# Assignment 1

Arda Eren Yilmaz

2025-10-06

```
# Installing necessary package(s)
install.packages("ggplot2")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```
install.packages("tidyverse")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.1     v stringr   1.5.2
## v lubridate 1.9.4     v tibble    3.3.0
## v purrr     1.1.0     v tidyr     1.3.1
```

```
## -- 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 errors
```

## Dataset

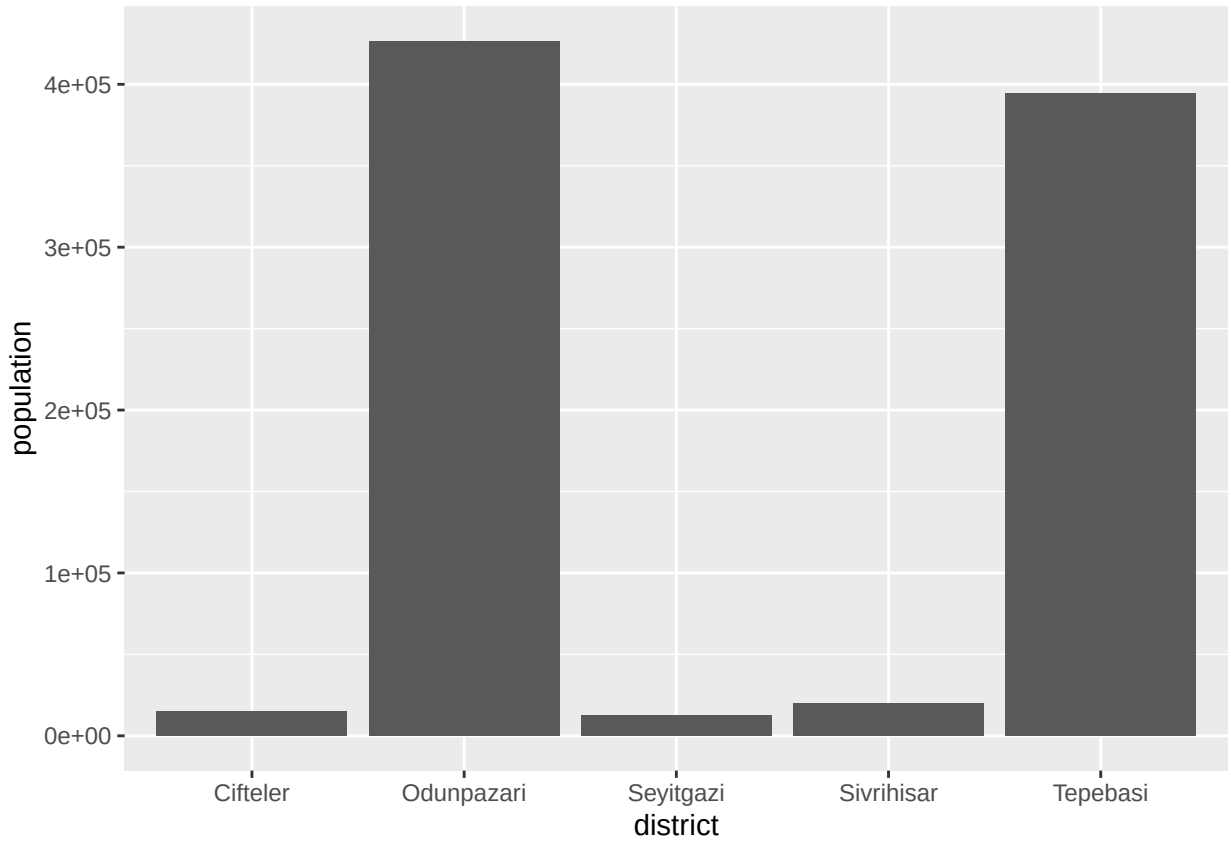
The dataset `eskisehir` contains the total population for the top-5 crowded districts of Eskisehir in 2024.

```
eskisehir <- data.frame(
  district   = c("Odunpazari", "Tepebasi", "Sivrihisar", "Cifteler", "Seyitgazi"),
  population = c(426581, 394734, 20258, 14814, 12878))
```

## Drawing a barplot

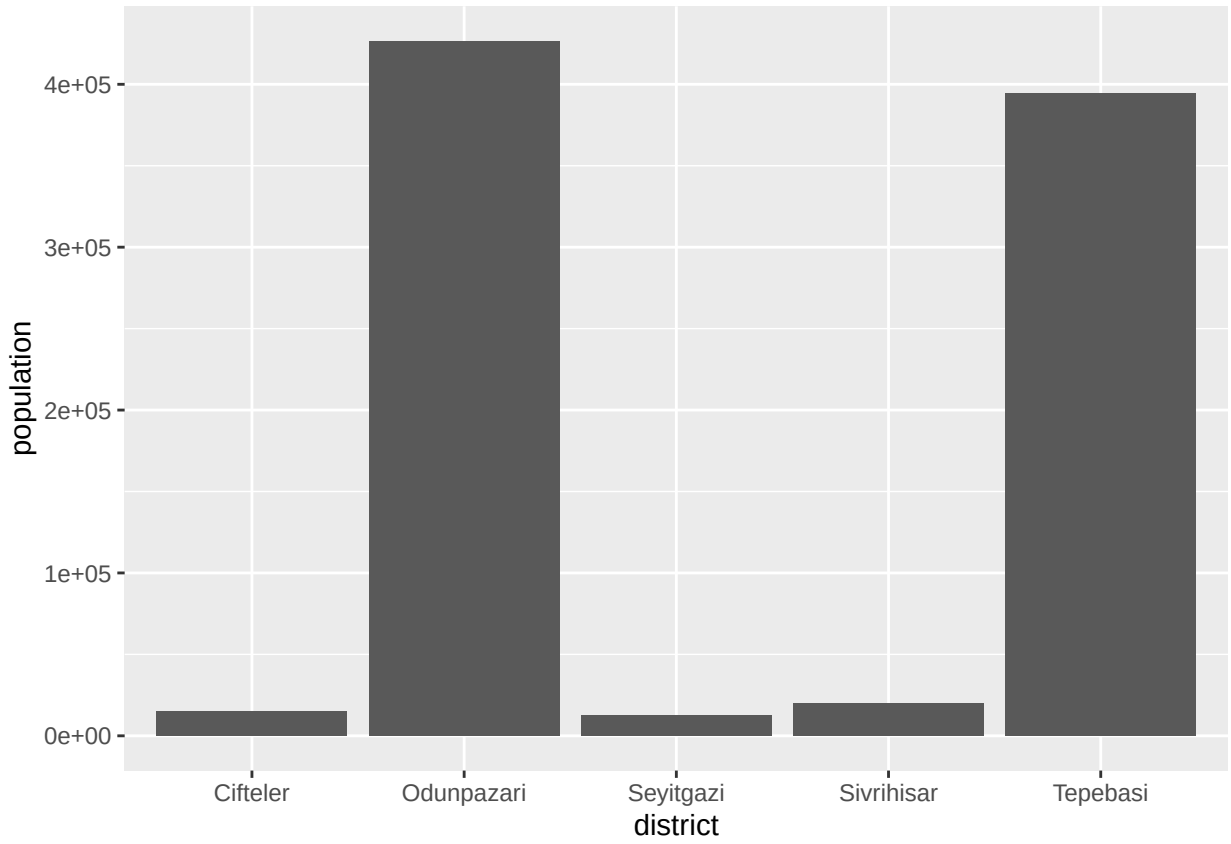
1. Please draw a boxplot to visualize the population of districts (20 pts).

```
ggplot(eskisehir, aes(x = district, y = population)) +
  geom_bar(stat = "identity")
```



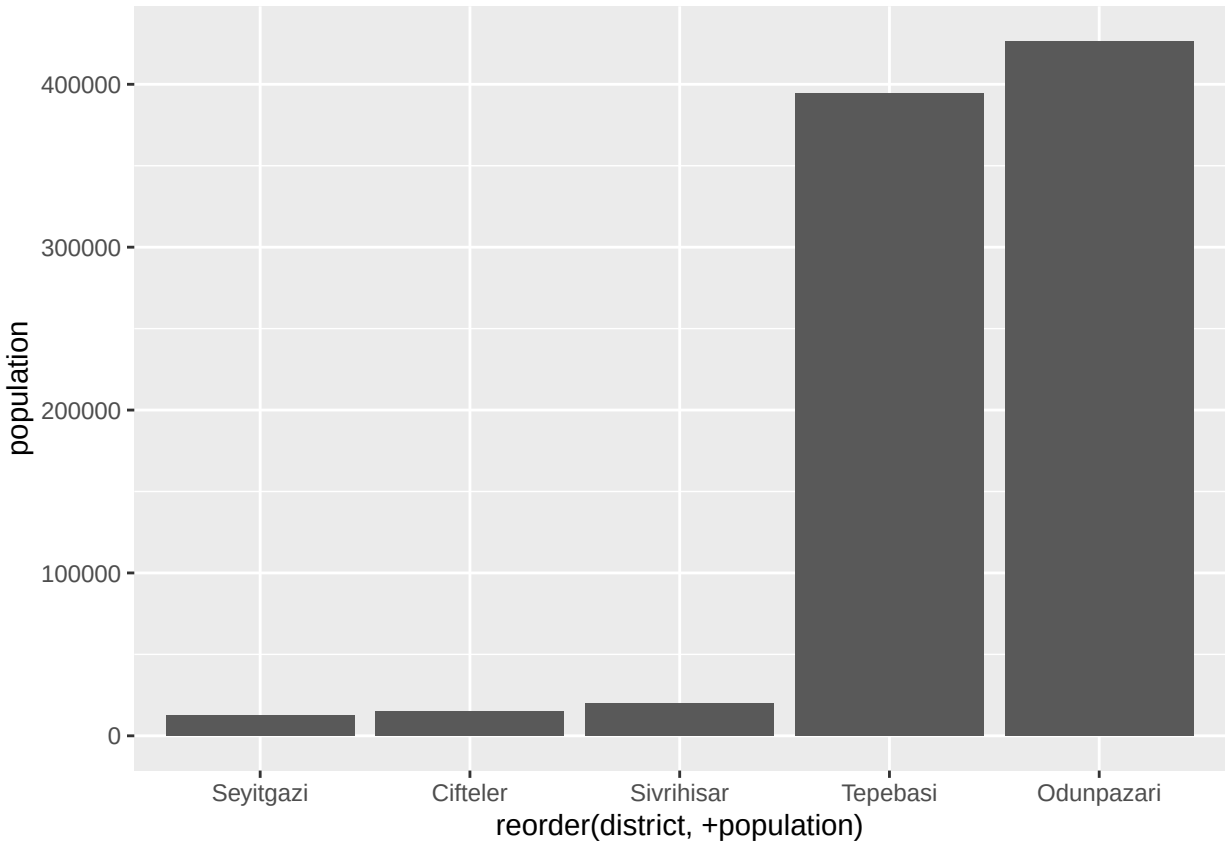

2. Please solve the problems in the plot to make it better (20 pts).

```
ggplot(eskisehir, aes(x = district, y = population)) +
  geom_bar(stat = "identity")
```




3. Then configure the plot to present it in a better way (40 pts). You can use the tricks that you learned in the lecture.

```
ggplot(eskisehir, aes(x = reorder(district, +population), y = population)) +
  geom_bar(stat = "identity") +
  scale_y_continuous(labels = function(x) format(x, scientific = FALSE))
```




4. Interpret the plot (20 pts). It must be about the information of the district populations not the technical part of the plot.

This data show how much density the districts are containing. Tepebasi and Odunpazari districts are the main populated (urban districs) districs as the data shows. The data contains positioning and ordered bar plot data. ....