**Milestone 4**

**QUANTITATIVE DATA INTERPRETATION**

Category: INSURANCE STOCKS

**A screenshot of a computer

Description generated with very high confidence**

**A picture containing screenshot

Description generated with very high confidence**

By choosing 9 stocks from the insurance category which are Allianz, LPI, MAA, Manulife, MNRB, MPHBCAP, P&O, TAKAFUL, and TUNEPRO the sentiment scores based on the crawled news of these stocks were computed and left joined with the stock data which consist of the date, closing price, opening price, high, low, volume, and percentage change. These 2 tables were joined by stock name and news date. Exploring the data through SAS showed that LPI and Allianz stocks had the largest closing prices in terms of sum. For percentage change, MNRB and TUNEPRO stocks were most volatile with the largest change

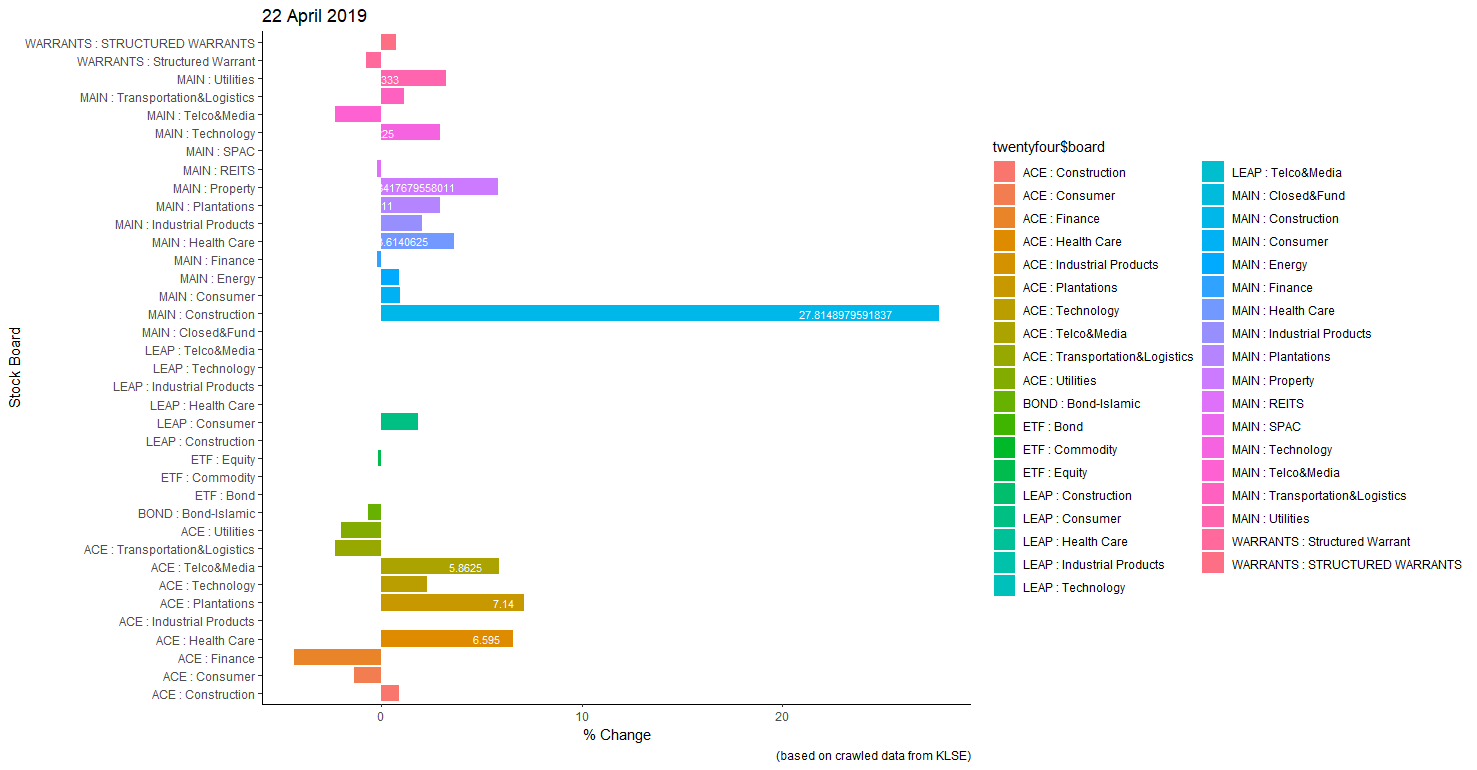


Figure 1: Mean percentage change for KLSE stock categories

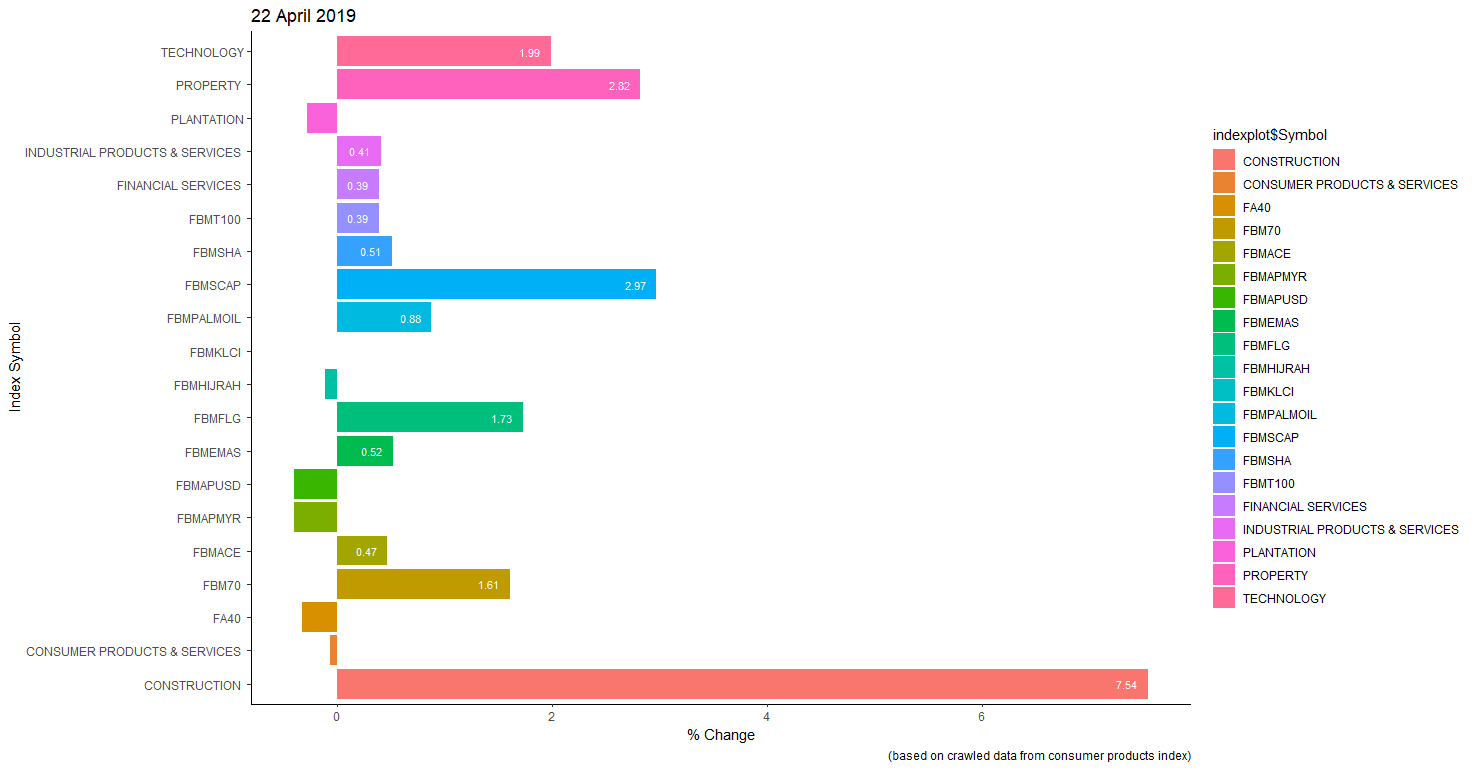
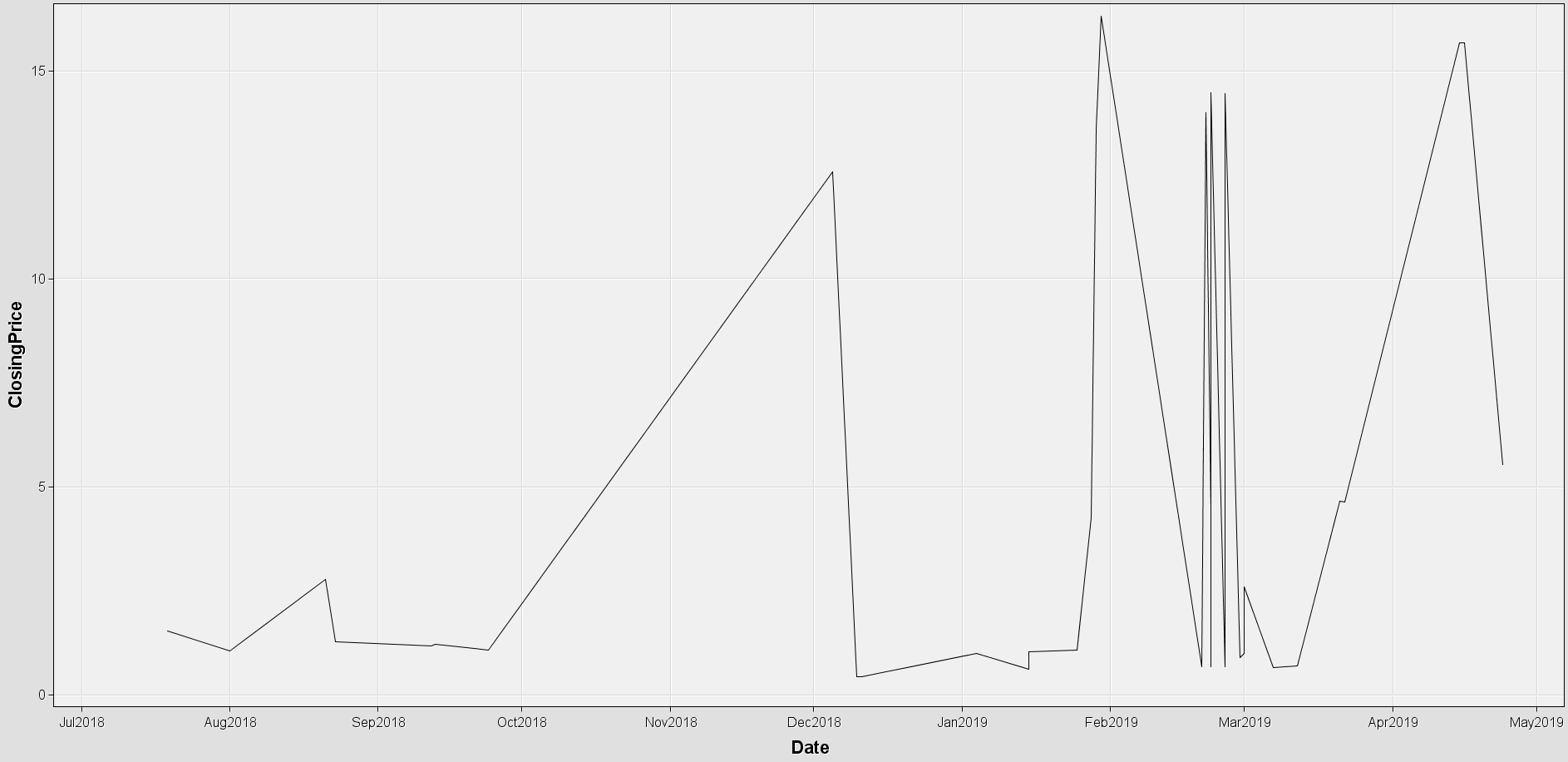
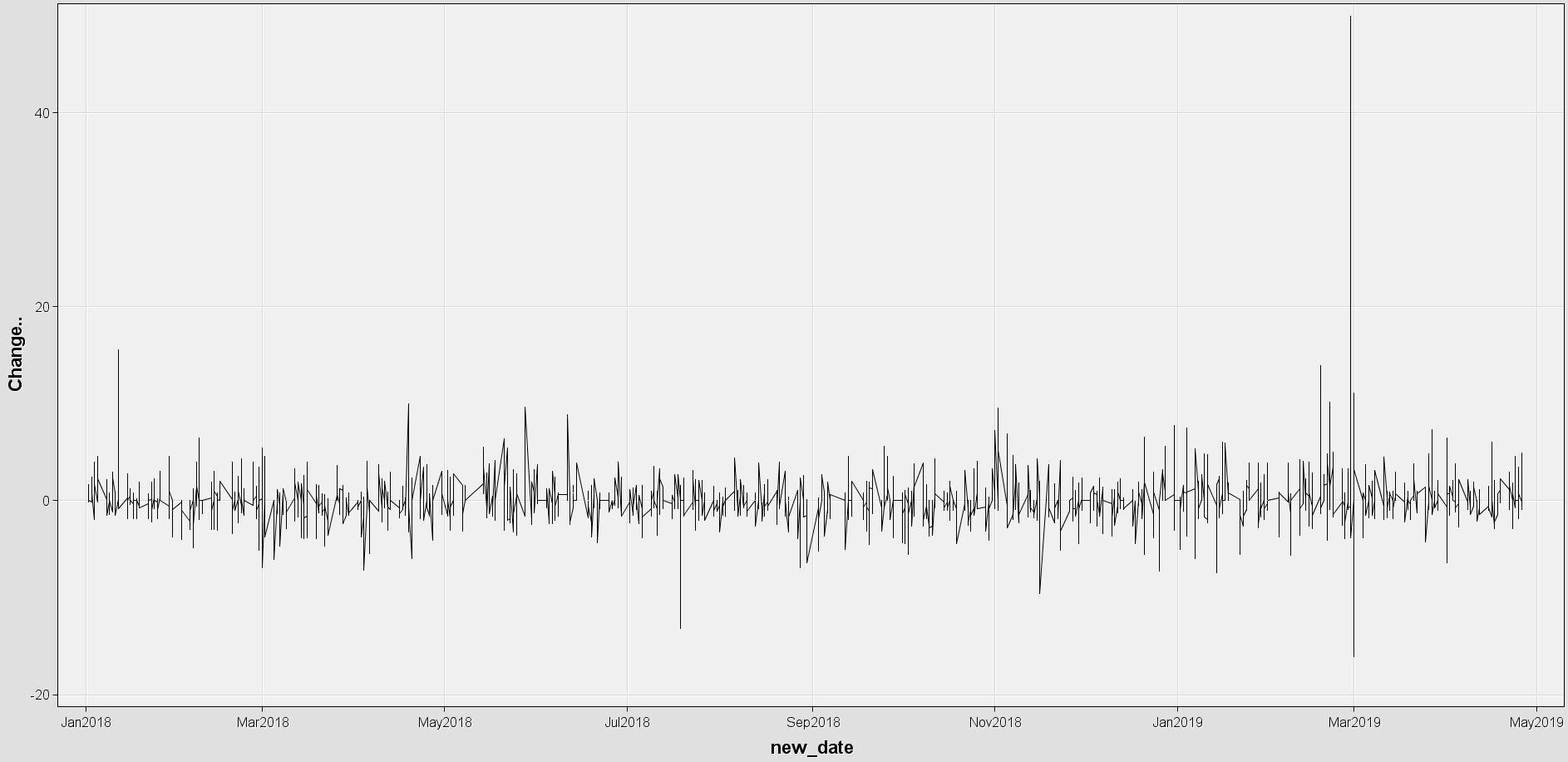


Figure 2: Percentage change for consumer products index

Based on crawled data from 22 April 2019 from KLSE in Figures 1 and 2 showed that stocks under the construction category had the largest percentage change in terms of mean percentage change and stock indexes with 27.8% followed by plantations, health care, and property with 7.1%, 6.6%, and 5.8% respectively. This can also be seen for consumer products index for construction and property with 7.5% and 2.8% change.

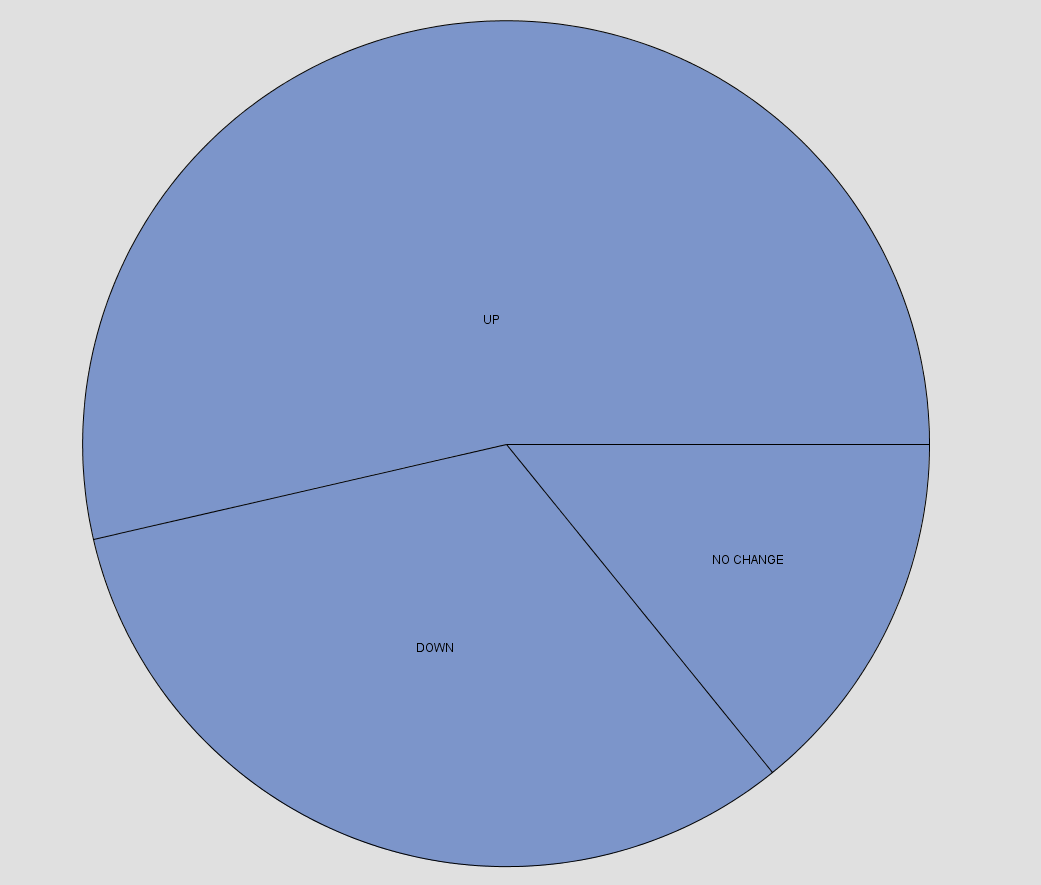


We found that the closing price against the dates are very inconsistent with impulsive spikes for insurance stocks only

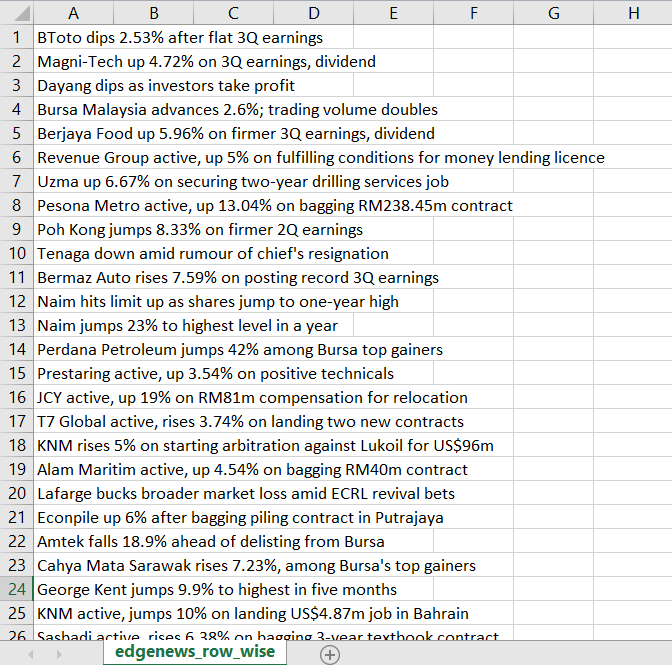


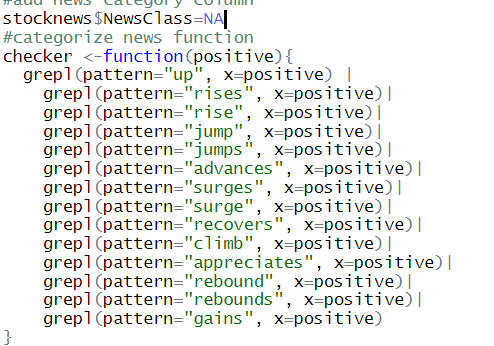
For all stocks, we can see the percentage change will vary in a certain pattern and range in the y axis. Only in March 2019 we saw spikes with very large percentage change. This indicates Malaysian stocks are not very volatile and the trends are quite controlled.

**QUALITATIVE DATA INTERPRETATION**



After categorizing the stocks into UP, DOWN, and NO CHANGE based on the percentage change %, we found that most stocks go towards the positive direction, and there are less stocks who go in the negative direction.





**A screenshot of text

Description generated with high confidence**

By using R Studio the crawled news data from the edge was labelled into negative and positive news based on the patterns such as ‘rise’, ‘jumps’, ‘surges’, ‘rebounds’, ‘gains’, etc. and the length of text was calculated. Above figure shows that there were more negative than positive news. However, the length of text are similar between positive and negative news.

**A screenshot of a cell phone

Description generated with very high confidenceA screenshot of a cell phone

Description generated with very high confidence**

The ‘udpipe’ package in R was able to classify words into nouns and adjectives and the text analytics found that the most occurring nouns were ‘rises’ and ‘earnings’ followed by most occurring adjectives which are ‘active’ and ‘positive’. These nouns and adjectives fall under the positive news category.

**A screenshot of a cell phone

Description generated with high confidenceA close up of a map

Description generated with very high confidence**

News data also showed that words related positively reoccurs within the sentence and these words follow each other. Based on the interpretation on words itself we can correlate them to the related stocks as positive and good quality performing stocks

**A close up of a newspaper

Description generated with high confidence**

Word Cloud