

10.

Financial Analysis

Dr. W.J.B. Mattingly

Smithsonian Data Science Lab and United States Holocaust Memorial Museum

August 2021

```
import spacy
import pandas as pd
```

```
INFO:tensorflow:Enabling eager execution
INFO:tensorflow:Enabling v2 tensorshape
INFO:tensorflow:Enabling resource variables
INFO:tensorflow:Enabling tensor equality
INFO:tensorflow:Enabling control flow v2
```

```
df = pd.read_csv("data/stocks.tsv", sep='\t')
```

df

	Symbol	CompanyName	Industry	MarketCap
0	A	Agilent Technologies	Life Sciences Tools & Services	53.65B
1	AA	Alcoa	Metals & Mining	9.25B
2	AAC	Ares Acquisition	Shell Companies	1.22B
3	AACG	ATA Creativity Global	Diversified Consumer Services	90.35M
4	AADI	Aadi Bioscience	Pharmaceuticals	104.85M
...
5874	ZWRK	Z-Work Acquisition	Shell Companies	278.88M
5875	ZY	Zymergen	Chemicals	1.31B
5876	ZYME	Zymeworks	Biotechnology	1.50B
5877	ZYNE	Zynerba Pharmaceuticals	Pharmaceuticals	184.39M
5878	ZYXI	Zynex	Health Care Equipment & Supplies	438.33M

5879 rows x 4 columns

```
symbols = df.Symbol.tolist()
```

```
companies = df.CompanyName.tolist()
```

```
print (symbols[0])
print (companies[0])
```

```
A
Agilent Technologies
```

```
df2 = pd.read_csv("data/indexes.tsv", sep="\t")
df2
```

	IndexName	IndexSymbol
0	Dow Jones Industrial Average	DJIA
1	Dow Jones Transportation Average	DJT
2	Dow Jones Utility Average Index	DJU
3	NASDAQ 100 Index (NASDAQ Calculation)	NDX
4	NASDAQ Composite Index	COMP
5	NYSE Composite Index	NYA
6	S&P 500 Index	SPX
7	S&P 400 Mid Cap Index	MID
8	S&P 100 Index	OEX
9	NASDAQ Computer Index	IXCO
10	PHLX Semiconductor Index	SOX
11	PHLX Gold/Silver Index	XAU
12	NYSE Arca Oil Index	XOI
13	Russell 2000 Index	RUT

```
indexes = df2.IndexName.tolist()
index_symbols = df2.IndexSymbol.tolist()
```

```
df3 = pd.read_csv("data/stock_exchanges.tsv", sep="\t")
df3
```

	BloombergExchangeCode	BloombergCompositeCode	Country	Description	ISOM
0	AF	AR	Argentina	Bolsa de Comercio de Buenos Aires	XBL
1	AO	AU	Australia	National Stock Exchange of Australia	XNE
2	AT	AU	Australia	Asx - All Markets	XAS
3	AV	NaN	Austria	Wiener Boerse Ag	XWE
4	BI	NaN	Bahrain	Bahrain Bourse	XBA
...
97	UR	US	USA	NASDAQ Capital Market	XNC
98	UV	US	USA	OTC markets	OOT
99	UW	US	USA	NASDAQ Global Select	XNC
100	VH	VN	Vietnam	Hanoi Stock Exchange	HS
101	VM	VN	Vietnam	Hochiminh Stock Exchange	XST

102 rows × 8 columns

```
exchanges = df3.ISOMIC.tolist()+df3["Google Prefix"].tolist()
descriptions = df3.Description.tolist()
```

```
stops = ["two"]
nlp = spacy.blank("en")
ruler = nlp.add_pipe("entity_ruler")
patterns = []
letters = "ABCDEFGHIJKLMNOPQRSTUVWXYZ"
#List of Entities and Patterns
for symbol in symbols:
    patterns.append({"label": "STOCK", "pattern": symbol})
    for l in letters:
        patterns.append({"label": "STOCK", "pattern": symbol+f".{l}"})

for company in companies:
    if company not in stops:
        patterns.append({"label": "COMPANY", "pattern": company})
        words = company.split()
        if len(words) > 1:
            new = " ".join(words[:2])
            patterns.append({"label": "COMPANY", "pattern": new})

for index in indexes:
    patterns.append({"label": "INDEX", "pattern": index})
    versions = []
    words = index.split()
    caps = []
    for word in words:
        word = word.lower().capitalize()
        caps.append(word)
    versions.append(" ".join(caps))
    versions.append(words[0])
    versions.append(caps[0])
    versions.append(" ".join(caps[:2]))
    versions.append(" ".join(words[:2]))
    for version in versions:
        if version != "NYSE":
            patterns.append({"label": "INDEX", "pattern": version})

for symbol in index_symbols:
    patterns.append({"label": "INDEX", "pattern": symbol})

for d in descriptions:
    patterns.append({"label": "STOCK_EXCHANGE", "pattern": d})
for e in exchanges:
    patterns.append({"label": "STOCK_EXCHANGE", "pattern": e})

ruler.add_patterns(patterns)

print (len(patterns))
```

169694

#source: <https://www.reuters.com/business/futures-rise-after-biden-xi-call-oil-bounce-2021-09-10/>

text = ''

Sept 10 (Reuters) – Wall Street's main indexes were subdued on Friday as signs of higher inflation and a drop in Apple shares following an unfavorable court ruling offset expectations of an easing in U.S.–China tensions.

Data earlier in the day showed U.S. producer prices rose solidly in August, leading to the biggest annual gain in nearly 11 years and indicating that high inflation was likely to persist as the pandemic pressures supply chains. [read more](#) .

"Today's data on wholesale prices should be eye-opening for the Federal Reserve, as inflation pressures still don't appear to be easing and will likely continue to be felt by the consumer in the coming months," said Charlie Ripley, senior investment strategist for Allianz Investment Management.

Apple Inc (AAPL.O) fell 2.7% following a U.S. court ruling in "Fortnite" creator Epic Games' antitrust lawsuit that stroke down some of the iPhone maker's restrictions on how developers can collect payments in apps.

Sponsored by Advertising Partner

Sponsored Video

Watch to learn more

Report ad

Apple shares were set for their worst single-day fall since May this year, weighing on the Nasdaq (.IXIC) and the S&P 500 technology sub-index (.SPLRCT), which fell 0.1%.

Sentiment also took a hit from Cleveland Federal Reserve Bank President Loretta Mester's comments that she would still like the central bank to begin tapering asset purchases this year despite the weak August jobs report. [read more](#)

Investors have paid keen attention to the labor market and data hinting towards higher inflation recently for hints on a timeline for the Federal Reserve to begin tapering its massive bond-buying program.

The S&P 500 has risen around 19% so far this year on support from dovish central bank policies and re-opening optimism, but concerns over rising coronavirus infections and accelerating inflation have lately stalled its advance.

Report ad

The three main U.S. indexes got some support on Friday from news of a phone call between U.S. President Joe Biden and Chinese leader Xi Jinping that was taken as a positive sign which could bring a thaw in ties between the world's two most important trading partners.

At 1:01 p.m. ET, the Dow Jones Industrial Average (.DJI) was up 12.24 points, or 0.04%, at 34,891.62, the S&P 500 (.SPX) was up 2.83 points, or 0.06%, at 4,496.11, and the Nasdaq Composite (.IXIC) was up 12.85 points, or 0.08%, at 15,261.11.

Six of the eleven S&P 500 sub-indexes gained, with energy (.SPNY), materials (.SPLRCM) and consumer discretionary stocks (.SPLRCD) rising the most.

U.S.–listed Chinese e-commerce companies Alibaba and JD.com , music streaming company Tencent Music (TME.N) and electric car maker Nio Inc (NIO.N) all gained between 0.7% and 1.4%

Report ad

Grocer Kroger Co (KR.N) dropped 7.1% after it said global supply chain disruptions, freight costs, discounts and wastage would hit its profit margins.

Advancing issues outnumbered decliners by a 1.12-to-1 ratio on the NYSE and by a 1.02-to-1 ratio on the Nasdaq.

The S&P index recorded 14 new 52-week highs and three new lows, while the Nasdaq recorded 49 new highs and 38 new lows.

'''

doc = nlp(text)

```
for ent in doc.ents:
    print (ent.text, ent.label_)
```

Print to PDF ►

Apple COMPANY
Apple COMPANY
AAPL.O STOCK
Apple COMPANY
Nasdaq COMPANY
S&P 500 INDEX
S&P 500 INDEX
ET STOCK
Dow Jones Industrial Average INDEX
S&P 500 INDEX
Nasdaq Composite INDEX
S&P 500 INDEX
JD.com COMPANY
Tencent Music COMPANY
TME.N STOCK
NIO.N STOCK
Kroger COMPANY
KR.N STOCK
NYSE STOCK_EXCHANGE
Nasdaq INDEX
S&P INDEX
Nasdaq INDEX

```
#source: https://www.reuters.com/companies/AAPL.O
text2 = '''
Apple Inc. designs, manufactures and markets smartphones, personal computers, tablets,
wearables and accessories, and sells a variety of related services. The Company's
products include iPhone, Mac, iPad, and Wearables, Home and Accessories. iPhone is the
Company's line of smartphones based on its iOS operating system. Mac is the Company's
line of personal computers based on its macOS operating system. iPad is the Company's
line of multi-purpose tablets based on its iPadOS operating system. Wearables, Home and
Accessories includes AirPods, Apple TV, Apple Watch, Beats products, HomePod, iPod
touch and other Apple-branded and third-party accessories. AirPods are the Company's
wireless headphones that interact with Siri. Apple Watch is the Company's line of smart
watches. Its services include Advertising, AppleCare, Cloud Services, Digital Content
and Payment Services. Its customers are primarily in the consumer, small and mid-sized
business, education, enterprise and government markets.
'''
```

```
doc2 = nlp(text2)
```

```
for ent in doc2.ents:
    print (ent.text, ent.label_)
```

Apple STOCK
Apple STOCK
TV STOCK
Apple STOCK
Apple STOCK
Apple STOCK

By William Mattingly
© Copyright 2021.