STOP COMPLAINING, IT SOLVES NOTHING.

My project will provide insights about **waste management**, which is of vital importance to the world and humanity. I hope to convey the awareness and perspective I wish to share with you, since change begins with awareness, and **we have to change our way** so that our children can live in the world they deserve!

1. Are we aware?

Important fact: Are we aware that when we do not recycle waste or use it for energy consumption, it pollutes our groundwater, our soil, and the air through greenhouse gases emitted by the waste, ultimately reducing the quality of the food produced in our soil and our overall life quality?

We just complain, don't we? Strawberries used to smell like strawberries, tomatoes used to taste different… right?

Unfortunately, complaining doesn't fix anything, and it won't. If we want to deserve to live in this world, we must work hard for our generation. The effort we do not put into our waste will heavily come back to haunt us and our children in this universe created with karma. **Let's quit complaining and start acting!**

The scope: In my project, I will attempt to extract insights from several data collections on waste management in Turkiye. I will particularly focus on municipal waste statistics. Initially, I will analyze data on a national level, then move on to a provincial basis. Later, I will use time series methods to forecast future waste amount trends in Turkiye. Additionally, using clustering methods, I will group provinces based on their waste amounts behaviors.

The aim: In this project, my main goals are to increase awareness about waste issues, determine future waste quantities, investigate the waste levels of the provinces, and discuss both prevention strategies and proper disposal and recycling methods for unavoidable waste.

2. Data 

I plan to use multiple data sources for analysis. The data I will utilize include waste quantities, waste types and population figures for Turkiye and its provinces.

2.1 Data Source 

The references from which I have gathered the data may include:

- Waste Statistics, TURKSTAT
- Environment Regional Data, Biruni TURKSTAT
- Waste Data, Ankara Municipality

2.2 General Information About Data 

- The municipal waste amount data of Turkiye, which includes information such as the population of Turkiye and its 81 provinces' municipalities, total waste amounts for the year 2022, the average waste amount per person, etc.:

```
library(openxlsx)
```

```
Warning: package 'openxlsx' was built under R version 4.3.3
```

```
municipal_waste <- read.xlsx("project/data/municipal_wasteamount_2022.xlsx")
str(municipal_waste)
```

```
'data.frame':   88 obs. of  8 variables:
 $ Belediye.atık.hizmeti.istatistikleri,.2022.Municipal.waste.services.statistics,.2022: chr
 $ X2                                                                                    : chr
 $ X3                                                                                    : chr
 $ X4                                                                                    : chr
 $ X5                                                                                    : chr
 $ X6                                                                                    : chr
 $ X7                                                                                    : chr
 $ X8                                                                                    : chr
```

- Data including the amounts of collected municipal waste that are sent to municipal landfills, waste processing facilities (the waste sent to landfill sites, incineration plants and all the waste recovery facilities), and disposed of using other methods (disposals by burning in an open area, dumping into river/onto land and burying.):

```
where_to_municipal_waste <- read.xlsx("project/data/municipal_waste_stat.xlsx")
str(where_to_municipal_waste)
```

```
'data.frame':   94 obs. of  10 variables:
 $ Belediye.atık.yönetimi.istatistikleri,.2022.Municipal.waste.management.statistics,.2022: ch
 $ X2                                                                                       : ch
 $ X3                                                                                       : ch
 $ X4                                                                                       : ch
 $ X5                                                                                       : ch
 $ X6                                                                                       : ch
 $ X7                                                                                       : ch
 $ X8                                                                                       : nu
 $ X9                                                                                       : ch
 $ X10                                                                                      : ch
```

- Time series data including municipal waste amounts, waste per capita, waste sent to processing facilities, etc., for the years 1994-2022:

```
time_series_municipal_waste <- read.xlsx("project/data/municipal_waste_timeseries.xlsx", colNa
str(time_series_municipal_waste)
```

```
'data.frame':   43 obs. of  19 variables:
 $ X1  : chr  "Türkiye nüfusu\nTurkey population" "Toplam belediye sayısı\nTotal number of mun
 $ 1994: chr  "62810111" "2740" "47597657" "1985" ...
 $ 1995: chr  "62810111" "2801" "47774543" "2126" ...
 $ 1996: chr  "62810111" "2827" "47843698" "2172" ...
 $ 1997: chr  "62810111" "2835" "47865511" "2275" ...
 $ 1998: chr  "62810111" "2834" "47862511" "2579" ...
 $ 2001: chr  "67803927" "3227" "53407613" "2921" ...
 $ 2002: chr  "67803927" "3227" "53421379" "2984" ...
 $ 2003: chr  "67803927" "3227" "53430733" "3018" ...
 $ 2004: chr  "67803927" "3225" "53935050" "3028" ...
 $ 2006: num  70586256 3225 58581515 3115 57451562 ...
 $ 2008: num  70586256 3225 58581515 3129 57800347 ...
 $ 2010: num  73722988 2950 61571332 2879 60946131 ...
 $ 2012: num  75627384 2950 63743047 2894 63105474 ...
 $ 2014: num  77695904 1396 72505107 1391 70843913 ...
 $ 2016: num  79814871 1397 74911343 1390 73854880 ...
 $ 2018: num  82003882 1399 76888607 1395 75952539 ...
 $ 2020: num  83614362 1389 78920614 1387 78204213 ...
 $ 2022: num  85279553 1391 80785141 1389 80319403 ...
```

- Data on the types and amounts of waste collected annually in Ankara:

```
Ankara_waste_type_year <- read.xlsx("project/data/Ankara_wastetype_year.xlsx")
str(Ankara_waste_type_year)
```

```
'data.frame':   6 obs. of  5 variables:
 $ kg./.yıl: chr  "20 01 34 Atık Pil " "08 03 17 Kartuş Toner atıkları " "Kağıt Karton Ambalaj
 $ 2019    : num  125 2631 19040 8710 5985 ...
 $ 2020    : num  52 2420 16040 7386 13763 ...
 $ 2021    : num  53 1860 23405 13794 8455 ...
 $ 2022    : num  0 250 1300 1210 1010 1020
```

2.3 Reason of Choice

This topic was chosen because it was realized that **waste management is not given enough importance** in Turkiye, and it is believed that carelessness should not continue in this matter. The importance of the subject is **indisputable**. By using the data sets mentioned above, it is aimed to reveal and analyze the current situation of waste management, to derive knowledge, and to contribute to the literature and our country.

2.4 Preprocessing

- **For "municipal_waste" dataset:**

Downloadable dataset in .RData version

```
library(tidyverse)
```

```
Warning: package 'stringr' was built under R version 4.3.2
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.3     v readr     2.1.4
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.4.4     v tibble    3.2.1
v lubridate 1.9.3     v tidyr     1.3.0
v purrr     1.0.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
# remove unnecessary columns
municipal_waste <- select(municipal_waste, -X5)
municipal_waste <- select(municipal_waste, -X6)
# rename columns
municipal_waste <- rename(municipal_waste, "Provinces" ="Belediye.atık.hizmeti.istatistikleri,
municipal_waste <- rename(municipal_waste, "Total municipal population" = "X2")
municipal_waste <- rename(municipal_waste, "Total number of municipalities" ="X3")
municipal_waste <- rename(municipal_waste, "Number of municipalities providing waste services"
municipal_waste <- rename(municipal_waste, "Amount of waste collected (Tonnes)
" ="X7")
municipal_waste <- rename(municipal_waste, "Amount of waste per capita (Kg/capita-day)
" ="X8")
# remove unnecessary rows
municipal_waste <- municipal_waste[-c(1, 2, 85, 86, 87, 88), ]
```

```r
# reorder row names that is disordered
row.names(municipal_waste) <- NULL
municipal_waste <- municipal_waste %>%
  mutate(row_id = row_number())
municipal_waste <- municipal_waste %>%
  select(row_id, everything())
municipal_waste <- municipal_waste[,-c(1)]
# adjust necessary columns as numbers
municipal_waste <- municipal_waste %>%
  mutate(across(-Provinces, ~as.numeric(as.character(.))))
sapply(municipal_waste,class)
```

```
                                     Provinces
                                   "character"
                       Total municipal population
                                     "numeric"
                    Total number of municipalities
                                     "numeric"
Number of municipalities providing waste services
                                     "numeric"
             Amount of waste collected (Tonnes) \n
                                     "numeric"
    Amount of waste per capita (Kg/capita-day) \n
                                     "numeric"
```

```r
head(municipal_waste)
```

```
        Provinces Total municipal population Total number of municipalities
1         Türkiye                   80785141                            1391
2           Adana                    2274106                              16
3        Adıyaman                     487642                              23
4 Afyonkarahisar                     588048                              60
5            Ağrı                     314539                              12
6          Amasya                     256679                               8
  Number of municipalities providing waste services
1                                              1389
2                                                16
3                                                22
4                                                60
5                                                12
6                                                 8
  Amount of waste collected (Tonnes) \n
1                           30283756.6
2                             665694.6
3                             179724.1
4                             198272.6
5                             181116.0
6                             111099.1
  Amount of waste per capita (Kg/capita-day) \n
1                                     1.0329907
2                                     0.8068814
3                                     1.0190094
```

```
4                                   0.9282235
5                                   1.5796783
6                                   1.1915254
```

```r
summary(municipal_waste)
```

```
  Provinces         Total municipal population Total number of municipalities
 Length:82          Min.   :   41120           Min.   :   4.00
 Class :character   1st Qu.:  242100           1st Qu.:  11.00
 Mode  :character   Median :  417945           Median :  16.00
                    Mean   : 1970369           Mean   :  33.93
                    3rd Qu.: 1139026           3rd Qu.:  21.00
                    Max.   :80785141           Max.   :1391.00
 Number of municipalities providing waste services
 Min.   :   4.00
 1st Qu.:  11.00
 Median :  16.00
 Mean   :  33.88
 3rd Qu.:  21.00
 Max.   :1389.00
 Amount of waste collected (Tonnes) \n
 Min.   :   21392
 1st Qu.:   81725
 Median :  151235
 Mean   :  738628
 3rd Qu.:  415882
 Max.   :30283757
 Amount of waste per capita (Kg/capita-day) \n
 Min.   :0.6498
 1st Qu.:0.8730
 Median :0.9672
 Mean   :1.0679
 3rd Qu.:1.1887
 Max.   :1.9962
```

- **For "where_to_municipal_waste" dataset:**

Downloadable dataset in .RData version

```r
library(tidyverse)
# remove unnecessary columns
where_to_municipal_waste <- select(where_to_municipal_waste, -X2)
where_to_municipal_waste <- select(where_to_municipal_waste, -X4)
where_to_municipal_waste <- select(where_to_municipal_waste, -X6)
where_to_municipal_waste <- select(where_to_municipal_waste, -X8)
where_to_municipal_waste <- select(where_to_municipal_waste, -X9)
# rename columns
where_to_municipal_waste <- rename(where_to_municipal_waste, "Provinces" = `Belediye.atık.yöne
where_to_municipal_waste <- rename(where_to_municipal_waste, "Total amount of waste collected
where_to_municipal_waste <- rename(where_to_municipal_waste, "Municipality's dumping sites" ="
where_to_municipal_waste <- rename(where_to_municipal_waste, "Waste treatment facilities"= "X7
where_to_municipal_waste <- rename(where_to_municipal_waste, "Other disposal methods"= "X10")
# remove unnecessary rows
```

```r
where_to_municipal_waste <- where_to_municipal_waste[-c(1, 2, 85:94), ]
# reorder row names that is disordered
row.names(where_to_municipal_waste) <- NULL
where_to_municipal_waste <- where_to_municipal_waste %>%
  mutate(row_id = row_number())
where_to_municipal_waste <- where_to_municipal_waste %>%
  select(row_id, everything())
where_to_municipal_waste <- where_to_municipal_waste[,-c(1)]
# adjust necessary columns as numbers
where_to_municipal_waste <- where_to_municipal_waste %>%
  mutate(across(-Provinces, ~as.numeric(as.character(.))))
sapply(where_to_municipal_waste,class)
```

```
                              Provinces
                            "character"
Total amount of waste collected  (Tonnes)
                              "numeric"
             Municipality's dumping sites
                              "numeric"
               Waste treatment facilities
                              "numeric"
                   Other disposal methods
                              "numeric"
```

```r
head(where_to_municipal_waste)
```

```
       Provinces Total amount of waste collected  (Tonnes)
1        Türkiye                                  30283756.6
2          Adana                                    665694.6
3       Adıyaman                                    179724.1
4 Afyonkarahisar                                    198272.6
5           Ağrı                                    181116.0
6         Amasya                                    111099.1
  Municipality's dumping sites Waste treatment facilities
1                   4092721.21                26016987.70
2                         0.00                  663894.61
3                       178453.00                 1271.15
4                        20089.93                 175951.63
5                       131116.00                  50000.00
6                         1200.00                109680.14
  Other disposal methods
1               174047.7
2                 1800.0
3                    0.0
4                 2231.0
5                    0.0
6                  219.0
```

```r
summary(where_to_municipal_waste)
```

```
  Provinces         Total amount of waste collected  (Tonnes)
```

```
 Length:82          Min.    :   21392
 Class :character   1st Qu.:   81725
 Mode  :character   Median :  151235
                    Mean   :  738628
                    3rd Qu.:  415882
                    Max.   :30283757
 Municipality's dumping sites Waste treatment facilities Other disposal methods
 Min.   :      0              Min.   :       0           Min.   :     0.0
 1st Qu.:   1260              1st Qu.:   52770           1st Qu.:     0.0
 Median :  17376              Median :  109595           Median :     0.0
 Mean   :  99822              Mean   :  634561           Mean   :  4245.1
 3rd Qu.:  56501              3rd Qu.:  333389           3rd Qu.:   557.5
 Max.   :4092721              Max.   :26016988           Max.   :174047.6
```

- **For "time_series_municipal_waste" dataset:**

Downloadable dataset in .RData version

```
library(tidyverse)
# remove unnecessary rows
time_series_municipal_waste <- time_series_municipal_waste[-c(2:6,10,14:43), ]
time_series_municipal_waste <- rename(time_series_municipal_waste, "Waste/Year" = "X1")
# rename rows
time_series_municipal_waste[1,1] <- "Turkey population"
time_series_municipal_waste[2,1] <- "Amount of municipal waste generated (Thousand tonnes/year
time_series_municipal_waste[3,1] <- "Amount of municipal waste collected (Thousand tonnes/year
time_series_municipal_waste[4,1] <- "Average amount of municipal waste per capita (Kg/capita-d
time_series_municipal_waste[5,1] <- "Waste treatment facilities"
time_series_municipal_waste[6,1] <- "Municipality's dumping sites"
time_series_municipal_waste[7,1] <- "Other disposal methods"
# reorder row names that is disordered
row.names(time_series_municipal_waste) <- NULL
# adjust necessary columns as numbers
time_series_municipal_waste <- time_series_municipal_waste %>%
  mutate(across(-`Waste/Year`, ~as.numeric(as.character(.))))
sapply(time_series_municipal_waste,class)
```

```
 Waste/Year        1994        1995        1996        1997        1998
"character"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
       2001        2002        2003        2004        2006        2008
  "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
       2010        2012        2014        2016        2018        2020
  "numeric"   "numeric"   "numeric"   "numeric"   "numeric"   "numeric"
       2022
  "numeric"
```

```
head(time_series_municipal_waste)
```

```
                                                Waste/Year       1994
1                                          Turkey population 62810111.0
2   Amount of municipal waste generated (Thousand tonnes/year)    23448.0
3   Amount of municipal waste collected (Thousand tonnes/year)    17757.0
4 Average amount of municipal waste per capita (Kg/capita-day)        1.1
```

```
5                                 Waste treatment facilities     1001.0
6                               Municipality's dumping sites    14479.0
         1995        1996        1997        1998        2001        2002
1 62810111.00 62810111.00 62810111.00 62810111.00 67803927.00 67803927.00
2    27234.13    29347.98    31943.77    32972.89    31030.87    30999.26
3    20910.00    22483.00    24180.00    24945.00    25134.00    25373.00
4        1.27        1.37        1.46        1.51        1.35        1.34
5     1603.00     3026.00     4544.00     5424.00     8522.00     7430.00
6    17175.00    17520.00    16805.00    16853.00    14570.00    16310.00
         2003        2004        2006        2008        2010        2012
1 67803927.00 67803927.00 70586256.00 70586256.00 73722988.00 75627384.00
2    31081.37    29736.10    30081.82    28454.00    29733.00    30786.00
3    26118.00    25014.00    25280.00    24361.00    25277.00    25845.00
4        1.38        1.31        1.21        1.15        1.14        1.12
5     7758.00     7353.00     9683.00    11223.00    13941.00    15639.00
6    16567.00    16416.00    14941.00    12678.00    11001.00     9772.00
         2014         2016         2018         2020         2022
1 77695904.00 7.981487e+07 8.200388e+07 8.361436e+07 8.527955e+07
2    31230.00 3.376346e+04 3.453265e+04 3.475776e+04 3.242218e+04
3    28011.00 3.158355e+04 3.220922e+04 3.232447e+04 3.028376e+04
4        1.08 1.171626e+00 1.161837e+00 1.132422e+00 1.032991e+00
5    17933.00 2.243038e+04 2.561468e+04 2.670724e+04 2.601700e+04
6     9935.00 9.094906e+03 6.520657e+03 5.492803e+03 4.093000e+03
```

- **For "Ankara_waste_type_year" dataset:** There is no need to preprocess the data

Downloadable dataset in .RData version

3. Analysis 

In the first phase of this section, *Exploratory Data Analysis (EDA)*, the data prepared for analysis is visualized to enable discoveries that are not immediately apparent at first glance. Later, relationships will be determined through *regression analysis*, future predictions will be made using *time series methods*, and provinces will be clustered according to their patterns using *clustering methods*.

3.1 Exploratory Data Analysis 

"The greatest value of a picture is when it forces us to notice what we never expected to see." John W. Tukey

- **For "municipal_waste" dataset:**

```
library(tidyverse)
library(ggthemes)
```

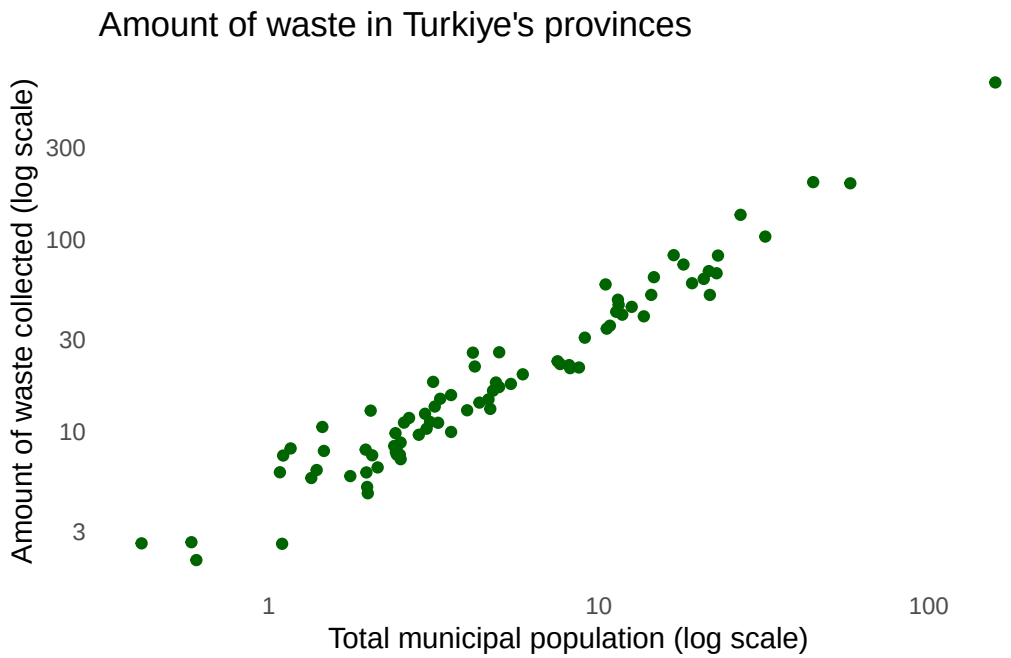
```
Warning: package 'ggthemes' was built under R version 4.3.2
```

```
library(ggrepel)
```

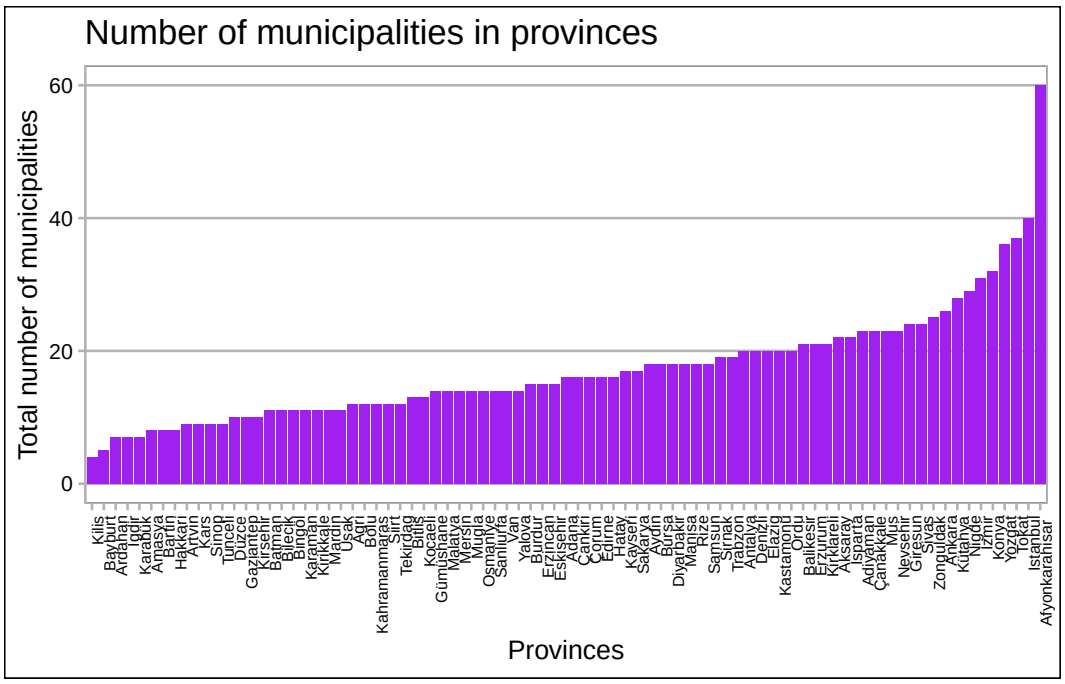
```
Warning: package 'ggrepel' was built under R version 4.3.2
```

```
municipal_waste <- municipal_waste[-c(1), ]

p <- municipal_waste |> ggplot(aes(`Total municipal population`/10^5,
`Amount of waste collected (Tonnes)
`/10^4))
p + geom_point(color = "darkgreen") +
scale_x_continuous(trans = "log10") +
scale_y_continuous(trans = "log10") +
xlab("Total municipal population (log scale)") +
ylab("Amount of waste collected (log scale)") +
ggtitle("Amount of waste in Turkiye's provinces")+
 theme(line = element_blank(), rect = element_blank())
```

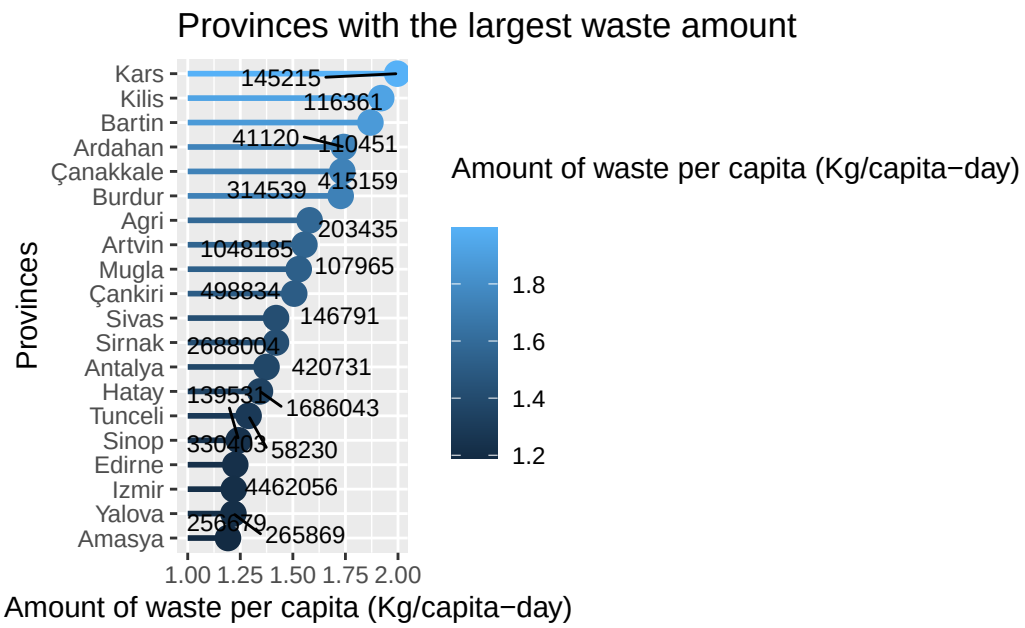



```
p <- ggplot(municipal_waste, aes(x = reorder(Provinces, `Total number of municipalities`, FUN
p + geom_bar(stat = "identity", fill= "purple") +
  xlab("Provinces") +
  theme_calc() +
  ggtitle("Number of municipalities in provinces") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, size = 6))
```




```
# The provinces that produce largest amount of waste
The_largest <- municipal_waste |> arrange(desc(`Amount of waste per capita (Kg/capita-day)
`)) |> head(n = 20)
p <- ggplot(The_largest, aes(x = reorder(Provinces, `Amount of waste per capita (Kg/capita-day
`, FUN = sum),
                             y = `Amount of waste per capita (Kg/capita-day)
`))
ggplot(The_largest, aes(x =`Amount of waste per capita (Kg/capita-day)
`, y = reorder(Provinces, `Amount of waste per capita (Kg/capita-day)
`, FUN = sum), color = `Amount of waste per capita (Kg/capita-day)
`)) +
  geom_point(size = 4) +
  geom_segment(aes(xend = 1, yend = Provinces), size = 1) +
  ylab("Provinces") +
  ggtitle("Provinces with the largest waste amount")+
  geom_text_repel(aes(label =`Total municipal population`), color = "black", size = 3)
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```



- **For "where_to_municipal_waste" dataset:**
- **For "time_series_municipal_waste" dataset:**
- **For "Ankara_waste_type_year" dataset:**

3.2 Trend Analysis

3.3 Model Fitting

3.4 Results

4. Discussion and Key Takeaways