Seriously, How Deadly Is It to Join the Avengers?

Jessie Miller

In this work, I reproduce the results of Walt Hickey's "Joining the Avengers is as Deadly as Jumping Off a Four-Story Building." I report that roughly 40% of Marvel's Avengers have died at least once, ⅔ of which are likely to return from the dead. This discovery suggests that an Avenger conquering death for the first time is just as likely as any civilian finding lasting love. Furthermore, Avengers who died more than once have a 50% chance of returning from the dead. I also investigate the relationship between death and gender, character popularity, experience as an Avenger, and Avenger status, finding no significant predictors regarding which Avengers will die. Finally, in an attempt to predict which characters die, I created logistic regression and LDA models using the variable `Death1` as the response.

1. Introduction

In this work, I investigate just how deadly it is to be a member of Marvel's Avengers. The basis for this report originates from Walt Hickey's *FiveThirtyEight* report titled "Joining the Avengers is as Deadly as Jumping Off a Four-Sotry Building," which was released on May 12, 2015[1]. The dataset used in this report (`avengers.csv`) is the exact same data used in the original analysis, and is easily accessible via GitHub[2].

The primary goals of this investigation are as follows: (1) recreate the findings of Hickey's report, (2) identify possible correlations between death, character popularity, Avenger's status, and gender, and (3) attempt to predict which characters will die.

The `avengers.csv` dataset details the deaths of Marvel comic book characters between the time they joined the Avengers and April 30, 2015 (one week before the release of "Secret Wars #1"). Deaths that occured after this date - including those in Marvel's *Infinity War*, *Avengers: Endgame,* and other character deaths in the Marvel Cinematic Universe - are not represented in this dataset. The detailed contents of this dataset are examined in §6.

2. Related Work

In this publication, Hickey reported that 69 out of the 173 Avengers (~40%) died at least once, leading to 89 total deaths. Of these 89 deaths, on 57 occasions the individual made a "comeback" after dying. Additionally, Hickey found that there is a ⅔ chance that a member of the Avengers returned from their first death, but only a 50% chance they will recover from a second or third death. Finally, in the 53 years since the Avengers operation began, based on the overall mortality rate, one of the Avengers will die every seven or so months, with a permanent death occurring roughly every 20 months.

It's important to analyze how Hickey defines "death" in his analysis. A character is declared dead if (1) they are shown or directly implied to be killed, and (2) all of their allies and closest friends sincerely believe they are dead. For clarification, the following are instances in which the general public believed a character was dead, but in reality the character was alive and well:

1. Iron Man fakes his death to take a vacation; his allies are aware he's not really dead
2. The Hulk is sent into the future and his comics take place there
3. If all of the Avengers are sent to a pocket dimension after seemingly being killed by Onslaught, yet we follow their adventures on Counter-Earth

These are actual events that take place in the Marvel comics in which the character is not truly dead.

Using Hickey's same definition of "character death," I was able to successfully recreate his results. I found that exactly 69 characters died at least once, 46 of which somehow returned after their first death. This yields a ⅔ chance that a member of the Avengers will return after their first death, just as Hickey reported. Of the 46 Avengers who returned after their first death, 16 were resurrected and killed a second time. Only 8 of these 16 came back after their second death. Of the 8 Avengers who managed to beat death twice, two were murdered a second time; one Avenger was resurrected a third time. This is the same Avenger who has been killed and resurrected on five different occasions. Nevertheless, these results yield a 50% chance that an Avenger will return after a second or third death, as anticipated. Additionally, since 46 Avengers were resurrected after their first death, 8 after their second death, and 1 after their third, fourth, and fifth death, we can conclude that there have been a total of 57 resurrection incidents.

By comparing the length of time since the Avengers initiative began to the frequency of Avengers deaths, we can conclude that every 7.15 months an Avenger death takes place. If we neglect character resurrections, we find that a permanent character death occurs every 19.9 months. Both of these findings are inline with expectations put forth by Hickey's work. Therefore we can claim that we successfully recreated the findings of Hickey's publication.

3. Additional Work

In expanding upon Hickey's work, the first point of interest I wanted to address centered on any correlations in the data. Hickey's report comprised an excellent statistical summary of the Avenger's data, but he did not attempt to understand any potential relationships and correlations that may exist in the data. Therefore, this is where I chose to begin my original analysis.

I chose to investigate potential correlations in the following sets of variables using the Pearson correlation coefficient as an estimate for correlation: popularity and multiplicity of deaths, gender and death, Avenger's status and death, and length of time as an Avenger and death. The results for this analysis are reported in §4.

After completing this rudimentary correlation investigation, I next began to create original models in an attempt to predict character deaths based on gender and popularity. I first attempted this using a logistic regression model with `Death1` as the response variable and `Appearances` and `Gender` as the predictors. I first tried this using 50% of the data as training and 50% as testing. After discovering the results of this fit (see §4), I chose to redo this model with 70% of the data as training data and 30% as testing data.

I chose to design a second model using linear discriminant analysis to investigate the same relationships that the logistic regression models did. As with the logistic regression models, I performed two variations of this model, the first with 50% training data and the second with 70% training data. All results for my additional work are reported in the following section.

4. Results

I began by investigating the relationship between character popularity and deaths. The central question: "Are more popular characters likely to die multiple times?" Here, the term "popular character" is defined by the number of appearances an individual character makes; the more popular a character is, the more comics they appear in. I discovered there is no apparent

relationship between these two variables (Pearson correlation coefficient ρ=0.105). These results, as well as those that follow below, are reported in Table 1.

Next I chose to investigate the relationship between character death (singular) and gender in an attempt to determine if any potential gender biases resulted in character deaths. Fortunately, with a Pearson correlation coefficient of only 0.0533, we can claim that this gender bias does not exist.

I also investigated the relationship between Avengers status (represented by `Honorary` in the dataset) and death. This was an attempt to answer the question "Are Avengers who completed their training less likely to die?" Interesting, I found a negative correlation between these two components (ρ = -0.0923), which tantalizingly could suggest that inexperienced superheroes lead a safer life than trained superheroes.

Finally, the last question I chose to answer was as follows: "Are newer Avengers more likely to die than long-time Avengers?" Here, "new Avengers" are defined as members who have joined an associated Avengers team for 25 or fewer years. This question is unique from the previous question in that here, I try to examine the relationship between Avengers with real field experience (not just training at the Avenger's Academy) and death. I found that there does not seem to be any correlation between length of time as an Avenger and character death.

With the investigation of potentially correlated variables completed, I will now discuss the original models I developed as a part of my additional research. Unfortunately, my logistic regression models were both insufficient; both logistic regression models reported a testing error rate of 100%, suggesting that `Appearance` and `Gender` are not good predictors of death.

Relationship	Pearson Correlation Coefficient, ρ
Popularity & Multiplicity of Deaths	0.105
Gender & Death	0.0533
Avenger's Status & Death	-0.0923
Time as an Avenger & Death	-0.135

Table 1. An investigation of potentially correlated variables.

Fortunately, my LDA model did produce more promising results. When I used 50% of the `avengers.csv` data as training data, the LDA model reported a testing error rate of 39.1% which, although is relatively high, is much better than that of the logistic regression model. The LDA model improved even more (as anticipated) when I reran the model with 70% of the original data as training data, reporting a test error rate of 26.9%.

5. Discussion and Conclusion

Hickey reported that joining the Avengers is as deadly as falling off of a four-story building. For detail, falling from a four-story building (48 ft) results in death in 50% of patients who experience such trauma[3]. Since only ~40% of the Avengers died once, I believe Hickey's claim is referring to Avengers who die two or three times, as these Avengers only have a 50% chance of returning from the dead. Instead, I believe it's more accurate to claim that dying (once) as a result of joining the Avengers is just as likely as getting away with murder in America[4]. If an Avenger was killed for the first time, the odds of returning from the

dead are just as likely as a woman losing her attraction for her romantic partner after they share a first kiss[5]. So what's more difficult - beating death or finding love?

Unfortunately, I wasn't able to find any strong correlations in the data regarding death, popularity, gender, Avenger's status, and experience. This could suggest that (1) character deaths are random or (2) character death relies more heavily on an individual character's personality and willingness to take risks, which is not quantified in this dataset.

I was also unable to create a successful logistic regression model to predict character death. Fortunately, my LDA model yielded decently promising results (test error rate 26.9% with 70% of original data as training data).

I think this analysis could be strengthened with a wider Avenger's dataset. Perhaps if we had more information about character origins (e.g. Are characters who grew up in cities more likely to take risks than those who grew up in suburban areas, and therefore more likely to die? How might acquisition of Avenger superpower effect character deaths? Are Avengers in a relationship less likely to die? Which Avengers died in battle? etc.) or story arcs (e.g. Do Avengers with one single long-time nemesis have a better chance at survival than Avengers who fight large-scale battles?) we might be able to perform a better predictive analysis. Nevertheless, lack of unique data and size of the `avengers.csv` dataset definitely contribute to the downfalls of this project.

6. About the Data

The columns contained in the `avengers.csv` dataset are as follows:

- URL - the URL of the comic character on the Marvel Wikia
- Name/Alias - the full name or alias of the character
- Appearances - the number of comic books that the character appeared in as of April 30, 2015
- Current? - Is the member currently active on an avengers affiliated team?
- Gender - the recorded gender of the character
- Probationary - if the character was given probationary status as an Avenger, this information will be listed here
- Full/Reserve - the month and year the character was introduced as a full or reserve member of the Avengers
- Year - the year the character was introduced as full or reserve member of the Avengers
- Years since joining - 2015 minus the Year
- Honorary - the status of the Avengers: if they were given "Honorary" Avengers status, if they are simply in the "Academy," or "Full" otherwise
- Death1 - Yes if the Avenger died, No if not.
- Return1 - Yes if the Avenger returned from their first death, No if they did not, blank if not applicable.
- Death2 - Yes if the Avenger died a second time after their revival, No if they did not, blank if not applicable.
- Return2 - Yes if the Avenger returned from their second death, No if they did not, blank if not applicable.
- Death3 - Yes if the Avenger died a third time after their second revival, No if they did not, blank if not applicable.
- Return3 - Yes if the Avenger returned from their third death, No if they did not, blank if not applicable.
- Death4 - Yes if the Avenger died a fourth time after their third revival,

No if they did not, blank if not applicable.

- Return4 - Yes if the Avenger returned from their fourth death, No if they did not, blank if not applicable.
- Death5 - Yes if the Avenger died a fifth time after their fourth revival, No if they did not, blank if not applicable.
- Return5 - Yes if the Avenger returned from their fifth death, No if they did not, blank if not applicable.
- Notes - Descriptions of deaths and resurrections.

Bibliography

[1] Hickey, Walt. "Joining The Avengers Is As Deadly As Jumping Off A Four-Story Building." *FiveThirtyEight*, FiveThirtyEight, 12 May 2015, fivethirtyeight.com/features/avengers-death-comics-age-of-ultron/.

[2] Fivethirtyeight. "Fivethirtyeight/Data." *GitHub*, github.com/fivethirtyeight/data/tree/master/avengers.

[3] Smith, Charles. "Trauma Anesthesia." *Cambridge University Press*, 2015, https://books.google.com/books?id=xGwaCAAAQBAJ&pg=PA3&lpg=PA3&dq=ld50+48+feet&source=bl&ots=i04cd_BrOK&sig=AjzlZHftKsRAL1l-fBXTTDe8wz0&hl=en&sa=X&ved=0ahUKEwjYjuHAoIzLAhUUwWMKHZFCBMgQ6AEIMzAD#v=onepage&q=ld50%2048%20feet&f=false.

[4] Lopez, German. "There's a nearly 40 percent change you'll get away with murder in America." *Vox*, 24 Sept. 2018, https://www.vox.com/2018/9/24/17896034/murder-crime-clearance-fbi-report.

[5] Post Staff Reporter. "Are the odds of everyday life in your favor?" *New York Post*, 4 Jan. 2014, https://nypost.com/2014/01/04/what-are-the-odds-youll-die-in-a-year-you-may-be-surprised/.