Exercises: Part 2

Please create a dataframe with the data below.

Land	**increase of the index (x)**	**unemployment (y)**
Belgium	2.8	9.4
Denmark	1.2	10.4
France	2.1	10.8
GB	1.6	10.5
Ireland	1.5	18.4
Italy	4.6	11.1
Luxembourg	3.6	2.6
Holland	2.1	8.8
Portugal	6.5	5.0
Spain	4.6	21.5
USA	3.0	6.7
Japan	1.3	2.5
Deutschland	4.2	5.6

Please save the dataframe into the file with following type: txt, csv, and xls (or xlsx). Read the txt, csv, and xls (or xlsx) files that you just created.
Next, calculate or create the following commands.

1. max, min (and corresponding country)
2. range $range = \max x - \min x$
3. quantiles

$$\tilde{x}_p = \begin{cases} x_{([np]+1)}, & np \notin Z \\ (x_{(np)} + x_{(np+1)})/2, & np \in Z \end{cases}$$

4. median $\tilde{x}_{0.5}$
5. quartiles difference $\tilde{x}_{0.75} - \tilde{x}_{0.25}$
6. mean $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
7. Median absolute deviation ($MAD =$ median of $|x_i - \tilde{x}_{0.5}|,\ i = 1,\dots,n$)
8. variance $\tilde{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.
9. standard deviation $\tilde{s}$
10. plot
    - box-plot

- histogram
- kernel smoothed density estimator
- empirical distribution function

11. covariance $\tilde{s}_{xy} = \frac{1}{n-1}\sum_{i=1}^{n} x_i y_i - \frac{n}{n-1}\bar{x}\bar{y}$
12. correlation $r_{xy} = \frac{\tilde{s}_{xy}}{\sqrt{\tilde{s}_x^2 \tilde{s}_y^2}}$
13. ranks of `index` and `unemp`
14. rank correlation coefficient

$$R_{XY} = \frac{\sum_{i=1}^{n}(R_i^X - \overline{R}^X)(R_i^Y - \overline{R}^Y)}{\sqrt{\sum_{i=1}^{n}(R_i^X - \overline{R}^X)^2 \sum_{i=1}^{n}(R_i^X - \overline{R}^Y)^2}}$$

15. calculate the confidence intervals ($\alpha = 0.05$), assuming that variances are known and unknown

$$KI = \left[\bar{X} - \frac{z_{1-\alpha/2}\tilde{s}}{\sqrt{n}}; \bar{X} + \frac{z_{1-\alpha/2}\tilde{s}}{\sqrt{n}}\right]$$

16. run the one sample test for the mean (f.e. index)

$$H_0 : \mu \leq \mu_0 \text{ vs } H_1 : \mu > \mu_0$$

$$\begin{aligned} T &= \frac{\bar{X} - \mu_0}{s}\sqrt{n} \sim t_{n-1} \\ KI &= \left[\bar{X} - \frac{1}{\sqrt{n}} t_{1-\alpha,n-1}\tilde{s}; \infty\right) \end{aligned}$$

17. run the two sample test for means ($\sigma_1 = \sigma_2$)

$$H_0 : \mu_1 = \mu_2 \text{ vs } H_1 : \mu_1 \neq \mu_2$$

$$\begin{aligned} T &= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{n_1+n_2}{n_1 n_2}\frac{(n_1-1)\tilde{s}_1^2 + (n_2-1)\tilde{s}_2^2}{n_1+n_2+2}}} \sim t_{n_1+n_2-2} \\ AB &: -t_{n_1+n_2-2,1-\alpha/2} \leq T \leq t_{n_1+n_2-2,1-\alpha/2} \end{aligned}$$

n1+n2-2

18. run the two sample test for variances

$$H_0 : \sigma_1 = \sigma_2 \text{ vs } H_1 : \sigma_1 \neq \sigma_2$$

$$T = \frac{\tilde{s}_1^2}{\tilde{s}_2^2} \sim F_{n_1-1,n_2-1}$$