PS5

Opal Fraser

March 6, 2023

0.1 Question 3

The first data set is the world records in swimming from Wikipedia. There was a lot of data to start with. I chose to write the csv with only the 1, 2 and 4th columns of the first 97 rows. This represents the event, time, and the name of thw swimmer. What interested me about it was the amount of random words that came out at first. I had to wrangle it around to get it the way it was at the end and that was pretty exciting. I love to swim and that is why I chose to look at swimming records. The data won't be useful for me in the future. Learning the process is helpful and getting me closer to the data that I am looking for to use for my project. I used ChatGBT. I tried this website: And, I followed Grant R. McDermott's Lecture 6.

0.2 Question 4

I looked at the unemployment rate and the civilian labor force participation rate. I chose the columns for date and value. I took it a bit further and I graphed them. I saw a dip in labor force participation and an increase in unemployment rate around 2020. It seems to show the effects of the COVID pandemic on the US labor force. The packages I used are fredr, tidyverse, purrr, dplyr, and ggplot2.