BOREHOLES

Marta Fioravanti

Visual Analytics

prof. Salvatore Rinzivillo

a.a. 2019 - 2020

UNIVERSITÀ DI PISA

Introduction

The project visualizes the italian geotermal database realized by the Institute of Geosciences and Earth Resources (CNR, Pisa, Italy).

The aim of the visualization is to allow the user to explore the data and make simple researches using filters or simply navigating the map.

The technologies used for the project are:

- python (numpy, pandas...) to clean and reorganize the datasets
- vue.js to write the web application
- d3 to make the visualizations
- mapbox to display the geographical data

Preprocessing

The original data consisted in 3 .csv divided in: geographical reference, litology and temperature scans. Since data were quite raw and unprocessed, a cleaning has been necessary. In all the files, the rows with unknown well ID were removed, and the columns with an huge amount of missing data were dropped. In addition, the string data were prettified in order to have a coherent use of capitals; numerical data were initially compiled with comma: they have been transformed in dot-separated floats or in integers.

Geographical data
The first dataset has been converted in a geojson file in order to simplify the visualization process. Two columns were added to display the extreme temperature misurations inside the well.

The attribute kept were:

- key: the ID
- name: the human-readable ID
- lat
- long
- quote
- province: the italian region of the well or 'Mare' if it is under the sea
- type: the typology of the well
- use: the actual function of the borehole
- status: whether is closed or in use
- outcome: the evaluation of the observation
- min_temp: the minimum temperature registered
- max_temp: the maximum temperature registered

Litology data
A column named lito_init was added keeping only the first word of the 'litologia' one to have a fast even if quite inaccurate classification; a series of regex substitutions was performed in order to uniform the strings and avoid that plurals or hortographical mistakes lead to class duplicates. The amount of unique ground-types was brought to 103. The age attribute was translated in english with the aim to allow a future utilisation, for example with an API that can transpose the string in a numerical data (e.g. the ancientness).

The final dataset provided the following attributes:

- key
- daprof: the starting depth of the stratus
- aprof: the ending depth of the stratus
- litologia: the complete geological description
- lito_init: the obtained geological classification
- etarel: the original era of the stratus (translated)

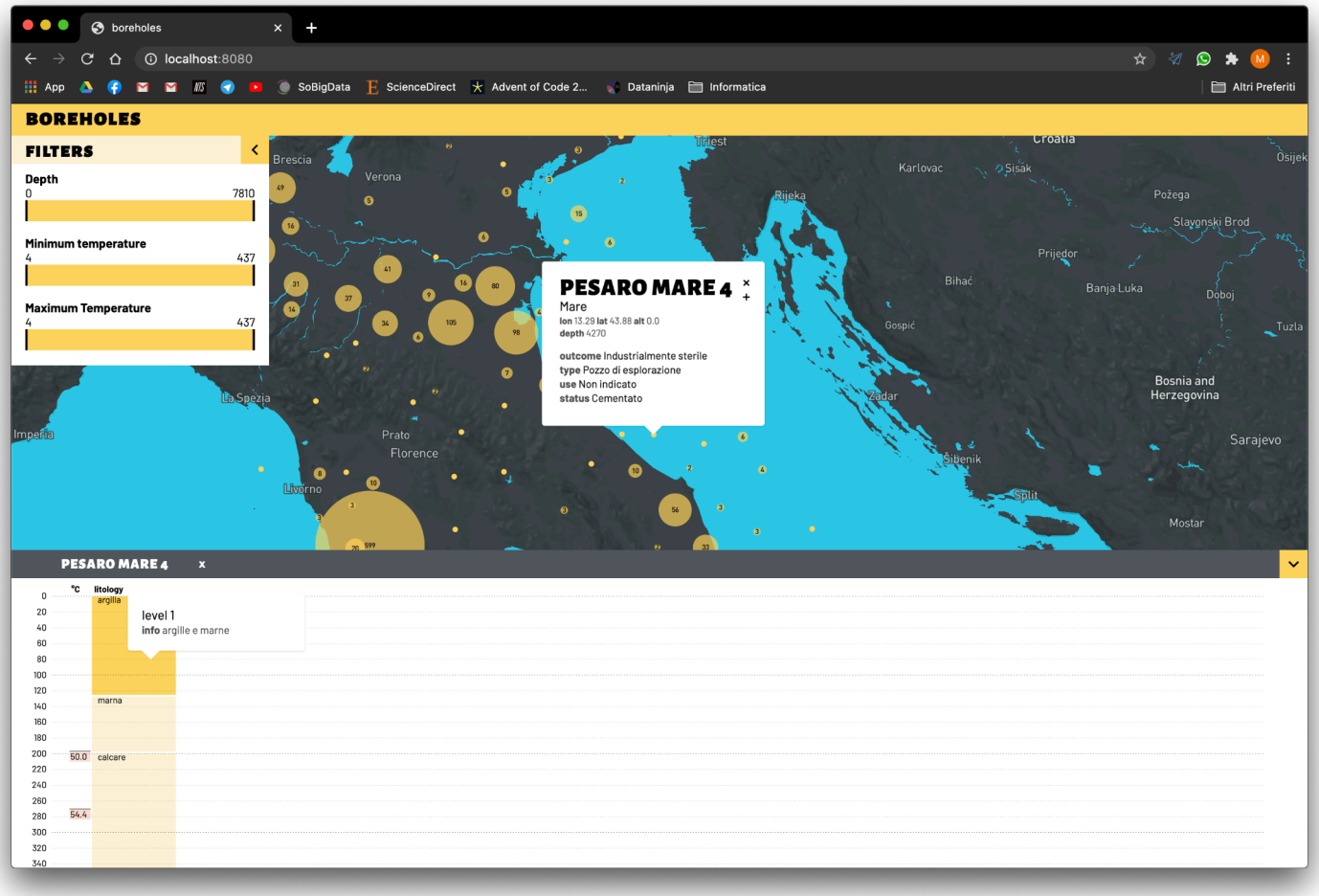
Temperature data
The dataset was reduced to 3 columns:

- key
- prof: the depth of the recording
- temp: the temperature in C°

Other smaller processings were performed directly in D3, like the removal of the temperature data at a depth greater than the depth of the well and the relativisation of the measures of the undersea wells, whose origin was negative.

Visualization

The visualization consists in a map, a filter section and a well-visualization section.



The map

The map's tileset was designed on mapbox studio and it contains only the essential informations: the reliefs, the water elements and the city centres. The visualization is limited to the surrounding area of Italy, in order to make the user experience less dispersive.

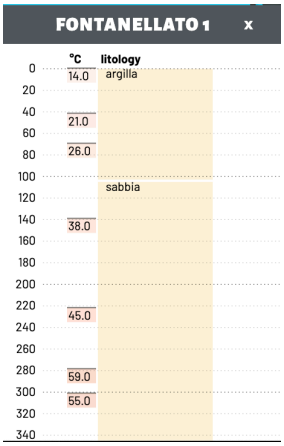
The data points are clustered until a certain detail level is reached, so that the visualization passes from a more general to a particular one. The dimension of clusters is obtained counting the records contained in each cluster and modifying the area of the cirlces accordingly. If at a lower zoom we can get an idea of the amount of wells in a certain area, at a greater one we can observe the real records.

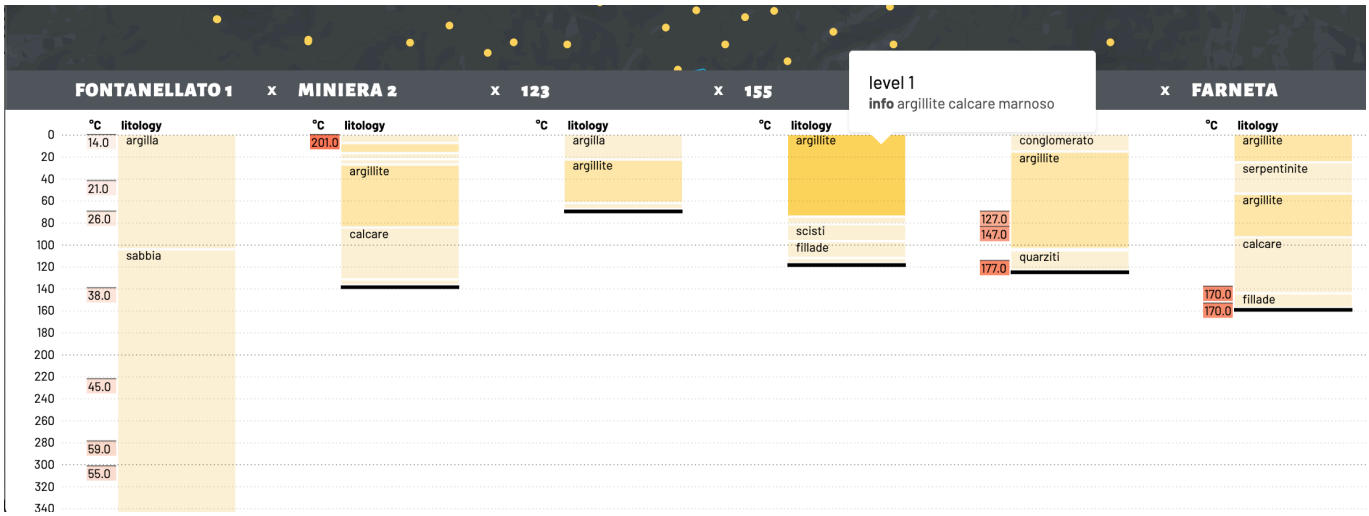
Clicking on a cluster, the zoom increases to display the points or clusters inside it, while when we click on an unclustered point a baloon appears revealing some basic

informations about it. If we want to know more, we can click on the '+' button and trigger the display of the litological and thermal data in the dedicated section.

The well-visualization

Inside the well-visualization section is displayed the whole well, stratus after stratus. The lito_init label is immediately visible over the rect represeting each litological component, unless it isn't enough deep. It's complete description is available when the mouse is on hover; when this occours, the rect acquire a more saturated color. If in the well visualization are present other levels with the same label, they are lighted. This strategy allows to distinguish similar data with a huge amount of labels: assigning each one a distinct color would have been disorienting.

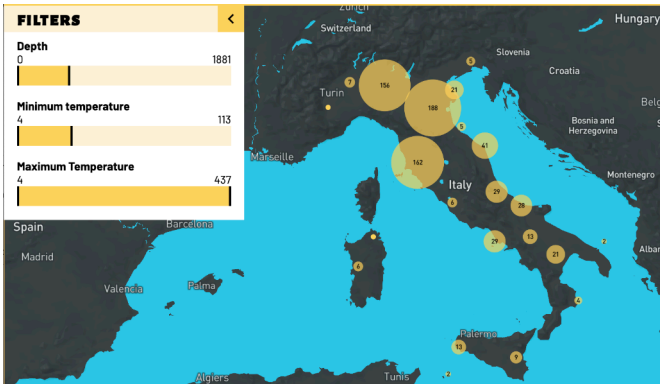




For what concerns the thermical data, a color scale whas dinamically created. Since the records are point-like, a flag-like visualization has been choosed: in fact, creating rects from one observation to another could have lead to misleading informations.

The filter section

The filter sections contains some double sliders and allow to select the data according to the well-depth and it's range of the temperature. Since the component is built with reusable snippets, many other search options can be added, like the altitude or similar. Every time a slider is used, the clusters and the points are updated according to the filtering options.



Proposed implementations

As further implementations I propose more filters and the possibility to visualize the temporal data. It would also be interesting a function that finds similar wells according to certain parameters, like litological composition. Also, the possibility to delete and move the wells in the well visualization and the possibility to resize the panel.