Project 1 Feedback Report

--- What are the prevailing attributes of UFO

--- Jiajun Yu (a1806320)

This feedback report of Grand Challenges in Computer Science aims at explaining reflections from different perspectives for project 1 in order to address issues and modify the research. All feedback have been categorized, grouped, and divided into five parts which are ‘Research Benefits’, ‘Visualizations’, ‘Research Methods’, ‘Research Contents’, and ’Presentation Skills’. All points listed below will be discussed.

- Research Benefits:
  - The research should identify some benefits. (8 repetitions)
- Visualizations:
  - Graphs may not illustrate patterns clearly. (No repetition)
  - Try to explore other types of graphs to be used. (1 repetition)
  - Try to show the results of this research in a more interactive way. (No repetition)
- Research Methods:
  - Lacking Python skills is challenging. (No repetition)
  - Analysis process might be difficult. (1 repetition)
  - Multiple expressions might lead to some misinformation. (No repetition)
  - Replacing unknowns in the dataset can be challenging. (No repetition)
  - Try to deal with multiple variables. (No repetition)
  - Giving some more solid data and evidence is important. (No repetition)
- Research Contents:
  - Try to find any other trends presented in the dataset. (4 repetitions)
  - Try to find some correlations for UFO events. (No repetition)
- Presentation skills:
  - Handling question-and-answer session well is important. (No repetition)

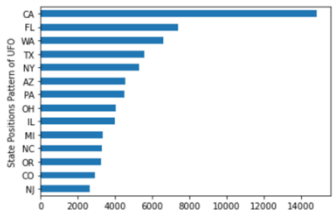
Discussion of Points Raised:

Research Benefits:

The feedback that the benefits of this research should be elaborated is an excellent advice. This issue has already been addressed in the first cut demo presentation. More concrete benefits will be included in the final presentation and the final written report.

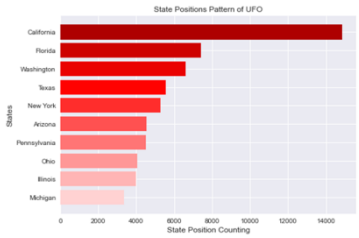
Visualizations:

Feedback on improving visualizations are enlightening. Previous research has shown several patterns of UFO by three bar plots. However, it appears that there was a lot of information in each single graph.



Previous bar plot

With the aim of helping people keep track of the key points, redundant information was deleted, and new implementations of Python code were adopted to beautify those graphs. As shown in the example on the right, only ten pieces of information left in the graph and abbreviated USA state names were replaced by their full names to avoid confusion.



Modified bar plot

Considering the time constraint in the final presentation, bar plots will be demonstrated in seconds. A classmate suggested the use of multiple visualization types. Using other intuitive visualization techniques combined with bar plots is an excellent suggestion to help audience have a better understanding of the research outcomes.

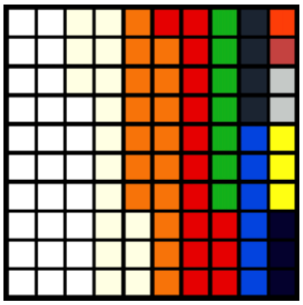


Figure 4

Figure 5

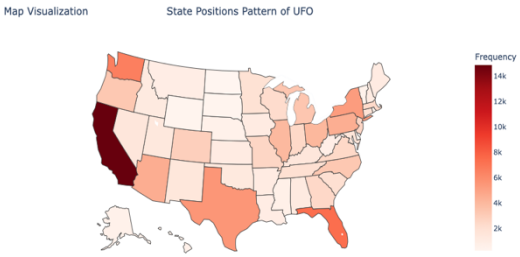


Figure 6

The three graphs above were implemented by Python. Figure 4 can not only show the prevailing colors of UFO, but also the ratio of them in a more intuitive way than pie charts. Figure 5 is able to show the frequencies for all words related to shapes by plotting them with various sizes. Figure 6 is a map visualization showing the distribution of UFO events in the USA as well as using color gradient to represent the overall number of UFO events in each state that witnessed from 1946 to 2021.

Utilizing interactive visualizations helps deliver more information in a subtle way and can draw attention from people easily by showing fancy animations. However, the outcomes of this research can be well illustrated by current visualization techniques, and due to the time limitation, it is hard to demonstrate interactive visualizations given the fact that there will be more than eight graphs to be explained in the final presentation.

Research Methods:

As one classmate mentioned in the feedback, it is true that lacking Python skills is challenging for this project. Searching targeted information on the Internet is beneficial for developing specific algorithms in a short time.

The analysis process is not as complicated as some students think. Using different Python scripts and Microsoft Excel can draw some interesting conclusions. More specific descriptions about the methods of this research will be specified in the final written report.

One of the classmates mentioned that multiple expressions might lead to misinformation. It is undeniable that the dataset does exist multiple expressions in term of describing time, durations, and positions. Handling this issue requires high Python proficiency, which is almost unattainable for a Python novice. Hence, utilizing the ‘Text Filter’ from Microsoft Excel combined with some

manual work to categorize and count various outputs is an appropriate approach to deal with this problem.

Someone mentioned that replacing unknowns in the dataset can be challenging. To investigate the appearance patterns by year and time of UFO, the research time span will be set from 1990 to 2020 in order to avoid unknown or incomplete records before 1990 (the information thereafter is almost intact). Apart from exploring these two patterns, the method of ignoring missing records will be adopted to explore the rest patterns of UFO since this project is closer to a qualitative research; therefore, the results of this research will not be as sensitive to some unknown information as the results of quantitative research.

Dealing with multiple variables to find some potential patterns is an excellent suggestion since all previous studies were based on double variables. To implement this idea, the appearance pattern of UFO in each month from 2008 to 2020 will be explored and a heatmap will be utilized to illustrate the relationship among appearance month, year, and the number of UFO events.

Giving more solid data and evidence is a good point. However, it is hard to find another dataset with more solid data and evidence. Nevertheless, referring to some academic articles from prominent researchers in this area can also improve the credibility of the research.

Research Contents:

Exploring any other trends presented in the dataset is another good suggestion. As mentioned above, the appearance pattern of UFO in each month from 2008 to 2020 will be investigated.

It will be interesting to find connections between some factors with UFO events by executing some correlation tests. However, the information needed is not included in the dataset. In addition, presenting huge number of findings in the final presentation is extremely challenging. Hence, finding correlations between UFO events with other factors will not be considered for this project.

Presentation skills:

The question-and-answer session is designed for audience to ask relevant questions about the presentation instead of letting the presenter elaborate what missed previously. Arranging the question-and-answer session well is of vital importance. Hence, presenters should be prepared for this part before giving their presentations.

Impact:

Appreciation should be given to people who gave feedback for this project. Constructive suggestions are of great help to the adjustment of the research. Visualizations were improved a lot by beautifying bar plots and utilizing intuitive visualization techniques. A new question was raised to enrich the contents of the research. In addition, the research method was adjusted to suit the actual situation. Peer given feedback changed some parts of the research and made my life harder. Adjusting visualizations is challenging especially for a Python novice because coding is

time-consuming and laborious. However, doing reflections can boost my research ability and sharpen time management skills.

Other reflections:

During this research, problems from Python scripts were discovered and fixed. The python script ‘Single word frequency counting' which counts how many descriptions of UFO events contains a unique word, would count repeated target words in each single description, which could make the result larger. The same problem existed in the ‘All words frequencies counting’ code, which would return inaccurate outcomes. In terms of presentation, overcoming tension and fear should be focused on. In conclusion, getting feedback from people is beneficial to this research.