Interactive Visualization

1. Setting

You are a Data Scientist working for the Port Authority of New York and New Jersey. The Port Authority is, among other things, in charge of operating the airports in the New York area. Your task is to create a prototype of an **interactive flight delays dashboard** for flights departing from these airports.

For your prototype you will use 6 months of **flights** data from the Bureau of Transportation Statistics (BTS) for the year 2024, along with the corresponding metadata on **airports** (e.g. John F Kennedy International Airport), **aircrafts** (AIRBUS A321), and **airlines** (American Airlines).

The Port Authority wants to inspire your creativity as a data scientist, which means that you are relatively free to make choices. However, there are also certain binding requirements:

- Put the main focus on flight delay patterns. Additional aspects (e.g. number of flights, cancelations) may be included as well, as long as the main focus remains on flight delays.
- The dashboard should contain at least one of each of the following elements: map, bar chart, line chart, a KPI (text element)
- The dashboard should allow for user interaction, such as filtering by airport or airports.
- Do not overwhelm the user with too much information, but display those aspects that you consider most relevant.

2. Tool choice

You can use one of the following three tools, depending on your preferences: Tableau, Power BI, or Python Streamlit.

2.1 Tableau

Tableau is a business analytics tool by Salesforce. Tableau allows you to process data and create interactive dashboards with a drag-and-drop interface. The use of Tableau Desktop Public Edition requires registration, but can be used for free. Note that there is also a dedicated website for students.

Recommended resources: 1. Tableau Training Videos or 2. Datacamp: Introduction to Tableau [1]

2.2 Streamlit

Streamlit is a Python library that allows you to create interactive web applications, based on Python scripts. Streamlit integrates well with other Python libraries such as Pandas, Plotly, and Matplotlib. And also, you can integrate it well into your normal Python Data Science workflow.

Recommended resources: 1. Streamlit Website, or 2. Streamlit Youtube Channel

2.3 Power BI

Power BI is a business analytics tool by Microsoft. Power BI allows you to process data and create interactive dashboards with a drag-and-drop interface. Note that Power BI Desktop can only be installed on Windows systems. Mac users can optionally use the free version of Power BI for the web, but it comes with limited functionality compared to the desktop version.

Recommended resources: 1. Datacamp: Introduction to Power BI or 2. Microsoft Power BI Learning Path

3. Deliverables

3.1 Dashboard

No matter which tool you use, submit a fully self-contained reproducible dashboard:

- **Tableau**: submit a **twbx file**
- **Power BI**: submit a **pbix file**
- **Streamlit**: Submit a **zip folder** containing the following files: **app.py, requirements.txt, the data files** (placed in a data directory), (optionally: a link to your Streamlit app hosted on Streamlit Community Cloud)

3.2 Documentation (2-page pdf file)

Page 1:

- Include a screenshot of your dashboard.
- Explain your most important choices, e.g. related to design, interactivity, data processing, etc.

[1] Access to Datacamp resources is available for free for FH Kiel Data Science students, see here. Notify me if I need to renew the Datacamp invite link.

Page 2:

- What was your biggest challenge, and how did you deal with it?
- What do you consider the best aspect of the visualization tool that you used?
- What do you consider its most annoying aspect?
- What would you do differently, if you had more time?

4. Grading

The following criteria are considered for the grading

- Reproducibility
- Requirements of the Port Authority
- Creativity
- Design: Try - where possible - to follow the design principles discussed in class, such as "CRAP" (Contrast, Repetition, Alignment, Proximity). Note however that this is often more difficult in an interactive setting, and therefore won't be always possible.
- Honest, critical reflection

Note: you are not expected to produce a comprehensive and flawless product, but to grow through experimentation: try out features of your visualization tool, apply your creativity, and reflect on your learning process. You will get a very good score, if you show very good engagement and creativity, and critically reflect on your project.