PSYR6003 Assignment 1

Chelsea Ingham

PSYR6003: Fundamentals of Applied Statistics and Research Design

Dr. Selena Singh

26 February 2026

PSYR6003 Assignment 1

This study explored data collected from 814 avengers during two final battles with Thanos, which took place at two separate battlefields. The primary analysis explored the combat effectiveness of the overall sample and between the two battlefields. The secondary analysis examined differences in IQ between participants with superpowers versus those without superpowers. It was hypothesized that participants with superpowers would have a higher IQ than those without a superpower, to be explored with an independent samples t-test. However, a power analysis revealed that the sample was under-powered to detect a medium effect size due to the large discrepancy in sample sizes between the two groups. Future studies should gather larger samples, if possible, to detect the effect size of interest.

Methods

The dataset was cloned from a GitHub repository created by Igor Yakovenko, accessed at the following link: https://github.com/iyakoven/PSYR6003-Assignment-1. Data was collected from two battlefields during a final battle with Thanos. Participants were 814 avengers. A retrospective analysis was conducted on 15 variables encompassing a variety of physical and psychological characteristics and battle statistics. As the primary analysis was interested in only those participants who were deceased and did not have any superpowers, data was cleaned to remove any participants who did not fit these criteria. All participants were included in the secondary analysis examining the relationship between IQ and presence of a superpower.

Statistical Analysis

Participants with missing values for any variables (n = 2) were removed from the dataset, resulting in a sample of n = 812. A new construct was created to operationalize overall *combat effectiveness*, calculated using the sum of *agility, speed, strength,* and *willpower* for each participant. For the primary analysis, the dataset was filtered to include only participants that were both deceased and possessed no superpower (n = 101). Descriptive statistics including mean, standard deviation, and range were calculated for the variables of *combat effectiveness*, *kills*, and *injuries*. Descriptive statistics were also calculated independently for each battlefield (north and south).

For the secondary analysis, an independent samples t-test was used to compare the IQs of participants with or without superpowers. Two methods of calculating the necessary sample size were considered; the first being a power analysis, and the second being to examine previous literature and determine a sample size that has been sufficient to detect a meaningful effect in

existing studies. As the literature on the topic is minimal, a power analysis was determined to be the best way to proceed.

A power analysis using the pwr package in R was conducted to determine if the sample was adequately powered to detect a medium effect size (d = .50). A medium effect size was selected because this is the smallest effect size of interest. Any effects smaller than this are not clinically meaningful, and therefore not of interest in the current context. The target statistical power was .80, meaning there is an 80% chance of detecting a true difference in IQ between avengers with and without superpowers, while accepting a 20% chance of a Type 2 error. This level balances the need to detect meaningful differences while acknowledging practical constraints of working with an existing dataset. An alpha level of .05 was selected, reflecting a 5% probability of committing a Type 1 error. This is consistent with standards in behavioural research, allowing these results to be comparable to previous literature. The sample was determined to be underpowered to detect the desired effect size, with a minimum necessary sample size of n=64 per group (the 'with superpower' group had a sample of n = 32 while the 'no superpower' group had a sample of n = 780). The current sample had power = .50 to detect the desired effect size. Another power analysis was conducted to determine if the sample was powered to perform two one-sided tests (a TOST test). It was determined the sample was also underpowered for the TOST test, which would require n = 70 per group. An independent samples t-test was conducted using the metafor package in R.

Results

Primary Analysis

Descriptive statistics for the kills, injuries, and combat effectiveness of subsample 1 are presented in Table 1.

Table 1

Descriptive statistics for deceased participants without superpowers (n=101)

Variables	*M*	*SD*	Min	Max
Kills	2.50	8.80	0.00	79.00
Injuries	4.50	0.74	2.00	5.00
Combat Effectiveness	497.50	177.60	67.20	946.90

Descriptive statistics for deceased participants without superpowers divided by battlefield are presented in Table 2.

Table 2

Descriptive statistics of deceased participants without superpowers (n=101) by battlefield

Battlefield	Variables	*M*	*SD*	Min	Max
North	Kills	1.71	4.60	0.00	34.00
	Injuries	4.60	0.68	2.00	5.00
	Combat Effectiveness	499.80	174.10	130.70	897.10
South	Kills	4.75	14.90	0.00	79.00
	Injuries	4.40	0.87	2.00	5.00
	Combat Effectiveness	491.70	189.50	67.20	946.90

The North battlefield demonstrated the highest combat effectiveness and the highest mean number of injuries.

Secondary analysis

The independent t-test analysis revealed an effect size of $d = .76$, or a medium to large effect size, 95% CI [.41, 1.11]. This is not a precise estimate as the bounds of the confidence spanned small, medium, and large effect sizes.

Discussion

Participants fighting on the North battlefield were found to be slightly more effective in combat ($M = 499.8$, $SD = 174.1$) than those fighting on the South battlefield ($M = 491.7$, $SD = 189.5$). Participants on the North battlefield also sustained slightly more injuries ($M = 4.6$, $SD = 0.7$) than those on the South battlefield ($M = 4.4$, $SD = 0.9$). The mean number of kills appears to be the most erroneous in this model, as this variable presented the widest range and largest standard deviation, displaying considerable variability compared to the other variables.

The sample was underpowered to detect our effect size of interest ($d = .50$), as well as to carry out a TOST test. An independent samples t-test revealed a medium to large effect size of $d = .76$, but this was not a precise estimate as the bounds of the confidence interval spanned from small to medium to large effect sizes. This is likely due to the large discrepancy in sample sizes between the two groups of interest, with the smaller group limiting the statistical power.