

The *Industries.csv* file contains information on the turnover of industrial activities at the locality level. The information is as follows:

Siruta - Siruta Code of the locality;

City - name of the city;

Food, Textile, Wood, Chemical, Pharmaceutical, Metallurgical, Construction, Machinery, Computers, Electronics, Furniture, Energy, Industrial activities with turnover.

The *PopulationLocalities.csv* file contains the population by localities and the county codes for each locality.

1. Save in the *Request\_1.csv* file the turnover per inhabitant for each activity, at the locality level. For each locality, the Siruta code, the name of the locality, and the turnover per inhabitant for each activity will be saved. **(2 points)**

2. Calculate and save in the *Request\_2.csv* file the dominant industrial activity (with the highest turnover) at the county level. For each county, the county code, the dominant activity, and the corresponding turnover will be displayed. **(1 point)**

3. *DataSet\_34.csv* file contains records for 27 European countries related to the consumption and production of various types of meat. The observed variables are as follows:

1. Pork meat production (1,000 tons);
2. Beef meat production (1,000 tons);
3. Production of sheep and goat meat (1,000 tons);
4. Poultry production (1,000 tons);
5. Pork meat consumption (kg/person/year);
6. Beef meat consumption (kg/person/year);
7. Consumption of sheep and goat meat consumption (kg/person/year);
8. Poultry consumption (kg/person/year).

To perform canonical correlation analysis, standardise the value of the variables and split the initial data set into two data subsets analysis as follows:

Pork meat production, Beef meat production, Sheep and goat meat production, Poultry production – *set X*;

Pork meat consumption, Beef consumption, Sheep and goat meat consumption, Poultry consumption – *set Y*.

Save the 2 standardised data sets in files *Xstd.csv* and *Ystd.csv*. **(1 point)**

4. Determine and save the canonical scores  $z$  and  $u$ , corresponding to the  $X$  and  $Y$  data sets, *Xscore.csv* and *Yscores.csv*, respectively **(1 point)**

5. Determine and save the factor loadings corresponding to variables from the  $X$  and  $Y$  data sets in the files *Rxz.csv* and *Ryu.csv*, respectively **(1 point)**

6. Plot the distribution of observations in the space of  $z_1, z_2$  and  $u_1, u_2$  canonical roots - biplot graphic. **(3 points)**