Data visualization - Milestone 2
Seasonal trends in hotel reviews

Hilda Abigél Horváth, Adrien Schurger-Foy, Julian Schnitzler

May 6, 2022

1 Hotel Map with Reviews

1.1 The idea

Usually websites only show the reviews of a specific hotel and we should navigate ourselves to another page to be able to compare with other hotels. This map will help the users compare different hotels for different season based on reviews. People will be able to easily navigate from hotel to hotel, and see the most important summary of the reviews.

- Map: the map shows Europe with points for hotel locations, it is zoom-able and pan-able.
- Time-frame: below the Europe map, a time axis will be added, where the viewer can select via a brush selector a time-frame (adjustable width) for which period he/she would like to see the data. This axis should display a line plot of a rolling window average of ratings for all the currently visible hotels. In other words, if you zoom into just one hotel, the line plot will be relevant to only that hotel.
- Colored points: the hotels are color coded based on the average score given in the chosen time period. It should be updated on click of the aforementioned brush selector and will map ratings to color. Clicking on a hotel point should display the tooltip and update the word clouds.
- Tooltip: the average rating, name, address, number of reviews. The tooltip should be persistent and the information it displays should be updated based on the selected time period.
- Word clouds: Updated based on the chosen time-frame and the chosen hotel. One wordcloud for negative reviews and one for positive.

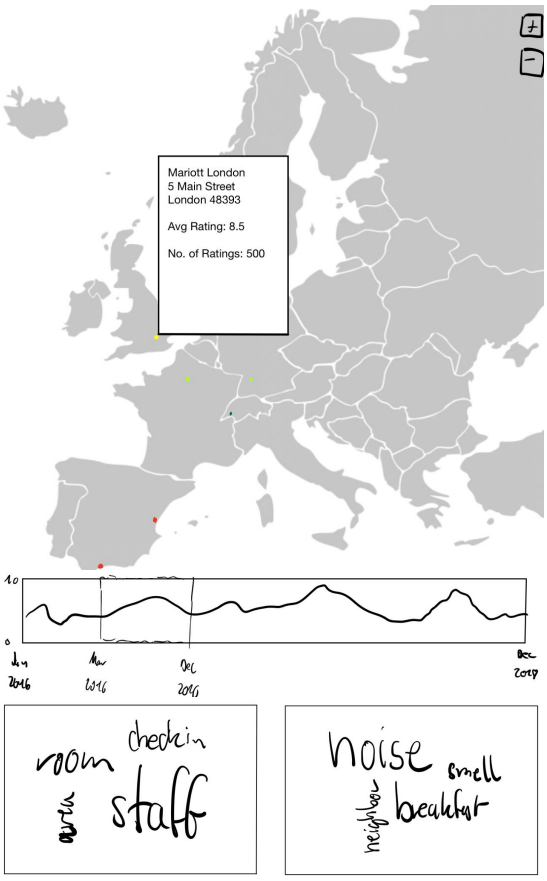

Figure 1: Sketch of the first visualization

1.2 Steps of delivery

MVP

The core visualization includes the following parts:

- Map: the map shows Europe with points for hotel locations, it is zoom-able and pan-able.
- Seasons: in the initial state, the 4 seasons can be chosen (buttons), which will determine which data to use for the following parts. The button should be highlighted and update the selected subset of data on click.
- Colored points: in the chosen seasons the hotels are color coded based on the average score given in the chosen seasons. It should be updated on click of the aforementioned button and will map rating to color.
- Tooltip: by clicking the locations of hotels, a tooltip is shown where the average score, name and address can be found. The tooltip should be persistent and the information it displays should be updated on click of the season buttons.

Next steps

Time-frame The next step will be to include the adjustable time selection.

Word clouds The last part of the visualization would be to include the word clouds. For that, we first need to preprocess the data and filter for nouns, to get only the topics that people were mentioning. *(Further possible advancements: We can add a section which shows on click on the word the number of reviews containing the word, and example reviews containing that word.)*

Further ideas
We could think of adding more color schemes for the user to choose from: Not only the average score, but also let the user choose a color based on whether the current "trend" of the ratings, i.e. fitting a linear regression on a sliding window of current reviews and determine whether its slope is positive or negative. Another idea would be to let the user replay his chosen period, so hit a "play button", run through the days in that period and then show a ping in a color relative to the rating of the review on the day it was published. The datastory in the end could be built by letting the users click several links to set pre-defined time-frames and level of zoom to let the user interactively be guided through the datastory, always with the possibility to also find own patterns.

Tools and Lectures needed
To generate the map, we need topojson together with some geoprojection. We might as well choose leaflet.js, we will probably try both and see which version makes it better, since both have pros and cons (topojson adds less unnecessary background information, whereas leaflet.js allows to better differentiate between hotels in the same city). We then add d3.zoom to make it zoomable and with event listeners we can allow users to pan it to move around in the zoomed map. For the time-frame, we use d3.line, add some d3.axis, and then add a d3.brushX to allow the user to choose the timeframe to look at (the end event should move it to closest month as the finest granularity to choose from). This should then influence the map, and the wordclouds. To our surprise, we apparently do not any additional d3 tool for the tooltips, and can simply generate them using the mouseover events. For the wordclouds we can use the d3-cloud plugin.
Furthermore, we need the following lectures: For the map, we need the lectures on Maps and Practical maps. To include the timeframe and connect the brush with the map, we need the lecture on interactions. The colored points are added accordingly to the lectures on colors and marks. For the wordclouds, we need the lecture on Text visualization.

2 Connection Map

2.1 The idea

Some people want to go to hotels where their nationality is present, or where a wide range of nationalities are. This map will help the users to identify the nationality of other guests (assuming people from different nationalities are equally likely to review a hotel), and choose the right place for themselves.

- Map: the map shows the World with points for hotels, it is zoom-able and pan-able.
- Connections: on click on the hotel, arrows will appear connected to the origin of the reviewers, with width weighted by the number of reviewers.

Figure 2: Sketch of the 2. visualization

2.2 Steps of delivery

MVP
For this part the core visualization is the map with the points for hotels and arrows between all hotels and the countries of origin of their reviewers.

Next step
Filter arrows The next step will be filtering the arrows for a specific hotel, and only showing them when a user clicks on one of the hotels in the map. A small tooltip should tell the user which hotel he is currently looking at.

Further ideas
It would be great to connect both maps interactively: Whenever a new hotel is chosen above, the same hotel is chosen in this map, and vice versa. The arrows could then also be updated according to the chosen time frame, showing the users in which season they could expect visitors from which countries in the world.

Tools and Lectures needed
To deliver this part we need the following tools:
To draw the map, we again need topojson, together with a Geoprojection. Here, we assume we do not need to use leaflet.js since this map should be mostly used in the most distant zoom level, to actually be able to see the arrows. We then need to draw arrows, which can be done by adding a simple svg triangle to a d3.line. Finally, we would add a d3.zoom to make the map zoom-able.

Additionally, we need the following lectures. We again need the lecture on maps, to be able to draw the maps and draw arrows on top. Finally, we need the lecture on data stories to connect both topics well and match them together to get one concise story.

3 Prototype

Please find our website with the skeletons of the visualizations on https://com-480-data-visualization.github.io/datavis-project-2022-bottas/.