

The Question

Nowadays, we can get tons of information online within seconds. However, some of the news are fake and are misleading the people. The goal of this project is to develop a fake news model such as classification model. Hence, we can filter the fake news before they are released to the public.

The Data

The dataset we found now is from the Kaggle “Getting Real about Fake News”. The dataset contains text and metadata from 244 websites and represents 12,999 posts in total from the 30 days by the BS Detector Chrome Extension by Daniel Sieradski. Before cleaning, there are 170949 row and 20 columns.

The Tools

The project plans to use excel perform the data cleaning and exploratory analysis. Tableau will be used to visualize the data.