Data Visualization Concepts

Prof. Dr. Renato Pajarola

Exercise and Homework Completion Requirements

1. Exercises and reading assignments are **mandatory.** They must be completed successfully to finish the class and get a sufficient passing final grade.

2. Exercises are graded coarsely into categories **pass** or **fail***.*

   - *A **fail** is given to failed submissions and incomplete solutions, and no points are awarded.*

   - *A **pass** indicates that the exercise is sufficiently good to receive the corresponding points.*

   - *Late submissions (up to one day) will result in a reduction of one point. Submissions after more than one day will not be accepted or graded.*

3. The four exercises give ascending points in the following distribution: 2 – 3 – 5 – 5.

   - *A **minimum of 7 points** from all four exercises must be achieved to pass the module. Failure to achieve this minimum will result in a failing grade for the entire module.*

   - *Thus at least two exercises have to be correctly solved, and one has to be from the more advanced ones.*

4. We give **bonus points** to students who have completed more than 8 points from all the exercises.

   - *Thus **7 points** from the exercises are required, **8 points** are still a normal pass, and **9 and above** would give 1 or more extra bonus points.*

   - *Only the bonus points can and will be added directly to the final grade.*

5. Do not copy assignments, tools to detect copying and plagiarism will be used.

   - *The exercise results are an integral part of the final course grade, therefore the handed-in solutions to the exercises **must be your personal work**.*

Submission Rules

- *The deadline for submitting Exercise 4 is Sunday, 28 May 2023 at 23:59h.*

- Please submit your solution code via OLAT with the file name **'dvc_ex4_MATRIKELNUMBER.py'**.

- If additional packages are used other than the ones in the 'environment.yml' file, please specify them in a **'readme.txt'** file.

- The code should run without errors and generate the bokeh app that contains the expected plots, the interaction, and the animation.

Exercise 4

In this exercise, you will build an interactive map to show the geographic distribution of tech companies in the US. Each circle in the map represents the total market cap and number of employees of the tech companies in a city, which are encoded in the color and size respectively. The map has the following functions:

1. The user can tap on a circle to show in the subplot the market cap and number of employees of each company in that city.

2. The user can use the slider to change the lower bound of the market cap to filter out the companies with a market cap smaller than this value.

3. The user can click the play button to see the animation of the changes in the market cap and the number of employees over the years.

Please read the comments and references in the code skeleton for more details and instructions.

Task 1: Data Processing

1.1   Define a function to create the data frames for the main plot and subplot.

Task 2: Visualization

2.1   Define a function to draw the main plot.

2.2   Define a function to draw the subplot.

Task 3: Interaction

3.1   Define a callback function for the tap tool in the main plot.

3.2   Add a slider and define a callback function for it to filter companies by the market cap.

Task 4: Animation

Add a play button to show the animation of changes over the years in the market cap and the number of employees on the main plot and the subplot.

4.1   Define a function to update the elements that change along with the year.

4.2   Define a function to wrap the update function in a periodic callback.

Run your app with the bokeh server. An example app looks like this demo (not necessarily the same):
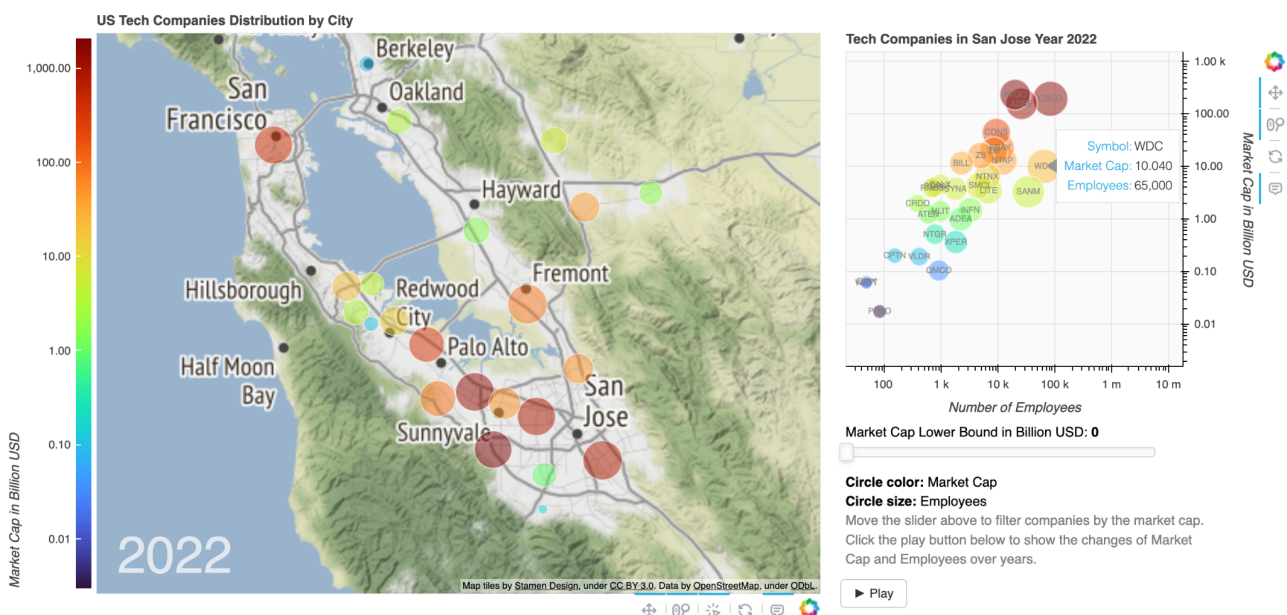
Remarks:

- The code skeleton is structured into sections corresponding to the tasks. You are free to change the skeleton and rewrite the code in your own way. The comments and references in the skeleton code are important for understanding and completing the tasks.

- The part of plotting the map is not required in the tasks. If interested in learning more about it, you are recommended to go through the contents in [Bokeh Tutorial 09. Geographic Plots](#).

- You can first try to implement the data processing and visualization in a notebook, then try some simple examples of the callback functions in sketch scripts, and in the end integrate all the parts together. Please note that the final submission of your code should be a .py file.

- Try Google first for errors. Chances are good that someone else has solved the problem. If Google cannot help, please use the OLAT forum to post technical questions regarding the exercise.

- The Q&A session for Exercise 4 will be on Monday, 22 May 2023, from 17:00 to 18:00. The place will be announced in time.