**Analysis of Combat Effectiveness and the Impact of Superpowers on IQ in the Avengers Dataset: An Examination Using a Fictional Dataset**

Author: Alireza Aleali

Affiliation: Dalhousie University

Course: PSYR 6003: Advanced Topics in Psychology

Instructor: Dr. Igor Yakovenko

Date: 6th of Mar 2024

**Abstract**

This assignment explores a fictional dataset named "Avengers," focusing on combat effectiveness, the impact of superpowers on cognitive abilities, and associated outcomes. As part of an educational exercise designed to assess the application of skills learned in class, the project emphasizes data management using Tidyverse packages, computational reproducibility principles, and statistical analysis techniques. The objectives include data cleaning, creation of new variables, subset analysis, sample size estimation, and power analysis to understand the nuances of conducting psychological research with complex datasets. The findings, while based on fictional data, serve as a practical examination of the ability to apply statistical methods and reproducibility principles in research, showcasing the importance of rigorous data management and analysis in understanding hypothetical scenarios within the Avengers universe. This assignment is an educational tool meant to evaluate proficiency in database management and sample size estimation as taught in the PSYR 6003 course.

**Introduction**

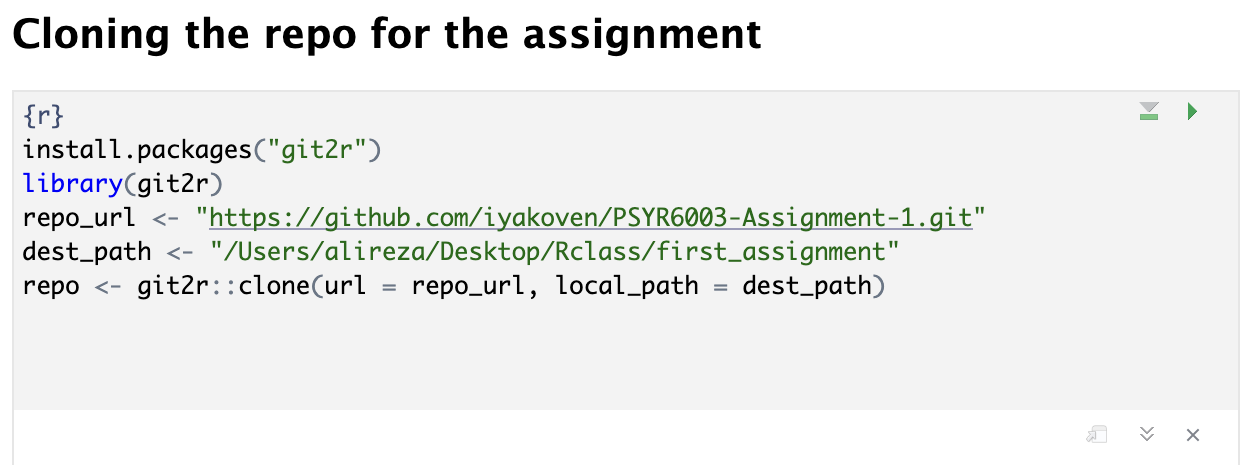
This assignment utilizes a fictional "Avengers" dataset to apply and evaluate data management and statistical analysis skills learned in the PSYR 6003 course. The dataset, while hypothetical, offers a complex set of variables that simulate real-world data analysis scenarios, allowing for a hands-on approach to learning. The focus is on assessing combat effectiveness, the hypothetical impact of superpowers on cognitive abilities, and other related outcomes, with an emphasis on employing Tidyverse packages for data manipulation and adhering to principles of computational reproducibility. The purpose of this exercise is to provide a practical application of course concepts, testing the ability to manage, analyze, and interpret data effectively.

**Methodology**

**Task 1: Data Cloning and Inspection**

Objective: Demonstrate ability to import and inspect a dataset for cleanliness and integrity.

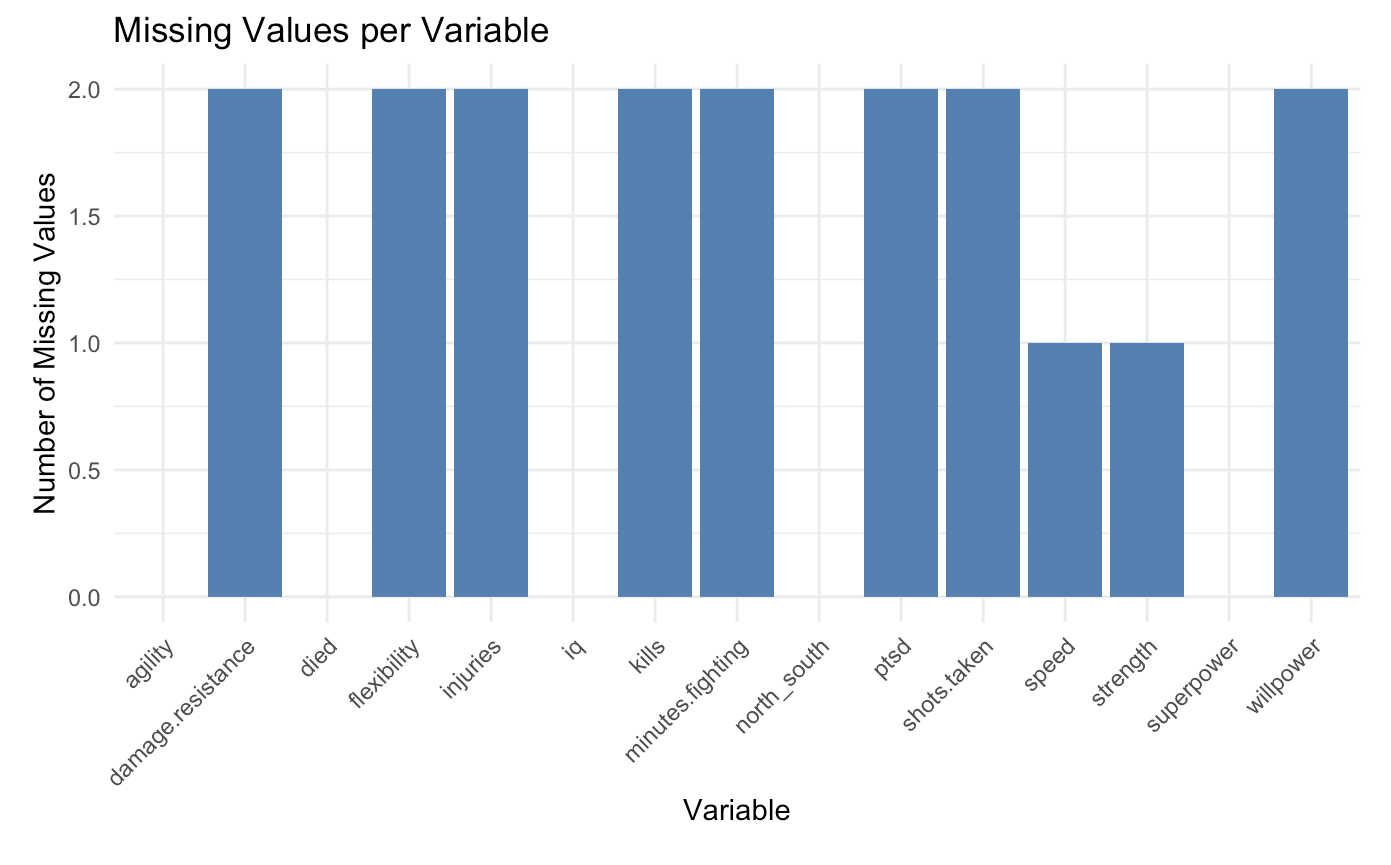
Approach: The dataset was cloned from a GitHub repository, followed by an inspection for missing or unusual values using appropriate functions within the R programming environment by using code below:





As you can see from the code, I wrote extra lines for showing the missing values per variable

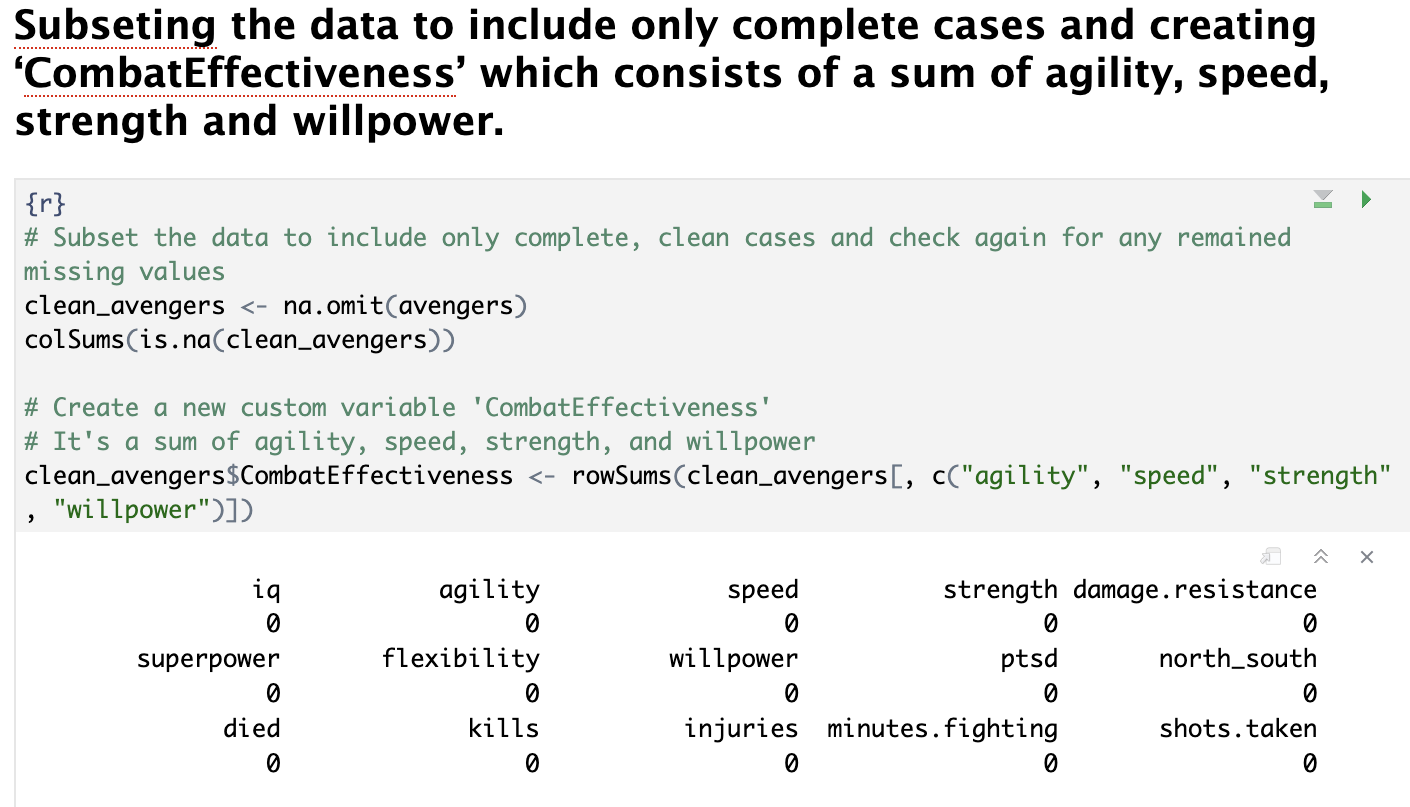
And by running that you can have this plot:



**Task 2: Data Cleaning and Variable Creation**

Objective: Apply data cleaning techniques and create a new variable representing combat effectiveness.

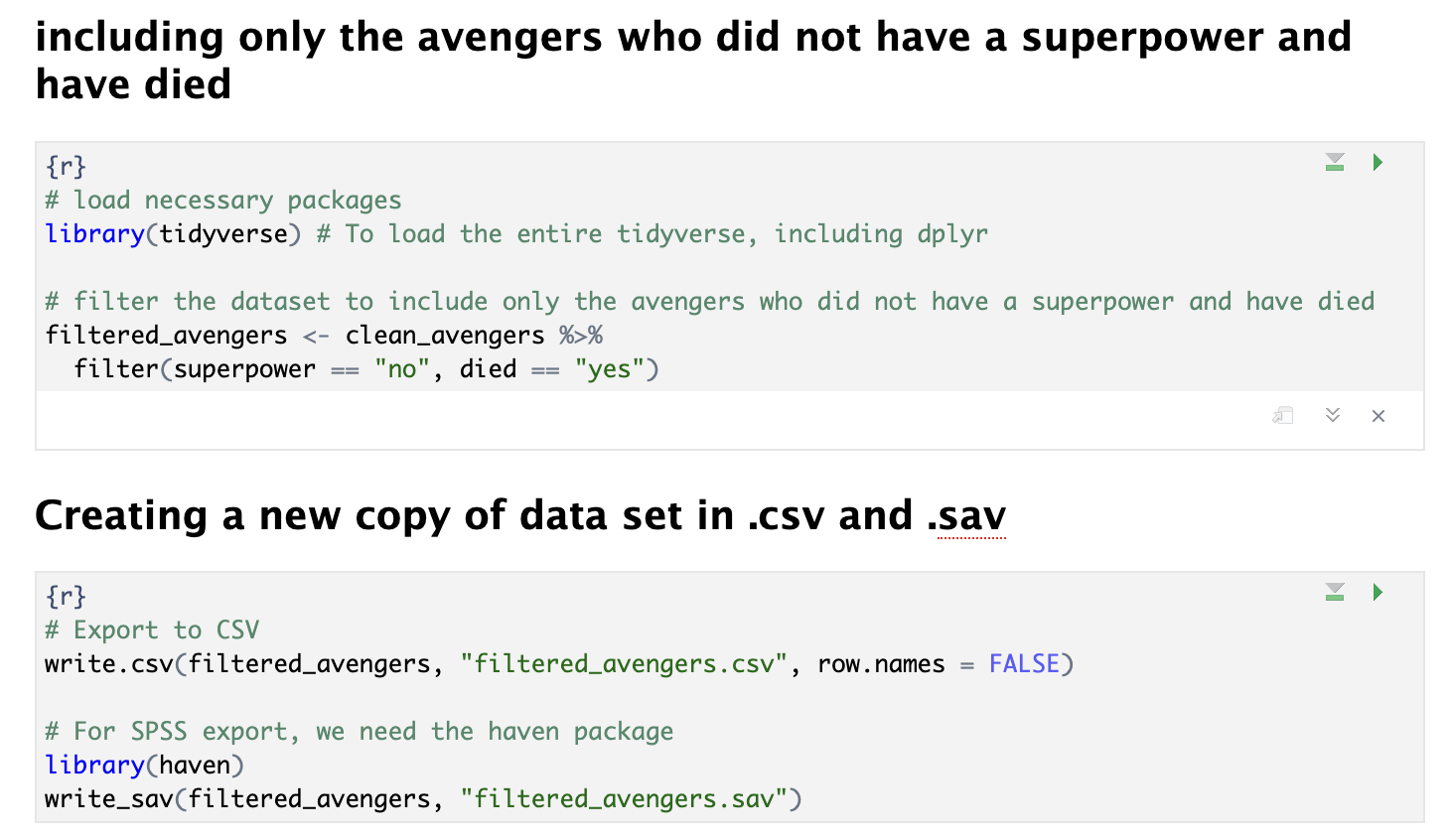
Approach: Data cleaning was conducted to remove incomplete cases. A new variable, CombatEffectiveness, was created as a sum of selected attributes, showcasing the ability to manipulate and engineer features within a dataset.



**Task 3: Data Export and Subset Analysis**

Objective: Practice data exporting techniques and analyze a specific subset of the data.

Approach: The dataset was filtered to include only avengers without superpowers who had died, then exported in both SPSS and CSV formats. Descriptive statistics were calculated for this subset to understand the properties of combat effectiveness, kills, and injuries.

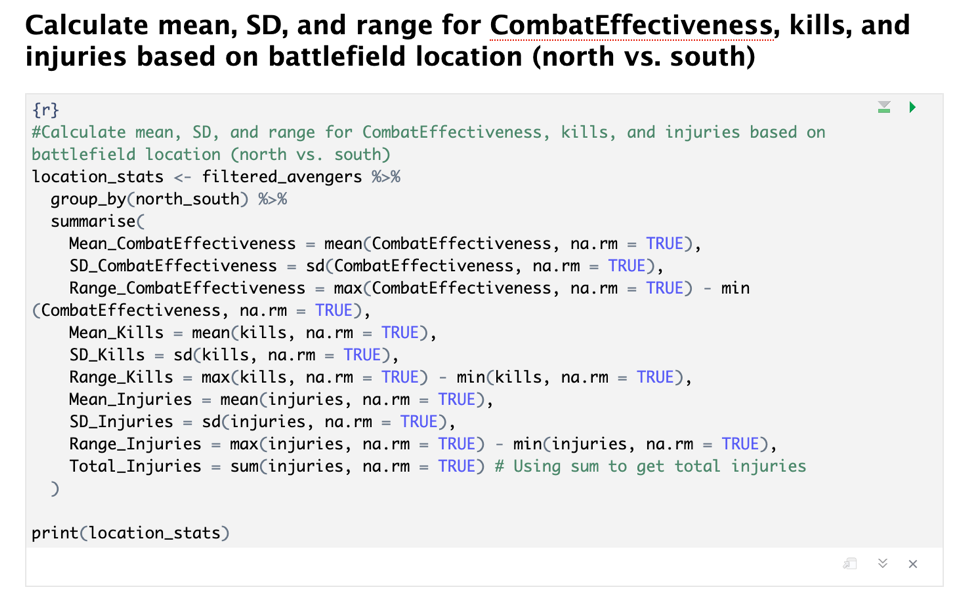


****

**Task 4: Combat and Injury Analysis**

Objective: Analyze the dataset to determine combat effectiveness, kills, and injuries based on battlefield location.

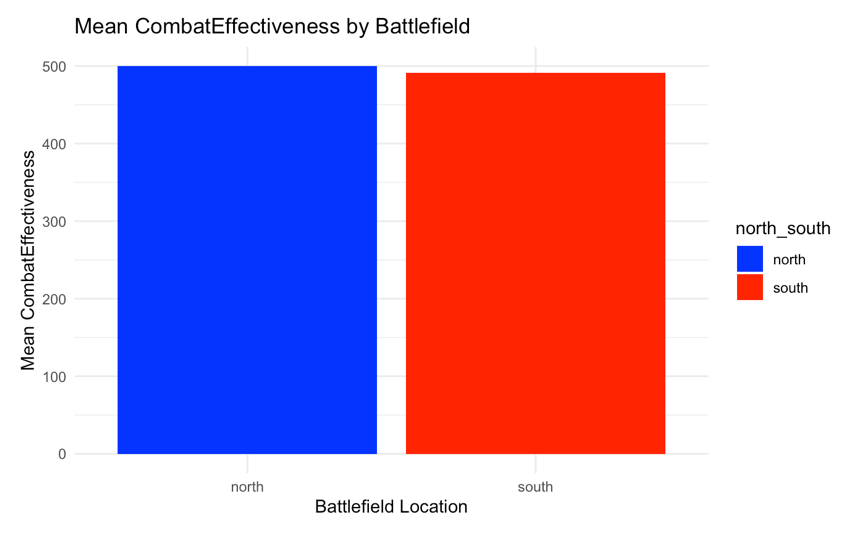
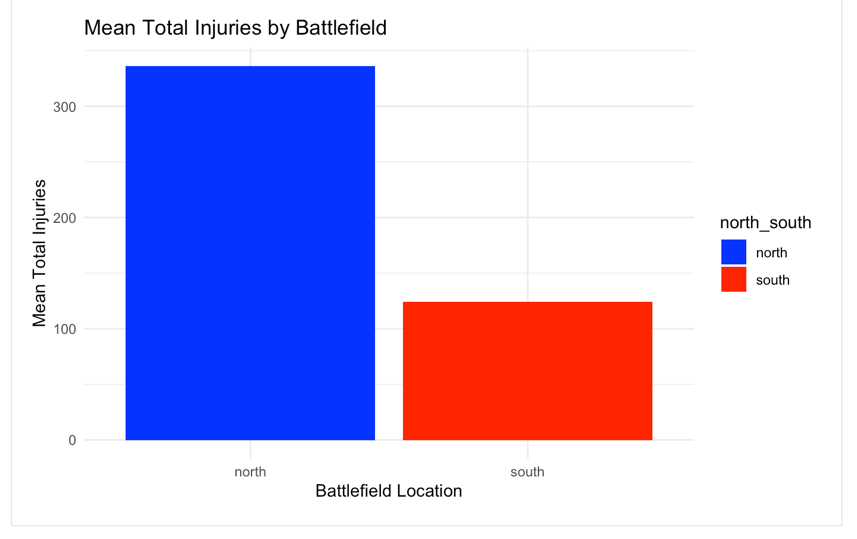
Approach: Comparative analysis was conducted to identify which battlefield was most effective in combat and which had the most injuries, demonstrating the ability to derive insights from the data.





From the last code we can understand that north battlefields were the most effective one and also, they had the most injuries. Also, I add extra section for just plotting them to make it easy to follow.





**Task 5: Identifying the Most Erroneous Variable**

Objective

Determine which variable among combat effectiveness, kills, and injuries demonstrates the highest level of error or variability in the mean model. This assessment helps understand the stability and reliability of these metrics in reflecting the Avengers' performance.

Approach

To determine which of the three variables—CombatEffectiveness, kills, and injuries—is the most erroneous in terms of contributing to the mean model's error, I would typically look at measures such as the standard deviation or the variance of these variables. These measures give us an idea of the spread or dispersion of the variable values around their mean, with a larger value indicating more variability or error from the mean.

However, without directly computing a specific model, such as a regression model where we could assess the error term or the residuals for each predictor variable, we're somewhat limited to basic statistical summaries to infer "error" or variability.

Assuming we're working with basic descriptive statistics, we can calculate the coefficient of variation (CV) for each variable as a measure of relative variability. The CV is defined as the standard deviation divided by the mean of the variable, and it allows us to compare the variability of variables that are on different scales. A higher CV indicates a higher level of dispersion relative to the mean, which might be interpreted as "more erroneous" in a loose sense.



**Task 6: Hypothesis, Sample Size Justification, and Preliminary Analysis**

Objective: Formulate a hypothesis about the relationship between superpowers and IQ, justify the needed sample size for the study, and conduct a preliminary analysis by creating a new data frame.

Approach:

**Sample Size Justification:** To estimate the required sample size for our study, which aims to compare the IQ levels of Avengers with and without superpowers using an independent samples t-test, we face an unconventional situation due to the dataset's fictional nature. In this context, traditional empirical justifications for sample size estimation, typically grounded in literature review or prior studies, are not applicable. Therefore, the most straightforward justification is "**no justification**," acknowledging the dataset's hypothetical basis.

However, for educational purposes and to illustrate standard statistical practices, we consider two common justifications: the smallest size of effect of interest (SSOEI) and statistical power. Focusing on SSOEI aligns with our interest in detecting even small yet meaningful differences between groups. Additionally, examining statistical power ensures our study is designed with a high probability of detecting an effect if it exists, crucial for avoiding type II errors.

This approach demonstrates a methodical way to sample size estimation that, while acknowledging the fictional scenario, still applies key statistical concepts. This not only illustrates the application of these principles in an educational setting but also highlights the adaptability of statistical reasoning to both empirical and theoretical research scenarios.

**1. Smallest Size of Effect of Interest (SSOEI)**

The concept of SSOEI refers to the smallest effect size that is practically significant for the research or study at hand. In other words, it's the minimum difference between groups (in this case, Avengers with superpowers vs. those without regarding their IQ scores) that we would consider meaningful or worth detecting. This could be based on previous research, theoretical considerations, or practical implications.

* 1. **Why SSOEI is important**: Estimating SSOEI is crucial because it helps in determining the required sample size to reliably detect an effect of that size or larger. If the effect size we're interested in is small, we'll need a larger sample size to detect it with sufficient statistical power. Conversely, for larger effects, a smaller sample might suffice.
  2. **How to use SSOEI for sample size estimation**: To estimate the sample size using SSOEI, I would first decide on the minimum effect size that I can consider meaningful. This could be a specific difference in IQ points between the two groups. Then, using R or sample size calculation formulas that incorporate effect size, desired statistical power, and alpha level (commonly set at 0.05 for a 5% chance of Type I error), I can estimate the necessary sample size.

**2. Statistical Power**

Statistical power is the probability that my study will detect an effect if there is one to be detected (i.e., the probability of not making a Type II error). A commonly accepted power level is 0.80, meaning there's an 80% chance of detecting an effect if it exists.

2.1. **Why statistical power is important**: Power analysis helps in determining the required sample size to achieve a certain level of confidence in the study's findings. A study with low statistical power might fail to detect a true effect, leading to false conclusions about the absence of a relationship.

2.2. **How to use statistical power for sample size estimation**: I will need to decide on the desired level of power (e.g., 0.80), the significance level (alpha, commonly 0.05), and the effect size (which could be informed by SSOEI). With these parameters, I can use power analysis tools or software to estimate the sample size needed (in our case by using R package).

**Choice Between SSOEI and Statistical Power**

Both SSOEI and statistical power are critical considerations for sample size estimation, and they are interrelated. SSOEI helps define what constitutes a meaningful effect, guiding the choice of effect size in the power analysis. Therefore, a comprehensive approach to estimating the required sample size would consider both the smallest effect size of interest to ensure the study is practical and meaningful, and the desired statistical power to ensure the study has a high probability of detecting that effect but by considering that there are no previous studies on this topic the priority is to focus on power.

**Preliminary Analysis:**

A new data frame was created to separate Avengers based on the presence or absence of superpowers. This step demonstrated competency in data manipulation techniques, facilitating targeted analyses. Descriptive statistics, specifically mean, SD, and also SEM (as the number of samples in each group are not equal) of IQ for each group (Avengers with and without superpowers), were computed. This provided a quantitative basis for understanding the distribution and central tendencies of the IQ scores within the dataset, showcasing the application of descriptive statistical methods to summarize data.



**Task 7: Hypothetical Effect Size Determination**

Objective: Learn how to choose and justify a hypothetical effect size for a planned study.

Approach: For a study investigating the relationship between having superpowers and IQ among Avengers, let's choose a hypothetical effect size. We'll say that we're looking for a medium effect size, quantified as a Cohen's d of 0.5. This effect size suggests that the mean IQ of Avengers with superpowers is 0.5 standard deviations higher than the mean IQ of those without superpowers.

Justification for Choosing a Medium Effect Size (Cohen's d = 0.5)

**Previous Research and Literature:** In a real study, the first step to choosing an estimated effect size would be to review existing literature on related topics. If previous studies have investigated similar variables (e.g., cognitive abilities and superhuman traits), those findings can inform a reasonable expectation for the effect size. In the absence of directly comparable studies, a medium effect size is a prudent assumption that balances detectability with realistic expectations, especially when exploring new or complex phenomena.

**Theoretical Considerations:** The theoretical framework underpinning the study might suggest that superpowers could confer significant cognitive advantages due to the necessity of managing complex tasks, strategic thinking in battles, or enhanced neurological development associated with superhuman abilities. However, recognizing that superpowers are diverse and not all directly linked to cognitive enhancement, a medium effect size assumes a balanced view that acknowledges significant yet not universal cognitive superiority among superpowered Avengers.

**Practical Implications:** Choosing a medium effect size reflects a consideration of the practical implications of the study's findings. Detecting a medium effect size would suggest that the difference in IQ between Avengers with and without superpowers is substantial enough to warrant attention in strategic planning, training, and deployment in battle scenarios, without overestimating the cognitive impact of superpowers.

**Statistical and Methodological Feasibility:** A medium effect size strikes a balance between the need for a large enough sample to detect smaller effects and the practical limitations of data collection in specialized populations, such as Avengers. It represents a realistic goal that can be achieved with a feasible sample size, ensuring the study's design is both statistically robust and practically achievable.

**Task 8: Power Analysis Execution**

Objective: Perform power analysis using the predetermined effect size to ensure the study's design is capable of detecting the intended effect.

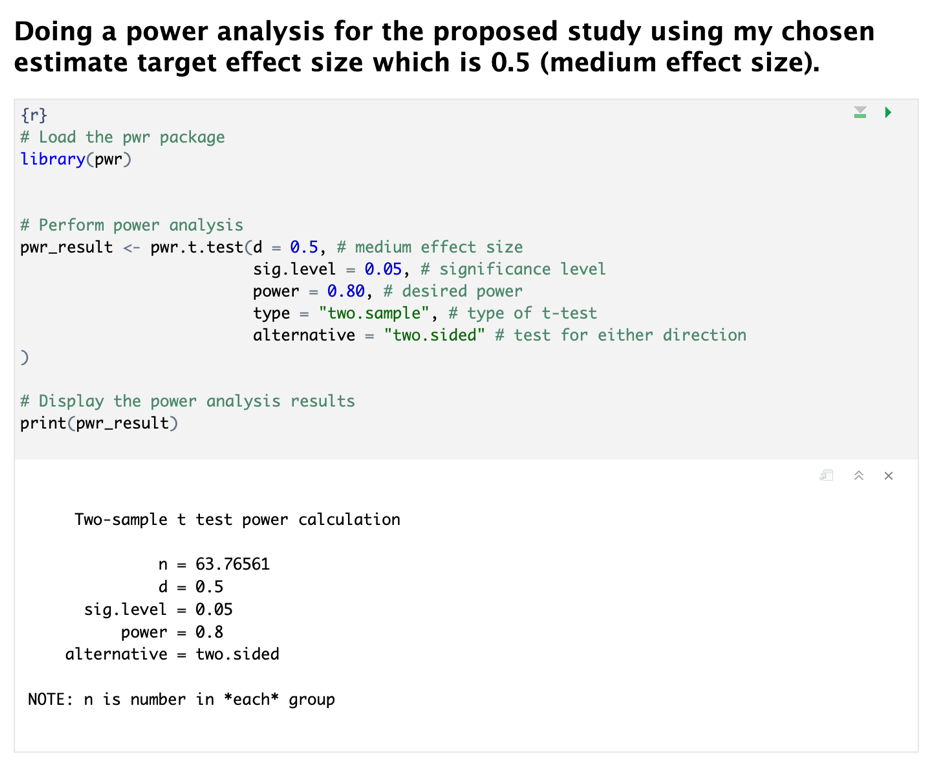
Approach: To conduct a power analysis for our proposed study, examining the relationship between having superpowers and IQ among Avengers, with a chosen hypothetical medium effect size of Cohen's d = 0.5, we'll specify several parameters: the effect size, significance level (alpha), power (1 - beta), and the allocation ratio of the sample sizes across groups. For this demonstration, I'll use R's pwr package, which provides functions for conducting power analysis across different study designs, including t-tests.

**Parameters for Power Analysis**

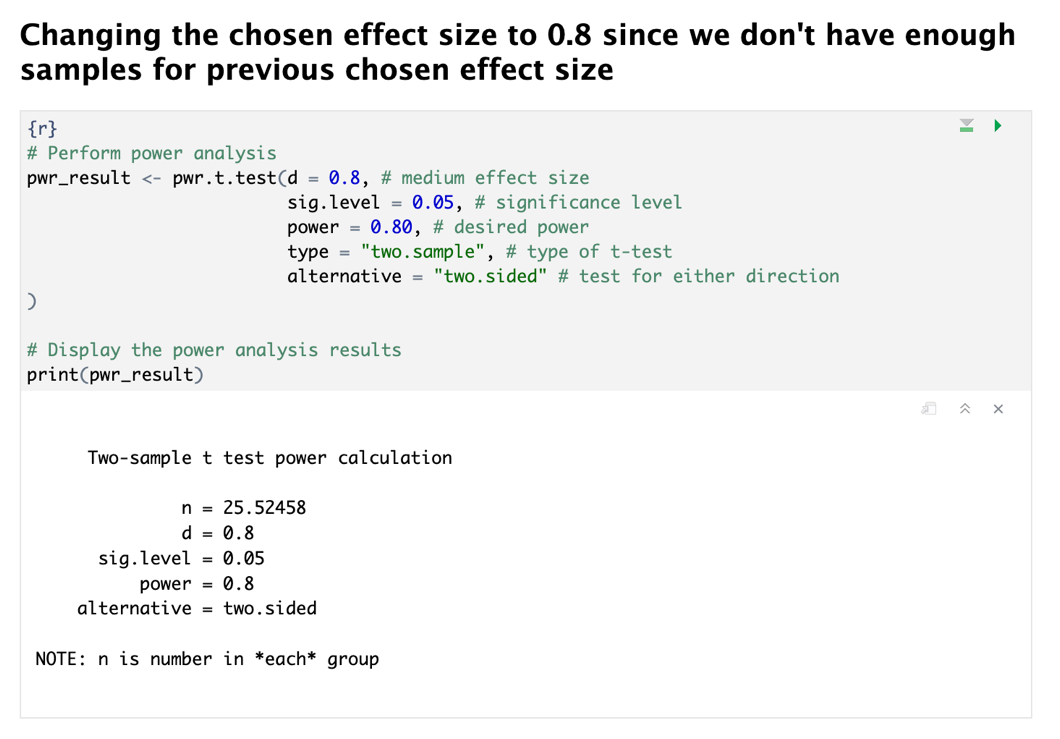
Effect Size (Cohen's d = 0.5): Chosen based on the rationale that the difference in IQ between Avengers with and without superpowers is expected to be medium. This size is sensible for initial investigations where theoretical and empirical evidence suggests a noticeable but not extreme difference.

Significance Level (Alpha = 0.05): A conventional choice for psychological and social sciences research, reflecting a 5% willingness to accept a Type I error, where we might wrongly conclude that there is an effect (difference in IQ) when there is none.

Power (1 - Beta = 0.80): This is the probability of correctly rejecting the null hypothesis when it is false (i.e., detecting an effect if there is one). A power of 0.80 means there's an 80% chance of detecting the effect size of 0.5 if it truly exists. This level is commonly used in research to balance between the risk of Type II errors and the feasibility of obtaining large sample sizes.



From the result we can see that for that justified condition I need 64 samples per group which is not feasible according to my data set (number of avengers with superpower = 32), after changing the chosen effect size from 0.5 to 0.8 which is large effect size and empirically is not acceptable for some of the studies, we can see that, we have enough sample for updated effect size:



**Task 9: Zero Effect Power Estimation**

Objective: Explore methods to estimate the power required to confirm the absence of an effect, enhancing understanding of null hypothesis significance testing.

Approach: The concept of estimating power to detect a zero effect was discussed, explaining how to adjust study parameters to ensure sufficient power to confidently claim no difference exists between groups but at the same time the objective of this task is ambiguous at least for me! Since, I didn't get the point that for this section whether the power estimation required or estimating enough sample size to do that?! So, I did this twice once for estimating proper sample size and once for power value.

**First approach**  
For Task 9, focusing on confirming the absence of a meaningful difference between groups in a fictional Avengers dataset, the choice of a +/- 1 IQ point equivalence margin is justified by a blend of narrative suitability, statistical precision, and educational value:  
  
**a. Narrative Suitability**

Superhuman Context: In the Avengers universe, where characters possess extraordinary abilities, a +/- 1 point margin reflects the narrative expectation that superhuman intelligence might exhibit minimal variability, emphasizing the exceptional nature of the subjects.

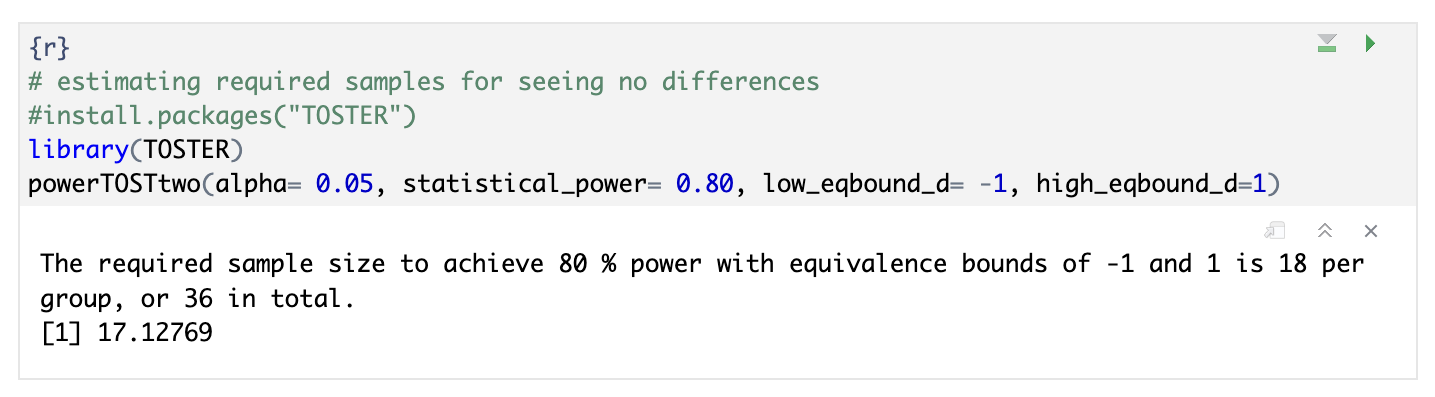
**b. Statistical Precision**

Stringent Testing: This tight margin demonstrates a high level of statistical rigor, aiming to detect only the most minimal, yet meaningful differences. It aligns with the objective of confirming a "zero effect" with robust statistical support, ensuring confidence in the null hypothesis validation.

**c. Educational Value**

Learning Opportunity: Selecting such a narrow margin provides an educational exploration of equivalence testing under conditions demanding extreme precision. It illustrates the impact of margin selection on study design, power analysis, and result interpretation, deepening understanding of statistical methodologies.

In conclusion, opting for a +/- 1 IQ point margin justifies the unique requirements of the fictional dataset, showcasing the need for precision in statistical analysis and offering a rich context for educational exploration in equivalence testing. This approach effectively balances narrative context with the demands of statistical rigor and learning objectives.



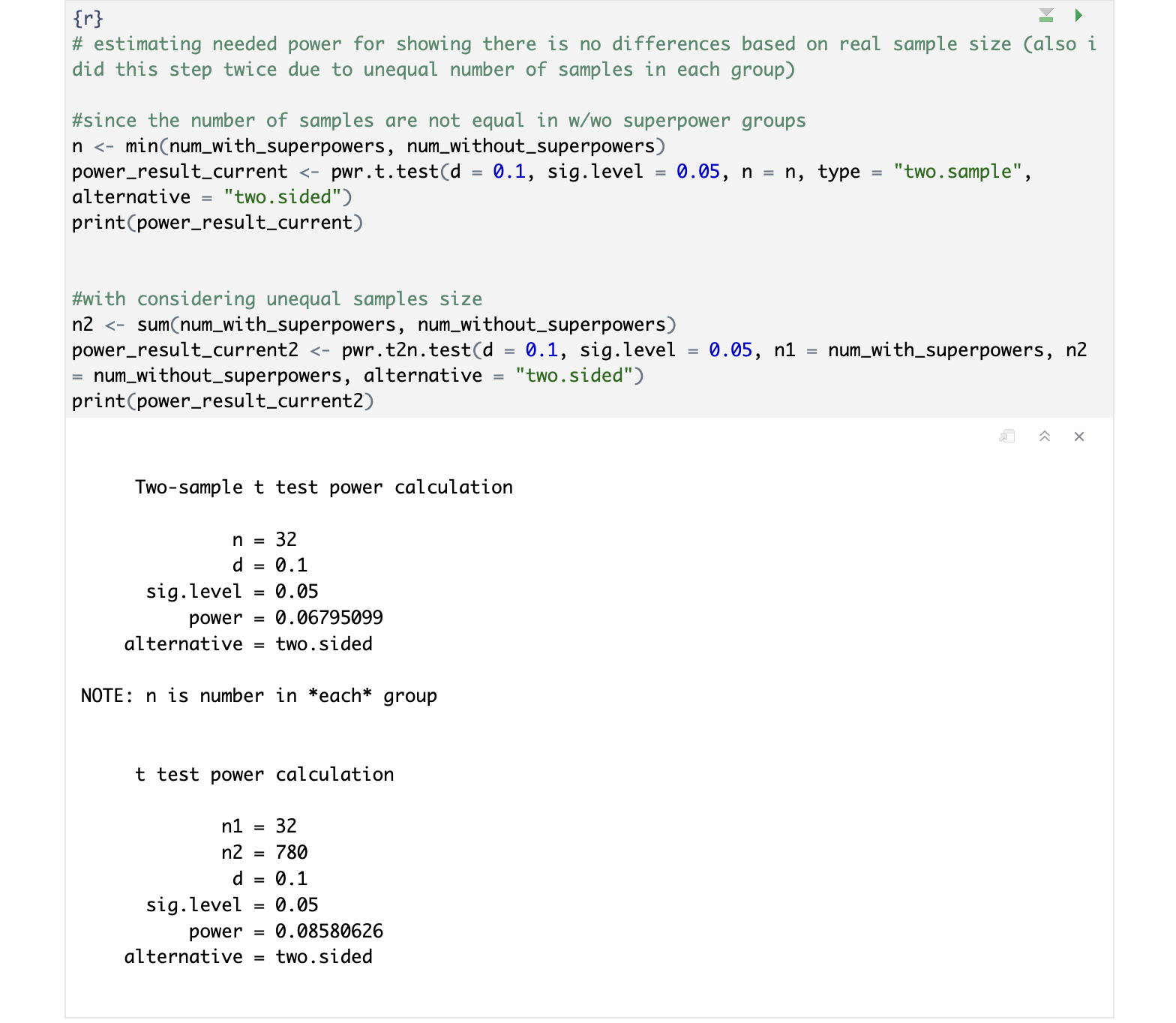
**Second approach**

In tackling Task 9, which aims to estimate the statistical power necessary to demonstrate the absence of significant differences between groups in a fictional dataset, an alternative approach is adopted to accommodate the practical challenges of unequal sample sizes between Avengers with and without superpowers. This nuanced methodology reflects the complexities encountered in real-world data analysis, where unequal group sizes can significantly impact statistical power and the interpretability of results. Here’s how this approach is pragmatically justified and executed:

**a. Adjusting for Unequal Sample Sizes**

First Calculation (Equalized Sample Size): Given the reality of unequal samples between the groups, the initial step pragmatically utilizes the size of the smaller group (n) for both, aiming for a conservative estimate of the study's power. This simplification allows for a straightforward assessment under the assumption of equal group sizes, using standard parameters for a two-sided test to detect minimal practical differences.

Second Calculation (Real Sample Sizes): Acknowledging the true unequal sample sizes, the analysis progresses to incorporate the actual numbers (n1 and n2) using pwr.t2n.test. This function is tailored for situations with disparate group sizes, offering a more precise estimation of the study's power reflective of the dataset's specific characteristics.



Rationale for the Approach

Realistic Scenario Modeling: This dual-step analysis method provides a comprehensive view of how sample size differences affect the study's capability to assert the non-existence of significant differences. It demonstrates adaptability in statistical methodology, aligning with Task 9's objective by illustrating various approaches to power calculation in the face of dataset idiosyncrasies.

Educational Value: Through this approach, the procedure serves as an instructional showcase on handling unequal sample sizes in statistical analyses. It underscores the importance of aligning statistical planning with the dataset's nuances and introduces the pwr package's functionalities for varied research needs.

Statistical Prudence: By choosing a conservative effect size and a standard significance level, the approach balances rigor with common research conventions, ensuring that the analysis remains grounded in established statistical practices despite the fictional dataset's unique context.

This introduction and methodology provide a detailed roadmap for navigating Task 9's requirements, emphasizing practical statistical solutions and educational insights into power analysis under conditions of unequal sample sizes. However, at the end I’m still not sure about the way I’m thinking and I hope all of this explanation and codes be correct!

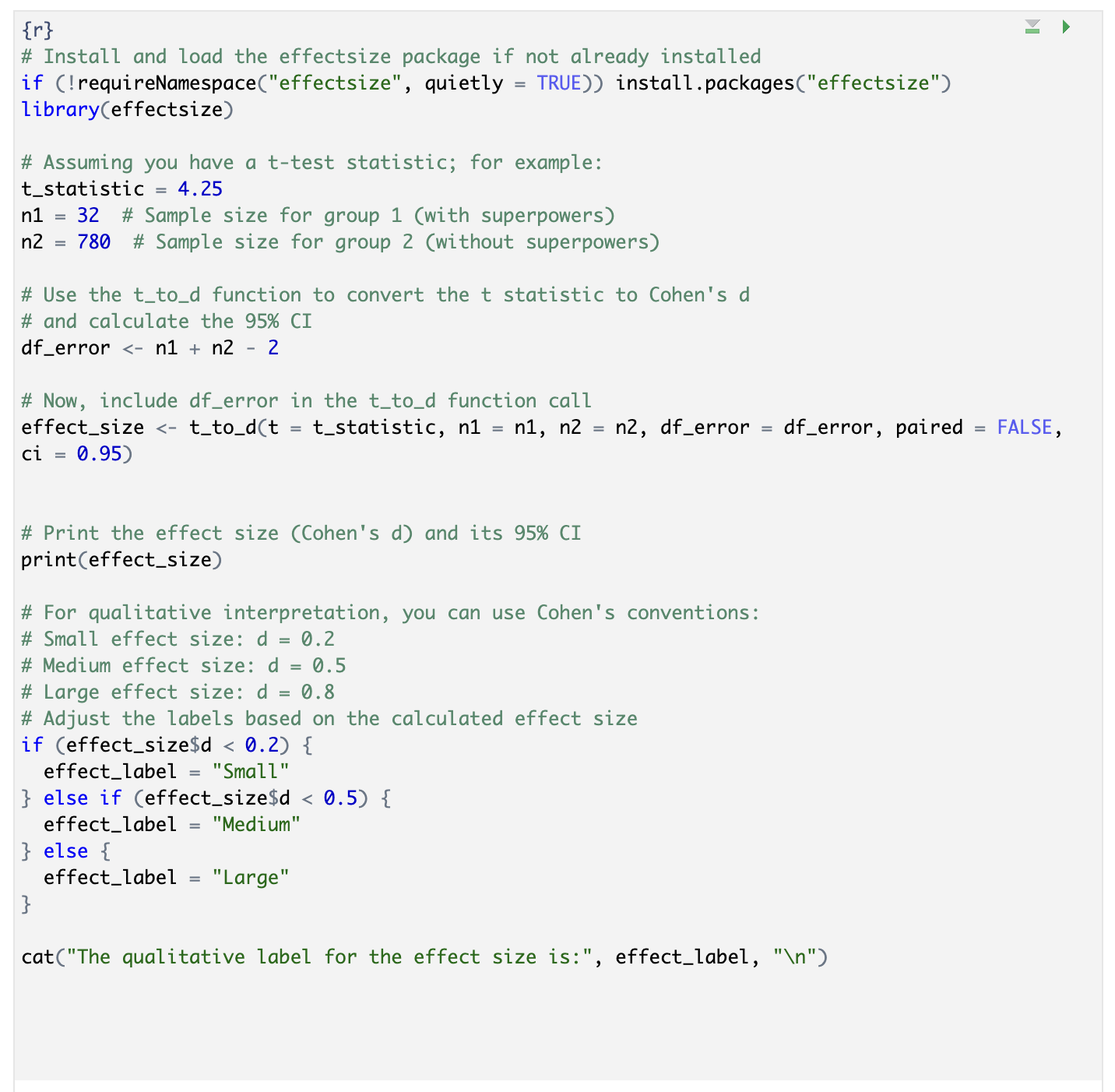
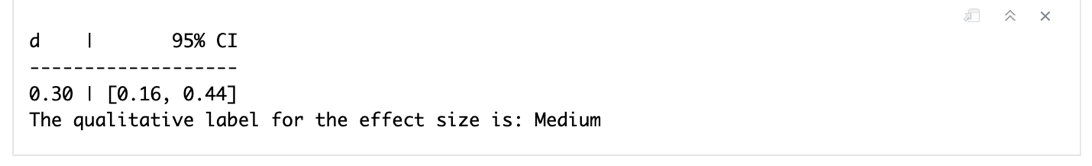
**Task 10: Calculating and Interpreting Effect Size from Test Statistic**

Objective

The objective of Task 10 is to translate a given t-test statistic into a meaningful effect size measure, Cohen's d, including its 95% confidence interval (CI), to assess the magnitude and precision of the difference between two groups. This task further involves evaluating the precision of the effect size estimate and categorizing its magnitude using Cohen's conventions.

Approach

I used ‘effectsize package’ for calculating Cohen’s d and CI directly from t statistic:



**Precision Statement**

The Cohen's d effect size of 0.30, with a 95% confidence interval ranging from 0.16 to 0.44, indicates a medium magnitude of effect between the two groups. The confidence interval's range is relatively narrow, suggesting a moderate level of precision in the effect size estimate.

**Precision Explanation**

Moderate Precision: The confidence interval's span (from 0.16 to 0.44) is relatively narrow, indicating that the estimated effect size of 0.30 is determined with a reasonable degree of precision. This means there's moderate confidence that the true effect size lies within this interval, minimizing the risk of large errors in the effect size estimation.

Implications for Research: The moderate precision of this estimate suggests that the sample provides a reliable indication of the difference in IQ scores between Avengers with and without superpowers. It reflects well on the study's design and sample size, indicating that they were adequate to capture the true effect size with a reasonable degree of certainty.

Future Research Directions: While the precision is moderate, future studies could aim for narrower confidence intervals by increasing the sample size or reducing data variability. This could further refine the understanding of the effect size and its implications.