HW 1

1) b) The population is drawn from the subset of voting-age US residents (not necessarily registered voters) that spend time doing online surveys. This necessarily makes it so that this sample is not random. Every number of the population does not have an equal chance of being selected to be part of the sample. We'd be sampling individuals with technical savvy — at least enough to be able to use the internet. This builds a skew into our sample which may weak im-balance in any statistical analysis we perform on the sample. e.g. Social Media analysis.

2) It appears that Trump articles were shared much more.

3) Non-exclusive:
{'snopes': 138, 'buzzfeed': 21, 'politifact': 13} ⟸ hashmap {database name: # of articles}

Exclusives:
{'snopes': 124, 'buzzfeed': 9, 'politifact': 9}

Snopes had the most comprehensive data set with the highest exclusive articles / total articles ratio of .899.
Buzzfeed had the least comprehensive dataset with a ratio of .429.

4) No, the vast majority of respondents time (45%) was not spent on social media.

5) Social media is fairly important, but not more widely-used than Cable TV or the Website category.