Assignment 1

Can Binay

2025-10-06

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

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

```r
install.packages("dplyr")
```

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

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

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

```r
library("ggplot2")
library("dplyr")
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library("tidyverse")
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## 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
## v readr     2.1.5

## -- 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))
```
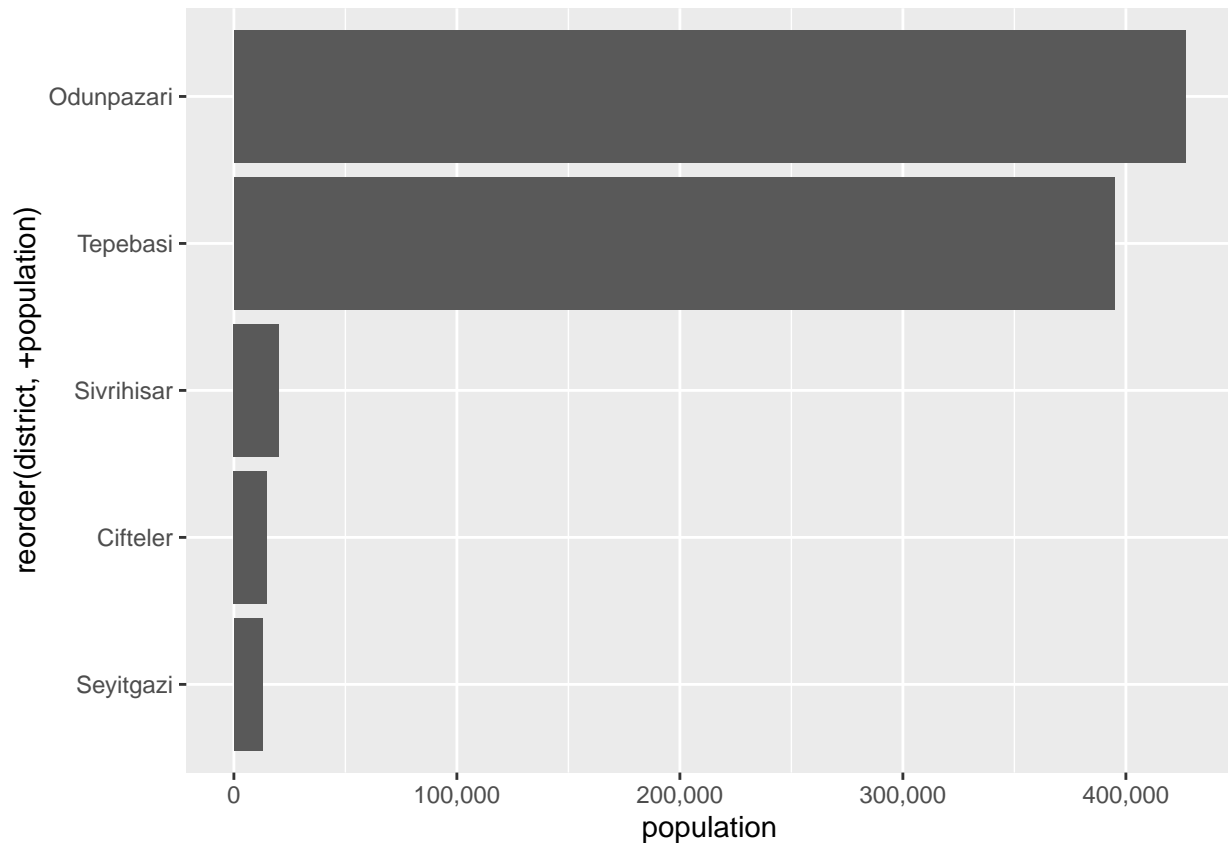
Drawing a barplot

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

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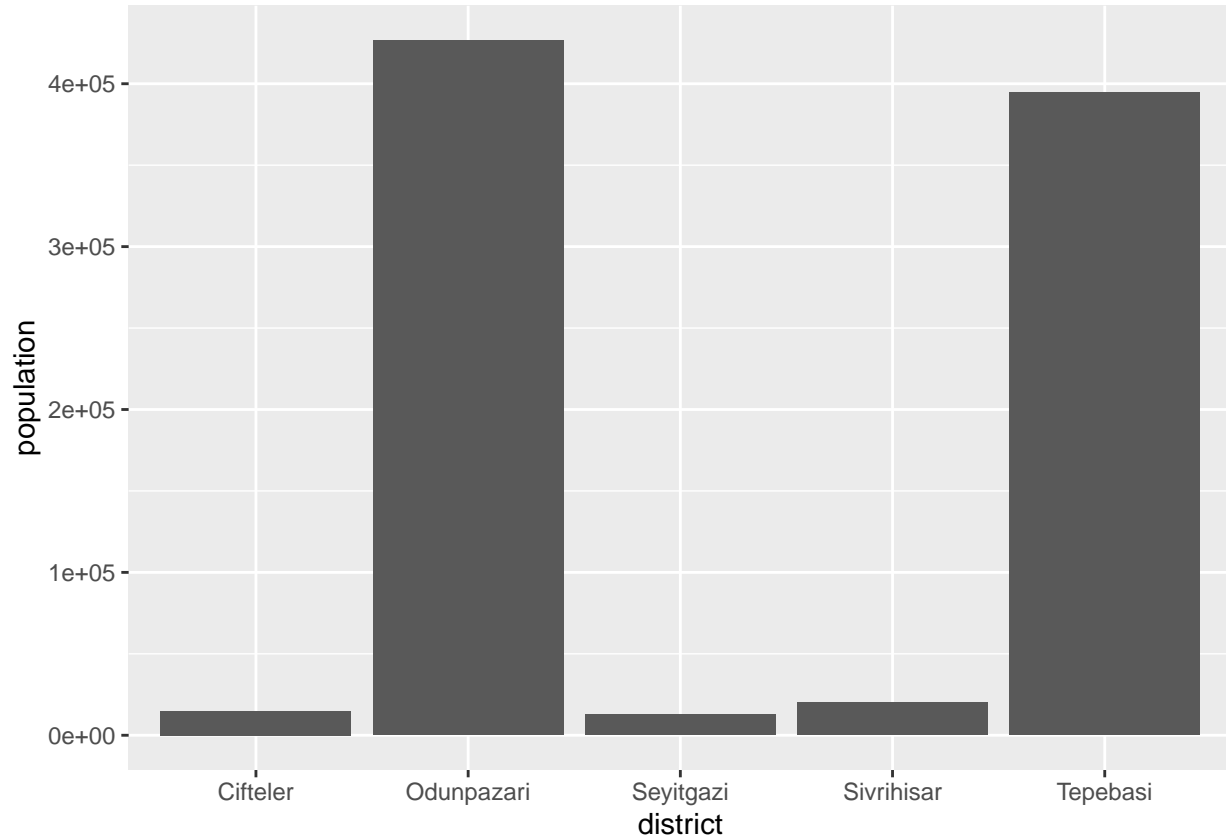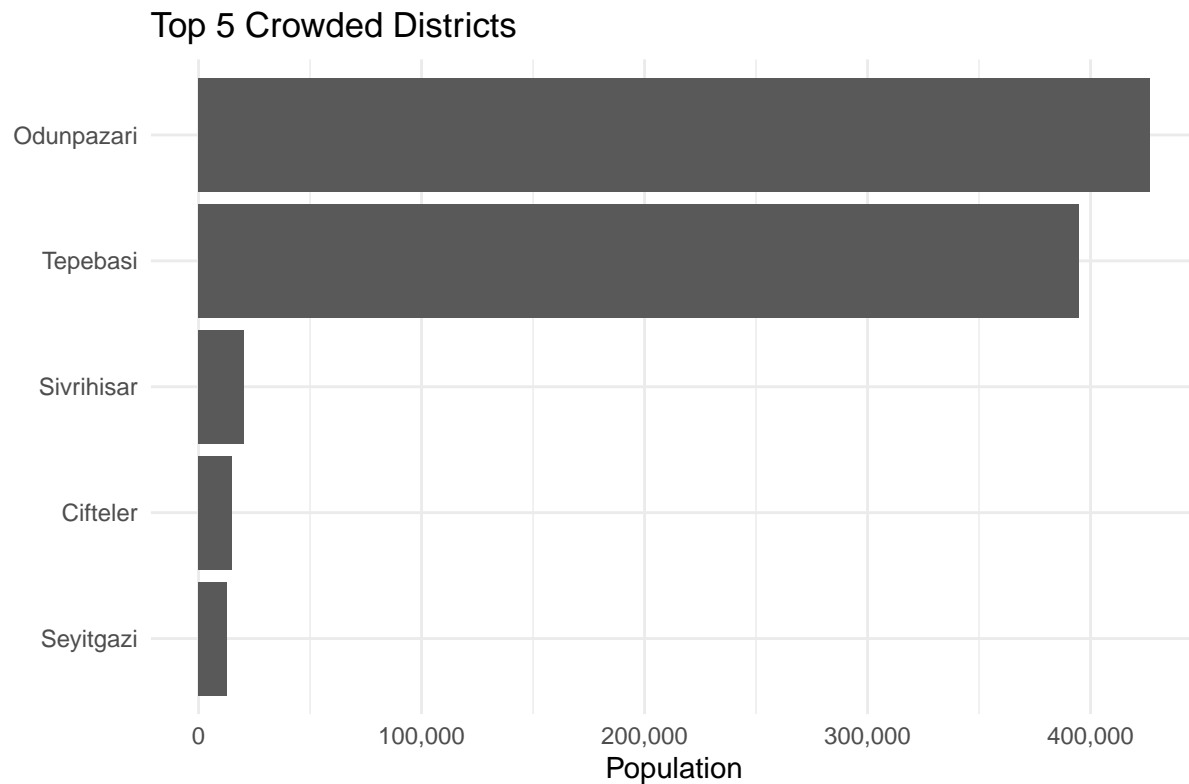

2. Please solve the problems in the plot to make it better (20 pts).

```
ggplot(eskisehir, aes(x = population, y = reorder(district, +population))) +
  geom_bar(stat = "identity") +
  scale_x_continuous(labels = scales::comma)
```

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 = population, y = reorder(district, +population))) +
  geom_bar(stat = "identity") +
  scale_x_continuous(labels = scales::comma) +
  labs(x       = "Population",
       y       = "",
       title   = "Top 5 Crowded Districts",
       caption = "The dataset contains the total population for the top-5 crowded
districts of Eskisehir in 2024.") +
  theme_minimal()
```

Top 5 Crowded Districts

Population distribution chart showing horizontal bars for Odunpazari, Tepebasi, Sivrihisar, Cifteler, and Seyitgazi districts, with Population on the x-axis ranging from 0 to 400,000.

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

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

Interpretation

The most crowded city in Eskisehir is Odunpazari. In Eskisehir most of population spreaded to Odunpazari and Tepebasi.