Data Visualization Concepts

BINF4234

Dr. Alexandra Diehl

Exercise and Homework Completion Requirements

1. Exercises and reading assignments are **mandatory** and they must be completed successfully to finish the class and get a sufficient passing final grade.
2. Exercises are graded coarsely into categories **incomplete** or **complete**.
   - *An **incomplete** is given to failed submissions and partial solutions, but no bonus points are awarded.*
   - *A **complete** indicates that the exercise is sufficiently good to receive the bonus points.*
3. A **minimum of 5 points** from the exercises 1, 2, and 4 must be achieved to pass the module, whereas exercises 3 and 5 are optional and can offer you bonus points as described in Point 4.
   - *Failure to achieve this minimum will result in a failing grade for the entire module.*
   - *Hence at least the first two, or the fourth exercise has to be fully solved.*
4. The five exercises give rise to the following point distribution: 2 – 3 – 2 (**optional bonus points**) – 5 - 3 (**optional 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 and therefore the handed in attempts and solutions to the exercises **must be your personal work**.*

Submission Rules

- Submitted code must compile and run without errors using the indicated Python environment, using the included libraries, packages and frameworks. If additional libraries/packages are needed, please specify in your ‘readme.txt’ file.
- The whole project source code must be zipped and submitted before the given deadline, including the output results (saved in .html file or a screenshot picture).
- Submit your .zip archive named *dvc_ex2_MATRIKELNUMBER.zip* (e.g. dvc_ex2_01234567.zip) through the OLAT course page.

- **Deadline is Wednesday, 30 October 2019 at 23:59h**

Exercise 2

In this exercise we will get a closer understanding of the eBird dataset and visualize some interesting attributes by applying Python-Bokeh visualization techniques. The general purpose of this exercise is to apply the visualization concepts and visual encoding schemas learned from the lecture to real dataset, and finally to see if visualization can help people better understand and analyze the dataset. For example, how much time and which time period within one day is commonly taken for observers to follow and record the birds' behavior, is it correlated with the species of the birds?

The whole exercise contains five mini-tasks as described in details below:

Task1: Set up necessary plotting components (the main chart area and the selector) after reading the data into dataframe.

— Generate proper y axis labels, which is consistent through all the conditions.

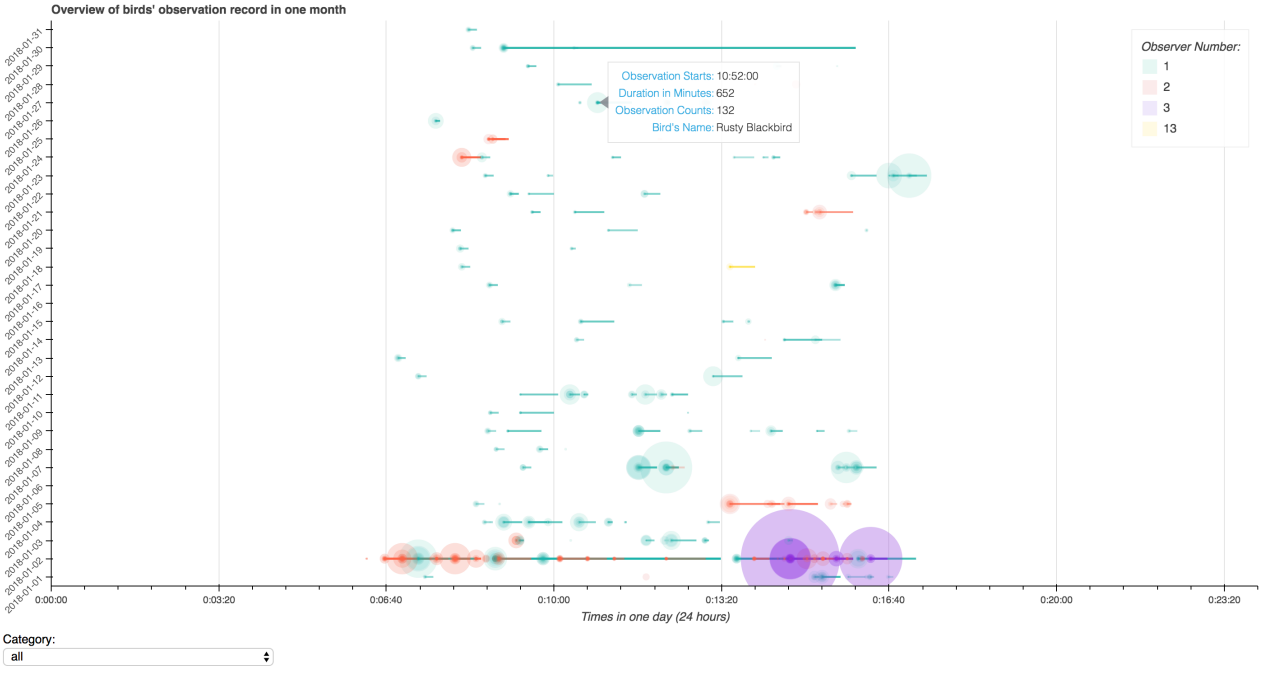
— Add an interactive selector for selecting birds based on the category (or name, locality, protocol, etc.)

Task2: Define the datasource construction function (which will be called when the datasource need to be updated and the output of this function will serve as the input of the plot function) by following the hints in the code skeleton.

Task3: Define the update source function accordingly (which will be called whenever the selector is triggered which means the data source need to be updated). Pay attention that the selector should always be able to select "all" the data points instead of only being able to select part of the data under certain category.

Task4: Define the plot function which will be called whenever the selector is triggered, as well as when running the server initially.

Task5: Draw the figure and add bokeh server for interactive visualization.



IMPORTANT: You will get **2 points** out of 3 if you can draw the plot (including the hover tool), and **3 points** after adding the selector successfully.

A **live demo** will be given in the lecture to demonstrate the final result, which should look **similar** as the picture shown below:

The deliverables of this exercise will be a clean version of your code with proper comments, any additional files necessary for executing it, a “readme.txt” file for your comments or remarks (if necessary), as well as an export of the final output result in .html or .jpg format. The absence of any required deliverable files will automatically lead to a **FAIL**.