Tidy Tuesday Adventures (3 Projects – March 2020)

Hi there! I am currently writing this blog while social distancing due to COVID-19. These are unprecedented times we are living in and our lives are changing daily. I hope, optimistically, that by the time anyone reads this, life will be back to being somewhat normal.

For me personally, this has been a tough time. 2.5 weeks ago, I was temporarily laid off from my full-time sales job at Town Sports International and am not guaranteed my position back when gyms are able to reopen. On the bright side, this epidemic has given me plenty of time to dive deeper in my learning and practice my data analysis and visualization with the #TidyTuesday projects from the R For Data Science community on Twitter. I wrote in my previous blog from February 8th that I was hoping to start participating in these projects.

My first ever Tidy Tuesday submission was posted on Thursday March 5th using a dataset of the top 250 goal scorers in NHL history. The fact that the project was based on sports data gave me the confidence to give it a shot. I decided to analyze and rank the individual goal scoring seasons of the top 10 goal scorers. Here is my plot:

My code can be found here:

<https://github.com/msurdek/Surdek-Tidy-Tuesday/blob/master/Goal%20Scoring%20Seasons.R>

Based on this plot,

My second Tidy Tuesday submission was posted on Wednesday March 18th using data from the scripts of the TV show, The Office. This was my first time analyzing text data using the {tidytext} package. Here is my plot:

My code can be found here:

<https://github.com/msurdek/Surdek-Tidy-Tuesday/blob/master/the%20office%20script.R>

I found that there were 16 characters who had significantly more lines than the rest of the secondary characters, so I considered them the main characters. I ranked the 16 characters

My third and most recent Tidy Tuesday submission was posted on Tuesday March 31st with data on US beer production from 2009-2018. This data included information about beer ingredients used over time and brewery production broken down by size and state. This was another dataset where I felt confident about my knowledge prior to analyzing. Here is my plot:

My code can be found here:  
<https://github.com/msurdek/Surdek-Tidy-Tuesday/blob/master/Beer%20Production.R>

Based on the results displayed in my graph,