

**How difficult was it to find a dataset**

Choosing the dataset was thankfully a straightforward affair. The group, during a Blackboard collaborate meeting, visited kaggle.com and spotted the Marvel Snap Decks dataset on the front page while perusing through all the available content. It included an eye-catching image and focused on information that we were all at least passingly familiar with, if not directly interested in, which made it a good choice for the assignment. We looked at a few other data sets but none of them caught our attention in the same way, so we reached an agreement that this was the one we wanted. This was a quick process that culminated in a unanimous decision.

**What challenges did you encounter on analyzing the data**

From a human analysis perspective, there weren't a lot of problems that we encountered with the data. It's clearly labeled, and the meaning of each column is self-evident, which datasets should be. From this side, it was easy to isolate the specific aspects of the dataset we wanted to use for the machine analysis portion. The real challenge came with the machine analysis portion, however, primarily in wrangling the code into doing what we wanted it to. There were several unforeseen issues when it came to writing the script which we were initially unprepared to answer, including several unknown errors, whose fixes were a little more inexplicable than we would have liked. One error, for example, was solved by appending 'numeric only' despite the data for that section consisting only of integers. Fortunately, we were able to draw on outside knowledge to solve the problems, which left the code and our analysis working as intended. The other issue that we encountered during this analysis was a result of assigning different portions of it to different people, so steps needed to be taken to ensure that there was a more unified naming scheme so that everything would work the way we expected. These issues were resolved promptly which resulted in the completion of everyone's assigned tasks with ample time for review by the group.

**What is the data showing you? Break this into "end-user" terminology**

The data shows information about the publicly accessible decks made by players of Marvel Snap, which is an online card battler, and sourced from the associated Marvel

Snapzone website. The data set consisted of the name of the user who created the deck, the name of the deck itself, the URL where the specific deck list is located, the associated card IDs collected into arrays, the average power of the cards in the deck, the average cost of each of the cards, the total likes that the given deck has received, and the last time the deck was updated. This provides a quick reference point for the relevant information that goes into creating the decks.