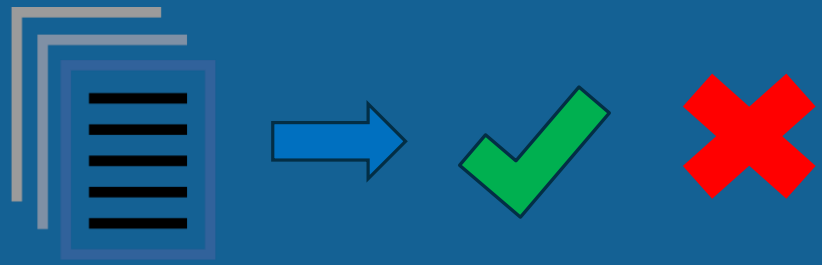
MODEL 3

CAN WE TURN TECHNICAL DOCUMENTS INTO DECISIONS?



SEC 10K item1A > **Buy** or **Sell**

```
]: ct_df.iloc[3]['risk']

1]: '[1] "ITEM 1A. RISK FACTORS.. Set forth below are the risk factors that we b
    carefully consider these risk factors, together with all of the other inform
    onsolidated financial statements and the related notes thereto, before you d
    ties. The occurrence of any of the following risks could adversely affect ou
    ition and value of our Common Shares. In such case, the value of our Common
    uld decline, and you may lose all or a significant part of your investment.
    d-looking statements. Refer to the explanation of the qualifications and lim
    ssed in the beginning of this Form 10-K.. The following risk factors are not
    clude additional factors that could adversely affect our business and financ
    ompetitive and rapidly changing environment. New risk factors emerge from ti
    ct all risk factors, nor can we assess the impact of all risk factors on our
    ombination of factors, may affect our business. Investors should also refer
    reports on Form 8-K for future periods for material updates to these risk fa
    PERTIES. There are risks relating to investments in real estate that may adv
    perty investments are subject to multiple risks. Real estate values are affe
    n the general economic climate, local conditions (such as an oversupply of s
    an area), the quality and philosophy of management, competition from other a
    de adequate maintenance and insurance and to control variable operating cost
```

Files

```
documents=[]
for i in range(0,1315):
    documents.append(riskdf.iloc[i]['Riskfactors'])

vectorizer = TfidfVectorizer(stop_words='english')
X = vectorizer.fit_transform(documents)
```

TF-IDF Vectorizer

Labels

Kmeans Clusters

```
true_k = 30
model = KMeans(n_clusters=true_k, init='k-means++', max_iter=100, n_init=1)
model.fit(X)
```

```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=100,
    n_clusters=30, n_init=1, n_jobs=None, precompute_distances='auto',
    random_state=None, tol=0.0001, verbose=0)
```

```
print("Top terms per cluster:")
order_centroids = model.cluster_centers_.argsort()[:, ::-1]
terms = vectorizer.get_feature_names()
for i in range(true_k):
    print("Cluster %d:" % i),
    for ind in order_centroids[i, :50]:
        print(' %s' % terms[ind]),
    print

print("\n")

print("Prediction")
```

```
Top terms per cluster:
Cluster 0:
 projects
 business
 operations
 financial
 contracts
 project
```

Method 3

Validate Correlation matrix

	topic	risk
	Topic 0: investment, property, distribution, i...	Shareholders interest risk
	Topic 1: regulation, change, law, financial, o...	*Regulation changes(accounting)
	Topic 2: gas, price, oil, natural, operation, ...	Input prices risks
	Topic 3: stock, price, share, market, future, ...	** Stakeholders profit
	Topic 4: cost, regulation, environmental, law,...	*Regulation changes(environment)

Clusters

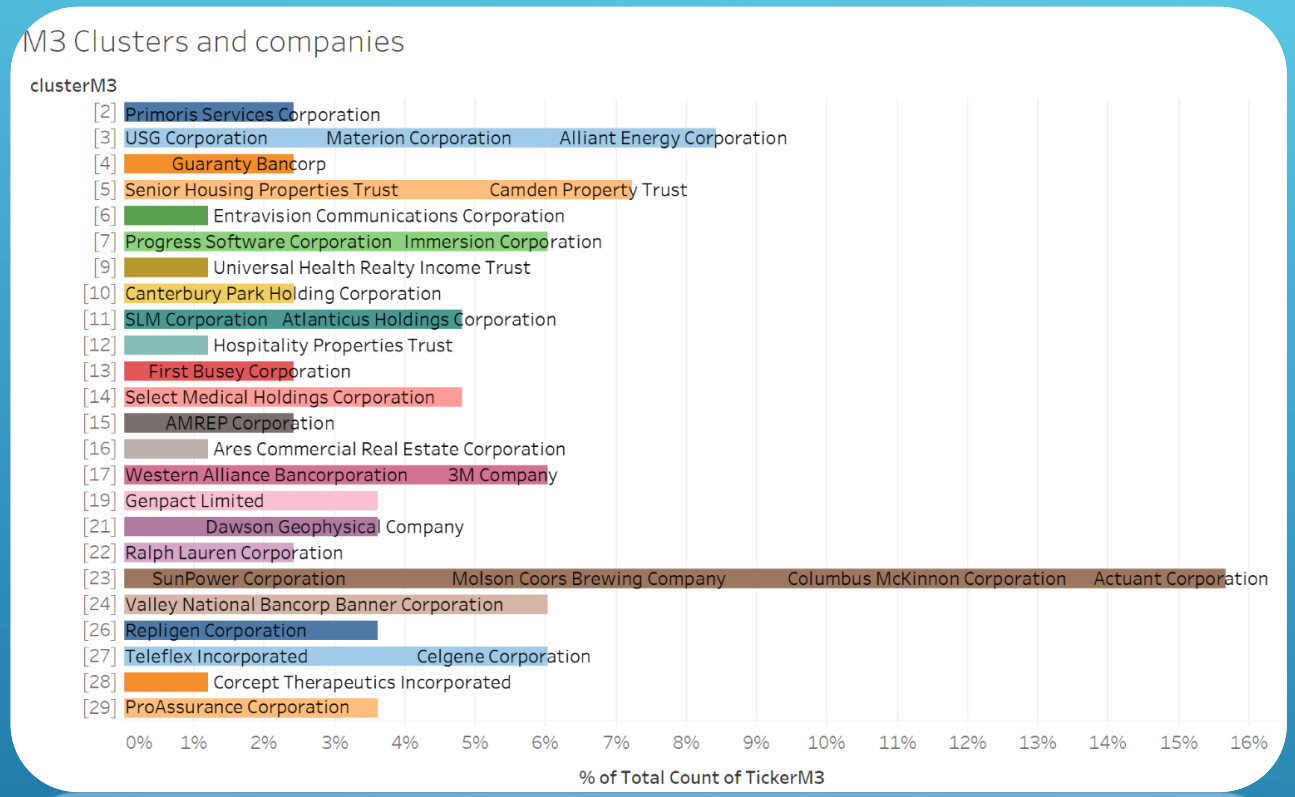
	date_x	adj_close_x	adj_volume_x	date_y	adj_close_y	adj_volume_y	companyname	ticker	cluster
	2018-01-02	5.30	1722138.0	2017-01-03	9.560000	959360.0	Western Alliance Bancorporation	FTK	[17]
	2018-01-02	29.10	306457.0	2017-01-03	37.980000	1077707.0	Valley National Bancorp	FN	[24]
	2018-01-02	95.73	1181040.0	2017-01-03	56.646023	690124.0	USG Corporation	FMC	[3]
	2018-01-02	193.87	905935.0	2017-01-03	144.650000	721903.0	Universal Health Realty Income Trust	FLT	[9]
	2018-01-02	30.43	4433114.0	2017-01-03	26.598418	6369018.0	Union Bankshares Corporation	FITB	[13]

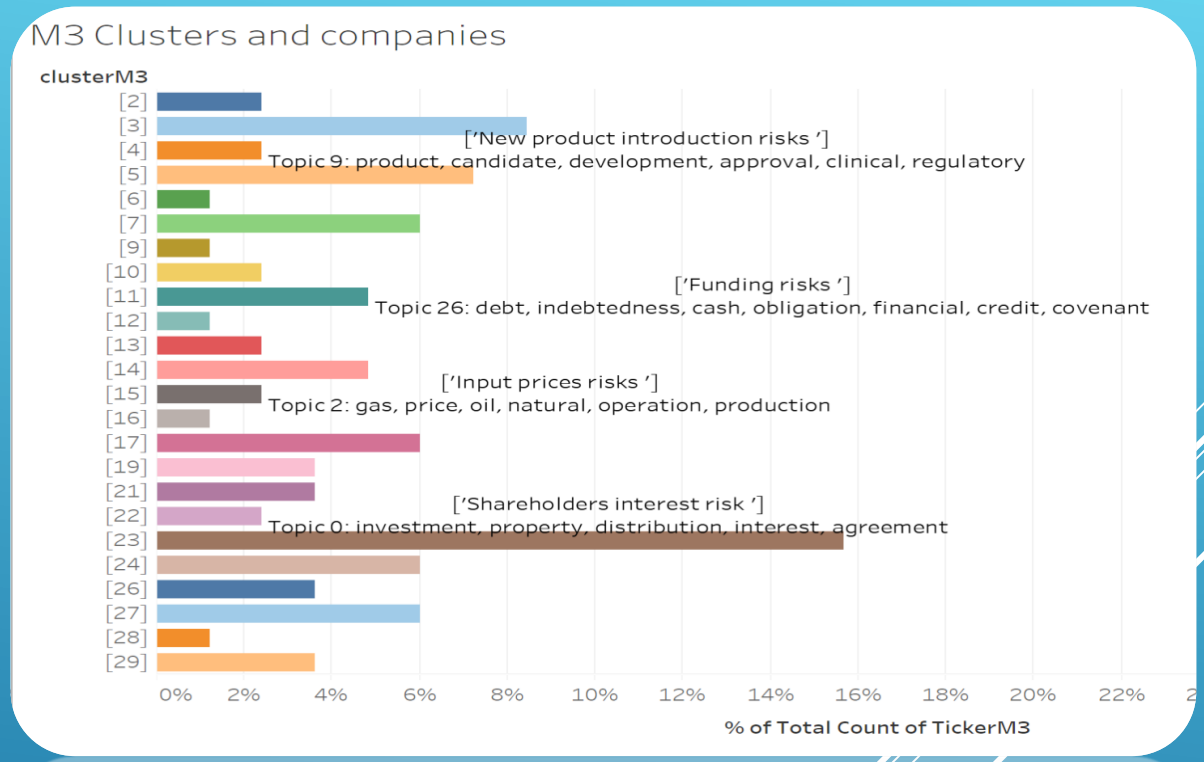
Classify

	topic	risklabel	cluster
	Topic 0: investment, property, distribution, i...	Shareholders interest risk	[14]
	Topic 1: regulation, change, law, financial, o...	*Regulation changes(accounting)	[14]
	Topic 2: gas, price, oil, natural, operation, ...	Input prices risks	[21]
	Topic 3: stock, price, share, market, future, ...	** Stakeholders profit	[14]
	Topic 4: cost, regulation, environmental, law,...	*Regulation changes(environment)	[14]

Cluster Classification

Method 3

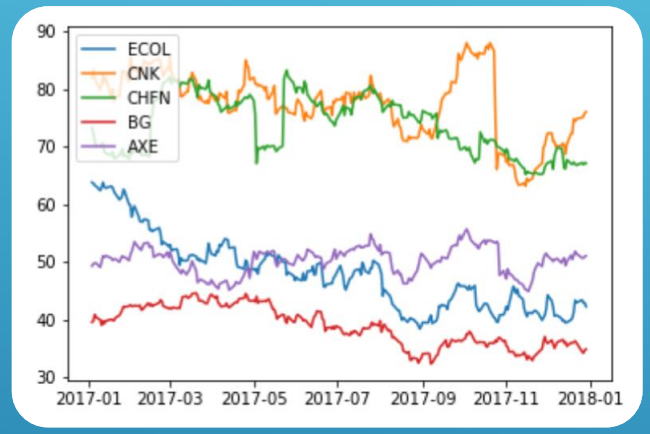




Validate via correlation matrix
Clusters in to market trend

Method 3

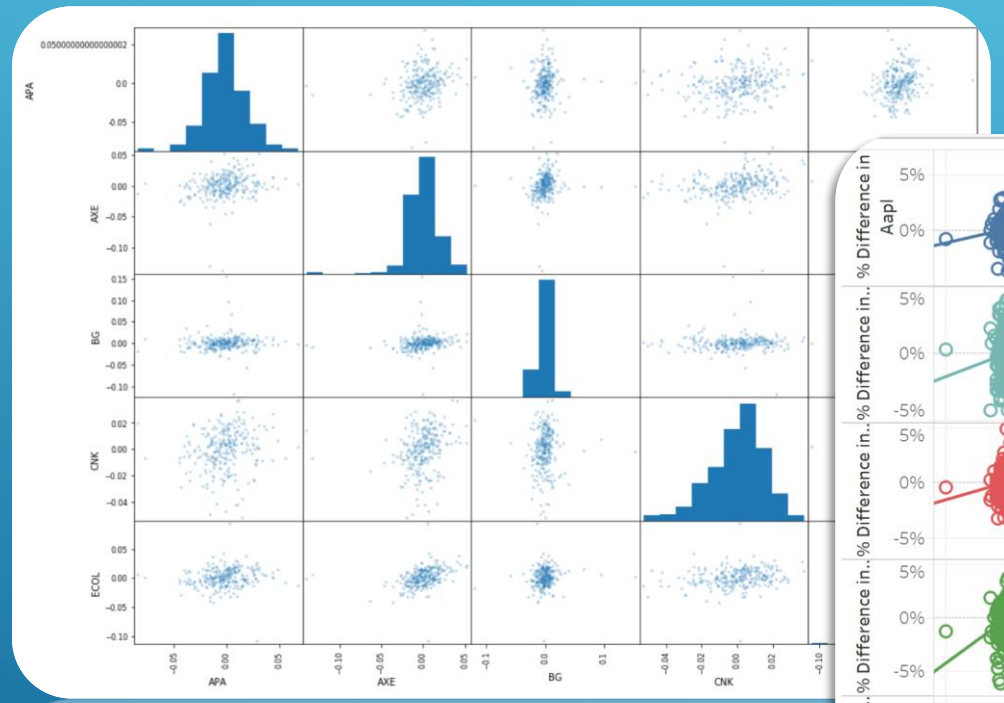
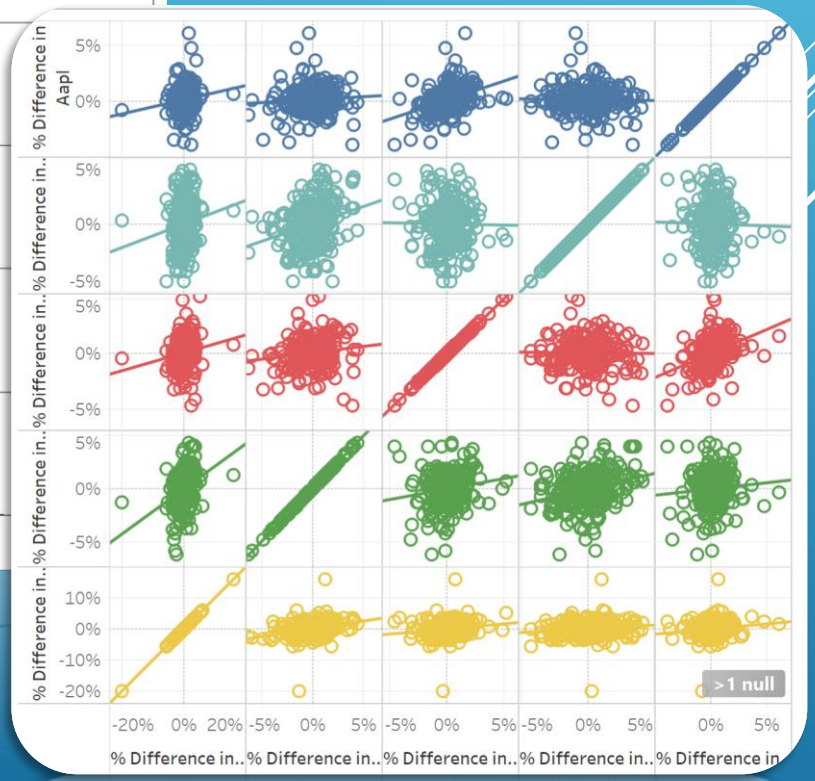
clusterM3 TickerM3 company M3 [5] APA Camden Property Trust [5] AXE CoreSite Realty Corporation [5] BG Duke Realty Corporation [5] CHFN Investors Real Estate Trust [5] CNK Liberty Property Trust [5] ECOL Senior Housing Properties Trust



Date	APA	AXE
2017-01-03	63.779999	81.900002
2017-01-04	63.590000	83.400002
2017-01-05	63.250000	82.199997
2017-01-06	63.119999	81.199997
2017-01-09	62.360001	79.900002

Date	APA	AXE	BG	CNK	ECOL
2017-01-04	-0.002983	0.018149	-0.013333	0.006054	0.008089
2017-01-05	-0.005361	-0.014493	-0.016322	0.026311	0.002012
2017-01-06	-0.002057	-0.012240	-0.013594	-0.011333	0.000000
2017-01-09	-0.012114	-0.016139	0.000000	-0.010461	-0.014170

	APA	AXE	BG	CNK	ECOL
APA	1.000000	0.207176	0.045573	0.147253	0.115081
AXE	0.207176	1.000000	0.148788	0.227729	0.356446
BG	0.045573	0.148788	1.000000	0.011784	0.105265
CNK	0.147253	0.227729	0.011784	1.000000	0.198824
ECOL	0.115081	0.356446	0.105265	0.198824	1.000000



Do the stocks prices trend as a group…. Yes: we have a useful cluster
No:….oh oh!