

Assignment 2

Nesta Ndungu

```
rm(list=ls())  
library(tidyverse)
```

```
-- Attaching packages ----- tidyverse 1.3.2 --  
v ggplot2 3.4.0      v purrr   1.0.1  
v tibble  3.1.8      v dplyr  1.1.0  
v tidyr   1.3.0      v stringr 1.5.0  
v readr   2.1.3      v forcats 1.0.0  
-- Conflicts ----- tidyverse_conflicts() --  
x dplyr::filter() masks stats::filter()  
x dplyr::lag()    masks stats::lag()
```

```
library(rvest)
```

Attaching package: 'rvest'

The following object is masked from 'package:readr':

guess_encoding

```
# Scraping the html code from the page source.  
  
motor <- minimal_html('<table class="" style="">  
  
  <thead>  
  <tr>  
  <td>
```



```

<tr><td>Cupra Born</td><td>395 km&#x2F;15,4 kWh</td><td>339 km</td><td>-14,18 %</td></tr>
<tr><td>Volvo C40 Recharge</td><td>437 km&#x2F;21,1 kWh</td><td>333 km</td><td>-23,80 %</td></tr>
<tr><td>Mercedes-Benz EQA 250</td><td>401 km&#x2F;17,7 kWh</td><td>331 km</td><td>-17,46 %</td></tr>
<tr><td>BMW iX xDrive40</td><td>402 km&#x2F;20,7 kWh</td><td>316 km</td><td>-21,39 %</td></tr>
<tr><td>Mercedes-Benz EQB 350 4matic</td><td>407 km&#x2F;18,1 kWh</td><td>315 km</td><td>-21,39 %</td></tr>
<tr><td>Opel Mokka-e</td><td>338 km&#x2F;16,2 kWh</td><td>263 km</td><td>-22,19 %</td></tr>
<tr><td>Peugeot e-2008</td><td>320 km&#x2F;15,6 kWh</td><td>228 km</td><td>-28,75 %</td></tr>
</tbody>
</table>')

```

```
# Making the dataframe.
```

```

df_motor <- motor %>%
  html_element("table") %>%
  html_table()

```

```
# Renaming the columns.
```

```
df_motor <- rename(df_motor, Modell = "X1", WLTP_tall = "X2", STOPP = "X3", Avvik = "X4")
```

```
# Removing the first row.
```

```
df_motor <- df_motor[-1,]
```

```
# Removing characters that are unnecessary.
```

```

df_motor$WLTP_tall <- substr(df_motor$WLTP_tall,1,nchar(df_motor$WLTP_tall)-12)
df_motor$STOPP <- substr(df_motor$STOPP,1,nchar(df_motor$STOPP)-3)

```

```
# Making the values in the columns as numeric.
```

```

df_motor$WLTP_tall <- as.numeric(as.character(df_motor$WLTP_tall))
df_motor$STOPP <- as.numeric(as.character(df_motor$STOPP))

```

```
# Making the plot.
```

```
# Found the code for the diagonal line at: https://statisticsglobe.com/add-diagonal-line-p
```

```
df_motor %>%
```

```

ggplot() +

geom_point(aes(x = WLTP_tall, y = STOPP)) +

geom_abline(intercept = 0, slope = 1) +

labs (title = "WLTP test of electric cars", y = "STOPP", x = "WLTP") +

theme_bw()

```

Warning: Removed 2 rows containing missing values (`geom_point()`).

