**Practicum 1 Progress Report – 2/3/2021**

**Week**: 4

**Name**: Harmon Hawkins

**Title**: Sentiment and Text Analysis of News Articles Relating to Airlines and Technology

**Description**:

This project is about looking at how different travel websites and news sources report on different airlines. Specifically, I want to see if there is different sentiment in the articles themselves or the article comments depending on what airline is being reported on, or the subject matter of the article.

**Proposed “TO DO’s” From Last Week:**

* Finish writing the last web scraper
* Finish data collection

**This Week’s Progress:**

* Finished coding the 2 web scrapers
* Gathered the URLs for around 1500 webpages into a csv for use by the web scraper
* Scraped the desired information from the webpages and stored the data into a csv
* Finished data collection

**Issues:**

* Due to the time it took to gather and scrape the web page URLs the scope of the project was reduced from 3 websites to 2.
* Found that there were emojis in some of my data, not an issue yet but still deciding how to handle that

**Next Week “To Do’s”:**

* Clean the data
* Start some EDA and initial data analysis