

Individual Reflection

Declaration of Authorship

I, [YouTian Peng], confirm that the work presented in this assessment is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

[YouTian Peng]

Date: 2022/01/18

Student Number: 22076982

1 What factors do you think help to explain what went well/not well in the project?

This section discusses what factors went well/not well in the project.

Factors that did well:

- The project explored the impact of covid on the Airbnb rental market from different perspectives, diversely interpreting the impact.
- We divided the rental market according to the type of housing rental (short-term and long-term), which makes the analysis more practical.
- The designed data analysis framework improved the efficiency of code optimization.
- The NLP was used to conduct sentiment analysis on Airbnb user comments, and the results represented a practical pattern.

Factors that did not well:

- The assumptions are too strict in the research direction, reducing results confidence. More factors should be considered to improve the analysis accuracy.
- Some possible biases due to the number difference of samples are ignored because the amount of data in the three (covid period classification) data sets differ. The team has adjusted some applied methods (such as resampling data in NLP). However, some potential errors cannot be avoided in data visualization.
- GitHub was not well utilized, which makes code integration less efficient.

2 How do you think the other students in your group would evaluate your contribution to each of these outcomes?

This session will discuss other team members' evaluation of my contribution to the project from two aspects.

- Determination and adjustment of research framework:

I cooperated with other teammates to identify this project's research direction, which focused on the impact of covid on the Airbnb rental market from two different perspectives.

In addition, considering the differences between long-term and short-term housing sources, I achieved the definition of the short/long-term, improving the result's explanatory power.

- Technical realization:

In order to increase the traceability of the project, I determined the scientific framework for the code implementation based on previous studies Neff *et al.* (2017). The reproducible framework includes collecting, analyzing, identifying patterns, hypothesizing explanations, and testing hypotheses. In addition, the framework makes the project's code more manageable for teammates to adjust and optimize.

Under the designed framework, I first collected data from Inside Airbnb's website; the data included listing information on Airbnb; I cleaned the data set and made adjustments to the data type, format, etc. Then, based on the pre-processed data, I explained the changing trend of Airbnb's listings occupancy rate in different covid stages, and the analysis showed remarkable patterns. What is more, for analyzing the impact on the neighborhood level, I assisted other teammates in adjusting the visualization results into a more straight-forward version. When investigating host service quality with listings score, I break the original overall listing score visualization result into more complex constructed score components, which presents a more detailed peek at how the score changes at different outbreak stages. For sentiment analysis implementation, I first collected the 'reviews' data provided by Inside Airbnb. Then, after data preprocessing, I used VADER to classify user reviews. Next, the epidemic's impact on user comments' attitudes was observed, showing that residents' positive comments increased after the epidemic. In addition, I found that this pattern is related to the previous neighborhood pattern, and the connection was interpreted.

However, in the part of technical implementation, because I was responsible for most of the technical implementation, this caused other teammates not to have a deep understanding of the project code, which limited the efficiency of group integration and reduced the explanatory power of the report analysis. Moreover, this deprives the possibility of more insightful implementation methods due to a lack of critical aspects.

3 Briefly describe *one* event or experience from the group assessment that gave you new insight into ‘doing’ data science, and explain how and why this will be useful to you in the future.

Before our group discussions, I had already prepared the general research direction and framework due to previous data analysis experience. However, I hope to hear the thoughts of people who have studied different fields and have different thinking logic in this group work. During discussions, I guided and motivated my teammates to have practical discussions. Nevertheless, the final result was unsatisfactory because I did not get more helpful help other than my ideas.

I am an active listener and contributor to ongoing debates during discussions. I primarily act as a leader in the group, especially when an argument does not seem to be leading to anything beneficial, guiding the discussion while making sure that every point is thoroughly understood. However, this requires patience, as everyone learns differently, and the other members must be persuaded to wait until everyone is comfortable with a final decision on the issue. Therefore, I listened to and understood others' perspectives before accepting or contradicting them. Even when I have to contradict them, I try to do so in a way that does not make them feel like a failure to maintain a respectful discussion environment.

However, as a project leader, I found it challenging to inspire intrinsic motivation in members. Because too much leading will make the team members rely too much on your judgment. Therefore, I had to consult more information based on facts before making judgments, which made my workload more than others. Moreover, other team members will partially lose the will to think. I did not find an excellent way to guide and motivate other team members to make targeted thinking and contributions in this group assignment.

During this group work, I learned that members must actively interact with each other. But the agreement on everything sometimes is not beneficial. Sometimes a disagreement among group members on an issue heightens the discussion, with everyone participating and making more suggestions. For future group assignments, it makes sense to create critical disagreements to deepen the discussion.

Word Count: 960

References

Demchenko, Y. *et al.* (2016) 'EDISON data science framework: A foundation for building data science profession for research and industry', in *2016 IEEE international conference on cloud computing technology and science (CloudCom)*. IEEE, pp. 620–626.

Neff, G. *et al.* (2017) 'Critique and contribute: A practice-based framework for improving critical data studies and data science', *Big data*, 5(2), pp. 85–97.