

Task 1: Corpus

Three corpora for the specific tasks:

| Corpus 1 (Task 2) | | Corpus 2 (Task 3) | | Corpus 3 (Task 4 and 5) | |
|--------------------|-------------|---------------------|--------------------|-------------------------|-------------|
| Metric | Value | Metric | Value | Metric | Value |
| Nr. of Documents | 929 | Nr. of Documents | 897 | Nr. Of Documents | 898 |
| Total Words | 10137679 | Total Words | 109059 | Total Words | 76483 |
| Unique Terms | 30894 | Unique Terms | 11719 | Unique Terms | 9597 |
| Avg. words per doc | 10912 | Avg. words per doc. | 122 | Avg. words per doc. | 85 |
| Top terms | Total count | Top terms | Total count | Top ten terms. | Total count |
| will | 132207 | nvidia | 3710 | nvidia | 4693 |
| us | 85532 | market | 1007 | market | 678 |
| think | 76660 | tech | 742 | tech | 515 |
| going | 64302 | stocks | 653 | chips | 499 |
| see | 60450 | company | 636 | earnings | 487 |
| get | 57502 | companies | 618 | apple | 487 |
| year | 56949 | earnings | 599 | stocks | 461 |
| market | 55951 | chips | 582 | companies | 454 |
| can | 54488 | lot | 529 | scanned | 446 |

525 company

Table 1. The three corpora.

49588 china

now





429

NHH

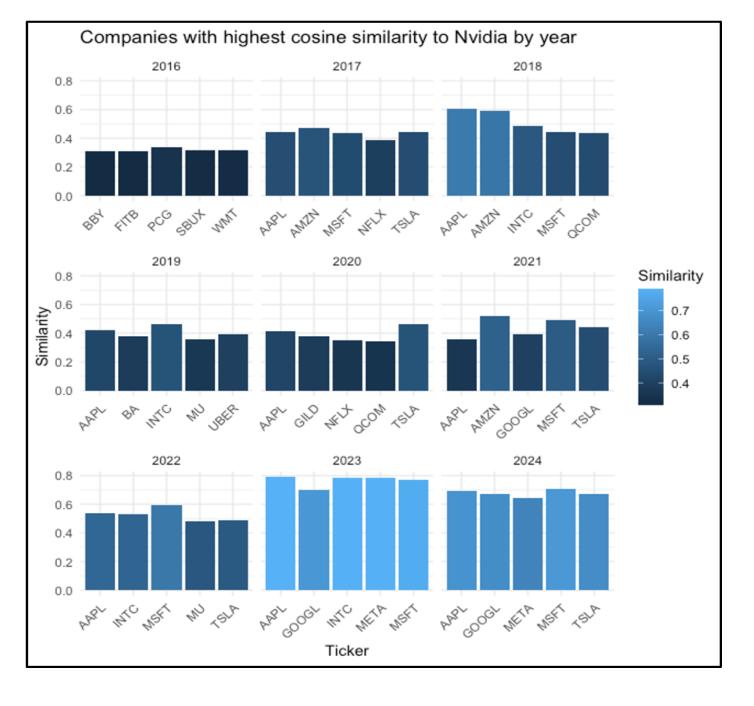
The most cosine similar companies during 2016-2024

| 父 | ₫ |
|---|---|
| 丰 | 美 |

| Term | Similarity | Ticker | Company | |
|-----------|------------|--------|----------------------------|--|
| apple | 0,836 | AAPL | APPLE INC | |
| intel | 0,831 | INTC | INTEL CORP | |
| tesla | 0,780 | TSLA | TESLA INC | |
| microsoft | 0,776 | MSFT | MICROSOFT CORP | |
| amazon | 0,712 | AMZN | AMAZON.COM INC | |
| micron | 0,710 | MU | MICRON TECHNOLOGY INC | |
| meta | 0,696 | META | META PLATFORMS INC CLASS A | |
| alphabet | 0,696 | GOOGL | ALPHABET INC CL A & CL C | |
| qualcomm | 0,620 | QCOM | QUALCOMM INC | |
| broadcom | 0,610 | AVGO | BROADCOM INC | |

Table 2. SP500 companies (Nov 2024) mentioned in the corpus ranked by cosine similarity. I with Corpus 1.

The changes in most similar or related companies throughout the years 2016-2024



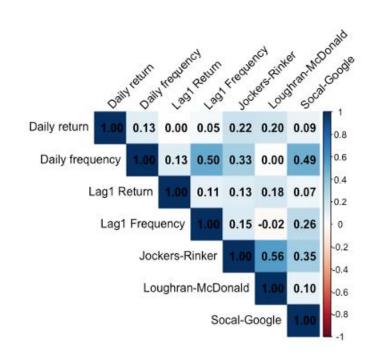


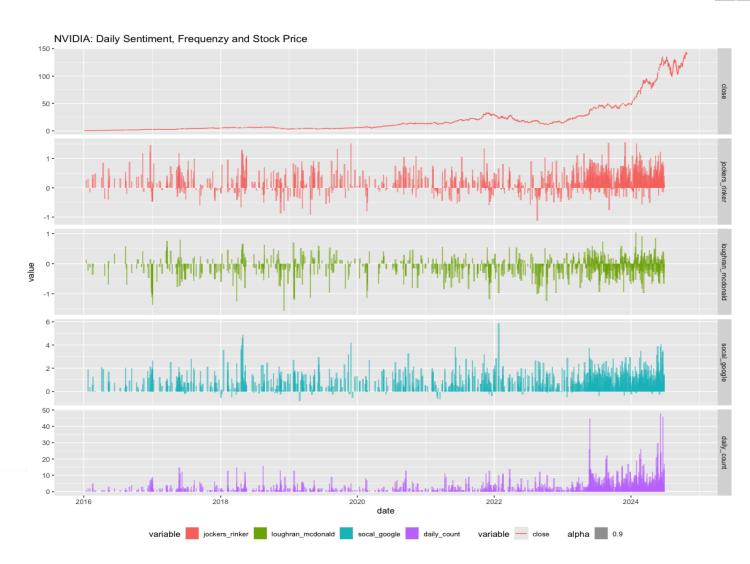






Task 2 NVIDIA's Daily Sentiment, Frequency and Stock Price and correlation





Daily return as a dependent variable of Sentiment and Frequency?

NHH



A linear regression with interaction terms.

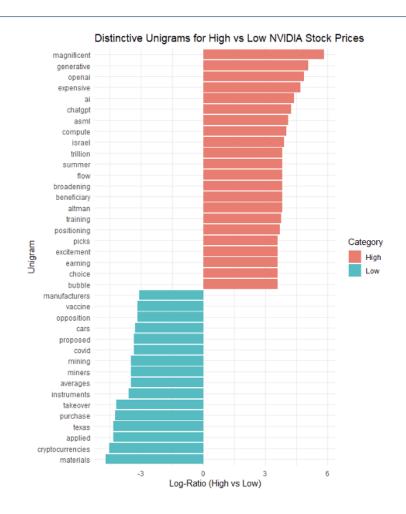
| Linear regression on Daily return (dependent variable) using Corpus 2. | |
|--|--|
| Including interation terms between daily frequency and sentiment. | |

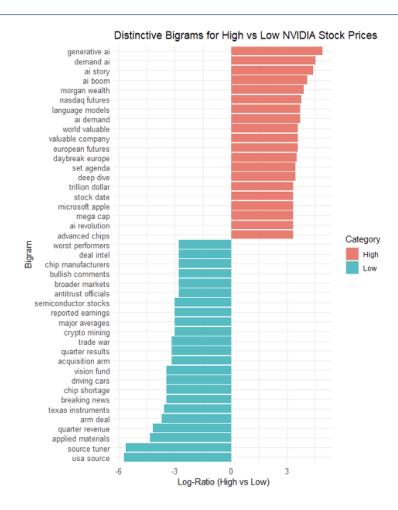
| Term | Estimate | Std.error | t statistic | P value | Sign. Level |
|--|----------|-----------|-------------|---------|----------------|
| (Intercept) | 0,00067 | 0,0030 | 0,23 | 0,82 | |
| Daily frequency | 0,00044 | 0,0007 | 0,64 | 0,52 | |
| Jockers-Rinker | 0,02834 | 0,0061 | 4,64 | 0,00 | *** |
| Loughran-Mcdonald | -0,00936 | 0,0070 | -1,33 | 0,18 | |
| Socal-Google | -0,00405 | 0,0023 | -1,75 | 0,08 | • |
| Daily frequency: Jockers-Rinker | -0,00269 | 0,0009 | -2,95 | 0,00 | ** |
| Daily frequency: Loughran- Mcdonald | 0,00456 | 0,0011 | 4,20 | 0,00 | *** |
| Daily frequency : Socal-Google | 0,00067 | 0,0003 | 1,95 | 0,05 | • |

Table 2. Significance codes: '***' = 0.001, '**' = 0.01, '*' = 0.05, '.' = 0.1, '' = 1.





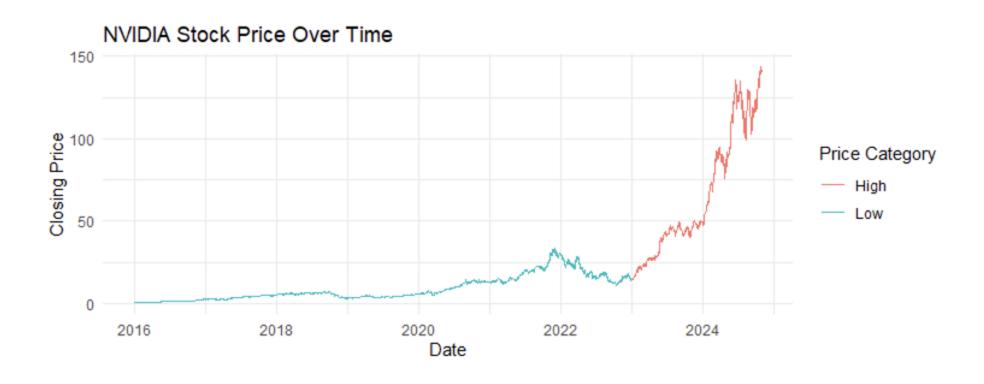




Stock price







NHH

Topics from LDA Topic Model (5 selected out of 8)

talkbuy usa china intel video stock start

chipmak francisco

chip san

ca growth tuner continu

tradetech

move app come

share nasdag



Topic 1 Topic 6 Topic 4 downgrad stowdown miss drive deal low mobil action Cryptoculr deal low mobil action Cryptoculr deal low mobil action Cryptoculr deal low mobil down miss drive deal low mobil deal low mob ludlow acceler manag ludlow acceler morn domin build tsmc drive gpu spend yield semiconductor dapi runtrillion week biden risk e earn microsoft stori manufactur valley acquisit mine Sourc scan takeov design to biggest bigg lisa month mandeep market ralli train equiti topic catch magnific shortag ■ conductor agreregulatori versus zoom semiconductorcomment Topic 8 Topic 5 microsoftpeopl amazon quarter liquid hike control tumbi export

collaps Warn plan leav walk ukrain ipo plan leav walk liniti interiori project supplied supplied by the project supplied supplied by the project suppl

tuner outlook surpris low

aggress impact musk

Shows clear

