

# Homework #6

İsmail SEZEN

*Due date Jan 12, 2022 00:00 EET*

**Note:** Use `pm10.csv` file for this homework.

**Q1:(20pts)** By using all data in `pm10.csv` file, plot a histogram of all  $PM_{10}$  data and mark median on the histogram by a vertical line.

**Q2:(20pts)** By using the histogram in Q1, decide a maximum  $PM_{10}$  threshold and set NA the values higher than the threshold. Calculate correlation between stations and plot boxplots of all stations in a single plot. Discuss results in terms of variance and correlation. Which ones are similar?

**Q3:(20pts)** Establish a linear model between `sta3` and `sta4`. Plot scatter between them and draw the regression line.

**Q4:(20pts)** Calculate monthly and yearly means for all stations. Plot boxplots using for each month and year using hourly data. Discuss your results.

**Q5:(20pts)** Calculate means for each hour for all stations. plot results for all stations in a single plot and discuss your results.

---

## Notes:

- Please, email your answers privately to `sezenismail at gmail com`.
- Answers/codes should be in a text file named `name_surname.r`.
- Do not use `for-loop` statements.
- Write your conclusions as *R comments* in the text file.
- If you didn't understand a question clearly, email your question to `ysb801e@googlegroups.com`.