Assignment Lab 1

Ceren Yıldız

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)
```

```
library(ggplot2)
install.packages("dplR")
```

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

```
library(dplR)
```

```
## This is dplR version 1.7.8.
## dplR is part of openDendro https://opendendro.org.
## New users can visit https://opendendro.github.io/dplR-workshop/ to get started.
```

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

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

```
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

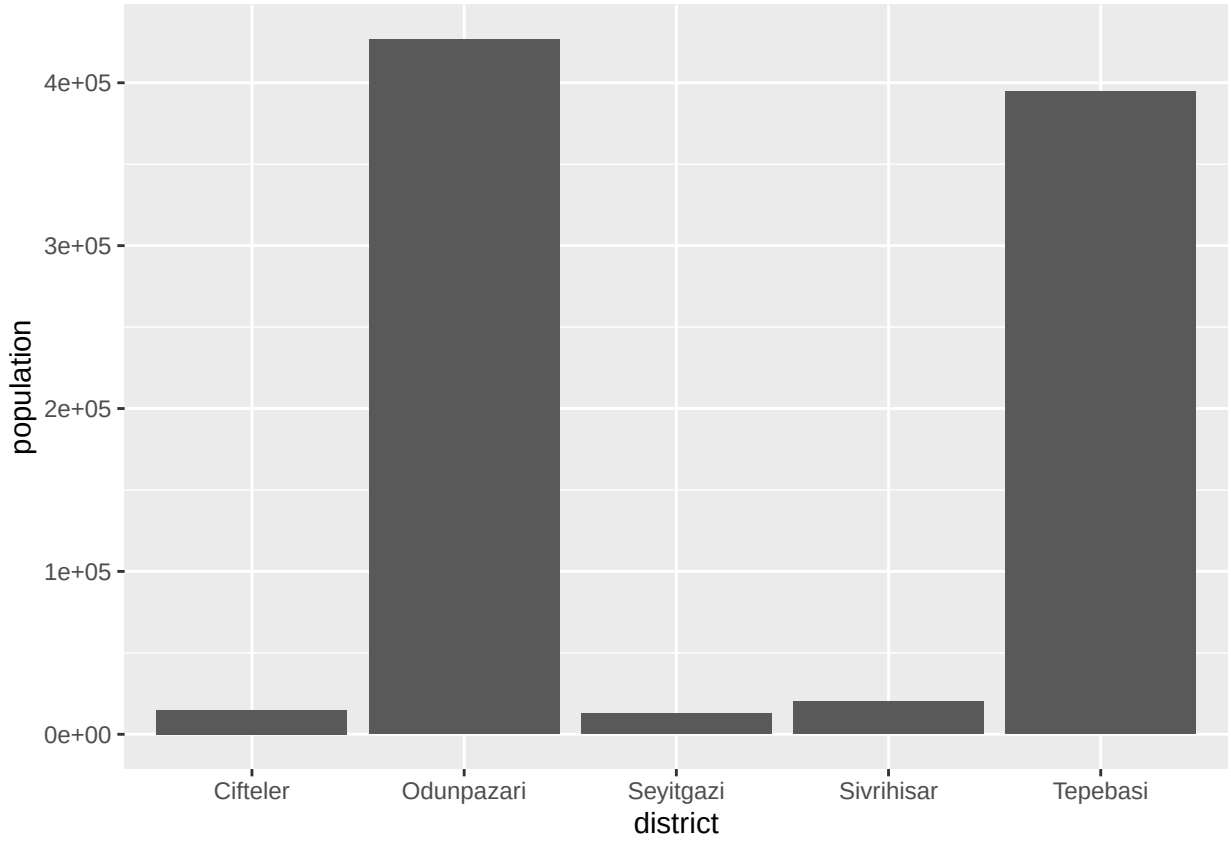
The dataset `eskisehir` contains the total population for the top-5 crowded districts of Eskisehir in 2024.

```
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).

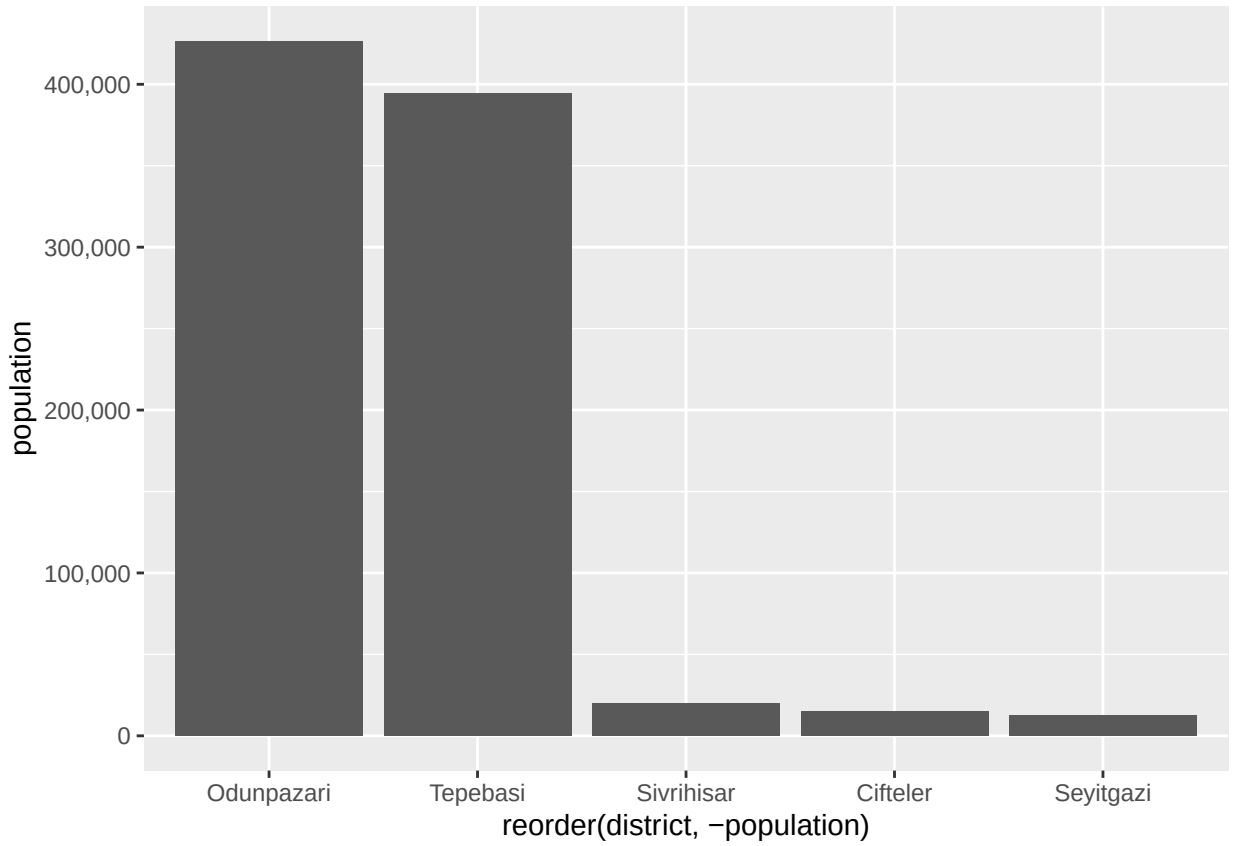
```
ggplot(eskisehir, aes(x = district, y = population)) +
  geom_bar(stat = "identity")  # if you use geom_bar() , you will see error. You have to  write geom_ba
```




```
#
```

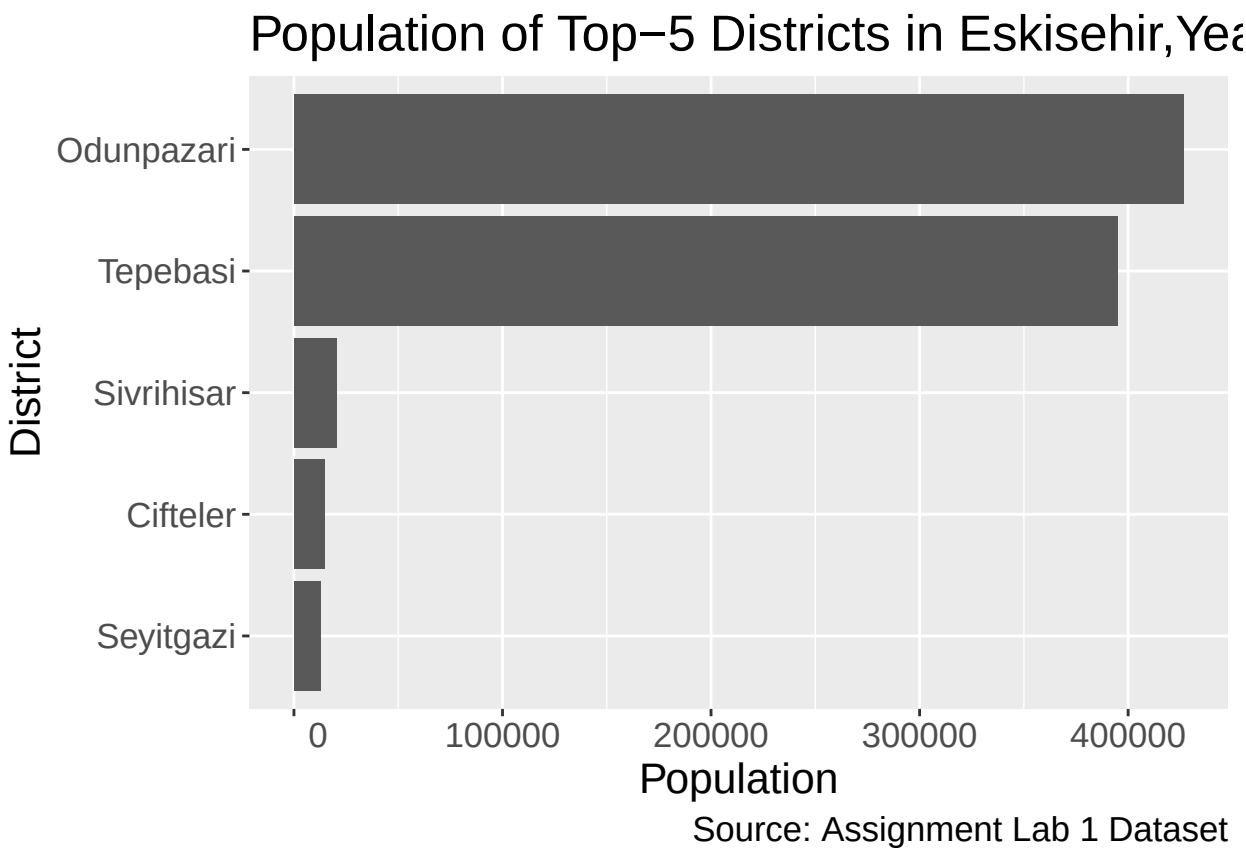
2. Please solve the problems in the plot to make it better (20 pts).

```
ggplot(eskisehir, aes(x = reorder(district, -population), y = population)) +
  geom_bar(stat = "identity") +
  scale_y_continuous(labels = scales::comma) #this graph is ordered data.Also the texts are not neatly
```



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)) +
  labs(
    x = "District",
    y = "Population",
    title = "Population of Top-5 Districts in Eskisehir,Year 2024",
    caption = "Source: Assignment Lab 1 Dataset"
  ) +
  coord_flip() +
  theme(text = element_text(size = 16))
```



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

Odunpazari has the largest population.Later followed by Tepebasi. There is a very large gap between these two districts and the remaining three (Sivrihisar, Cifteler, Seyitgazi). Each of the latter has fewer than 25,000 people, while Odunpazari and Tepebasi are around 400,000.