Karan Sarkar

Data Visualization

HW 2

1. Queries
   - What is the relationship between male labor participation rate and female labor participation?
   - Does the balance of fossil fuel usage to renewables usage correlate in any way to the GDP per capita of a country?
   - Is there a relationship between the level of inequality as measured by the share of wealth controlled by the top 1% and top 10% to the rate of GDP growth? What is the optimal level of inequality?
   - What factors are correlated with different ratios of internet usage to mobile usage? Internet usage without mobile is likely enterprise and PC. How does GDP per capita affect the balance between mobile's share of computing?
2. Insights
   - Total fertility rate is negatively correlated to the number of years of schooling.
   - Income per capita is positively correlated to percentage of internet users.
   - Fossil fuel usage is positively correlated to CO_2 emissions.
   - Life expectancy is positively correlated to GDP per capita.
3. In order to find relationships, I sorted the data table by a single column. Then, I looked to see if any other columns were roughly monotonic. These monotonic columns showed me that there were either positive or negative correlations. If the two columns went in the same direction, it is a positive correlation. If they went in opposite directions, it is a negative correlation.
4. I had a very hard time really visualizing the data. I was really only able to find monotonic linear relationships. I do not believe I would be able to spot clusters or any non-linearity. Moreover, only really obvious relationships could be spotted. It is very hard to judge noisy relationships by raw numbers.