Assignment 6: Testing Covid Hypothesis

Objective

To get understanding and hands-on practice in formulating and testing hypotheses by means of statistical instruments.

To get skills in solving Data Science tasks in multifeatured environment:

- working with real data set,
- pre-processing it in MS Excel,
- analyzing it in a programming language, like `Python`, and
- sharing the results in `Peergrade`.

Tasks

Test the following null hypothesis: `Covid` *infection spreads randomly and similarly in big Danish cities such as Copenhagen and Aarhus.*

Recommendations:

1. Download the file `municipality_cases.xslx` from our `Github` repo.
2. Copy the columns Copenhagen and Aarhus or any other two or more columns in another tab.
3. Create diagrams to compare the values visually.
4. Process the data in both `Excel` and `Python` or another programming language.
   a. Compute t-value for both data sets.
   b. Select p-value and degrees of freedom for the test.
   c. Make t-test using the t-table.
5. Share the result with us on Peergrade.

For t-test, use thee instructions below.

t-Test	Calculation of t-value
For each set compute □ mean $\overline{X}$ □ std deviation - S □ variance V = S^2 □ t-value	$t-value = \dfrac{signal}{noise}$ signal = $\|\overline{Xc} - \overline{Xa}\|$ noise = $\sqrt{\dfrac{Vc}{nc} + \dfrac{Va}{na}}$ Vc = Sc^2 Va = Sa^2