Information Visualisation

Assignment 2

Group 8:

CHESNAY Suzanne (00800188), JACOB Emily (12143768), ARYA Kartik (12402369)

Github Repository Link: https://github.com/suzannechesnay/ArtVis

Dashboard for 'ArtVis' Dataset

Assignment 1 Overview and Updates:

As part of the earlier assignment, we included a broad range of **target users** for our interactive visualization and decided to slightly narrow it down to visitors of an exhibition. Since the dataset focuses specifically on the period 1905-1915 and we focussed on the European continent (including western Russia), such an interactive display panel could be found among other panels at an exhibition about the European art scene at the beginning of the 20th century, as one can find in museums such as the Leopold Museum with its permanent exhibition *Aufbruch in die Moderne.*

Choice of Visualization - An interactive overview panel has been implemented, combining statistical charts and a map overview. This approach aligns with the discussions in earlier phases of the assignment. An overview is discussed below, as well as a description of the implementation and the difficulties encountered while building the dashboard.

Preprocessing Summary for Visualization Dashboard

The preprocessing of the data resulting in the final dataset `artVis_data_cleaned.csv` is found in the Jupyter Notebook `Data_cleaning_and_queries.ipynb` and can summarised as follows:

1. Geographic Filtering:
   Loaded and examined the dataset `artvis_dump_NEW.csv`.
   We focused exclusively on European countries based on predefined country codes.
   This part of the code can be commented to include the whole dataset.

2. Eliminating duplicates:
   We extracted unique artists and noticed, when we grouped the rows of the original file by (artist-ID, exhibition-ID) pairs, that there were multiple duplicates. While the artist entries had no duplicates (i.e., for one given artist ID there were unique values for the columns ‘firstname’, ‘lastname’, ‘gender’, ‘birthdate’, ‘deathdate’, ‘birthplace’, ‘deathplace’ and ‘nationality’), many exhibitions were listed multiple times with

different venues (sometimes listed under one name and listed again under ‘exact location unknown’).
We counted **1039 different exhibition IDs**, of which 340 came up more than once for a given title, start date and type. We therefore proceeded to eliminate superfluous entries: When multiple venues were listed, we favoured the names not containing substrings like '*(exact location unknown)*', '*Vereeniging*', '*Vereinigung*', '*Société*', '*Künstlerbund*', '*Verein*', '*Verband*', '*Association*' as names like these don’t usually refer to a spatial location but to groups of individuals, which were used as a replacement for unknown venues.

3. Output Preparation:
   Exported cleaned artist and exhibition data as separate CSV files for integration into the dashboard. We also saved in separate csv files respectively the list of all exhibitions, artists and venues with corresponding properties, and finally we defined python functions to query for selected time ranges.

Steps to run the Dashboard

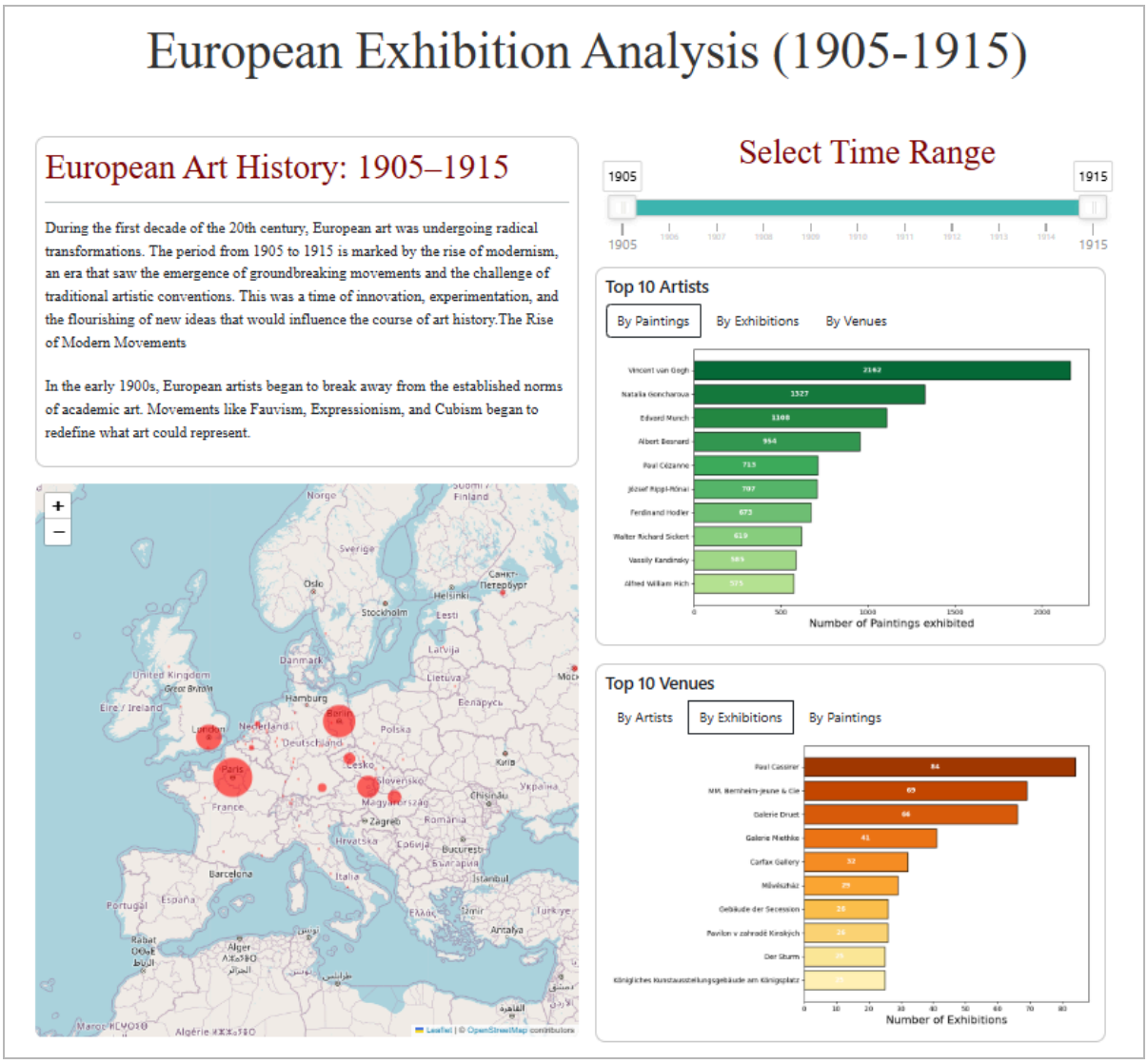
The ‘ArtViz’ package contains all the necessary files. After navigating to the current working directory, run the following using ‘Terminal’:

1. Install dependencies
   `pip install -r requirements.txt`
2. Start the Dashboard
   `open index.html`
3. Initialize the API
   `uvicorn plotting:app --reload`

Setup

- **OS:** Mac/Windows
- **Python Version:** 3.11.6

Dashboard Structure:



To get the desired information the user sees 4 interconnected sections:

The ‘input panel’ on the right lets the user plugin 3 parameters:

1. The time range
2. ‘Artists’ filter, based on ‘Paintings’, ‘Exhibitions’ and ‘Venues’
3. ‘Venues’ filter, based on ‘Artists’, ‘Exhibitions’ and ‘Paintings’

The map on the left outputs the cities in Europe in which exhibitions were hosted, shown by the coloured dots of varying size (depending on the number of exhibitions), allowing the

user to zoom in or out as they wish. Hovering the cursor on the city gives a count of exhibitions taking place there during the selected time range.

Description of each section:

Time Range Slider

A slider was used to dynamically change the years as time is a highly relevant variable in the context of the Artist Visualization dataset. This slider makes sure that the data is plotted on the temporal axis for all the 3 visualisations, thus making it easy to understand the Art scene in Europe from 1905-1915.

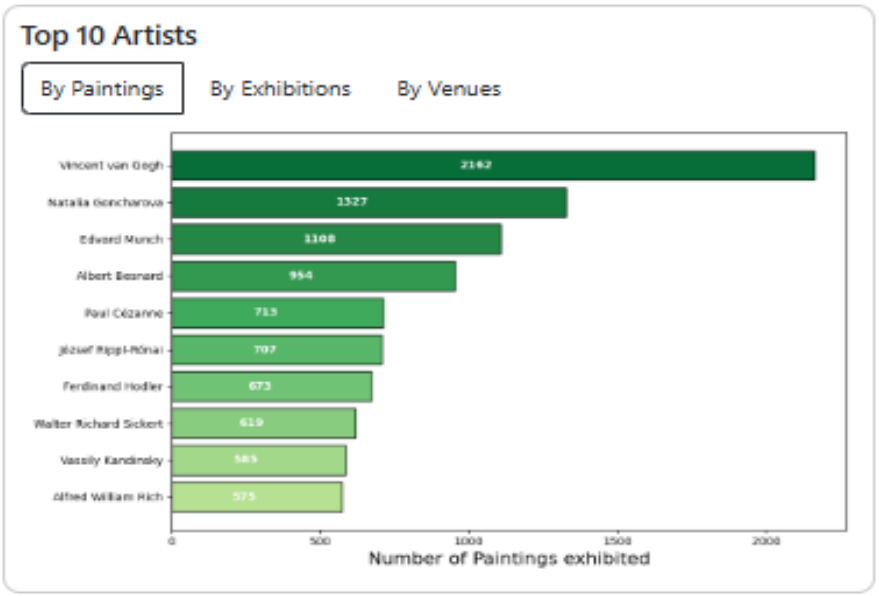
- Users can select the time range they wish to view. By default, it displays all years from 1905 to 1915.
- The slider can be adjusted from both ends. Once the time range is selected, subsequent visualisations automatically adjust to this range.

At first we tried implementing the charts section entirely in javascript, but for time reasons we settled for python charts, which are loaded into the html page every time the slider selection is updated by a user interaction.

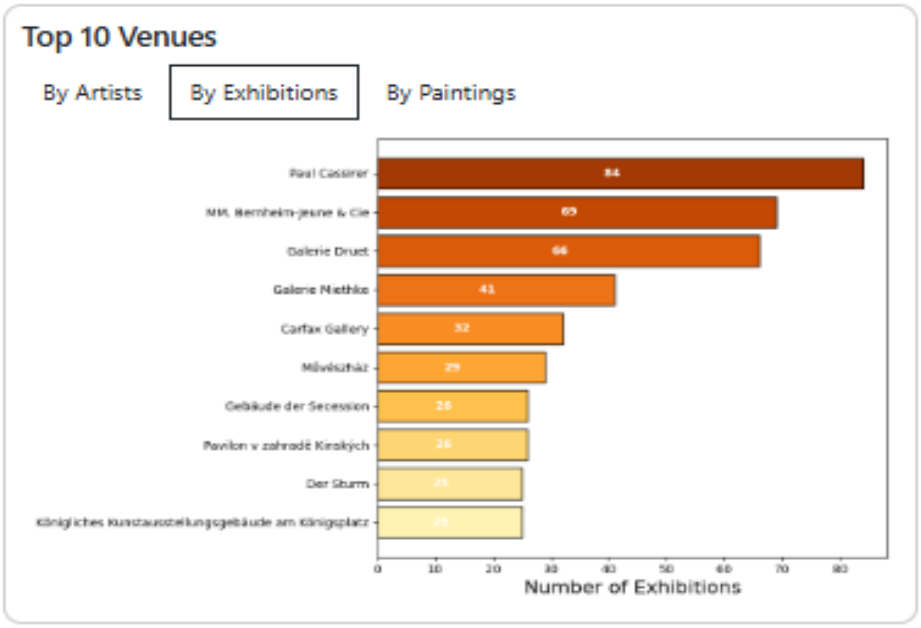
Filters

We used the number of exhibitions at which an artist was present, the number of paintings he/she exhibited there and the number of different venues where he/she was presenting his/her work during a given time range to quantize their presence in the art scene at the time. The popularity of the venues was measured similarly which is why We thought that bar charts would best represent this and enable comparisons. There are two filters on the sidebar to get the 'Top 10' stats of the data as follows:

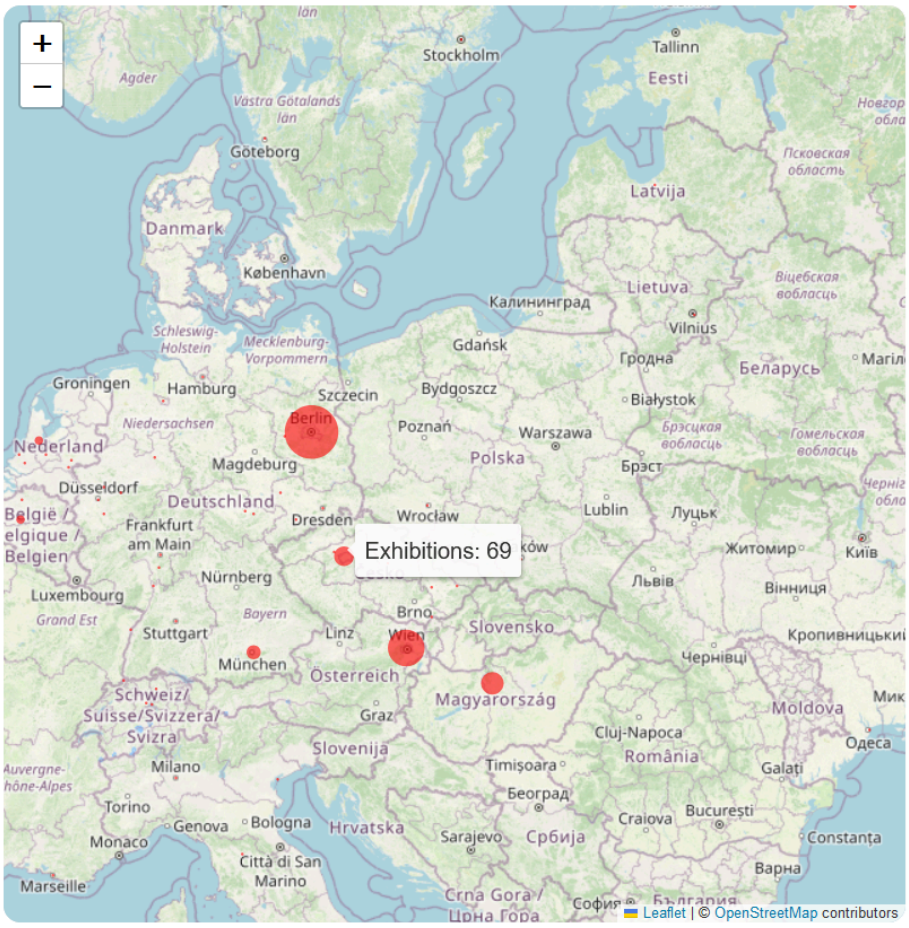
1. **Top 10 Artists:** This filter picks the top 10 Artists based on their no. of paintings, no. of exhibitions, and the venues where their art is displayed.



2. **Top 10 Venues:** Similar to the Artist filter, this filter gives the stats related to the Venues based on the no of artists, exhibitions and paintings



Map



- The world map takes the input values from the time slider and displays the exhibitions based on the cities.
- Users can hover over cities to see a pop-up summary of the number of exhibitions held in that location.

Together, these components can be very useful for exhibition viewers to get an overview of the activity and the geographical clusters of the art scene at the beginning of the 20th century.