Evaluation of Project: Golden Global Twitter Analysis

Hoang Duong Trong & Cenzhuo Yao & Lu Zhang & Luyun Zhao & Sicun Huang

1.A one sentence summary of the overall goal/findings of the project.

The Golden Globes Nominees have different Tweeting Pattern, summarized by gender, age, and the prize result.

2.What is the motivation?

Comparing winners and losers' twitter activities to see the difference; Probably will predict the winner by their twitter activities.

3.Does the title make the motivation and goal/findings clear? Suggest a better title.

Better one will be: 73rd Annual Golden Globes: Twitter Activity of the Winners

4.Provide a bulleted list of the 3-4 main (sub)questions addressed by the project.

* Twitter Influence (popularity, time joined) and Temporal Patterns
* Social Popularity of Winners and Nominees
* Sentiment Analysis
* Pre-Post-During Golden Globe Analysis

5.Do these questions fit together or do they appear unrelated? How do they relate to the motivation and goal/findings?

These questions touch on multiple aspects of golden globe nominees' twitter experiences; the answers to them provide an overview of the demographics of these nominees and their tweeting habits together with the general audiences' attitude towards them.

6.List 2 to 4 additional questions you think should be addressed in this study.

* Are celebrities' tweeting habits different from their audiences'?
* What type of celebrity is more popular amongst the general audience?
* What type is more popular amongst the celebrities?

7.Explain what the data was. How many variables? How many observations? What dates are included in the data? Anything else interesting about the data.

The data are Golden Globe Nominees' (Celebrates and Movies) information and timelines.

8. Identify the slides/graphics that most clearly addressed the main questions.

Personally, the tweets heat map is the most compelling to me. It clearly shows when do celebrities tweet and the outcome is really interesting. For each element in heat map, the sample is large enough to convince us the conclusion is not just a coincidence.

9. Identify the slides/graphics that least clearly addressed the main questions.

Partially because the sample size is too small and it may be just by coincidence and because that the data cleaning has not been done so the outcome seems not that interesting. For instance, the most frequent word in each category are all RT, and I is the runner-up. I think it may be improved if we can remove those stop words and meaningless words.

10.Identify the most interesting slides/graphics.

The plot in slide 9: Follower/Following Behavior Exhibits Distinct Groups.

11.How easy did the slides make answering the above questions? If it was difficult, please suggest ways to improve it.

According o the conclusion in the last slide, it is easy and clear.

12.Comment on the visual consistency? For example, did the team use a color code to track specific variables? Suggest possible ways to provide more coherent and consistent visuals.

There are some inconsistency between the plots: In slide 7 and 8, the color for Old and Weak, Young and Powerful are different; In slide 9,10 and 14 and 17,18, the color for L and M are different