WORLD'S MONUMENTS

Visual Analytics course 2020/21
MSc in Engineering in Computer Science
Sapienza University of Rome

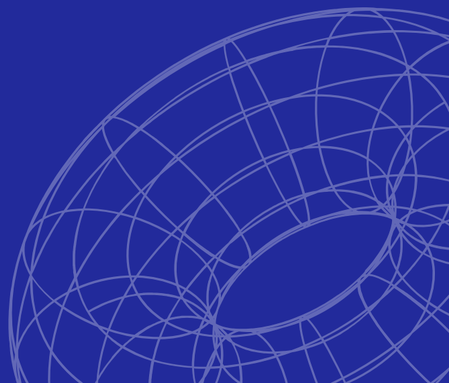
Samuele Olivieri Pennesi
Francesco Ottaviani

In this Presentation

Overview

General idea

We want to represent the **places** of interests all around the world through several **visualizations** which underline different characteristics, and to classify them using a **relevance** notion.



Goal of the project

We aim to let travellers discover new interesting places and also to **make people aware about the importance and the beauty of the cultural sites** all around the world.

In other words, we want to provide the **ultimate travel guide**, highlighting interesting aspects thanks to our visualizations.

Datasets used (1/2)

- **GeoNames**: the most complete source regarding the places of interests all around the world, and containing information such as **coordinates, name, category and country**.

- We have chosen the following subset of categories: **Church, Archaeological site, Historical site, Palace, Building, Museum, Castle, Monument, Amphitheater, Pyramid**.

Datasets used (2/2)

- **Wikipedia Pageviews**: it associates to each Wikipedia page the **number of visits** it receives in a given hour of a specific day. We will use this information as a **relevance** value (i.e. how much is important each site).
- We summed up the visits of 24 hours to obtain reliable data.
- We **removed all the pages with less than 50 visits per day** (in this way we considerably reduce the size of the dataset and we also cut off not relevant data).

Integration of data

We made an important **integration** process using **Talend** software, which can be summarized as follows:

- We **joined the data between Geonames and Pageviews on the name attribute.** The sites for which a correspondance has not been found, will have the relevance field empty.
- In order to make the data more user friendly, we provided a **mapping of the categories and country codes into their complete names**.

Final data structure

The output of the integration performed with Talend is a **TSV file** with **67282 rows** (one per site). Each entry has the following attributes: **name** (of the site), **longitude** (in decimal degrees), **latitude** (in decimal degrees), **country** (full name), **category** (full name), **relevance** (coming from Pageviews) and **country iso** (ISO-3166 2-letter country code).

Related works (1/2)

Papers with similar objective as ours:

- An Aspect of Archaeology's Recent Past and Its Relevance in the New Millennium
  https://link.springer.com/chapter/10.1007/978-0-387-72611-3_2

- UNESCO World Heritage sites and tourism attractiveness: The case of Italian provinces
  https://www.sciencedirect.com/science/article/abs/pii/S0264837718318155

Related works (2/2)

Papers with similar dataset as ours:

- YAGO: A Multilingual Knowledge Base from Wikipedia, Wordnet, and GeoNames
  https://link.springer.com/chapter/10.1007/978-3-319-46547-0_19

User input

In order to display the data in the best possible way, and to make the visualizations more interesting and dynamic, we **provide** to the user a **menu to filter the data**:

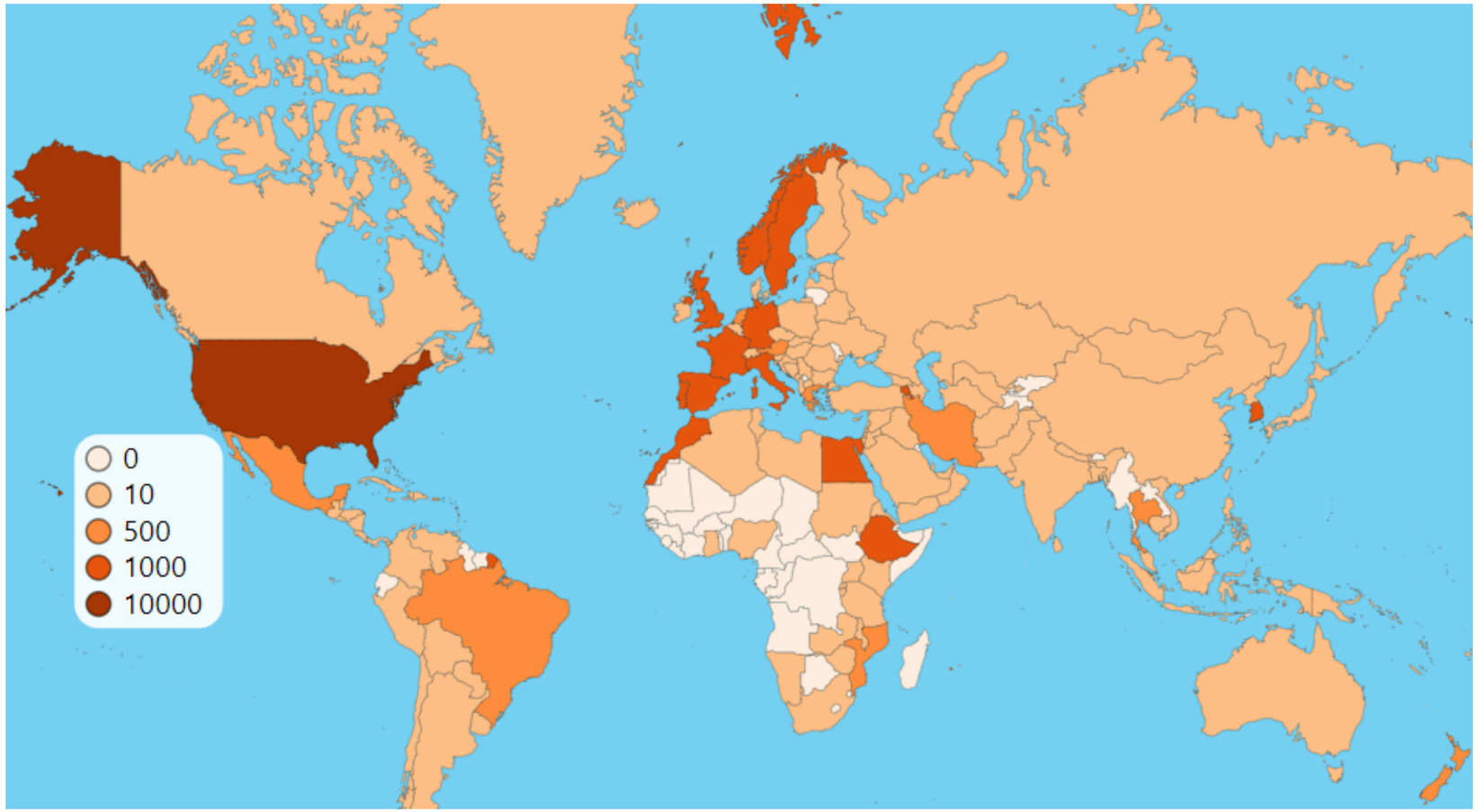
- **Country dropdown**: It is possible to select one country at a time, or the entire world.
- **Category dropdown checkboxes**: The user can select one or more categories.
- **Relevance slider**: Only the sites which respect the threshold relevance will be visualized.

Visualizations

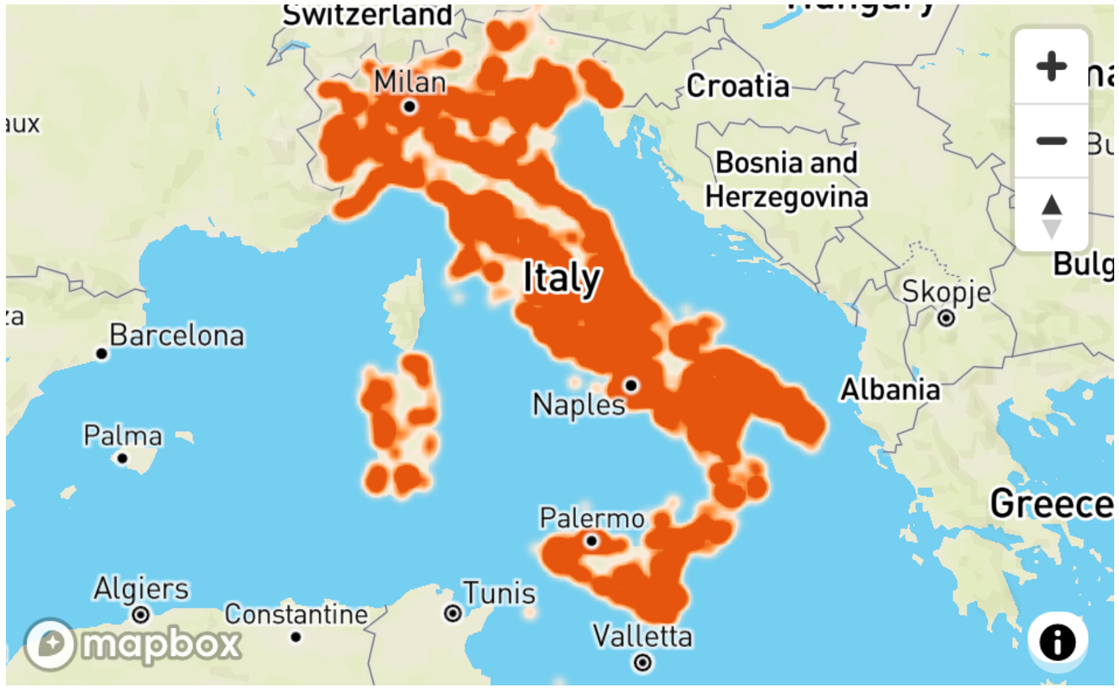
We have considered 5 different visualizations:

- Choropleth map
- Heatmap/Scatter plot
- Star plot
- Bar chart
- Scatter plot from PCA

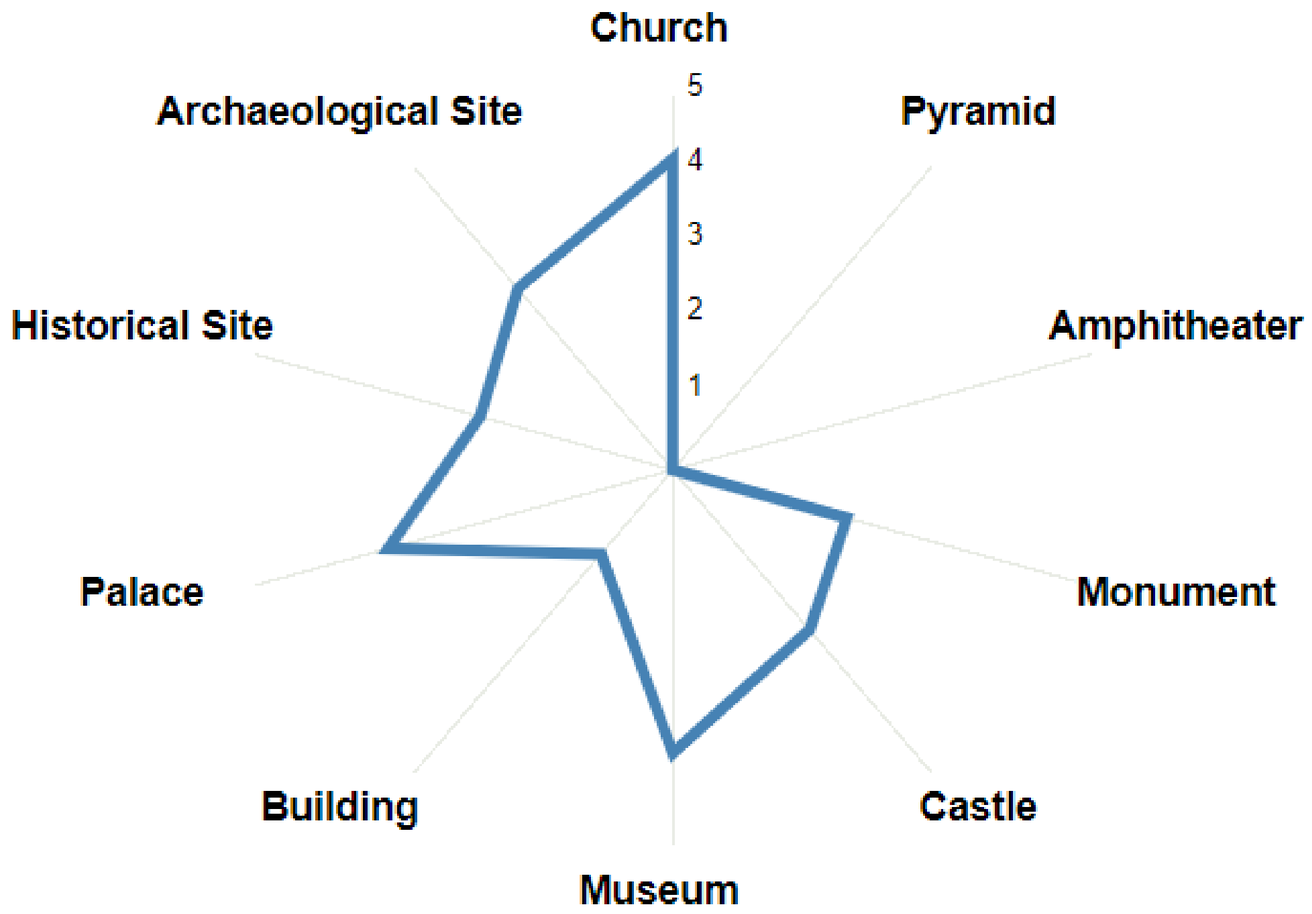
Choropleth Map



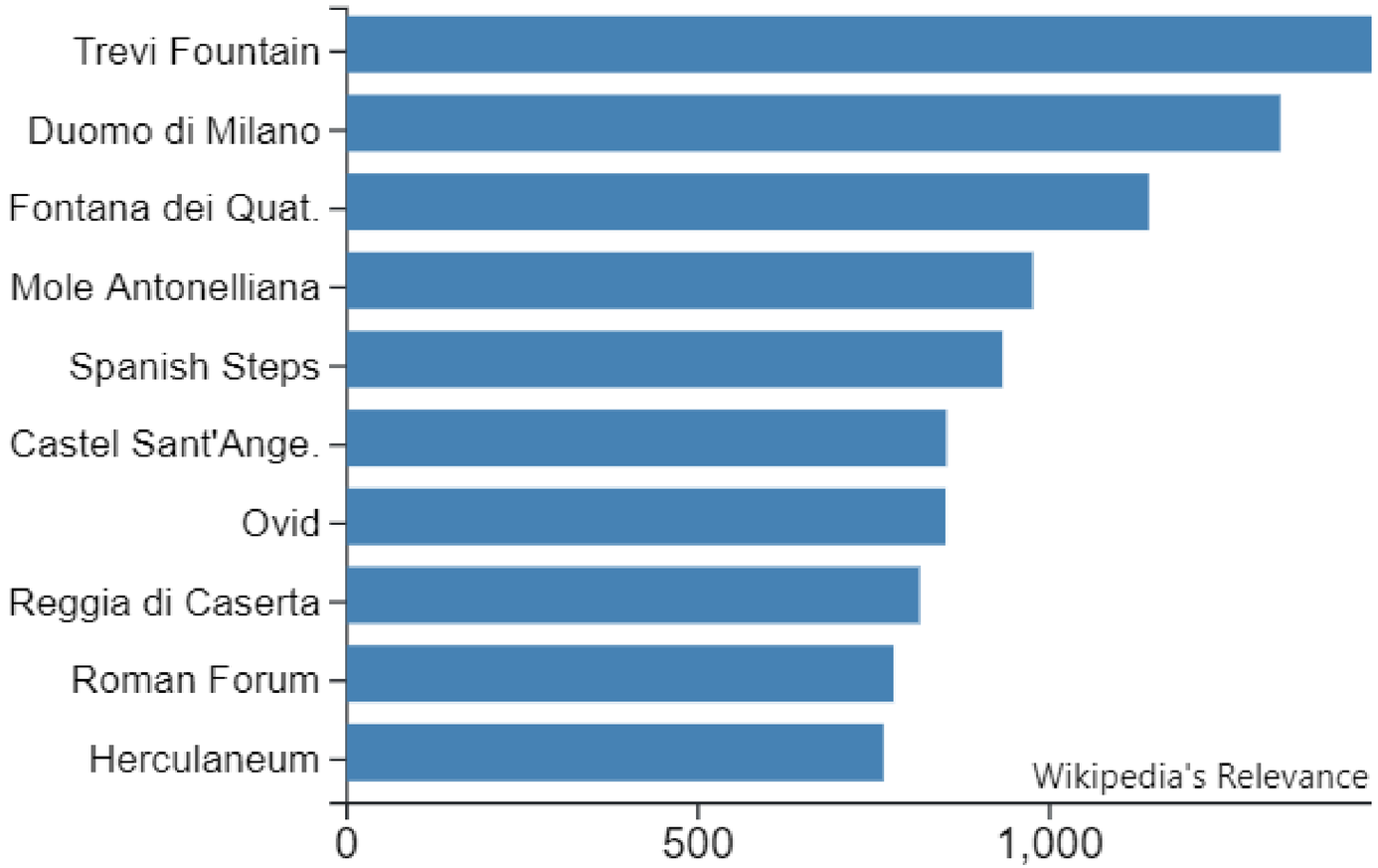
Heatmap/Scatter plot



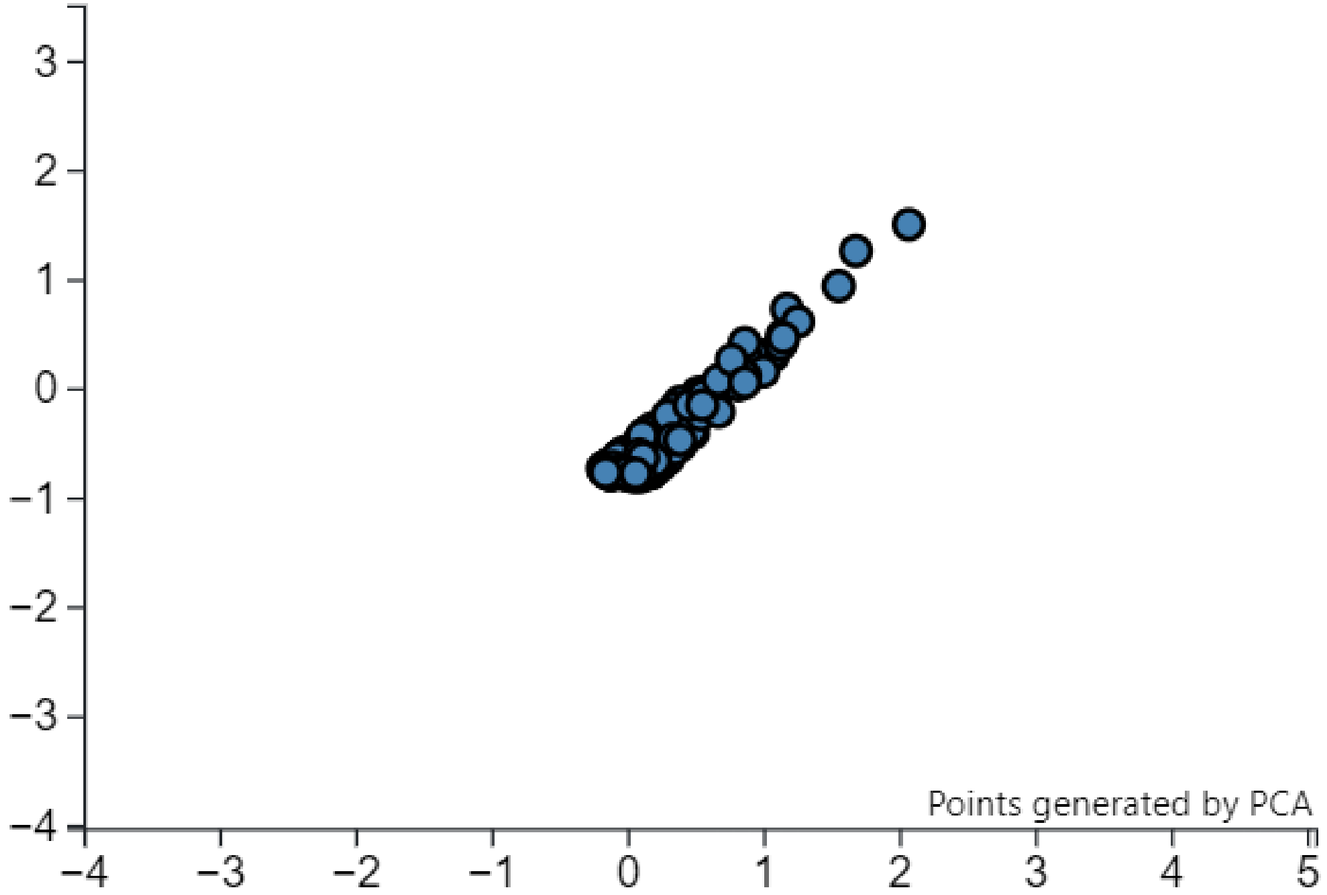
Star plot


Church
Pyramid
Amphitheater
Monument
Castle
Museum
Building
Palace
Historical Site
Archaeological Site
5
4
3
2
1

Bar chart


Trevi Fountain
Duomo di Milano
Fontana dei Quat.
Mole Antonelliana
Spanish Steps
Castel Sant'Ange.
Ovid
Reggia di Caserta
Roman Forum
Herculaneum
Wikipedia's Relevance
0
500
1,000

Scatter plot from PCA



WORLD'S MONUMENTS

Thank you for the attention!

Link to the project:
https://github.com/Programmer100th/Visual-Analytics-Project