MDS6212 Fintech Theory and Practice

Assignment for Week four

Data confidential requirements:

* Do not share datasets with others
* Do not post datasets on the Web site
* Datasets can only be used to work on assignments of MDS6212, M.Sc. in Data Science, CUHK-Shenzhen.

Datasets:

* MDS6212 Week 4 Tesla.xlsx

Questions 1-3:

For the following news article,

“Tesla‘s third-quarter sales jumped 44% as global demand for its electric vehicles outpaced that of most other automakers. The company reported Friday that it had delivered 139,000 SUVs and sedans from July through September, compared with 97,000 deliveries during the same period a year ago. The sales topped even some of the most optimistic projections coming from Wall Street. Analysts polled by data provider FactSet expected the company to sell closer to 137,000. Telsa has been rewriting the script throughout the year amidst a pandemic that has closed factories and scrambled supply lines. “This puts Musk & Co. in prime position to hit the area code of 500k units for the year which six months ago was not even on the map for the bulls,'" Daniel Ives of Wedbush wrote Friday. China was likely a major source of strength in the quarter, Ives said. Tesla could post its fifth consecutive quarter of profits later this month.”

1. Present word cloud
2. Calculate the news sentiment variable using Loughran and McDonald Sentiment Word Lists
3. Calculate the Fog index

Question 4: using the data set of Tesla, a) report the summary statistics of *sentiment*, *novelty*, and *impact*; b) present the correlation coefficient among *sentiment*, *novelty*, and *impact*; and c) show the frequency and fraction of top 10 news categories.

Deadline: Oct. 24, 2020