Minimum requirements (8 points):

The project must be a Data Analytics project, i.e., must include content of the Data Analytics module.

The project must integrate several topics presented in the module, at a minimum these are the following:

(1) Data collection using both, Web Scraping and a Web API.

(2) Data preparation (e.g. remove missing values and duplicates, create new variables, enrich the data with open data).

(3) Data storage in a database like sqlite or MySQL.

(4) Non-graphical and graphical exploratory data analysis (EDA).

(5) Use of either regression or classification as the modelling method.

(6) Model evaluation using suitable measures of fit.

(7) Correct interpretation of model results and measures of fit.

(8) Making the material (data, Jupyter notebooks, ...) available on Moodle.

Additional points, if the following criteria are met (max 5 points):

(1) Creativity of implementation (creative is anything not specified in the lessons and exercises).

(2) Use of a MySQL database for data storage and SQL-queries from within Python.

(3) Integration and visualization of geographical data.

(4) Use of a statistical test for the analysis of contingency tables or analysis of variance (ANOVA).

(5) Use of k-means clustering in addition to the regression or classification model.

Project presentation in the form of a video-recording (max. 3 points).

To be evaluated:

(1) Quality of the presentation

(2) Structure of the presentation

(3) Time keeping (length of presentation (min.) = number of students per group x 5 min.)

To be submitted on Moodle with the group number by each student:

1.) Video recording of the project presentation as mp4 file, e.g.: 'video\_recording\_group\_01.mp4'.

(Ensure that the video does not exceed the max. allowed file size on Moodle).

2.) Zip-File with all data and your Jupyter notebook(s), as, e.g.: 'materials\_group\_01.zip'

3.) Presentation as PDF file (as e.g. 'presentation\_group\_01.pdf') with the following minimal structure:

- Title and full name of each student

- Introduction (background, problem, objective, research question)

- Materials and methods (data collection method, content of data, exploratory data analysis methods, modelling methods)

- Results & Discussion

- Conclusions

- IMPORTANT! Please clearly state on a seperate page what you have done to get the points and where to find it in the presentation.