Famous World: Milestone 2

Anmol Prasad, Michal Bien, Ravinithesh Annapureddy

Project goal

This project aims to provide a one-stop web page that will show the famous people placed on a world map and give foundational information about the person.

Achieved goals

1) **Exploratory Data Analysis**: the knowledge from Lecture 4 and last semester's ADA course was used to properly load, explore, and extract initial insights about the dataset.
2) **Data preprocessing:** The data processing pipeline was built using python, pandas, GitHub actions and Amazon S3. In the process, only required columns were preserved in the dataset, and the data was enriched by adding biography summaries from Wikipedia and pictures from wikidata. Whenever the data processing script is changed, the new version of the preprocessed dataset which "fuels" visualization is deployed to the object store: Here
3) **Initial map with markers on Github Pages:** the resources of the 6 first lectures were used to create the first D3JS visualizations. The data from the processing pipeline is read by the script, and the first map with loadable markers is presented to the user. The final deployment (and functional prototype) URL of our visualization will be: Here.

Future project milestones (MVP)

A sketch of the functional prototype of the website is provided in appendix 1.
1) **Navigation Bar** (appendix 2): The visualization's navigation bar serves the basic information and interaction source. Therefore, we limit the buttons on it to the home page, project information pane, and eventually a search bar.
2) **Responsive Map and Markers** for each person (appendix 3): Our core value is a fast, responsive and interactive visualization of famous people through time and space. Therefore, the design of the methods of interaction and visualization on top of the map is the key part of our project. It's challenging for two reasons: first, keeping the information clean in the context of big clusters of people born in the top European cities will require us to limit the number of visible results when the map is fully visible, and increase this number when the user zooms in. This is not only a design but also a performance issue. Loading images of thousands of people takes a considerable amount of time and has to be done smoothly. We are considering several methods, such as downsampling, and tiling, to decrease the bandwidth and computational complexity of rendering our map.
3) **Person Information pop-up** (appendix 4): This pop-up is supposed to be available for the user on-click when the person visible on the map is selected. First, the most important information will be presented next to the image. This information will be followed by the extract from English Wikipedia.

4) **Timeline Filter** (appendix 5): One of the ways to make our visualization more useful, not only as a data story but also as a tool is by adding various filtering methods. Our MVP method, which we find the most useful, is the timeline filter. We want the users to have an option to select the exact time period, which interests them. Then, only the people which were born in this period will be displayed in the visualization.

5) **Occupation filter:** The second MVP filter, which is expected to catch the people by the category of their occupation. We propose the occupation list based on the WikiData ontological assignments. We think that filtering the visualization by occupation, and the possibility of performing the joint timeline and occupation filters will give our visualization flexibility of use.

Future work (nice to have)

1) **Country based filter**: In our MVP we decided to use only the spatial location of the people to create a map. Seeing the individuals filtered by the country of their birth could be useful, however has several limitations. The location of countries is not strict, they change in time, and there are many disputed areas, which are hard to map properly without starting a political discourse. Moreover, many countries didn't appear until recently or changed their names throughout history. This makes it difficult to map the famous person from the early medieval to the map of modern Europe in a meaningful way.

2) **Tile-based map** (mosaic of faces on the land) (appendix 6): this was our original idea, currently suspended and being reworked due to computational complexity. The faces would reside in an area where the person was born, but not exactly in the place. All the faces would be of the same minimum size, to make the visualization understandable, while the pictures could be bigger in the areas where there were not many famous people. On the other hand, in highly populated areas, only the subset of the most popular individuals (sorted by the decreasing popularity index) would be displayed at first. More tiles would appear in the zoomed areas.

3) **Search field:** - any good map service provides a full text search module. We would like to allow our user to search for the person by their first name, last name, and the biography summary from Wikipedia, and to return the most relevant search results on the dropdown list of the search input form. Then, by clicking on the search result, the user would be automatically moved to the relevant area of the map animatedly and the selected biographical information will be displayed.
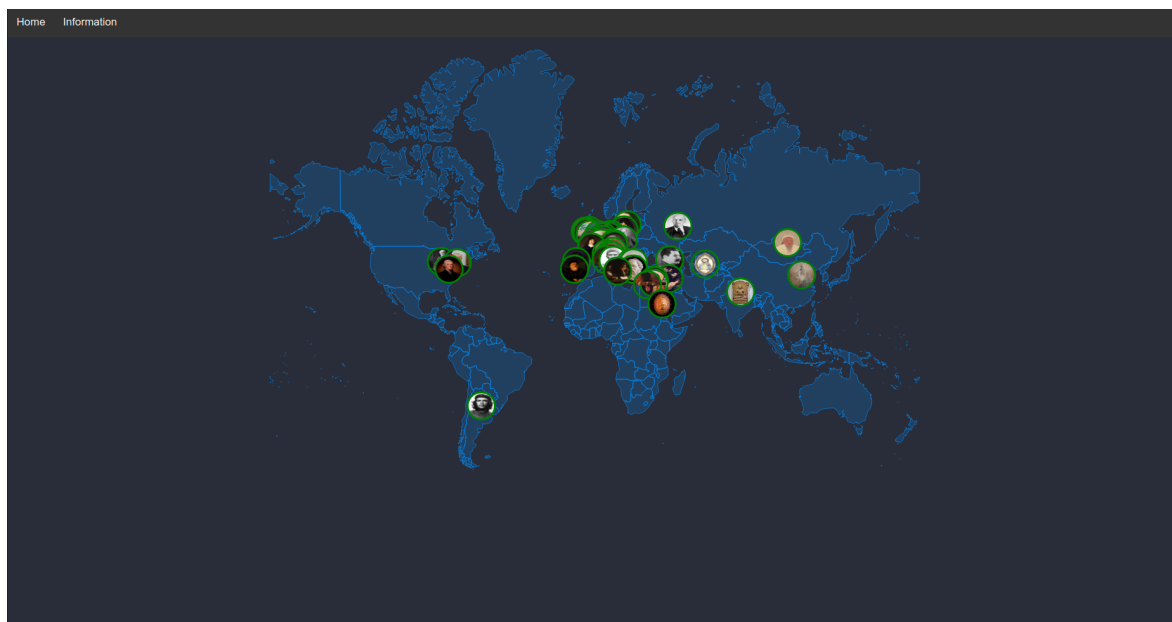
List of Tools

This project will require basic tools such as HTML, CSS (Lecture 1) to create the web pages. Then we shall require SVG (Lecture 1) and Javascript (Lecture 2, 3) to display vector graphics and handle events. We shall need the D3 library of Javascript (Lecture 4, 5) for dynamic and interactive data visualizations. As the base of our visualisation is a map we shall use the Leaflet library of Javascript and GeoJson (Lecture 8) to draw and create interactions on the map. Apart from the tools, we shall use the concepts related to Design (Lecture 7), Story Telling (Lecture 12) and Perception, colours, marks, channel (Lecture 6).
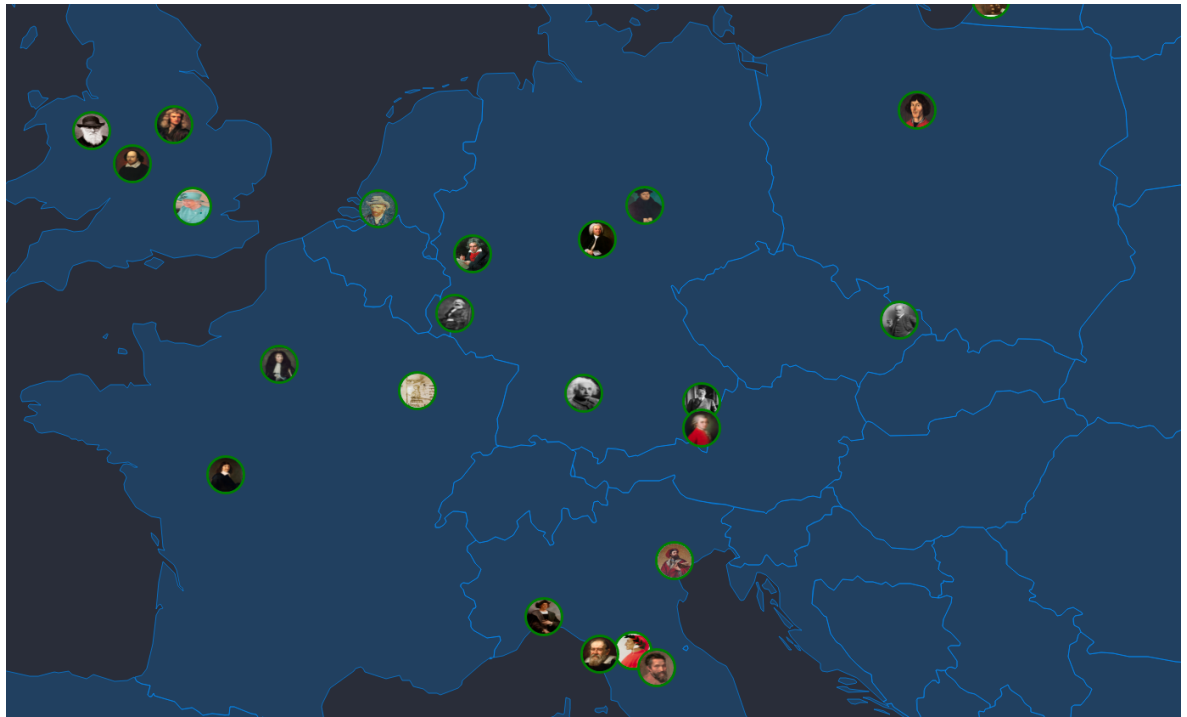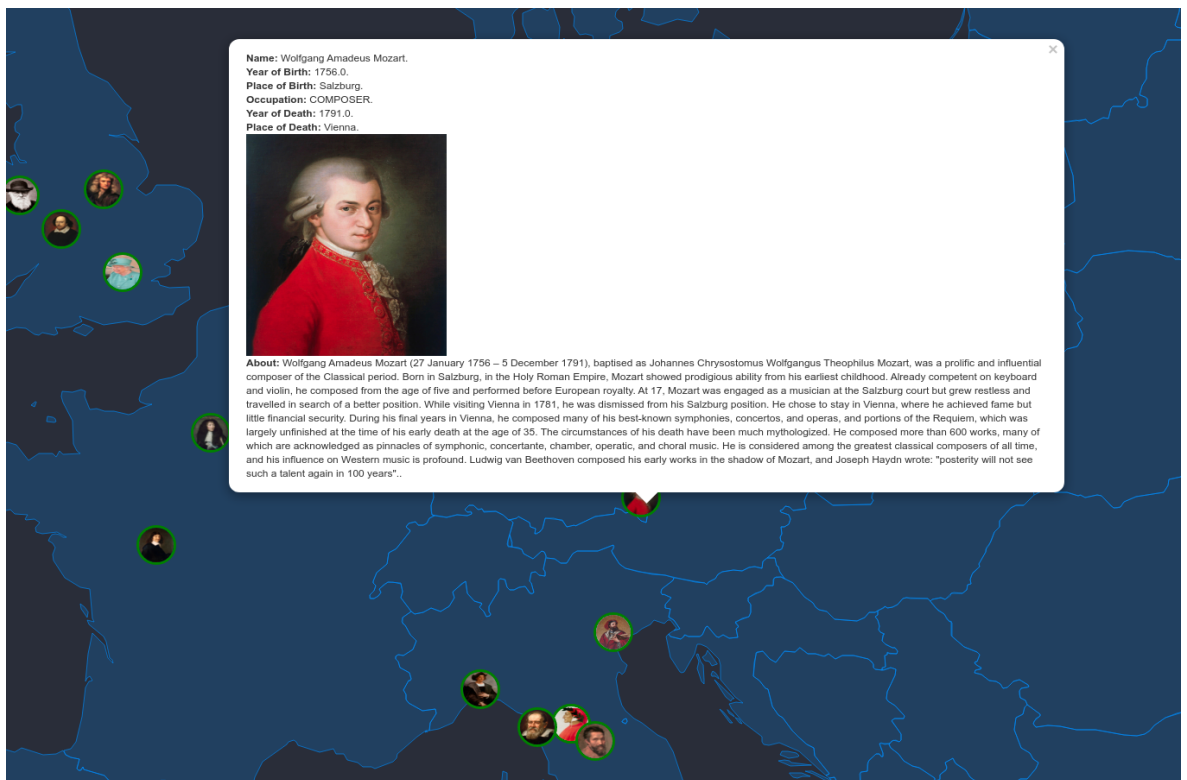
Appendix
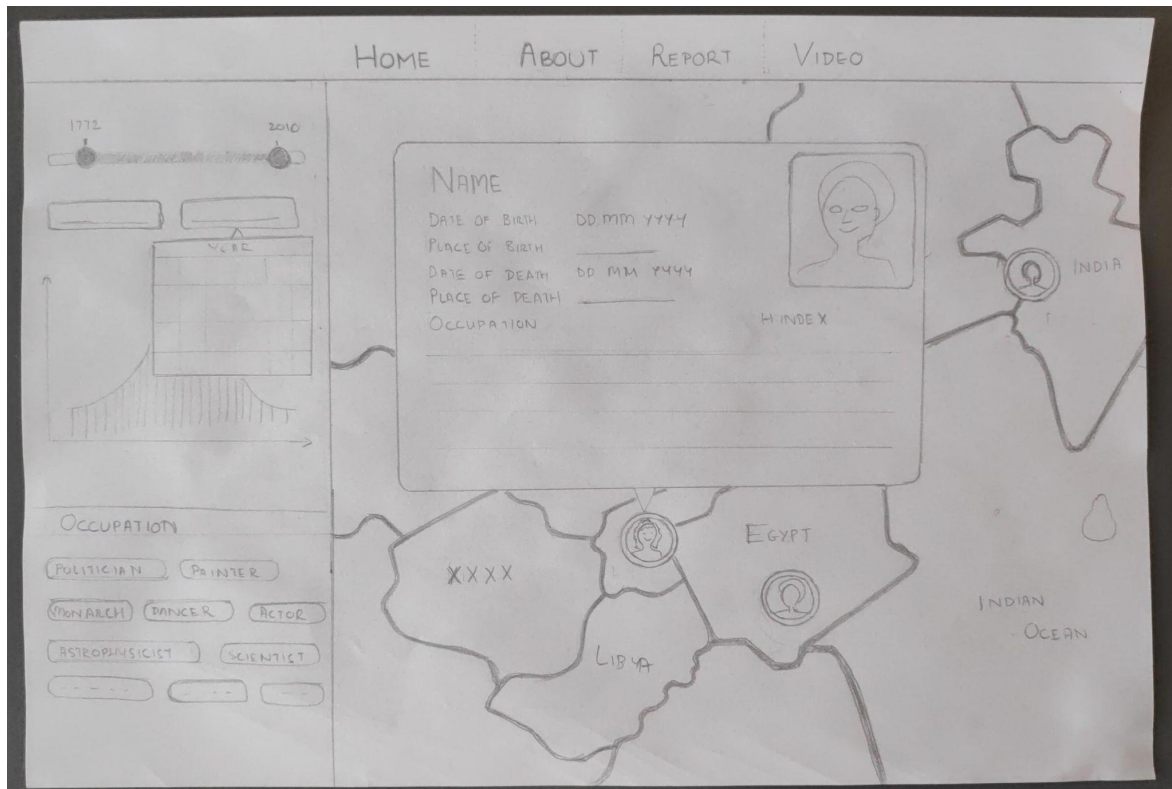
Appendix 1: Full map screen, with faces and navigation bar

Appendix 2: Full map screen, with faces and navigation bar
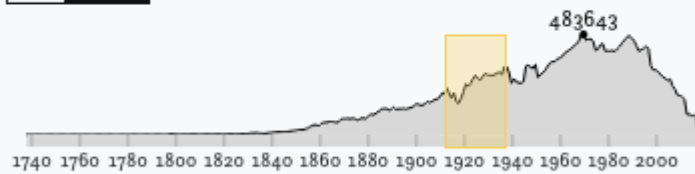
Appendix 3: Zoom to Europe

Name: Wolfgang Amadeus Mozart.
Year of Birth: 1756.0.
Place of Birth: Salzburg.
Occupation: COMPOSER.
Year of Death: 1791.0.
Place of Death: Vienna.

About: Wolfgang Amadeus Mozart (27 January 1756 – 5 December 1791), baptised as Johannes Chrysostomus Wolfgangus Theophilus Mozart, was a prolific and influential composer of the Classical period. Born in Salzburg, in the Holy Roman Empire, Mozart showed prodigious ability from his earliest childhood. Already competent on keyboard and violin, he composed from the age of five and performed before European royalty. At 17, Mozart was engaged as a musician at the Salzburg court but grew restless and travelled in search of a better position. While visiting Vienna in 1781, he was dismissed from his Salzburg position. He chose to stay in Vienna, where he achieved fame but little financial security. During his final years in Vienna, he composed many of his best-known symphonies, concertos, and operas, and portions of the Requiem, which was largely unfinished at the time of his early death at the age of 35. The circumstances of his death have been much mythologized. He composed more than 600 works, many of which are acknowledged as pinnacles of symphonic, concertante, chamber, operatic, and choral music. He is considered among the greatest classical composers of all time, and his influence on Western music is profound. Ludwig van Beethoven composed his early works in the shadow of Mozart, and Joseph Haydn wrote: "posterity will not see such a talent again in 100 years"..

Appendix 4: Person information pop-up

PUBLICATION DATE

Number of articles per year

% SUM (i)

483643

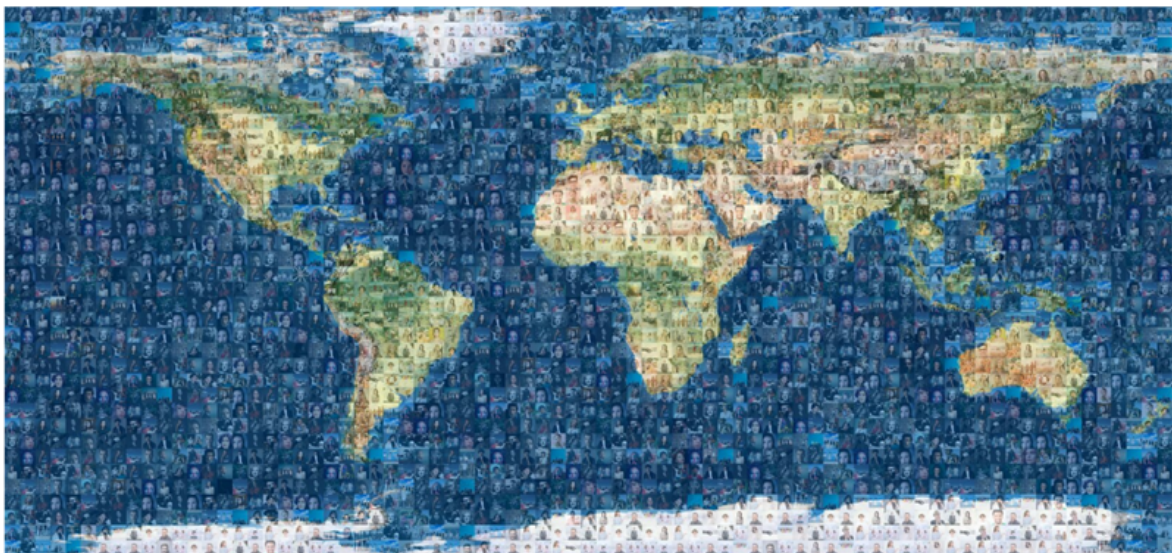1740 1760 1780 1800 1820 1840 1860 1880 1900 1920 1940 1960 1980 2000

ADD NEW DATE FILTER ...

Appendix 5: Example well-designed timeline filter from Impresso app

SEARCH

Appendix 6: Sketch of the tiled-mosaic map