualisations

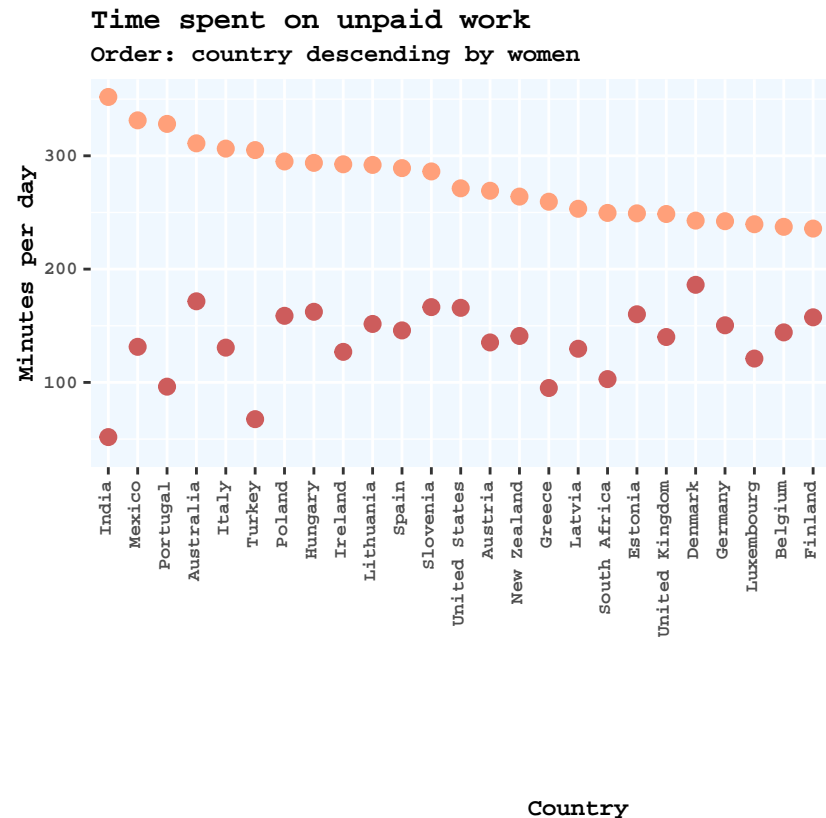
```
#delete unnecessary columns
timeuse <- timeuse_raw %>% select(Country, Sex, DESC, Minutes_per_day)

#1. Unpaid work
#create data frame including only unpaid work
upw <- timeuse %>% filter(DESC == "UPW")
#change the order of rows based on minutes_per_day for women
order_desc_women <- upw %>% filter(Sex == "Women") %>%
  arrange(desc(Minutes_per_day)) %>%
  mutate(Country = factor(Country))
#plot the data in descending order by values for women
upw %>% mutate(Country = factor(Country, levels = order_desc_women$Country, ordered = TRUE)) %>%
  ggplot(aes(x = Country, y = Minutes_per_day, color = Sex), group = Country) +
  #set point size
  geom_point(size = 2.5) +
  #set different colours for sexes
  scale_color_manual(values = c("#CD5C5C", "#FFA07A")) +
  #name axes; add title, subtitle and caption
  labs(x = "Country", y = "Minutes per day",
```

```

title = "Time spent on unpaid work",
subtitle = "Order: country descending by women",
caption = "Data source: Kaggle") +
#adjust the place and angle of the text on x axis
theme(axis.text.x = element_text(angle = 90, vjust = 0.25, hjust = 1),
#set background colour, customise text on the plot
panel.background = element_rect(fill = "aliceblue"),
text = element_text(family = "mono", face = "bold", size = 9))

```



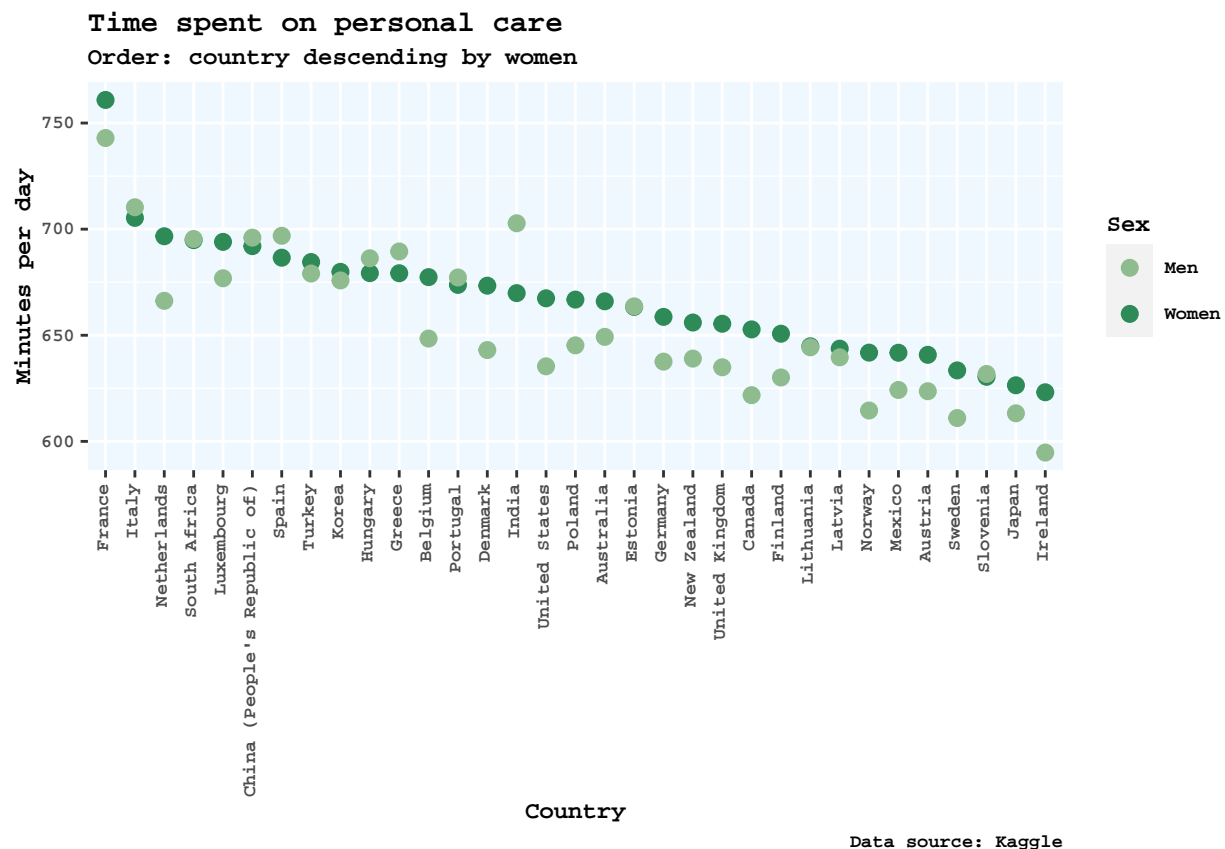
The amount of time spent on various activities

```

#2. Personal care
#create data frame with personal care only
pca <- timeuse %>% filter(DESC == "PCA")
#order rows by minutes_per_day of women
order_desc_women <- pca %>% filter(Sex == "Women") %>%
  arrange(desc(Minutes_per_day)) %>%
  mutate(Country = factor(Country))
#plot the data; countries in a descending order by women values
pca %>% mutate(Country = factor(Country, levels = order_desc_women$Country, ordered = TRUE)) %>%
  ggplot(aes(x = Country, y = Minutes_per_day, color = Sex), group = Country) +
  #set point size
  geom_point(size = 2.5) +
  #name axes, add title, subtitle and caption
  labs(x = "Country", y = "Minutes per day",
    title = "Time spent on personal care",
    subtitle = "Order: country descending by women",
    caption = "Data source: Kaggle") +

```

```
#customise colours for sexes
scale_color_manual(values = c("#8FBC8F", "#2E8B57")) +
#adjust text angle and position on x axis
theme(axis.text.x = element_text(angle = 90, vjust = 0.25, hjust = 1),
#choose background colour
panel.background = element_rect(fill = "aliceblue"),
#set text styling
text = element_text(family = "mono", face = "bold", size = 9))
```

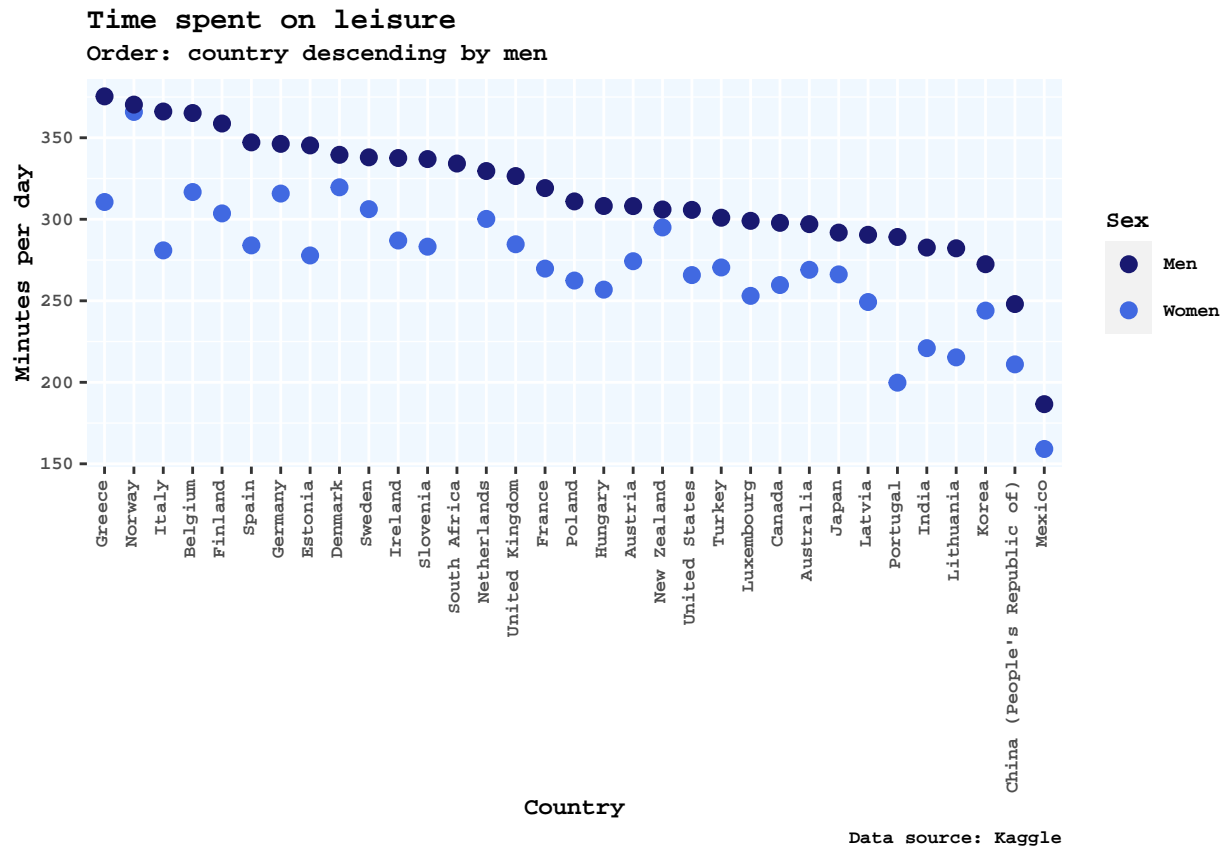


```
#3. Leisure
#create new data frame for leisure only
lei <- timeuse %>% filter(DESC == "LEI")
#order rows by minutes_per_day for men; descending order
order_desc_men <- lei %>% filter(Sex == "Men") %>%
  arrange(desc(Minutes_per_day)) %>%
  mutate(Country = factor(Country))
#plot the data; descending order of countries by values for men
lei %>% mutate(Country = factor(Country, levels = order_desc_men$Country, ordered = TRUE)) %>%
  ggplot(aes(x = Country, y = Minutes_per_day, color = Sex), group = Country) +
  #point size
  geom_point(size = 2.5) +
  #name axes
  labs(x = "Country", y = "Minutes per day",
    #add title, subtitle, caption
    title = "Time spent on leisure",
    subtitle = "Order: country descending by men",
```

```

caption = "Data source: Kaggle") +
#set different colours for sexes
scale_color_manual(values = c("#191970", "#4169E1")) +
#adjust angles and position of text on x axis
theme(axis.text.x = element_text(angle = 90, vjust = 0.25, hjust = 1),
#set background colour
panel.background = element_rect(fill = "aliceblue"),
#text styling
text = element_text(family = "mono", face = "bold", size = 9))

```



```

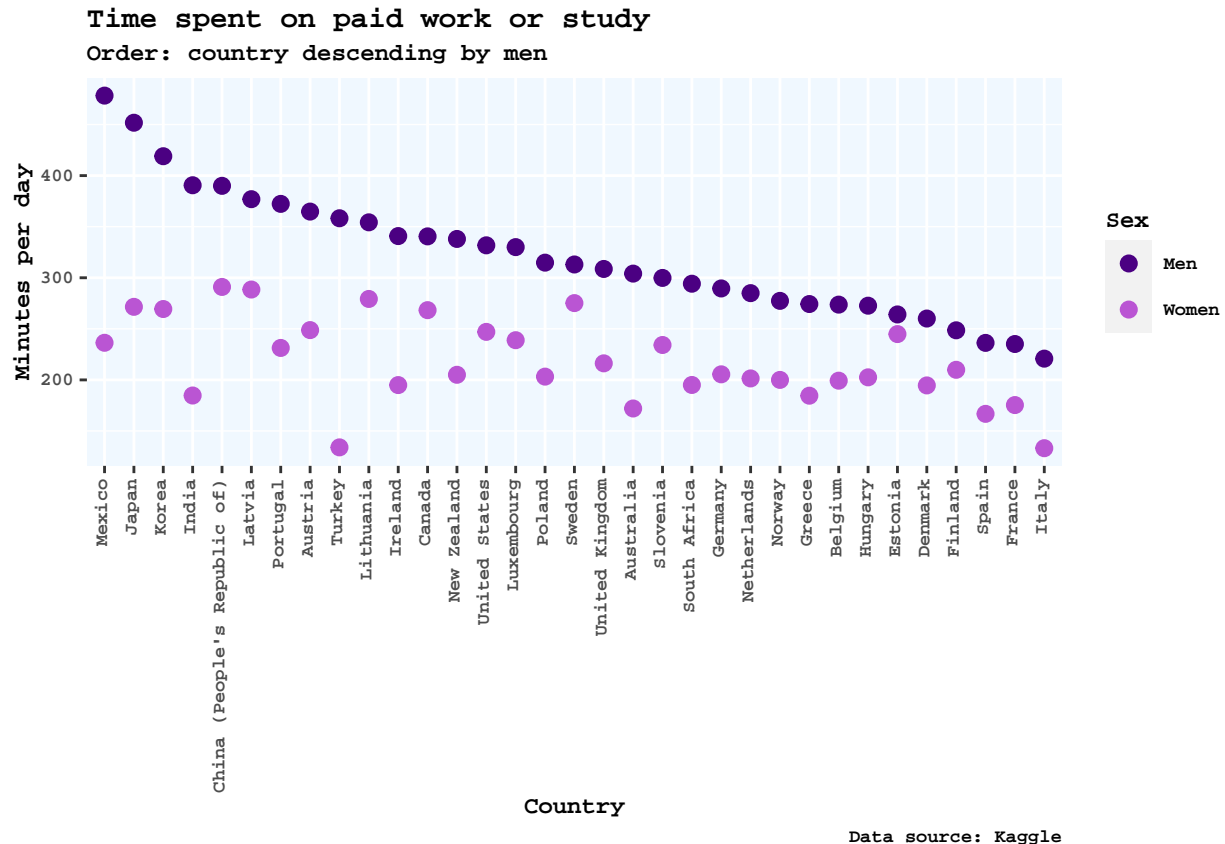
#4. Paid work or study
#new data frame including paid work or study only
paw <- timeuse %>% filter(DESC == "PAW")
#order rows (countries) by minutes_per_day values for men; descending order
order_desc_men <- paw %>% filter(Sex == "Men") %>%
  arrange(desc(Minutes_per_day)) %>%
  mutate(Country = factor(Country))
#plot the data; countries in a descending order by values for men
paw %>% mutate(Country = factor(Country, levels = order_desc_men$Country, ordered = TRUE)) %>%
  ggplot(aes(x = Country, y = Minutes_per_day, color = Sex), group = Country) +
  #set point size
  geom_point(size = 2.5) +
  #name axes; add title, subtitle and caption
  labs(x = "Country", y = "Minutes per day",
        title = "Time spent on paid work or study",
        subtitle = "Order: country descending by men",

```

```

caption = "Data source: Kaggle") +
#adjust point colours
scale_color_manual(values = c("#4B0082", "#BA55D3")) +
#x axis text position and angle; background colour; text styling
theme(axis.text.x = element_text(angle = 90, vjust = 0.25, hjust = 1),
      panel.background = element_rect(fill = "aliceblue"),
      text = element_text(family = "mono", face = "bold", size = 9))

```

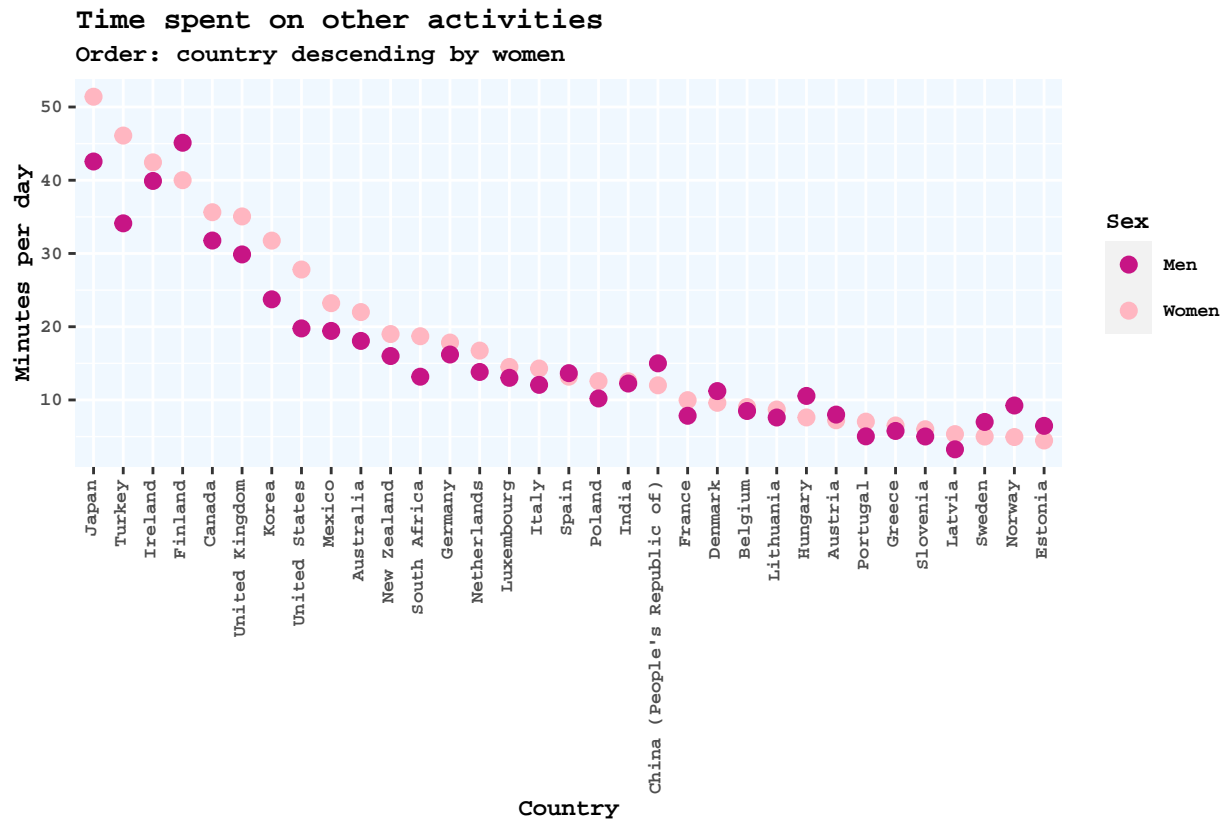


```

#5. Other activities
#data frame for other activities only
oth <- timeuse %>% filter(DESC == "OTH")
#order countries by minutes_per_day values for women
order_desc_women <- oth %>% filter(Sex == "Women") %>%
  arrange(desc(Minutes_per_day)) %>%
  mutate(Country = factor(Country))
#plot the data; countries in a descending order by values for women
oth %>% mutate(Country = factor(Country, levels = order_desc_women$Country, ordered = TRUE)) %>%
  ggplot(aes(x = Country, y = Minutes_per_day, color = Sex), group = Country) +
  #set point size
  geom_point(size = 2.5) +
  #name axes and add title, subtitle and caption
  labs(x = "Country", y = "Minutes per day",
       title = "Time spent on other activities",
       subtitle = "Order: country descending by women",
       caption = "Data source: Kaggle") +
  #set point colours

```

```
scale_color_manual(values = c("#C71585", "#FFB6C1")) +
#set the position and angle of text on x axis
theme(axis.text.x = element_text(angle = 90, vjust = 0.25, hjust = 1),
#adjust background colour and text styling
panel.background = element_rect(fill = "aliceblue"),
text = element_text(family = "mono", face = "bold", size = 9))
```



```
#data preparation by calculating time difference between sexes [minutes men - minutes women] in different countries
#delete unnecessary columns
timeuse <- timeuse_raw %>% select(Country, Description, Sex, Minutes_per_day)

#create unpaid work data frame and calculate differences
df1 <- timeuse %>% filter(Description == "Unpaid work") %>% group_by(Country) %>% mutate(Difference = Minutes_per_day_Men - Minutes_per_day_Women)

#create personal care data frame and calculate differences
df2 <- timeuse %>% filter(Description == "Personal care") %>% group_by(Country) %>% mutate(Difference = Minutes_per_day_Men - Minutes_per_day_Women)

#create leisure data frame and calculate differences
df3 <- timeuse %>% filter(Description == "Leisure") %>% group_by(Country) %>% mutate(Difference = Minutes_per_day_Men - Minutes_per_day_Women)

#create paid work or study data frame and calculate differences
df4 <- timeuse %>% filter(Description == "Paid work or study") %>% group_by(Country) %>% mutate(Difference = Minutes_per_day_Men - Minutes_per_day_Women)
```

```

#create other data frame and calculate differences
df5 <- timeuse %>% filter(Description == "Other") %>% group_by(Country) %>% mutate(Difference = Minutes_per_day - Minutes_per_day_women)

#merge all created data frames
df_list <- list(df1, df2, df3, df4, df5)
timeuse_diff <- df_list %>% reduce(full_join)

#delete sex column
timeuse_diff <- timeuse_diff %>% select(Country, Description, Minutes_per_day, Difference)

#delete rows with NA
timeuse_diff <- timeuse_diff %>% drop_na(Difference)

#show first few rows of timeuse_diff
head(timeuse_diff)

```

The difference in the time spent on various activities

```

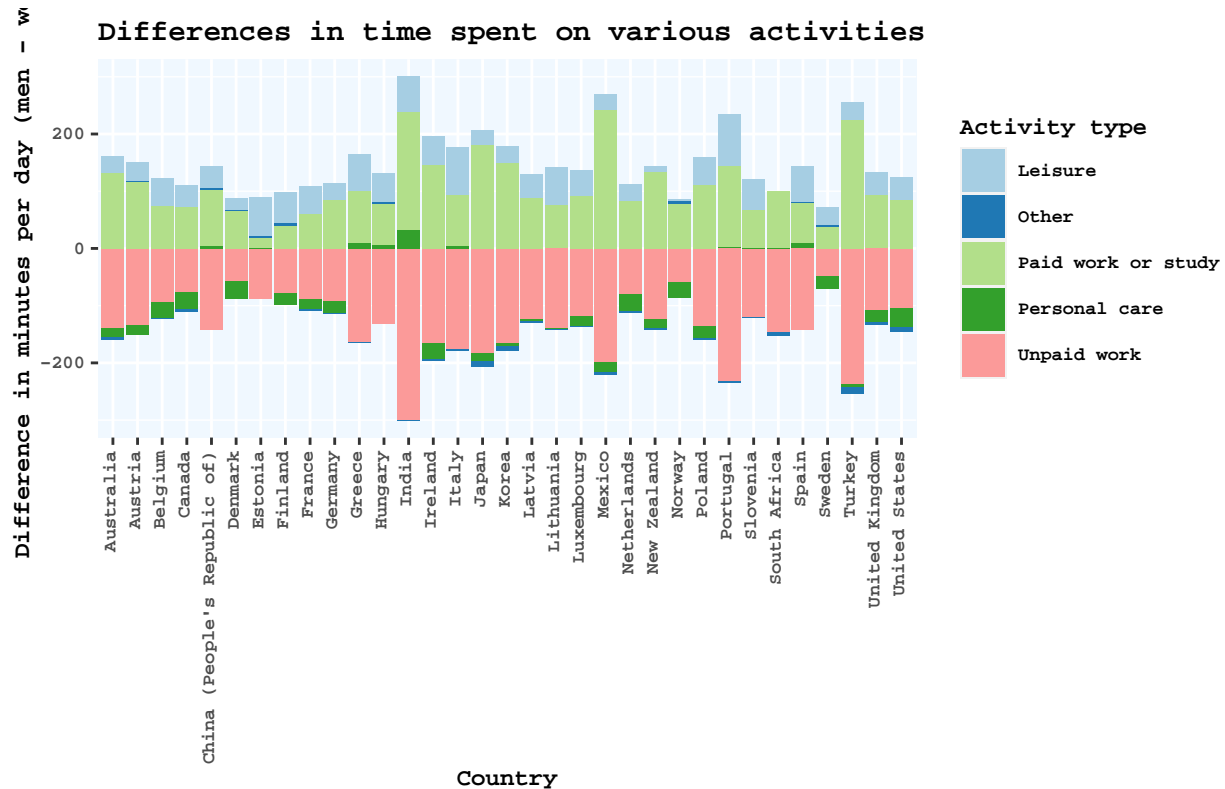
## # A tibble: 6 x 4
## # Groups:   Country [6]
##   Country Description Minutes_per_day Difference
##   <chr>      <chr>          <dbl>      <dbl>
## 1 Australia Unpaid work      172.      -139.
## 2 Austria   Unpaid work      135.      -134.
## 3 Belgium   Unpaid work      144.       -93.2
## 4 Canada    Unpaid work      148.       -75.5
## 5 Denmark   Unpaid work      186.       -56.7
## 6 Finland   Unpaid work      157.       -78.3

```

```

#visualise
timeuse_diff %>% ggplot(mapping = aes(x = Country, y = Difference)) +
  geom_bar(aes(fill = Description), stat = "identity") +
  #chose colour palette
  scale_fill_brewer(palette = "Paired") +
  #adjust position of the text on x axis for readability
  theme(axis.text.x = element_text(angle = 90, vjust = 0.25, hjust = 1),
        #chose background colour
        panel.background = element_rect(fill = "aliceblue"),
        #customise text - font, size, etc.
        text = element_text(family = "mono", face = "bold", size = 9)) +
  #colour fill by activity type
  labs(fill = "Activity type",
        #name axes, include title and caption
        x = "Country", y = "Difference in minutes per day (men - women)",
        title = "Differences in time spent on various activities",
        caption = "The differences were calculated by subtracting minutes spent by women from men.\nNegative values indicate more time spent by women than men.")

```

The differences were calculated by subtracting minutes spent by women from men.
 Negative values indicate that women spend more time a day on an activity.
 Positive values indicate that men spend more time on an activity.

Summary

Brief interpretation of graphs The graphs illustrate that there are indeed differences in the time women and men spent on different activities. The graphs clearly indicate that, in general, women spend more time on unpaid work, whereas men spend more time on leisure and paid work or study. In terms of personal care and other activities, it seems that the time is rather similar with women spending seemingly more time on personal care than men. There does not appear to be a clear pattern in the data set. Whether men or women spend more time on personal care or other activities seems to depend on a country.

Brief thoughts Overall, as someone who has never had any experience with RStudio before, I am slightly proud of what I have managed to accomplish with this project. Especially, considering how challenging this whole module was for me. If I had more time, I would try to make my graphs interactive. I would combine the first five showing the time spent on various activities into one and order countries by the time differences between men and women. This is something I have tried but failed to do, but hopefully will figure out how to do in the future.

What else can be done I think it would be interesting to group data by continents to find out whether cultural differences might be influential on the time difference between men and women. Moreover, the age range in this data set was extremely broad (age 15-64), therefore it may also be useful to find more data and find out whether time differences in sexes vary based on age and/or marital status.

Reference: Chapa, F. (2021). *World time use, work hours and GDP*. [Data File]. Kaggle. Retrieved from: https://www.kaggle.com/datasets/felipechapa/time-use-employment-and-gdp-per-country?select=Time_use_OECD.csv