**Assignment 1: Database Management and Sample Size Estimation**

Mikaela Ethier-Gagnon

Department of Psychiatry, Dalhousie University

PSYR6003: Fundamentals of Applied Statistics

Dr. Igor Yakovenko

March 6th, 2024

**Assignment 1: Database Management and Sample Size Estimation**

In the current investigation, data was collected from Avengers (*N*=814) in the final battle against Thanos across two battlefields (North vs. South). This dataset was obtained from Dr. Igor Yakovenko’s repository on GitHub (<https://github.com/iyakoven/PSYR6003-Assignment-1>). Data analysis was completed using a custom R script in R Studio (version 4.3.2; R Core Team, 2023). Data cleaning was performed to remove incomplete or erroneous cases from the dataset. Two participants were excluded as they were incomplete cases, resulting in a final sample of *N*=812.

Descriptive statistics for participants who were reported to not have superpowers and had died in battle (*N*=101) were calculated using the *tidyverse* package (Wickham et al., 2019). Outcomes of interest included combat effectiveness; representing a sum of scores for agility, speed, strength, and willpower; (*M=*497.53, *SD* = 177.56, Range = 67.25-946.89), kills (*M=*2.55, *SD* = 8.81, Range = 0-79), and injuries (*M=*4.55, *SD* = 0.74, Range = 2-5). Given the error in the mean model for combat effectiveness (*SD*=177.56), this indicates potentially great variability in the data, suggesting that the mean model may not best represent the observed data and is likely erroneous relative to the mean models for kills (*SD* = 8.81) and injuries (*SD* = 0.74) which have less error.

Amongst this subset of data including only participants who did not have superpowers and had died, descriptive statistics for avengers categorized in either the North or South battlefields can be found in Table 1. We aimed to determine the battlefield (North vs. South) which was most effective in combat, and which had the most injuries, using the *tidyverse* package (Wickham et al., 2019). Based on mean combat effectiveness scores, the North battlefield was most effective in combat (*M= 499.78*, *SD* = 174.07, Range = 130.68-897.06) relative to the South battlefield (*M=* 491.68, *SD* =189.53, Range = 67.25-946.89). On the other hand, the battlefield with the most injuries was the North battlefield (*M= 4.60*, *SD* = 0.68, Range = 2-5), as compared to the South battlefield (*M=4.43*, *SD* = 0.88, Range = 2-5).

In a secondary analysis of this sample, we aimed to understand the relationship between presence or absence of superpowers and IQ scores using an independent samples t-test. We hypothesized that avengers with superpowers had a higher average IQ than those without.

To estimate the required sample size, we decided to conduct an a-priori power analysis. Given a specified level of desired statistical power, this method enables the examination of specific effect sizes which can be statistically rejected. Therefore, using this calculation we can be relatively certain that the sample size attained from this population will provide us with reasonable power to detect the effect of interest, if a true effect is present. Alternatively, given that the sample of avengers in the battle against Thanos is finite, we could have instead chosen to measure the entire population, this way all data that could have possibly been collected is being examined. However, given potential resource constraints, the a-priori power analysis in this case was considered the optimal option to collect the minimum amount of data to achieve appropriate statistical power to detect effects of interest.

For the purpose of the current hypothesis, we chose a small Cohen’s d effect of 0.2. If this were a real study, this effect size would have been based on the smallest effect size of interest. Therefore, in this case, *d=*0.2 would be considered the smallest effect of practical or theoretical significance in the context of this study. For instance, perhaps only small differences in IQ amongst avengers with vs. without superpowers may justify the need for educational resources for either group, hence an effect of *d*=0.2 would have been deemed sufficiently significant in a practical sense. On the other hand, we may have considered the expected effect based on other peer-reviewed literature and theoretical predictions, ideally a meta-analytic effect size estimate examining the same or a similar research question. However, given that the literature examining IQ and superpowers in avengers is limited, we did not choose this option.

Using the *pwr* package (version 1.3-0; Champely, 2020), a-priori power analysis was conducted to determine the required sample size to conduct an independent samples t-test examining differences in IQ between two groups (avengers with superpowers vs. avengers without superpowers). For the purposes of this analysis, as described previously, we assumed that the smallest effect size of interest was *d=*0.2. We selected a standard two-sided alpha level of 0.05, though we anticipate that avengers with superpowers have a higher IQ than avengers without superpowers, literature examining these properties in avengers remains limited, and thus a two-sided alpha will be considered given the possibility of avengers without superpowers having a higher mean IQ than those with superpowers. Given resource constraints, the highest anticipated statistical power we selected was a standard of 80%. Therefore, we will have 80% power to detect between-group differences (i.e., IQ of avengers with superpowers vs. without superpowers) with a sample size of *N=788,* or *n*=394 per group.

We also aimed to determine whether we had enough power to confirm if there truly was no difference between the groups given the sample size calculated previously. Using the *TOSTER* package (Caldwell, 2022; Lakens, 2017), an equivalence test was conducted by performing two one-sided tests. Given the lower and upper bounds of -0.2 and 0.2, as specified by our smallest effect size of interest, and a sample size of *n*=394 per group as determined by the previous power analysis, we would have 75.49 % power to detect a zero effect. In other words, we would have 75.49% power to confirm that there is truly no difference of interest in IQ between avengers with superpowers compared to avengers without superpowers with a total sample size of *N*=788.

An independent samples t-test was performed to compare the means of the two groups from the existing pool of collected data, and the the t-statistic was found to be 4.25. Using the *effectsize* package (Ben-Shachar, Lüdecke & Makowski, 2020), a small Cohen’s effect size of *d*=0.30 (95%CI [0.16, 0.44]) was determined, indicating a mean difference in the IQ of avengers with and without superpowers with a small effect. Whether or not the CI is precise is dependent on the context of the research question, and whether the magnitude of the range of the CI is meaningful within this context. Given that the smallest effect size of practical significance was designated as *d*=0.2, though this value falls within the range of the CI, the lower bound of the CI is below the smallest effect size of practical significance, and thus, this confidence interval is not narrow enough to be considered precise and we cannot be certain that the true value in the population lies above the smallest effect size of practical significance.

**References**

Ben-Shachar, M., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of Effect Size Indices and Standardized Parameters. *Journal of Open Source Software*, *5*(56), 2815. https://doi.org/10.21105/joss.02815

Champely, S. (2020). pwr: Basic Functions for Power Analysis. R package version 1.3-0. https://CRAN.R-project.org/package=pwr

Lakens, D. (2017). Equivalence tests: A practical primer for t-tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*(4), 355-362. DOI: 10.1177/1948550617697177

R Core Team. (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

**Table 1**

*Sample characteristics for North and South battlefields*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Characteristic | North Battlefield | | South Battlefield | |
|  | Mean (SD) | Range | Mean (SD) | Range |
| Combat Effectiveness | 499.78 (174.07) | 130.68-897.06 | 491.68 (189.53) | 67.25-946.89 |
| Kills | 1.71 (4.57) | 0-34 | 4.75 (14.99) | 0-79 |
| Injuries | 4.60 (0.68) | 2-5 | 4.43 (0.88) | 2-5 |

*Note.* SD: Standard deviation; Combat Effectiveness refers to a sum of scores for agility, speed, strength, and willpower; Kills refers to how many bad guys the avenger killed; and Injuries refers to how many injuries the avenger sustained.