

Tutorial and Laboratories

Week 7

The purpose of this week's tutorial and lab exercises is to practice on the following topics:

- Tidy data using the grammar functions from '*tidyr*' package.
- Extracting the statistical summaries of different variables.

Exercises from previous week

Firstly, we will do practice exercises 2, 3, and 4 from last week. Then, we move to this week's exercises.

Tidy Data

There are three interrelated conditions to put a dataset into a tidy format:

1. Each variable must have its own column.
2. Each observation must have its own row.
3. Each value must have its own cell.

In the following exercises, we will use the "activities" dataset from this [link](#), please download it into your working directory. The data will be shown as following:

Id	trt	work.T1	play.T1	talk.T1	work.T2	play.T2	talk.T2
x1	cnt	0.652	0.865	0.536	0.275	0.354	0.0319
x2	cnt	0.568	0.615	0.0931	0.229	0.936	0.114
x3	tr	0.114	0.775	0.17	0.0144	0.246	0.469
x4	tr	0.596	0.356	0.9	0.729	0.473	0.397
x5	tr	0.358	0.406	0.423	0.25	0.192	0.834
x6	cnt	0.429	0.707	0.748	0.161	0.583	0.761
x7	tr	0.0519	0.838	0.823	0.017	0.459	0.573
x8	tr	0.264	0.24	0.955	0.486	0.467	0.448
x9	cnt	0.399	0.771	0.685	0.103	0.4	0.0838
x10	cnt	0.836	0.356	0.501	0.802	0.505	0.219

This data represents a sample of experimental data of how much time people spend on their phones, measured while doing three activities (work, play, or talk with someone), at two times. Each person has been randomly assigned to either treatment or control.

Exercise 1:

- 1- Import the data into a tibble or a data frame.
- 2- Discuss with your tutor, what are the problems in this data?
- 3- What would you do to fix these problems?

Exercise 2:

Use the data of the previous exercise and after you put them into a tidy format:

- 1- Extract the summary (count, mean, median, minimum, and maximum) of the following activities:
 - a. Play,
 - b. Talk, and
 - c. Work
-