Peng Yu
46635884

Peer Review

Group 3

In terms of presentation quality, the presentation follows every step in data science process, which is easy to keep up with the flow. Each slide has a key point and plot, which are simple, clear and easy to be understood. The group also used the plot to well explain the key point. The group did extremely well in the overall data presentation by using all different types of plots to make the data visible. The whole content is very logical and the data integration is comprehensive. In the part of data exploration, animation processing is used which is very creative. At the end, the recommendations were given based on the data analysis, which could definitely help the stream games to increase their game users and bring huge profits.

Good PPT layout.

The group 3 has done a perfect presentation. There is not much recommendations needed. The only recommendation is given as below:

Actionable：Although the overall content and data analysis are perfect, cross-validation of final results (88% players spent more than 99% of their game time in DOTA 2) needs to be performed rather than draw the conclusion from the scatter of unusual data analysis.

Group 1:

The presentation followed each step in data science process except the last part “enhance”. In the part of data fit for use, they analysed the relationship between influence factors including educational background, age & countries and the popularity of using language programming, which is easy to be understood. They conducted in-depth research on outlier in terms of age, excluding useless data. The overall layout can be improved by using some nice color pictures.

Actionable：In the part of storytelling with data, group 1 stated they would provide advice on students who was confusing about where to start leaning language programming skills. This is not well communicated and slightly deviated from the core theme because each language programming has its own function for different use. Instead of that, they can dip further in the part of storytelling with data: for example, to analyse why python is becoming popular? Drill down, zoom out and contrast can be used to tell why python is more popular.

Towards the end of presentation, the group mentioned “enhance”, R code with a few data is used but cannot be understood at all. The group should make this part simple and clear to allow the audiences to understand what they are presenting with using key point or simple sentence.

Specific: In the part of making data fit for use, a barplot for education background is given which is not specific to the main topic because is only reveals the amount of population for different educational levels but not relate it to programming languages users with different level of educational backgrounds, which is irrelevant to the main topic. However, a histogram of education vs popularity is presented later which well reveals different educational background population's preference with use of different programming languages.

Creative: the project is not creative because it does not include outside knowledge or information to enrich the content.

Group 5

The overall presentation is well designed and most diagrams are appropriate.
In the part of making the data confess, three-dimensional dataset and K-mean clustering are used to give a whole picture of relationship between in forest cover rate, rainfall and GDP. The graph is very creative as it makes the data clearly visible.

Actionable：In the part of storytelling with data, it gives an example of forest coverage rate which might be influenced by rainfall and GDP in Beijing and Anhui. It is good to give an example to reveal the relationship between forest coverage rate and the influence factors but cross-validations of the dataset are absent which make the example less convincing.

Specific: After giving an example of forest coverage rate and the influence factors (rainfalls and GDP), the author should further explain the relationship between GDP and forest coverage rate to help the audiences better understand why high GDP will increase forest coverage rate rather than present the reasons of inefficient forest cover and the strategies in Beijing.

Creative: in the part of recommendation, the project brings outside knowledge to help increase the rate of forest cover for Beijing including prevention of forest fire, reduce use of wood product and government support which slightly helps solve this global problem.