

**DATA WRANGLING & RETRIEVAL**

# **HEROES & VILLAINS**

**FIRST PROJECT IRONHACK:**  
**ÉLIO VIEIRA**

## PROJECT OVERVIEW

My data source contains a list of over 700 heroes, villains and fictional characters with a variety of information and attributes, which I'll explore next.

My project is divided into 3 hypotheses and opportunities, and I chose this theme because I'm a fan of hero and science fiction movies and series 😎

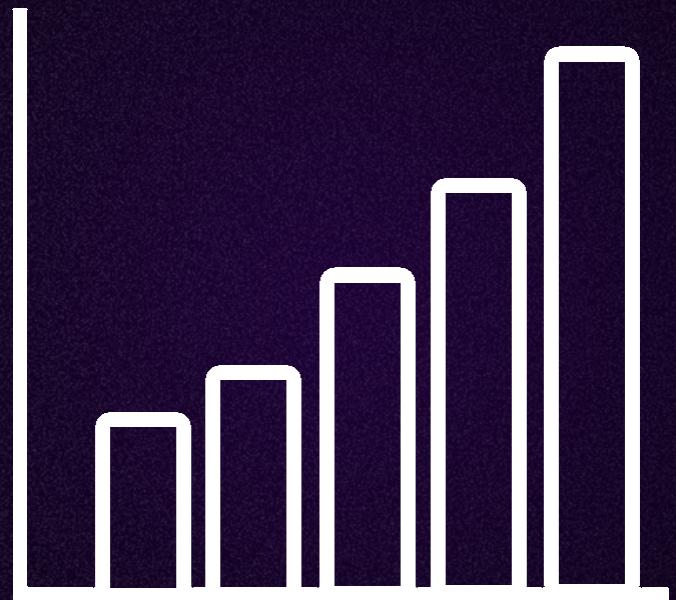
### SOURCES:



SUPERHERO API



SUPERHERO GLOSSARY



# DATA WRANGLING & CLEANING

- DROP COLUMNS
- CHECK NAN VALUES & REPLACE OR DROP
- CHECK DUPLICATES
- FORMAT HEIGHT AND WEIGHT TO KG AND CM
- FORMAT TEXT AND REPLACE NAMES
- CREATE MAPPING TO CHANGE NAMES
- SET INDEX



# DATA FRAME AFTER WRANGLING & CLEANING

	Name	Image	PS Intelligence	PS Strength	Gender	Race	Height	Weight	Full Name	First Appearance	Publisher	Alignment
id												
156	Captain Marvel		88	100	Male	Human	193 Cm	101 Kg	Billy Batson	Whiz Comics No. 2 (1940)	Dc Comics	
496	Nova		100	85	Male	Human	185 Cm	86 Kg	Richard Rider	Nova Vol. 1 #1 (september, 1976)	Marvel Comics	
620	Spider-man		90	55	75	74	85	446	Male Human	178 Cm	74 Kg	Peter Parker Amazing Fantasy #15

# HYPOTHESES & OPPORTUNITIES



**Hypothesis:** Provide valuable insights to create a superhero-themed game. Based on the insights gathered from the dataset, brainstorm potential game mechanics and character stats that leverage the diverse range of superhero attributes and characteristics.

## 1. Game Creation



**Hypothesis:** Analyzing data on the number of heroes per universe (studio) and first appearances can provide valuable information for optimizing the planning and organization of a successful Comic Con.

## 2. Comic Con Planning



**Hypothesis (fictional):** The selection of comic book characters for a defense team will be optimized based on a combination of power scores, alignment, and affiliation to effectively combat the alien invasion.

## 3. Earth's Last Defense

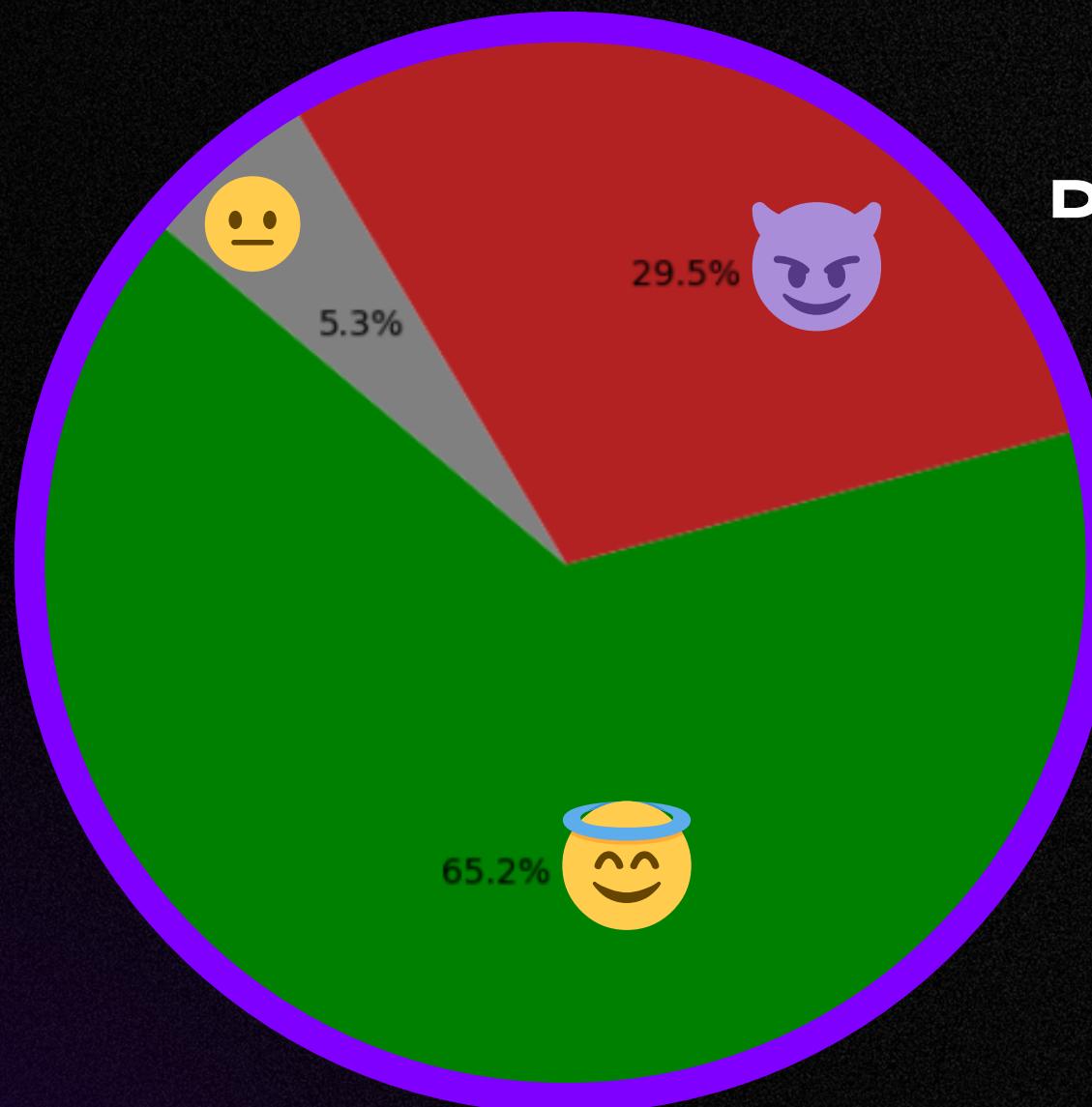


## 1. Game Creation

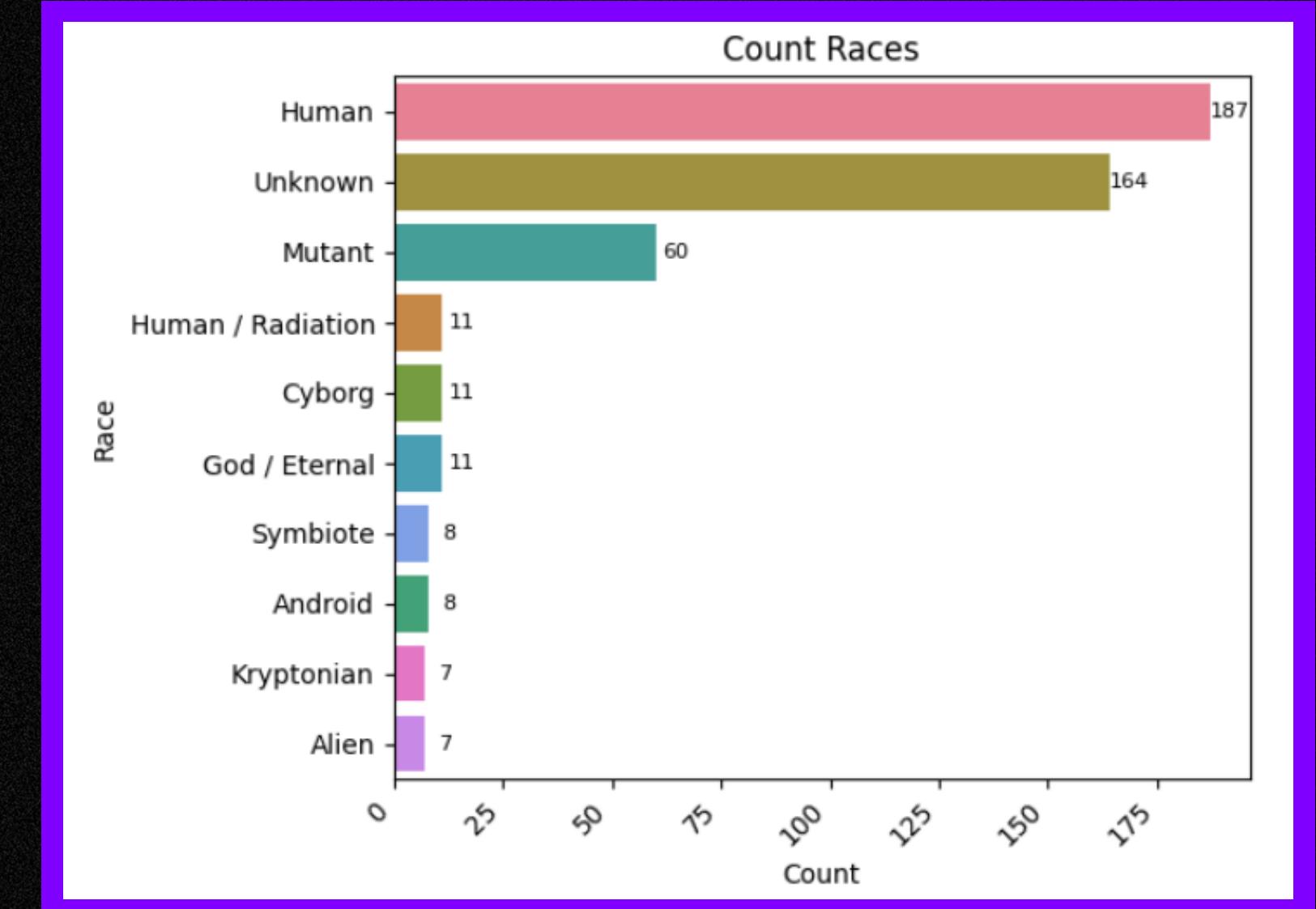
## HYPOTHESES & OPPORTUNITIES

### Expected Outcome:

By analyzing the available superhero dataset, developers can gain insights into the distribution and characteristics of various superhero attributes. This analysis can inform the design and development of a superhero-themed game.



### Distribution of Alignment



With access to PowerStats, alignment (Evil, Good or Neutral), Race and other data, we can create a game that incorporates all the fictional characters from several different universes.

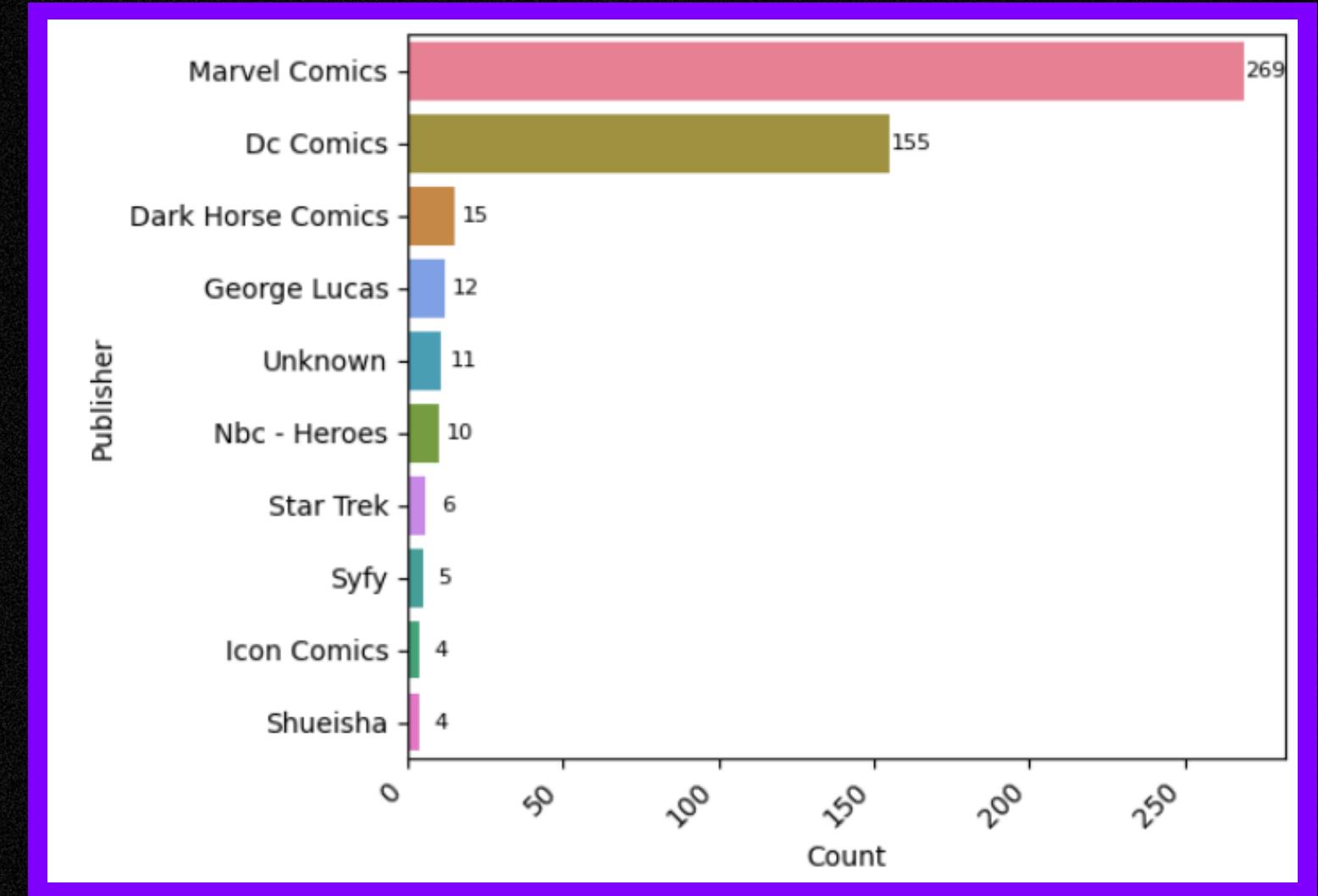
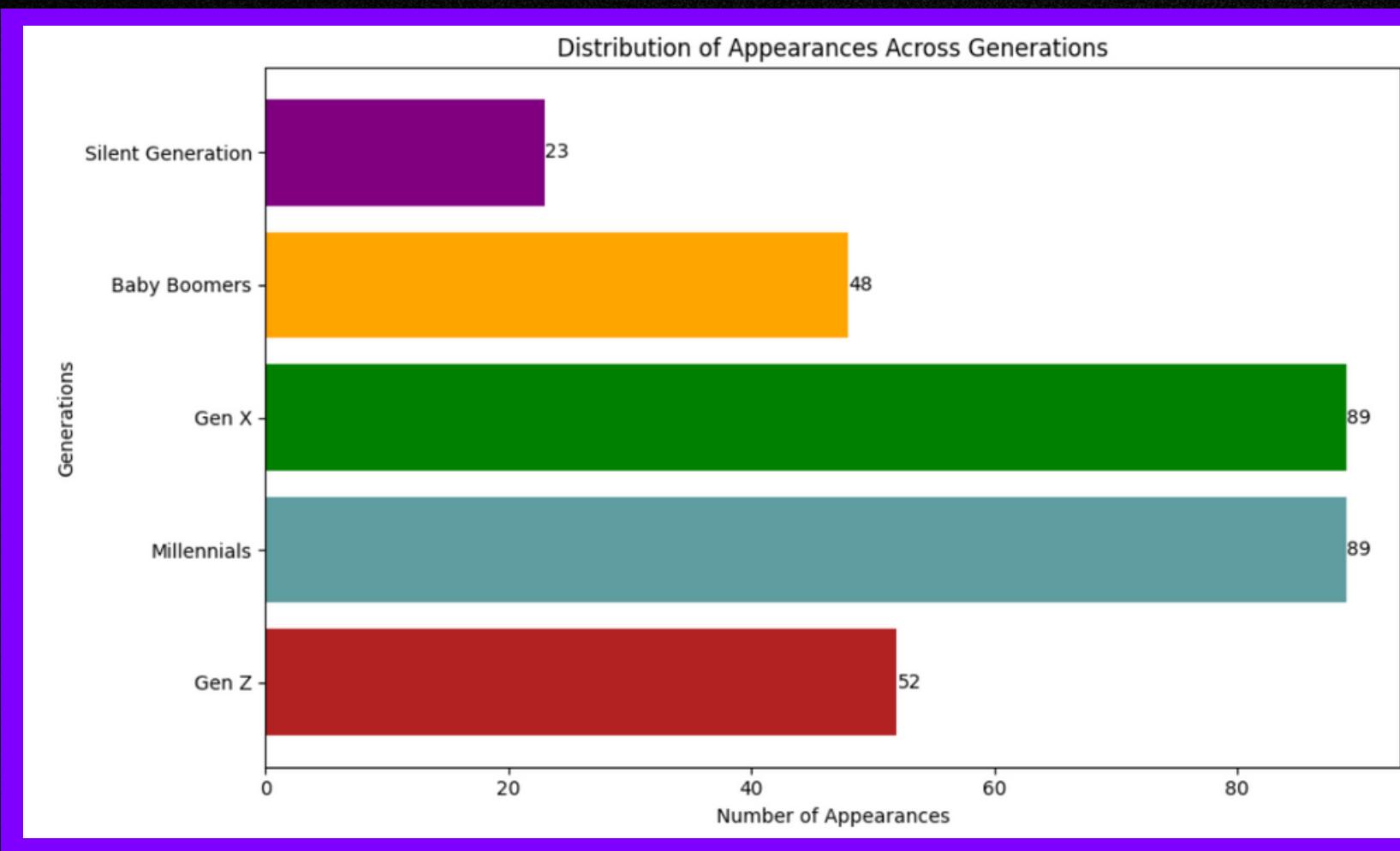


## 2. Comic Con Planning

## HYPOTHESES & OPPORTUNITIES

### Expected Outcome:

By leveraging data-driven insights, organizers can tailor the comicon experience to better meet the preferences and expectations of attendees, leading to increased attendance, engagement, and overall satisfaction with the event.



By analyzing the number of characters per publisher, and characterizing them by generation. It is possible to better organize and distribute the layout of the exhibition.

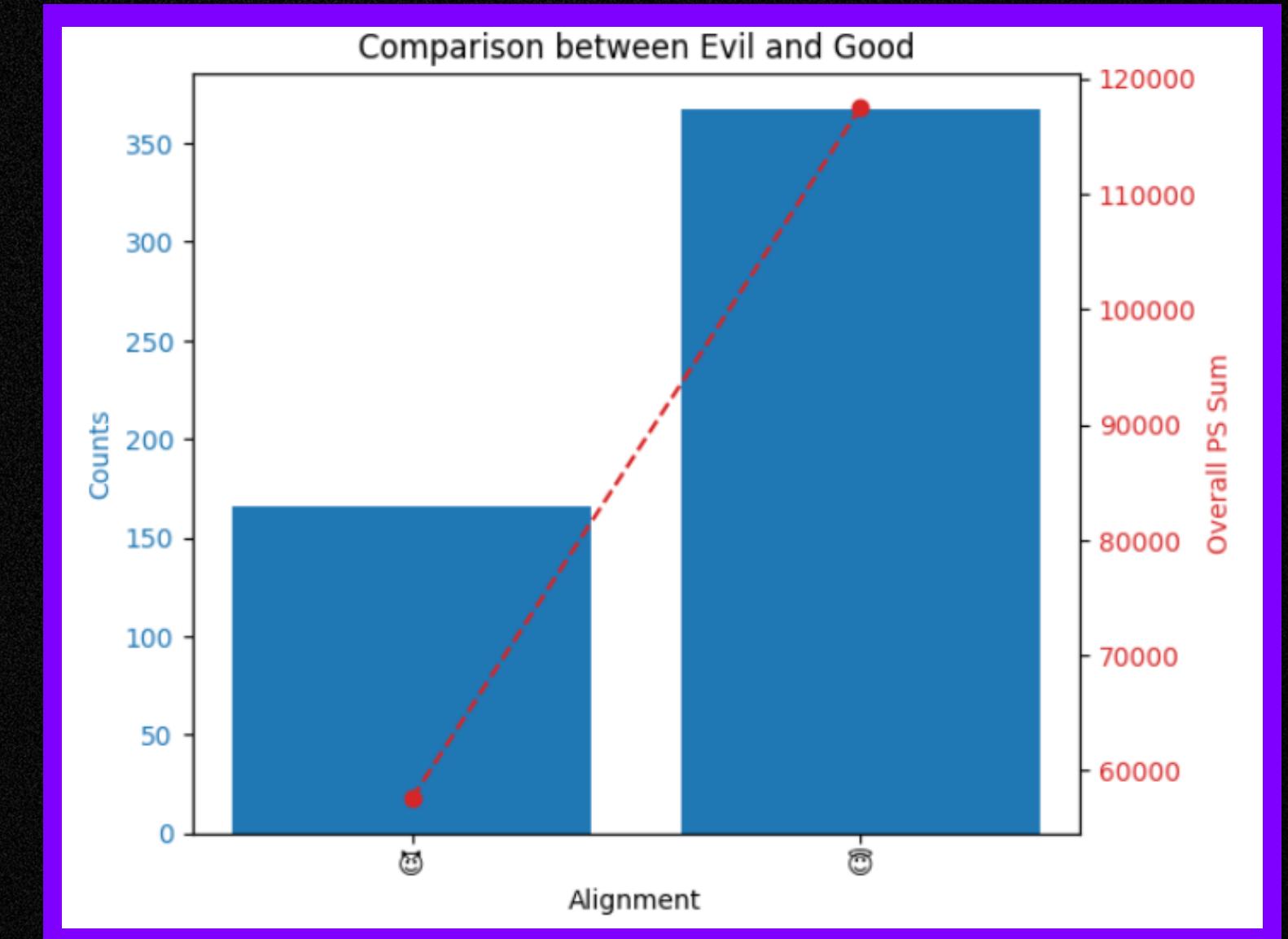
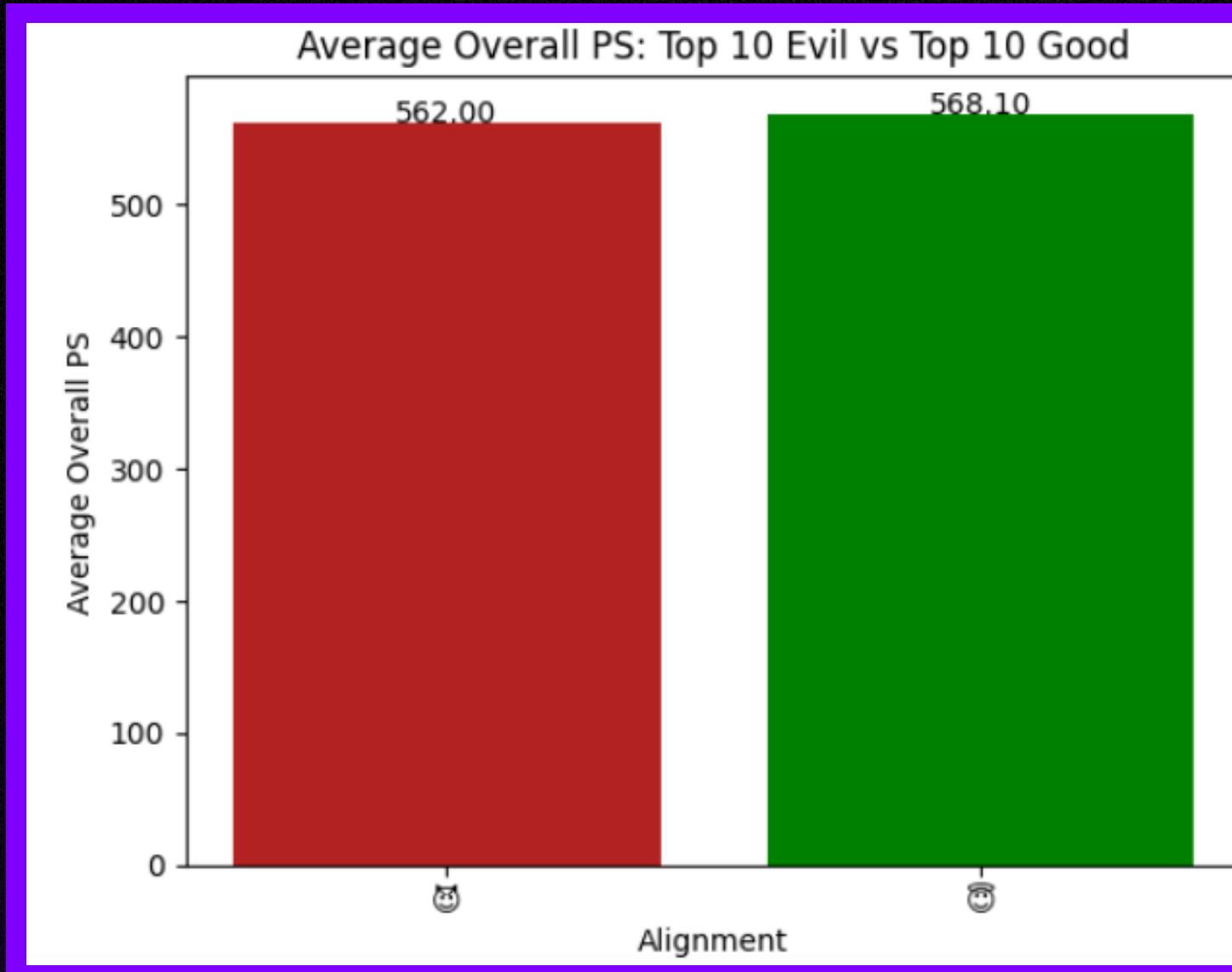


### 3. Earth's Last Defense

HYPOTHESES &  
OPPORTUNITIES

#### Expected Outcome:

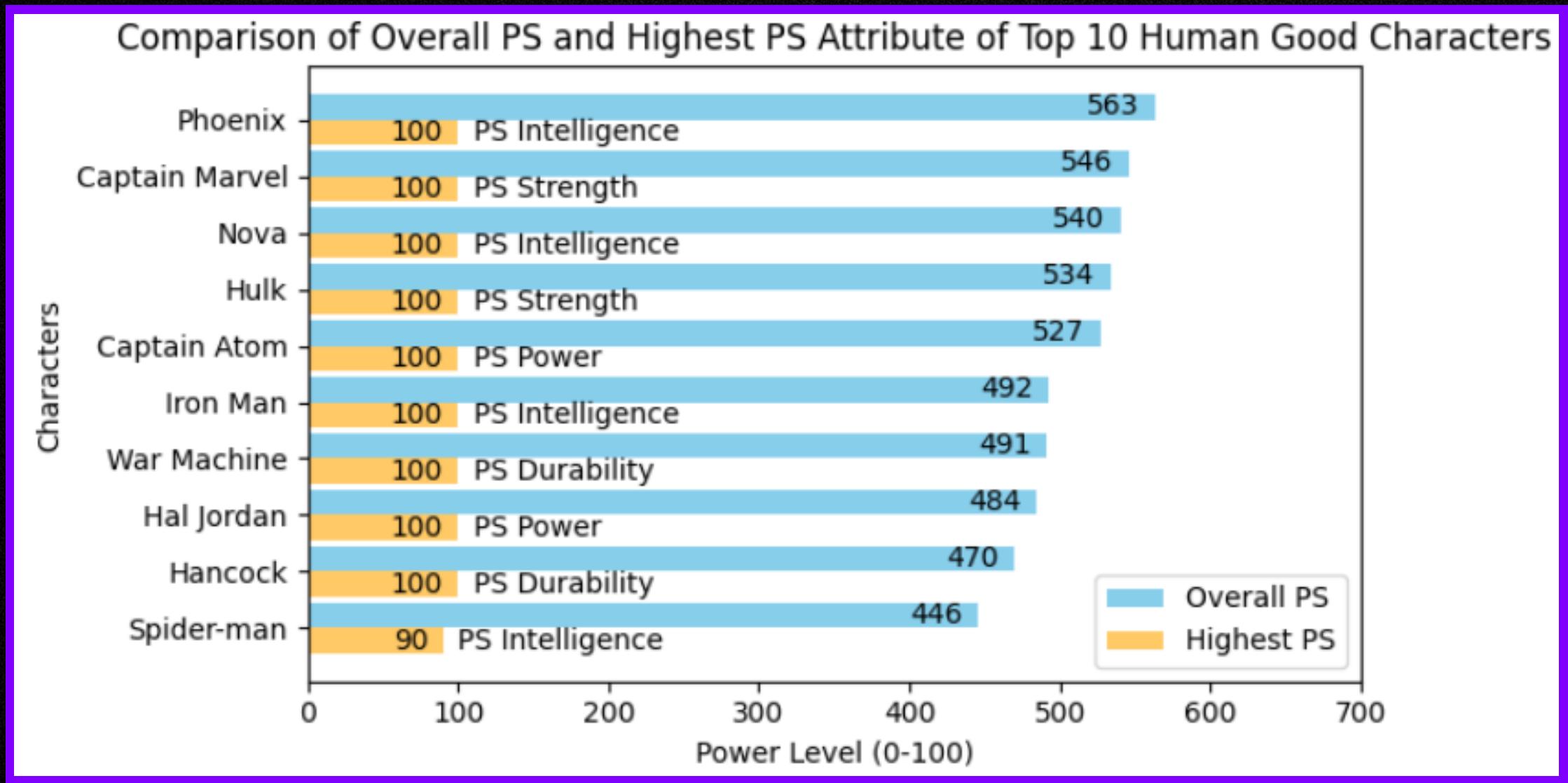
By forming a defense team with a strategic mix of powerful and diverse characters, the team will have a higher chance of success in repelling the alien invasion, as demonstrated through simulations or hypothetical scenarios.





### 3. Earth's Last Defense

## HYPOTHESES & OPPORTUNITIES



#### Race:

- Human
- Human / Radiation
- Mutant

#### Alignment:

- 😊



Analyze the dataset to identify characters with exceptionally high power scores across various attributes (e.g., Intelligence, Strength, Combat). Additionally, consider characters with diverse alignments and affiliations to ensure a balanced team composition.

AND BECAUSE I KNOW YOU WERE  
WONDERING ABOUT THE VARIETY OF THIS  
DATASET, HERE IT IS!

**THANK YOU**

