Final Project

Ambrose M Malagon Cordero

Bellevue University

DSC 540 – Data Preparation

Saturday February 29, 2020

**Marvel Comics API.**

For this term’s Final Project, I chose to pursue data from the Marvel Comics API (<https://developer.marvel.com/>). Registering for an API Access Key was easy and their online documentation is very solid; it even allows you to set up the queries and test them out via the website. There were a couple of things that were not explicitly called out on the website, as such the how to create the hash identifier, so I had to do some separate research on the side for that.

The next step was to identify which Python 3 modules I needed. For this project, I chose to use the following:

* Requests – To make the request to the API.
* Hashlib – In order to create the hash ID needed to make the call.
* Time – Needed to create the timestamp.
* Re – For regex purposes.

My first attempts at pulling the data did not go as smoothly as I would’ve wanted. I ended up losing an entire day trying to figure out what the problem was. As it turned out, the problem ended up being with the query I was trying to use. For some unexplained reason, the character name query name would continuously error. After some troubleshooting and changing the query to use Character ID’s, I managed to get a successful data pull.

Once I got the data, it was a matter of navigating the JSON file to see how extract the required data. There were some ideas that I wanted to do at first, but ultimately did not follow through with. First was to join all of the data into one Data Frame; However, not all of the nested data under specific subcategories was related to each other (Events had no way of being tied to Stories or Comics) so ultimately, I decided to just do separate joins where it made sense (Series to Comics). Another idea was to create a pivot/transposed view of the Events and Stories data. However, that was scrapped as the number of fields was not consistent among characters and it would’ve bloated the number of columns and made it a hassle to read.

**Conclusions and Final Thoughts.**

This first attempt at pulling data from the Marvel API has provided some context for a potential Marvel Comics character research by looking at other aspects of the API for more information to get more information regarding stories and events and join it to create a comprehensive character glossary.

**References:**

[1] Marvel. Marvel Developer. Documentation. Retrieved Saturday February 29, 2020 from <https://developer.marvel.com/>.