

# Chiebuka Onwuzurike MSBA '21

## BA870 – Final Project #4 (June 2021)

### Main Deliverables:

- IPYNB with code and analysis
- PDF of file
- Presentation
- PDF of Presentation
- CSV file of data

### Goals:

- Comparing predictive power of SIC, NAICS, and GICS classification to Early and Late 2020 Returns
- Determine most predictive industry classifier
- Determining which industries were hit the hardest and why.
- Determining which industries were bounced the most and why.
- Calculate metrics and ratios
- Determine most predictive metrics and ratios

### Importing Packages

```
In [301]: # pip install statsmodels
```

```
In [302]: import pandas as pd
import statsmodels.api as sm
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np
from scipy.stats import mstats
from sklearn.preprocessing import StandardScaler
```

### Importing Data

```
In [303]: returns2020 = pd.read_csv("returns2020.csv")
classifications = pd.read_csv("classifications.csv")
metrics = pd.read_csv("metrics.csv")
gic_sub_industry = pd.read_csv("gic_sub_industry.csv")
```

```
In [304]: returns2020.head()
```

```
Out[304]:
```

	tic	RetEarly2020	RetLate2020
0	JJSF	-0.340234	0.300340
1	PLXS	-0.290876	0.433469
2	HNGR	-0.435712	0.411425
3	ORCL	-0.083596	0.356166
4	MSFT	0.003212	0.421139

## Merging Data

```
In [305]: # Merging Data
stock_class_test = pd.merge(returns2020, classifications, on="tic", how="inner")
stock_class_test.to_csv(r'stock_class_test.csv', index = False, header=True)
```

## Classification Exploratory Data Analysis

```
In [306]: # stock_class_test.info()
# stock_class_test.describe().T
```

```

In [307]: classifier = {}

for classification in stock_class_test.iloc[:,3:].columns:
    classifier[classification] = len(stock_class_test[classification].unique
    ())
    # print(f'There are {len(stock_class_test[classification].unique())} unique
    e groups in {classification}')

classifier = pd.DataFrame(classifier, index = ["classifiers"])
classifier = classifier.T
classifier.sort_values("classifiers")
# classifier

```

Out[307]:

	classifiers
<b>gic_sector</b>	11
<b>naics_sector</b>	23
<b>gic_group</b>	24
<b>sic_sector</b>	64
<b>gic_industry</b>	69
<b>gic_sub_industry</b>	154
<b>naics_group</b>	232
<b>sic_group</b>	232
<b>sic_industry</b>	349
<b>naics_industry</b>	516

## Creating Dummy Variables

```
In [308]: df_gic_sub_industry = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.gic_sub_industry)], axis=1)
df_gic_industry = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.gic_industry)], axis=1)
df_gic_group = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.gic_group)], axis=1)
df_gic_sector = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.gic_sector)], axis=1)
df_naics_industry = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.naics_industry)], axis=1)
df_naics_group = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.naics_group)], axis=1)
df_naics_sector = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.naics_sector)], axis=1)
df_sic_industry = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.sic_industry)], axis=1)
df_sic_group = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.sic_group)], axis=1)
df_sic_sector = pd.concat([stock_class_test, pd.get_dummies(stock_class_test.sic_sector)], axis=1)
```

## Checking Adj R<sup>2</sup>

The best industry classification will have the highest predictive power. We will look at the Adj R<sup>2</sup> to see if more granular industry classifications systems help or hurt the model.

The Adj R<sup>2</sup> seems to differ depending if it is RetEarly2020 or RetLate2020 thus I will run both

```

In [309]: early_gic_sub_industry = sm.OLS(df_gic_sub_industry.iloc[:,1:2],df_gic_sub_ind
          ustry.iloc[:,13:]).fit()
          early_gic_industry = sm.OLS(df_gic_industry.iloc[:,1:2],df_gic_industry.iloc
         [:,13:]).fit()
          early_gic_group = sm.OLS(df_gic_group.iloc[:,1:2],df_gic_group.iloc[:,13:]).fi
          t()
          early_gic_sector = sm.OLS(df_gic_sector.iloc[:,1:2],df_gic_sector.iloc[:,13:]).
          fit()
          early_naics_industry = sm.OLS(df_naics_industry.iloc[:,1:2],df_naics_industry.
          iloc[:,13:]).fit()
          early_naics_group = sm.OLS(df_naics_group.iloc[:,1:2],df_naics_group.iloc[:,13
          :]).fit()
          early_naics_sector = sm.OLS(df_naicscs_sector.iloc[:,1:2],df_naicscs_sector.ilo
          c[:,13:]).fit()
          early_sic_industry = sm.OLS(df_sic_industry.iloc[:,1:2],df_sic_industry.iloc
         [:,13:]).fit()
          early_sic_group = sm.OLS(df_sic_group.iloc[:,1:2],df_sic_group.iloc[:,13:]).fi
          t()
          early_sic_sector = sm.OLS(df_sic_sector.iloc[:,1:2],df_sic_sector.iloc[:,13:]).
          fit()

          late_gic_sub_industry = sm.OLS(df_gic_sub_industry.iloc[:,2:3],df_gic_sub_indu
          stry.iloc[:,13:]).fit()
          late_gic_industry = sm.OLS(df_gic_industry.iloc[:,2:3],df_gic_industry.iloc[:,
          13:]).fit()
          late_gic_group = sm.OLS(df_gic_group.iloc[:,2:3],df_gic_group.iloc[:,13:]).fit
          ()
          late_gic_sector = sm.OLS(df_gic_sector.iloc[:,2:3],df_gic_sector.iloc[:,13:]).
          fit()
          late_naics_industry = sm.OLS(df_naics_industry.iloc[:,2:3],df_naics_industry.i
          loc[:,13:]).fit()
          late_naics_group = sm.OLS(df_naics_group.iloc[:,2:3],df_naics_group.iloc[:,13
          :]).fit()
          late_naics_sector = sm.OLS(df_naicscs_sector.iloc[:,2:3],df_naicscs_sector.ilo
          c[:,13:]).fit()
          late_sic_industry = sm.OLS(df_sic_industry.iloc[:,2:3],df_sic_industry.iloc[:,
          13:]).fit()
          late_sic_group = sm.OLS(df_sic_group.iloc[:,2:3],df_sic_group.iloc[:,13:]).fit
          ()
          late_sic_sector = sm.OLS(df_sic_sector.iloc[:,2:3],df_sic_sector.iloc[:,13:]).
          fit()

```

```

In [310]: adjR2_early = []
          for model in [early_gic_sub_industry, early_gic_industry, early_gic_group, early_gic_sector, early_naics_industry, early_naics_group, early_naics_sector, early_sic_industry, early_sic_group, early_sic_sector]:
              adjR2_early.append(model.rsquared_adj)

          adjR2_late = []
          for model in [late_gic_sub_industry, late_gic_industry, late_gic_group, late_gic_sector, late_naics_industry, late_naics_group, late_naics_sector, late_sic_industry, late_sic_group, late_sic_sector]:
              adjR2_late.append(model.rsquared_adj)

          classifier["adjR2_early"] = adjR2_early
          classifier["adjR2_late"] = adjR2_late
          classifier["adjR2_mean"] = (classifier.adjR2_early + classifier.adjR2_late)/2

```

```

In [311]: classifier.sort_values("adjR2_mean", ascending = False)

```

Out[311]:

	classifiers	adjR2_early	adjR2_late	adjR2_mean
<b>gic_sub_industry</b>	154	0.158047	0.175685	0.166866
<b>gic_industry</b>	69	0.146104	0.132098	0.139101
<b>naics_industry</b>	516	0.071038	0.186942	0.128990
<b>naics_group</b>	232	0.111412	0.144024	0.127718
<b>sic_industry</b>	349	0.107587	0.145237	0.126412
<b>sic_group</b>	232	0.106741	0.125953	0.116347
<b>gic_group</b>	24	0.122970	0.101845	0.112407
<b>sic_sector</b>	64	0.105715	0.109146	0.107431
<b>gic_sector</b>	11	0.115544	0.090513	0.103029
<b>naics_sector</b>	23	0.075383	0.072564	0.073974

## Analyzing Adj R<sup>2</sup> results

For Early 2020 Returns the top three were industry classifiers were:

- gic\_sub\_industry
- gic\_industry
- gic\_group

Very interesting to see as the GIS classifier got more granular the better it performed

For Late 2020 Returns the top three were industry classifiers were:

- naics\_industry
- gic\_sub\_industry
- sic\_industry

Very interesting to see Returns increase closer to normal the more granular classifiers started to perform better.

## Deciding which classifiers is best

Although naics\_industry was the highest for Late 2020 Returns which covered 9 months of the year, gic\_sub\_industry was more robust for all 2020 returns place highest for Early 2020 Returns and second highest Late 2020 Returns.

Additionally when looking at mean Adj R<sup>2</sup> for early and late Returns gic\_sub\_industry was the highest.

The industry classifier of choice is gic\_sub\_industry

## Linear Regression on GIC Sub Industry Classifier

```
In [312]: stock_class = pd.merge(stock_class_test.iloc[:,4], gic_sub_industry, on="gic_sub_industry", how="inner")
stock_class = pd.concat([stock_class, pd.get_dummies(stock_class.gic_sub_industry_name)], axis =1)
stock_class.to_csv(r'stock_class.csv', index = False, header=True)

stock_class.sample(3)
```

Out[312]:

	tic	RetEarly2020	RetLate2020	gic_sub_industry	gic_sub_industry_name	Aerospace & Defense	A
1450	VICR	-0.046661	1.070497	20104010	Electrical Components & Equipment	0	
2173	ANAT	-0.292998	0.204583	40301030	Multi-line Insurance	0	
22	CAG	-0.137576	0.260400	30202030	Packaged Foods & Meats	0	

3 rows × 151 columns

## Determining Industries that were Hit the Hardest

### Early 2020 Returns



```
In [313]: # Dependent Variables
X = stock_class.iloc[:,5:]

# Independent Variables
Y = stock_class.iloc[:,1:2]

# Fit Model
early_class = sm.OLS(Y, X).fit()

# Predict Model
# predictions = model.predict(X)

# Summary Statistics
print(early_class.summary(title="Early 2020 Returns with GIC Sub Industry"))
early_class_results = pd.read_html(early_class.summary().tables[1].as_html(),h
eader=0,index_col=0)[0]
```

## Early 2020 Returns with GIC Sub Industry

```

=====
=
Dep. Variable:          RetEarly2020    R-squared:                0.20
4
Model:                  OLS             Adj. R-squared:           0.15
6
Method:                 Least Squares   F-statistic:              4.29
8
Date:                   Thu, 08 Jul 2021 Prob (F-statistic):       7.90e-5
2
Time:                   14:36:09        Log-Likelihood:          -390.1
7
No. Observations:      2579            AIC:                      107
2.
Df Residuals:          2433            BIC:                      192
7.
Df Model:               145
Covariance Type:       nonrobust
=====
=====

```

```

=====
                                coef    std err          t
P>|t|    [0.025    0.975]
-----
Aerospace & Defense            -0.2897      0.053     -5.474
0.000    -0.393    -0.186
Agricultural & Farm Machinery   -0.2019      0.145     -1.394
0.164    -0.486      0.082
Agricultural Products          -0.2385      0.110     -2.177
0.030    -0.453     -0.024
Air Freight & Logistics         -0.2021      0.087     -2.313
0.021    -0.373     -0.031
Airlines                       -0.5674      0.084     -6.782
0.000    -0.731     -0.403
Airport Services               -0.3873      0.290     -1.336
0.182    -0.956      0.181
Alternative Carriers           -0.1040      0.118     -0.879
0.380    -0.336      0.128
Aluminum                      -0.5343      0.167     -3.193
0.001    -0.862     -0.206
Apparel Retail                 -0.5281      0.060     -8.739
0.000    -0.647     -0.410
Apparel, Accessories & Luxury Goods -0.5019      0.070     -7.141
0.000    -0.640     -0.364
Application Software           -0.1598      0.029     -5.458
0.000    -0.217     -0.102
Asset Management & Custody Banks -0.3173      0.052     -6.096
0.000    -0.419     -0.215
Auto Parts & Equipment          -0.4486      0.065     -6.922
0.000    -0.576     -0.322
Automobile Manufacturers       -0.3253      0.118     -2.750
0.006    -0.557     -0.093
Automotive Retail              -0.4363      0.075     -5.830
0.000    -0.583     -0.290
Biotechnology                  -0.1123      0.020     -5.671
0.000    -0.151     -0.073

```

Brewers			-0.2042	0.167	-1.221
0.222	-0.532	0.124			
Building Products			-0.2189	0.055	-3.997
0.000	-0.326	-0.112			
Casinos & Gaming			-0.5617	0.075	-7.506
0.000	-0.708	-0.415			
Coal & Consumable Fuels			-0.5318	0.118	-4.494
0.000	-0.764	-0.300			
Commercial Printing			-0.3657	0.130	-2.821
0.005	-0.620	-0.112			
Commodity Chemicals			-0.4284	0.075	-5.725
0.000	-0.575	-0.282			
Communications Equipment			-0.2465	0.052	-4.737
0.000	-0.349	-0.144			
Computer & Electronics Retail			-0.4841	0.145	-3.341
0.001	-0.768	-0.200			
Construction & Engineering			-0.3075	0.060	-5.089
0.000	-0.426	-0.189			
Construction Machinery & Heavy Trucks			-0.3618	0.065	-5.582
0.000	-0.489	-0.235			
Construction Materials			-0.3607	0.110	-3.293
0.001	-0.575	-0.146			
Consumer Electronics			-0.3389	0.130	-2.615
0.009	-0.593	-0.085			
Consumer Finance			-0.3838	0.060	-6.352
0.000	-0.502	-0.265			
Copper			-0.4054	0.205	-1.978
0.048	-0.807	-0.004			
Data Processing & Outsourced Services			-0.3124	0.049	-6.376
0.000	-0.408	-0.216			
Department Stores			-0.6250	0.145	-4.313
0.000	-0.909	-0.341			
Distillers & Vintners			-0.2976	0.167	-1.779
0.075	-0.626	0.031			
Distributors			-0.2993	0.118	-2.530
0.011	-0.531	-0.067			
Diversified Banks			-0.4359	0.118	-3.684
0.000	-0.668	-0.204			
Diversified Chemicals			-0.4338	0.167	-2.593
0.010	-0.762	-0.106			
Diversified Metals & Mining			-0.3851	0.205	-1.879
0.060	-0.787	0.017			
Diversified REITs			-0.3548	0.075	-4.742
0.000	-0.502	-0.208			
Diversified Real Estate Activities			-0.2261	0.167	-1.351
0.177	-0.554	0.102			
Diversified Support Services			-0.3646	0.097	-3.774
0.000	-0.554	-0.175			
Drug Retail			-0.1233	0.205	-0.602
0.547	-0.525	0.279			
Education Services			-0.2483	0.080	-3.089
0.002	-0.406	-0.091			
Electric Utilities			-0.1707	0.057	-3.004
0.003	-0.282	-0.059			
Electrical Components & Equipment			-0.2931	0.059	-4.955
0.000	-0.409	-0.177			
Electronic Components			-0.3190	0.092	-3.481

0.001	-0.499	-0.139			
Electronic Equipment & Instruments			-0.2977	0.065	-4.594
0.000	-0.425	-0.171			
Electronic Manufacturing Services			-0.2895	0.092	-3.159
0.002	-0.469	-0.110			
Environmental & Facilities Services			-0.3419	0.072	-4.719
0.000	-0.484	-0.200			
Fertilizers & Agricultural Chemicals			-0.3275	0.110	-2.990
0.003	-0.542	-0.113			
Financial Exchanges & Data			-0.1190	0.080	-1.480
0.139	-0.277	0.039			
Food Distributors			-0.3813	0.102	-3.721
0.000	-0.582	-0.180			
Food Retail			-0.0461	0.102	-0.450
0.653	-0.247	0.155			
Footwear			-0.3963	0.110	-3.618
0.000	-0.611	-0.181			
Forest Products			-0.4180	0.290	-1.442
0.149	-0.986	0.150			
Gas Utilities			-0.1526	0.087	-1.746
0.081	-0.324	0.019			
General Merchandise Stores			-0.3213	0.118	-2.716
0.007	-0.553	-0.089			
Gold			-0.3034	0.130	-2.341
0.019	-0.558	-0.049			
Health Care Services			-0.1722	0.057	-3.030
0.002	-0.284	-0.061			
Health Care Distributors			0.0298	0.097	0.309
0.758	-0.160	0.219			
Health Care Equipment			-0.0903	0.034	-2.626
0.009	-0.158	-0.023			
Health Care Facilities			-0.3508	0.075	-4.687
0.000	-0.497	-0.204			
Health Care REITs			-0.3149	0.072	-4.346
0.000	-0.457	-0.173			
Health Care Supplies			-0.1631	0.059	-2.757
0.006	-0.279	-0.047			
Health Care Technology			-0.0897	0.066	-1.350
0.177	-0.220	0.041			
Heavy Electrical Equipment			-0.2956	0.167	-1.767
0.077	-0.624	0.032			
Home Furnishings			-0.4471	0.102	-4.363
0.000	-0.648	-0.246			
Home Improvement Retail			-0.2751	0.130	-2.122
0.034	-0.529	-0.021			
Homebuilding			-0.4107	0.065	-6.338
0.000	-0.538	-0.284			
Homefurnishing Retail			-0.5350	0.110	-4.884
0.000	-0.750	-0.320			
Hotel & Resort REITs			-0.5776	0.077	-7.457
0.000	-0.729	-0.426			
Hotels, Resorts & Cruise Lines			-0.5550	0.075	-7.416
0.000	-0.702	-0.408			
Household Appliances			-0.3256	0.145	-2.247
0.025	-0.610	-0.041			
Household Products			-0.1140	0.092	-1.243
0.214	-0.294	0.066			

Housewares & Specialties	-0.4306	0.167	-2.574
0.010 -0.759 -0.103			
Human Resource & Employment Services	-0.4534	0.077	-5.854
0.000 -0.605 -0.302			
Hypermarkets & Super Centers	-0.0506	0.145	-0.349
0.727 -0.335 0.234			
IT Consulting & Other Services	-0.2342	0.080	-2.914
0.004 -0.392 -0.077			
Independent Power Producers & Energy Traders	-0.1881	0.145	-1.298
0.194 -0.472 0.096			
Industrial Conglomerates	-0.2449	0.118	-2.070
0.039 -0.477 -0.013			
Industrial Gases	-0.1639	0.205	-0.800
0.424 -0.566 0.238			
Industrial Machinery	-0.3464	0.037	-9.486
0.000 -0.418 -0.275			
Industrial REITs	-0.1359	0.080	-1.690
0.091 -0.293 0.022			
Insurance Brokers	-0.0784	0.102	-0.765
0.444 -0.279 0.123			
Integrated Oil & Gas	-0.5123	0.167	-3.062
0.002 -0.840 -0.184			
Integrated Telecommunication Services	-0.0103	0.110	-0.094
0.925 -0.225 0.204			
Internet & Direct Marketing Retail	-0.0977	0.066	-1.469
0.142 -0.228 0.033			
Investment Banking & Brokerage	-0.2534	0.063	-4.006
0.000 -0.377 -0.129			
Leisure Facilities	-0.5211	0.130	-4.020
0.000 -0.775 -0.267			
Leisure Products	-0.2456	0.070	-3.494
0.000 -0.383 -0.108			
Life & Health Insurance	-0.3642	0.070	-5.181
0.000 -0.502 -0.226			
Life Sciences Tools & Services	-0.1718	0.054	-3.193
0.001 -0.277 -0.066			
Managed Health Care	-0.1920	0.097	-1.987
0.047 -0.381 -0.003			
Marine	-0.4259	0.118	-3.600
0.000 -0.658 -0.194			
Metal & Glass Containers	-0.2323	0.092	-2.535
0.011 -0.412 -0.053			
Mortgage REITs	-0.5648	0.053	-10.673
0.000 -0.669 -0.461			
Motorcycle Manufacturers	-0.4808	0.290	-1.659
0.097 -1.049 0.088			
Multi-Sector Holdings	-0.0995	0.290	-0.343
0.731 -0.668 0.469			
Multi-Utilities	-0.1675	0.075	-2.239
0.025 -0.314 -0.021			
Multi-line Insurance	-0.3376	0.110	-3.082
0.002 -0.552 -0.123			
Office REITs	-0.2745	0.062	-4.442
0.000 -0.396 -0.153			
Office Services & Supplies	-0.3784	0.087	-4.330
0.000 -0.550 -0.207			
Oil & Gas Drilling	-0.7784	0.145	-5.372

0.000	-1.063	-0.494			
Oil & Gas Equipment & Services			-0.6312	0.059	-10.670
0.000	-0.747	-0.515			
Oil & Gas Exploration & Production			-0.6594	0.048	-13.652
0.000	-0.754	-0.565			
Oil & Gas Refining & Marketing			-0.5166	0.080	-6.427
0.000	-0.674	-0.359			
Oil & Gas Storage & Transportation			-0.3862	0.068	-5.653
0.000	-0.520	-0.252			
Other Diversified Financial Services			-0.3381	0.167	-2.021
0.043	-0.666	-0.010			
Packaged Foods & Meats			-0.1535	0.050	-3.089
0.002	-0.251	-0.056			
Paper Packaging			-0.2666	0.092	-2.909
0.004	-0.446	-0.087			
Paper Products			-0.3040	0.118	-2.569
0.010	-0.536	-0.072			
Personal Products			-0.3147	0.077	-4.062
0.000	-0.467	-0.163			
Pharmaceuticals			-0.2230	0.036	-6.155
0.000	-0.294	-0.152			
Property & Casualty Insurance			-0.2136	0.043	-4.998
0.000	-0.297	-0.130			
Railroads			-0.2079	0.145	-1.434
0.152	-0.492	0.076			
Real Estate Development			-0.3907	0.145	-2.696
0.007	-0.675	-0.106			
Real Estate Operating Companies			-0.2690	0.167	-1.608
0.108	-0.597	0.059			
Real Estate Services			-0.4420	0.092	-4.823
0.000	-0.622	-0.262			
Regional Banks			-0.3729	0.019	-20.142
0.000	-0.409	-0.337			
Reinsurance			-0.3250	0.118	-2.747
0.006	-0.557	-0.093			
Renewable Electricity			-0.0941	0.205	-0.459
0.646	-0.496	0.308			
Research & Consulting Services			-0.2127	0.058	-3.670
0.000	-0.326	-0.099			
Residential REITs			-0.3083	0.066	-4.636
0.000	-0.439	-0.178			
Restaurants			-0.4356	0.050	-8.634
0.000	-0.535	-0.337			
Retail REITs			-0.4928	0.056	-8.836
0.000	-0.602	-0.383			
Security & Alarm Services			-0.3964	0.167	-2.369
0.018	-0.725	-0.068			
Semiconductor Equipment			-0.2219	0.060	-3.672
0.000	-0.340	-0.103			
Semiconductors			-0.2292	0.046	-4.939
0.000	-0.320	-0.138			
Silver			-0.4624	0.290	-1.595
0.111	-1.031	0.106			
Soft Drinks			-0.1717	0.102	-1.675
0.094	-0.373	0.029			
Specialized Consumer Services			-0.3668	0.102	-3.580
0.000	-0.568	-0.166			

Specialized Finance			-0.4978	0.205	-2.429
0.015	-0.900	-0.096			
Specialized REITs			-0.1622	0.053	-3.066
0.002	-0.266	-0.058			
Specialty Chemicals			-0.3284	0.052	-6.309
0.000	-0.430	-0.226			
Specialty Stores			-0.4215	0.080	-5.244
0.000	-0.579	-0.264			
Steel			-0.4384	0.072	-6.050
0.000	-0.580	-0.296			
Systems Software			-0.1015	0.054	-1.887
0.059	-0.207	0.004			
Technology Distributors			-0.3520	0.102	-3.436
0.001	-0.553	-0.151			
Technology Hardware, Storage & Peripherals			-0.3164	0.077	-4.084
0.000	-0.468	-0.164			
Textiles			-0.5428	0.290	-1.873
0.061	-1.111	0.026			
Thrifts & Mortgage Finance			-0.3352	0.043	-7.759
0.000	-0.420	-0.250			
Tires & Rubber			-0.5253	0.205	-2.563
0.010	-0.927	-0.123			
Tobacco			-0.2186	0.130	-1.687
0.092	-0.473	0.036			
Trading Companies & Distributors			-0.3992	0.052	-7.669
0.000	-0.501	-0.297			
Trucking			-0.2319	0.068	-3.395
0.001	-0.366	-0.098			
Water Utilities			-0.0577	0.087	-0.660
0.509	-0.229	0.114			
Wireless Telecommunication Services			-0.1554	0.110	-1.419
0.156	-0.370	0.059			

=====

=

Omnibus:	4357.217	Durbin-Watson:	2.07
6			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	6401595.38
9			
Skew:	11.035	Prob(JB):	0.0
0			
Kurtosis:	246.076	Cond. No.	15.
7			

=====

=

## Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

```
In [314]: early_class_results.loc[early_class_results['P>|t|']<0.05,].sort_values("coef",
, ascending=False)
```

Out[314]:

	coef	std err	t	P> t	[0.025	0.975]
<b>Health Care Equipment</b>	-0.0903	0.034	-2.626	0.009	-0.158	-0.023
<b>Biotechnology</b>	-0.1123	0.020	-5.671	0.000	-0.151	-0.073
<b>Packaged Foods &amp; Meats</b>	-0.1535	0.050	-3.089	0.002	-0.251	-0.056
<b>Application Software</b>	-0.1598	0.029	-5.458	0.000	-0.217	-0.102
<b>Specialized REITs</b>	-0.1622	0.053	-3.066	0.002	-0.266	-0.058
...	...	...	...	...	...	...
<b>Hotel &amp; Resort REITs</b>	-0.5776	0.077	-7.457	0.000	-0.729	-0.426
<b>Department Stores</b>	-0.6250	0.145	-4.313	0.000	-0.909	-0.341
<b>Oil &amp; Gas Equipment &amp; Services</b>	-0.6312	0.059	-10.670	0.000	-0.747	-0.515
<b>Oil &amp; Gas Exploration &amp; Production</b>	-0.6594	0.048	-13.652	0.000	-0.754	-0.565
<b>Oil &amp; Gas Drilling</b>	-0.7784	0.145	-5.372	0.000	-1.063	-0.494

111 rows × 6 columns



## Early 2020 Return Class Explaintion

**Adj R<sup>2</sup>: 0.156**

The industries hit the hardest at the 95% statisitcal signifiant level were:

- Oil & Gas Drilling
- Oil & Gas Exploration & Production
- Oil & Gas Equipment & Services
- Department Stores
- Hotel & Resort REITs
- Airlines
- Mortgage REITs
- Casinos & Gaming
- Hotels, Resorts & Cruise Lines
- Homefurnishing Retail

This makes sense seeing as people stopped drive and demand for gas declined, people stopped going out to shop, travel and rereational activites were halted, and many people were behind on their rent.

The industries least affected at the 95% statisitcal signifiant level were:

- Health Care Services
- Life Sciences Tools & Services
- Electric Utilities
- Multi-Utilities
- Health Care Supplies
- Specialized REITs
- Application Software
- Packaged Foods & Meats
- Biotechnology
- Health Care Equipment

This makes sense seeing as mask and other health care and related tools were high in demand, everyone was in their homes using alot of utilities, and people need software and packaged good to work from home.

## Determining Industries that Bounced Back the Most

### Late 2020 Returns

```
In [315]: # Dependent Variables
X = stock_class.iloc[:,5:]

# Independent Variables
Y = stock_class.iloc[:,2:3]

# Fit Model
late_class = sm.OLS(Y, X).fit()

# Predict Model
# predictions = model.predict(X)

# Summary Statistics
print(late_class.summary(title="Late 2020 Returns GIC Sub Industry"))
late_class_results = pd.read_html(late_class.summary().tables[1].as_html(), header=0, index_col=0)[0]
```

## Late 2020 Returns GIC Sub Industry

```

=====
=
Dep. Variable:          RetLate2020    R-squared:                0.22
5
Model:                  OLS            Adj. R-squared:          0.17
9
Method:                 Least Squares   F-statistic:             4.88
2
Date:                   Thu, 08 Jul 2021 Prob (F-statistic):       3.51e-6
3
Time:                   14:36:10        Log-Likelihood:          -3531.
0
No. Observations:       2579           AIC:                     735
4.
Df Residuals:           2433           BIC:                     820
9.
Df Model:                145
Covariance Type:        nonrobust
=====
=====

```

```

=====
                                coef    std err          t
P>|t|    [0.025    0.975]
-----
-----
Aerospace & Defense           0.5213      0.179      2.915
0.004      0.171      0.872
Agricultural & Farm Machinery   0.7632      0.490      1.558
0.119     -0.197      1.724
Agricultural Products          0.4833      0.370      1.305
0.192     -0.243      1.209
Air Freight & Logistics         0.7299      0.295      2.471
0.014      0.151      1.309
Airlines                       0.6681      0.283      2.363
0.018      0.114      1.223
Airport Services               0.4871      0.980      0.497
0.619     -1.434      2.408
Alternative Carriers           0.6337      0.400      1.585
0.113     -0.150      1.418
Aluminum                       1.7536      0.566      3.101
0.002      0.645      2.863
Apparel Retail                 1.2599      0.204      6.169
0.000      0.859      1.660
Apparel, Accessories & Luxury Goods 1.1218      0.238      4.722
0.000      0.656      1.588
Application Software           1.2274      0.099     12.404
0.000      1.033      1.421
Asset Management & Custody Banks 0.6512      0.176      3.701
0.000      0.306      0.996
Auto Parts & Equipment          1.4651      0.219      6.689
0.000      1.036      1.895
Automobile Manufacturers       3.3169      0.400      8.294
0.000      2.533      4.101
Automotive Retail              1.4670      0.253      5.800
0.000      0.971      1.963
Biotechnology                  0.7958      0.067     11.884
0.000      0.664      0.927

```

Brewers			0.7644	0.566	1.352
0.177	-0.345	1.873			
Building Products			0.9148	0.185	4.942
0.000	0.552	1.278			
Casinos & Gaming			1.9196	0.253	7.590
0.000	1.424	2.416			
Coal & Consumable Fuels			1.0025	0.400	2.507
0.012	0.218	1.787			
Commercial Printing			0.3018	0.438	0.689
0.491	-0.557	1.161			
Commodity Chemicals			1.0489	0.253	4.147
0.000	0.553	1.545			
Communications Equipment			0.8414	0.176	4.782
0.000	0.496	1.186			
Computer & Electronics Retail			2.1936	0.490	4.479
0.000	1.233	3.154			
Construction & Engineering			0.9135	0.204	4.473
0.000	0.513	1.314			
Construction Machinery & Heavy Trucks			0.7429	0.219	3.392
0.001	0.313	1.172			
Construction Materials			0.8005	0.370	2.162
0.031	0.074	1.526			
Consumer Electronics			1.6776	0.438	3.830
0.000	0.819	2.537			
Consumer Finance			0.8213	0.204	4.021
0.000	0.421	1.222			
Copper			2.1096	0.693	3.046
0.002	0.751	3.468			
Data Processing & Outsourced Services			0.7080	0.166	4.276
0.000	0.383	1.033			
Department Stores			1.2105	0.490	2.472
0.014	0.250	2.171			
Distillers & Vintners			0.6014	0.566	1.063
0.288	-0.508	1.710			
Distributors			0.6017	0.400	1.505
0.133	-0.182	1.386			
Diversified Banks			0.4700	0.400	1.175
0.240	-0.314	1.254			
Diversified Chemicals			1.3004	0.566	2.299
0.022	0.191	2.409			
Diversified Metals & Mining			0.7497	0.693	1.082
0.279	-0.609	2.108			
Diversified REITs			0.3497	0.253	1.383
0.167	-0.146	0.846			
Diversified Real Estate Activities			0.6844	0.566	1.210
0.226	-0.425	1.793			
Diversified Support Services			0.7309	0.327	2.238
0.025	0.091	1.371			
Drug Retail			-0.0209	0.693	-0.030
0.976	-1.379	1.337			
Education Services			0.4606	0.272	1.695
0.090	-0.072	0.993			
Electric Utilities			0.1351	0.192	0.704
0.482	-0.242	0.512			
Electrical Components & Equipment			1.5759	0.200	7.882
0.000	1.184	1.968			
Electronic Components			0.7691	0.310	2.483

0.013	0.162	1.377			
Electronic Equipment & Instruments			0.9607	0.219	4.386
0.000	0.531	1.390			
Electronic Manufacturing Services			0.4826	0.310	1.558
0.119	-0.125	1.090			
Environmental & Facilities Services			0.5324	0.245	2.174
0.030	0.052	1.013			
Fertilizers & Agricultural Chemicals			0.8314	0.370	2.246
0.025	0.105	1.557			
Financial Exchanges & Data			0.5419	0.272	1.995
0.046	0.009	1.075			
Food Distributors			0.6577	0.346	1.899
0.058	-0.021	1.337			
Food Retail			0.2268	0.346	0.655
0.513	-0.452	0.906			
Footwear			1.0199	0.370	2.755
0.006	0.294	1.746			
Forest Products			1.1957	0.980	1.221
0.222	-0.725	3.117			
Gas Utilities			0.0047	0.295	0.016
0.987	-0.574	0.584			
General Merchandise Stores			1.2193	0.400	3.049
0.002	0.435	2.004			
Gold			0.6329	0.438	1.445
0.149	-0.226	1.492			
Health Care Services			0.7569	0.192	3.940
0.000	0.380	1.134			
Health Care Distributors			0.9316	0.327	2.853
0.004	0.291	1.572			
Health Care Equipment			0.6433	0.116	5.534
0.000	0.415	0.871			
Health Care Facilities			1.1691	0.253	4.622
0.000	0.673	1.665			
Health Care REITs			0.3856	0.245	1.575
0.115	-0.095	0.866			
Health Care Supplies			0.6661	0.200	3.332
0.001	0.274	1.058			
Health Care Technology			0.8632	0.225	3.841
0.000	0.422	1.304			
Heavy Electrical Equipment			2.5883	0.566	4.577
0.000	1.479	3.697			
Home Furnishings			2.1649	0.346	6.251
0.000	1.486	2.844			
Home Improvement Retail			3.6681	0.438	8.373
0.000	2.809	4.527			
Homebuilding			1.2077	0.219	5.514
0.000	0.778	1.637			
Homefurnishing Retail			3.0359	0.370	8.200
0.000	2.310	3.762			
Hotel & Resort REITs			0.7721	0.262	2.949
0.003	0.259	1.285			
Hotels, Resorts & Cruise Lines			1.0180	0.253	4.025
0.000	0.522	1.514			
Household Appliances			0.8836	0.490	1.804
0.071	-0.077	1.844			
Household Products			0.3675	0.310	1.186
0.236	-0.240	0.975			

Housewares & Specialties	7.1362	0.566	12.618
0.000 6.027 8.245			
Human Resource & Employment Services	1.0585	0.262	4.043
0.000 0.545 1.572			
Hypermarkets & Super Centers	0.4639	0.490	0.947
0.344 -0.497 1.424			
IT Consulting & Other Services	0.6837	0.272	2.517
0.012 0.151 1.216			
Independent Power Producers & Energy Traders	0.6427	0.490	1.312
0.190 -0.318 1.603			
Industrial Conglomerates	0.4205	0.400	1.052
0.293 -0.364 1.205			
Industrial Gases	0.4656	0.693	0.672
0.501 -0.893 1.824			
Industrial Machinery	0.8078	0.123	6.546
0.000 0.566 1.050			
Industrial REITs	0.3810	0.272	1.402
0.161 -0.152 0.914			
Insurance Brokers	0.6170	0.346	1.782
0.075 -0.062 1.296			
Integrated Oil & Gas	0.2907	0.566	0.514
0.607 -0.818 1.400			
Integrated Telecommunication Services	0.2135	0.370	0.577
0.564 -0.512 0.940			
Internet & Direct Marketing Retail	2.3715	0.225	10.553
0.000 1.931 2.812			
Investment Banking & Brokerage	0.7473	0.214	3.496
0.000 0.328 1.167			
Leisure Facilities	1.6090	0.438	3.673
0.000 0.750 2.468			
Leisure Products	1.7170	0.238	7.227
0.000 1.251 2.183			
Life & Health Insurance	0.6807	0.238	2.865
0.004 0.215 1.147			
Life Sciences Tools & Services	1.1373	0.182	6.253
0.000 0.781 1.494			
Managed Health Care	0.4790	0.327	1.467
0.142 -0.161 1.119			
Marine	0.4576	0.400	1.144
0.253 -0.327 1.242			
Metal & Glass Containers	0.5500	0.310	1.776
0.076 -0.057 1.157			
Mortgage REITs	0.9187	0.179	5.137
0.000 0.568 1.269			
Motorcycle Manufacturers	0.9432	0.980	0.963
0.336 -0.978 2.864			
Multi-Sector Holdings	0.3219	0.980	0.329
0.742 -1.599 2.243			
Multi-Utilities	0.1016	0.253	0.402
0.688 -0.394 0.598			
Multi-line Insurance	0.4874	0.370	1.316
0.188 -0.239 1.213			
Office REITs	0.0937	0.209	0.449
0.654 -0.316 0.503			
Office Services & Supplies	0.5621	0.295	1.903
0.057 -0.017 1.141			
Oil & Gas Drilling	1.2062	0.490	2.463

0.014	0.246	2.167			
Oil & Gas Equipment & Services			1.1361	0.200	5.682
0.000	0.744	1.528			
Oil & Gas Exploration & Production			1.3508	0.163	8.274
0.000	1.031	1.671			
Oil & Gas Refining & Marketing			0.8435	0.272	3.105
0.002	0.311	1.376			
Oil & Gas Storage & Transportation			0.4028	0.231	1.744
0.081	-0.050	0.856			
Other Diversified Financial Services			0.6559	0.566	1.160
0.246	-0.453	1.765			
Packaged Foods & Meats			0.2308	0.168	1.374
0.170	-0.099	0.560			
Paper Packaging			0.5859	0.310	1.892
0.059	-0.021	1.193			
Paper Products			0.4224	0.400	1.056
0.291	-0.362	1.207			
Personal Products			0.7858	0.262	3.002
0.003	0.272	1.299			
Pharmaceuticals			0.3722	0.122	3.040
0.002	0.132	0.612			
Property & Casualty Insurance			0.1916	0.144	1.326
0.185	-0.092	0.475			
Railroads			0.5913	0.490	1.207
0.227	-0.369	1.552			
Real Estate Development			0.5021	0.490	1.025
0.305	-0.458	1.463			
Real Estate Operating Companies			0.4232	0.566	0.748
0.454	-0.686	1.532			
Real Estate Services			1.7138	0.310	5.533
0.000	1.106	2.321			
Regional Banks			0.4264	0.063	6.814
0.000	0.304	0.549			
Reinsurance			0.2297	0.400	0.574
0.566	-0.554	1.014			
Renewable Electricity			1.9111	0.693	2.759
0.006	0.553	3.269			
Research & Consulting Services			0.4729	0.196	2.414
0.016	0.089	0.857			
Residential REITs			0.3750	0.225	1.669
0.095	-0.066	0.816			
Restaurants			1.1377	0.171	6.672
0.000	0.803	1.472			
Retail REITs			0.4587	0.189	2.433
0.015	0.089	0.828			
Security & Alarm Services			0.2746	0.566	0.486
0.627	-0.834	1.384			
Semiconductor Equipment			1.2910	0.204	6.321
0.000	0.890	1.692			
Semiconductors			1.1758	0.157	7.496
0.000	0.868	1.483			
Silver			2.5711	0.980	2.625
0.009	0.650	4.492			
Soft Drinks			1.8464	0.346	5.332
0.000	1.167	2.526			
Specialized Consumer Services			0.6285	0.346	1.815
0.070	-0.051	1.308			

Specialized Finance			0.3916	0.693	0.565
0.572	-0.967	1.750			
Specialized REITs			0.3052	0.179	1.706
0.088	-0.046	0.656			
Specialty Chemicals			0.8974	0.176	5.101
0.000	0.552	1.242			
Specialty Stores			1.6473	0.272	6.063
0.000	1.115	2.180			
Steel			0.8779	0.245	3.585
0.000	0.398	1.358			
Systems Software			0.9220	0.182	5.069
0.000	0.565	1.279			
Technology Distributors			0.6220	0.346	1.796
0.073	-0.057	1.301			
Technology Hardware, Storage & Peripherals			0.9428	0.262	3.601
0.000	0.429	1.456			
Textiles			0.5359	0.980	0.547
0.584	-1.385	2.457			
Thrifts & Mortgage Finance			0.4694	0.146	3.214
0.001	0.183	0.756			
Tires & Rubber			1.1918	0.693	1.721
0.085	-0.166	2.550			
Tobacco			0.3819	0.438	0.872
0.383	-0.477	1.241			
Trading Companies & Distributors			0.9415	0.176	5.352
0.000	0.597	1.287			
Trucking			0.8479	0.231	3.673
0.000	0.395	1.301			
Water Utilities			0.0928	0.295	0.314
0.753	-0.486	0.672			
Wireless Telecommunication Services			0.6435	0.370	1.738
0.082	-0.083	1.369			

=====

=

Omnibus:	2211.150	Durbin-Watson:	2.10
1			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	122160.47
2			
Skew:	3.746	Prob(JB):	0.0
0			
Kurtosis:	35.874	Cond. No.	15.
7			

=====

=

## Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.



```
In [316]: late_class_results.loc[late_class_results['P>|t|']<0.05,].sort_values("coef",
ascending=False)
```

Out[316]:

	coef	std err	t	P> t	[0.025	0.975]
<b>Housewares &amp; Specialties</b>	7.1362	0.566	12.618	0.000	6.027	8.245
<b>Home Improvement Retail</b>	3.6681	0.438	8.373	0.000	2.809	4.527
<b>Automobile Manufacturers</b>	3.3169	0.400	8.294	0.000	2.533	4.101
<b>Homefurnishing Retail</b>	3.0359	0.370	8.200	0.000	2.310	3.762
<b>Heavy Electrical Equipment</b>	2.5883	0.566	4.577	0.000	1.479	3.697
...	...	...	...	...	...	...
<b>Research &amp; Consulting Services</b>	0.4729	0.196	2.414	0.016	0.089	0.857
<b>Thrifts &amp; Mortgage Finance</b>	0.4694	0.146	3.214	0.001	0.183	0.756
<b>Retail REITs</b>	0.4587	0.189	2.433	0.015	0.089	0.828
<b>Regional Banks</b>	0.4264	0.063	6.814	0.000	0.304	0.549
<b>Pharmaceuticals</b>	0.3722	0.122	3.040	0.002	0.132	0.612

84 rows × 6 columns

## Late 2020 Return Class Explaintion

**Adj R<sup>2</sup>: 0.179**

The industries bounced back the most at the 95% statisitcal signifiant level were:

- Housewares & Specialties,
- Home Improvement Retail
- Automobile Manufacturers
- Homefurnishing Retail
- Heavy Electrical Equipment
- Silver
- Internet & Direct Marketing Retail
- Computer & Electronics Retail
- Home Furnishings
- Copper

Putting money from stimulus checks into home and auto makes sense espically because of the low interest rates. Additonally focus on electrical, computer, and internet marketing makes sense with more people online. I can only attribute the rise in silver and copper becasue the are materials needed for tech and appliances

The industries bounced back the least at the 95% statisitcal signifiant level were:

- Asset Management & Custody Banks
- Health Care Equipment
- Financial Exchanges & Data
- Environmental & Facilities Services
- Aerospace & Defense
- Research & Consulting Services
- Thrifts & Mortgage Finance
- Retail REITs
- Regional Banks
- Pharmaceuticals

Some of these industries such as health care were already high and stayed high. With businesses suffering it was hard to justify paying for consultants, facility improvements, and other related business expenses. Some of the other categories I don't have great explanation for.

## Evaluating Metrics

```

In [317]: metrics["book_value"] = metrics.book_value_share * metrics.shares_out # Book Value

metrics["acid_ratio"] = (metrics.current_assets - metrics.inventory)/metrics.current_liabilities # Acid Ratio
metrics["current_ratio"] = metrics.current_assets / metrics.current_liabilities # Current Ratio

metrics["coverage_ratio"] = metrics.interest_exp / metrics.ebit # Interest Coverage Ratio
metrics['bep'] = metrics.ebit / metrics.total_assets # Basic Earnings Power Ratio

metrics["inventory_turnover"] = (metrics.inventory/365) / metrics.cogs # Inventory Turnover
metrics["dso"] = metrics.receivables / (metrics.sales/365) # Days Sales Outstanding

metrics["roa"] = metrics.net_income / metrics.total_assets # Return on Assets
metrics["roe"] = metrics.net_income / metrics.total_equity # Return on Equity

metrics['debit_asset_ratio'] = metrics.total_liabilities / metrics.total_assets # Debit to Asset Ratio
metrics["debit_equity_ratio"] = metrics.total_liabilities / metrics.total_equity # Debit to Equity Ratio

metrics["book_price"] = metrics.total_equity / metrics.mkt_value # Price to Book
metrics["equity_price"] = metrics.net_income / metrics.mkt_value # Price to Equity
metrics["ebit_price"] = metrics.ebit / metrics.mkt_value # Price to EBIT
metrics["ebitda_price"] = metrics.ebitda / metrics.mkt_value # Price to EBITDA
metrics["sales_price"] = metrics.sales / metrics.mkt_value # Price to Sales Ratio

metrics["cash_assets"] = metrics.cash / metrics.total_assets
metrics["cash_current_liab"] = metrics.cash / metrics.current_liabilities
metrics["cash_total_liab"] = metrics.cash / metrics.total_liabilities
metrics["cash_interest"] = metrics.interest_exp / metrics.cash
metrics["fixed_total"] = metrics.ppe / metrics.total_assets
metrics["fixed_employees"] = metrics.ppe / metrics.employees
metrics["working_coverage"] = metrics.working_cap * metrics.coverage_ratio

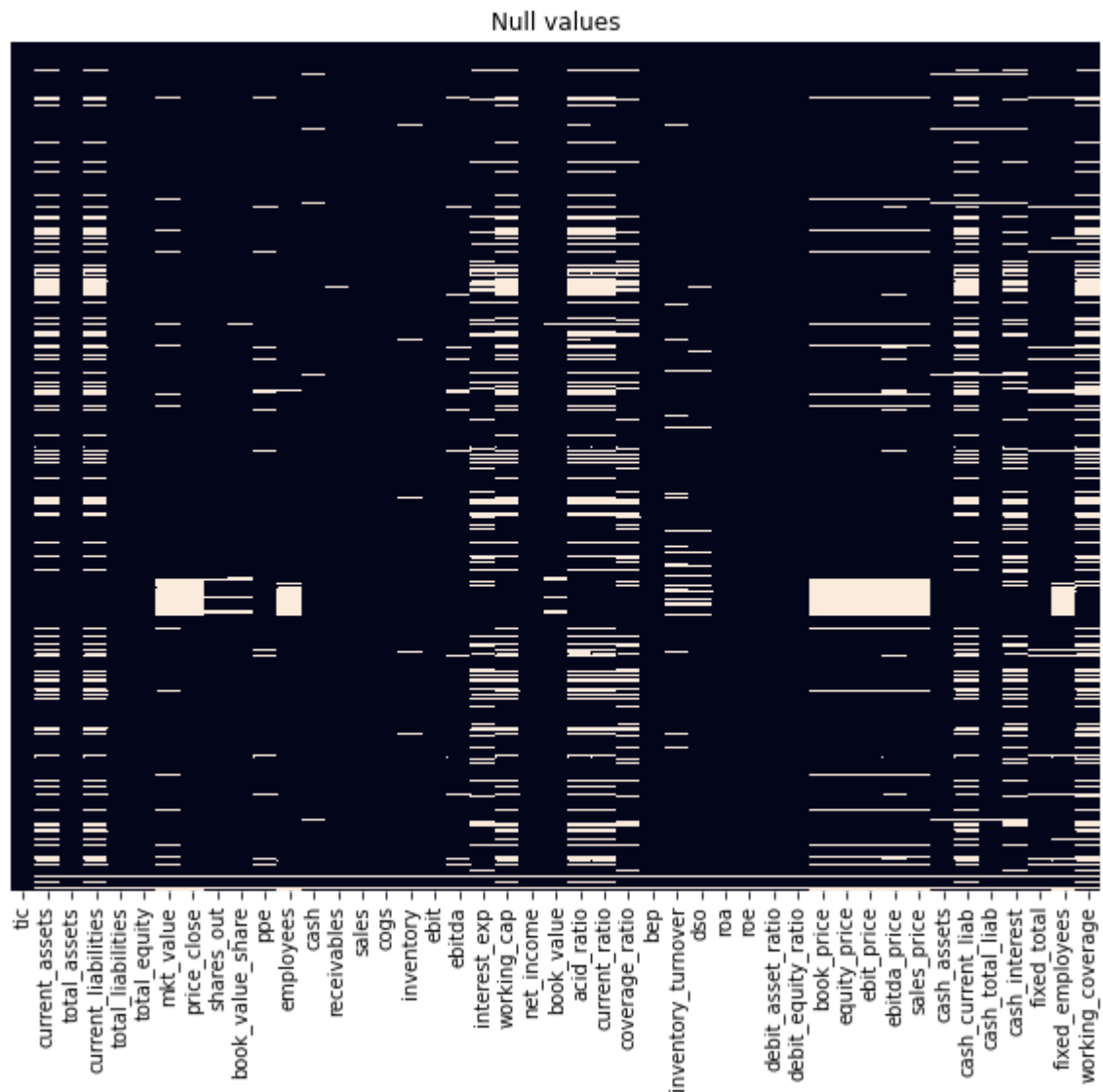
```

## Metric Exploratory Data Analysis & Cleaning

In [318]: `metrics.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2839 entries, 0 to 2838
Data columns (total 45 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   tic                                   2839 non-null   object
1   current_assets                       2147 non-null   float64
2   total_assets                         2831 non-null   float64
3   current_liabilities                 2147 non-null   float64
4   total_liabilities                   2825 non-null   float64
5   total_equity                        2831 non-null   float64
6   mkt_value                           2588 non-null   float64
7   price_close                         2702 non-null   float64
8   shares_out                          2810 non-null   float64
9   book_value_share                    2803 non-null   float64
10  ppe                                 2698 non-null   float64
11  employees                           2731 non-null   float64
12  cash                               2790 non-null   float64
13  receivables                         2808 non-null   float64
14  sales                              2830 non-null   float64
15  cogs                               2830 non-null   float64
16  inventory                           2793 non-null   float64
17  ebit                               2830 non-null   float64
18  ebitda                             2681 non-null   float64
19  interest_exp                       2367 non-null   float64
20  working_cap                         2147 non-null   float64
21  net_income                         2830 non-null   float64
22  book_value                         2803 non-null   float64
23  acid_ratio                         2122 non-null   float64
24  current_ratio                      2147 non-null   float64
25  coverage_ratio                     2367 non-null   float64
26  bep                                2830 non-null   float64
27  inventory_turnover                 2676 non-null   float64
28  dso                                2735 non-null   float64
29  roa                                2830 non-null   float64
30  roe                                2830 non-null   float64
31  debit_asset_ratio                  2825 non-null   float64
32  debit_equity_ratio                 2825 non-null   float64
33  book_price                         2588 non-null   float64
34  equity_price                       2588 non-null   float64
35  ebit_price                         2588 non-null   float64
36  ebitda_price                       2475 non-null   float64
37  sales_price                        2588 non-null   float64
38  cash_assets                        2790 non-null   float64
39  cash_current_liab                  2117 non-null   float64
40  cash_total_liab                    2784 non-null   float64
41  cash_interest                      2325 non-null   float64
42  fixed_total                        2698 non-null   float64
43  fixed_employees                    2579 non-null   float64
44  working_coverage                    1992 non-null   float64
dtypes: float64(44), object(1)
memory usage: 998.2+ KB
```

```
In [319]: plt.figure(figsize=(10, 8))
plt.title("Null values")
sns.heatmap(metrics.isnull(),yticklabels=False,cbar=False)
plt.show()
```



```
In [320]: metrics.dropna(axis =0, inplace = True)
```

In [321]: `metrics.describe().T`

Out[321]:

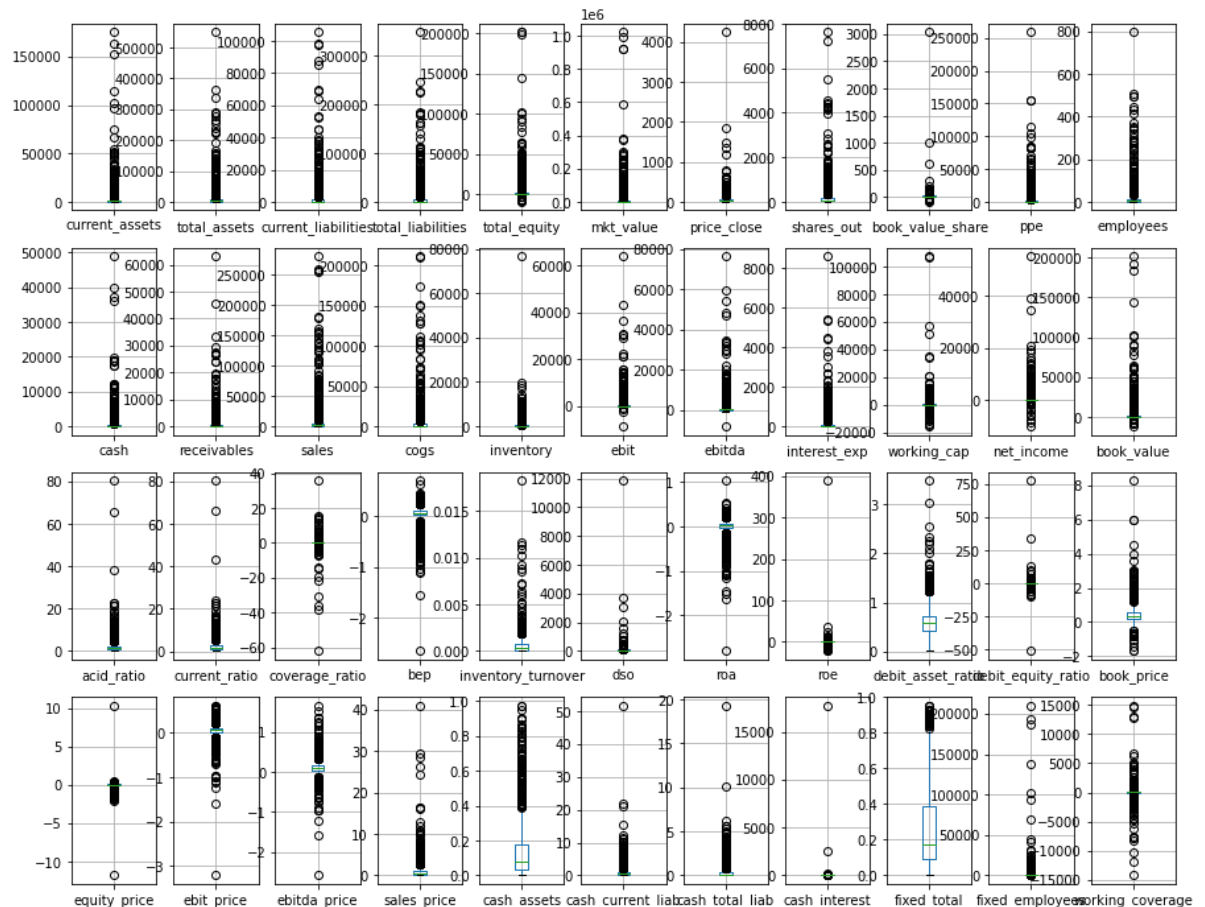
	count	mean	std	min	25%	50%
<b>current_assets</b>	1729.0	2.896381e+03	10001.534838	1.584000	233.156000	605.760000
<b>total_assets</b>	1729.0	1.060675e+04	31775.684472	2.215000	602.209000	1898.400000
<b>current_liabilities</b>	1729.0	2.211117e+03	7503.377387	0.328000	98.422000	327.051000
<b>total_liabilities</b>	1729.0	6.994001e+03	21207.294580	0.478000	283.654000	1135.307000
<b>total_equity</b>	1729.0	3.592888e+03	12544.886272	-9599.000000	217.790000	697.480000
<b>mkt_value</b>	1729.0	1.476748e+04	58300.651959	14.267600	726.896000	2217.796700
<b>price_close</b>	1729.0	6.742545e+01	143.860453	0.249000	16.780000	37.210000
<b>shares_out</b>	1729.0	1.932751e+02	499.394579	0.621000	34.185000	65.033000
<b>book_value_share</b>	1729.0	1.917380e+01	80.447175	-87.732000	4.299800	11.538500
<b>ppe</b>	1729.0	3.433339e+03	12515.756940	0.000000	74.676000	380.049000
<b>employees</b>	1729.0	1.832565e+01	49.641686	0.004000	0.938000	4.200000
<b>cash</b>	1729.0	6.989715e+02	2571.215788	0.000000	45.733000	133.667000
<b>receivables</b>	1729.0	9.419549e+02	3154.067390	0.000000	45.733000	176.573000
<b>sales</b>	1729.0	6.869483e+03	21079.508935	0.000000	390.255000	1450.201000
<b>cogs</b>	1729.0	4.431139e+03	15099.612535	0.113000	182.505000	820.279000
<b>inventory</b>	1729.0	5.815589e+02	2393.352546	0.000000	0.332000	68.905000
<b>ebit</b>	1729.0	8.308371e+02	3081.518678	-8596.000000	10.522000	115.300000
<b>ebitda</b>	1729.0	1.253013e+03	4326.569586	-8124.000000	33.234000	198.716000
<b>interest_exp</b>	1729.0	1.459368e+02	438.991654	0.000000	4.700000	26.300000
<b>working_cap</b>	1729.0	6.852642e+02	4650.789305	-15588.000000	50.600000	199.283000
<b>net_income</b>	1729.0	5.196862e+02	2523.704246	-10137.000000	-18.423000	47.536000
<b>book_value</b>	1729.0	3.454362e+03	12144.743184	-11576.947926	210.894530	673.271754
<b>acid_ratio</b>	1729.0	2.183379e+00	3.471982	0.056552	0.871466	1.364203
<b>current_ratio</b>	1729.0	2.621867e+00	3.539935	0.146189	1.214888	1.782883
<b>coverage_ratio</b>	1729.0	1.088364e-01	2.671752	-62.025641	0.002324	0.120408
<b>bep</b>	1729.0	2.163803e-02	0.190137	-2.661400	0.017459	0.058467
<b>inventory_turnover</b>	1729.0	5.588160e-04	0.001029	0.000000	0.000006	0.000328
<b>dso</b>	1729.0	inf	NaN	0.000000	35.737811	55.033126
<b>roa</b>	1729.0	-1.573134e-02	0.203434	-2.797291	-0.021692	0.031079
<b>roe</b>	1729.0	1.845371e-01	9.544946	-22.083071	-0.071754	0.078394
<b>debit_asset_ratio</b>	1729.0	5.955645e-01	0.286379	0.015731	0.419848	0.585466
<b>debit_equity_ratio</b>	1729.0	2.004778e+00	24.997584	-509.432203	0.585953	1.255491
<b>book_price</b>	1729.0	4.361267e-01	0.541483	-1.686420	0.154452	0.318920
<b>equity_price</b>	1729.0	-1.864512e-02	0.423149	-11.813662	-0.019258	0.026737

	count	mean	std	min	25%	50%
<b>ebit_price</b>	1729.0	4.044955e-02	0.161888	-3.217889	0.014204	0.053977
<b>ebitda_price</b>	1729.0	9.802007e-02	0.204254	-2.578188	0.035128	0.084778
<b>sales_price</b>	1729.0	1.104076e+00	2.202830	0.000000	0.220271	0.508346
<b>cash_assets</b>	1729.0	1.340508e-01	0.157263	0.000000	0.030306	0.076944
<b>cash_current_liab</b>	1729.0	9.241053e-01	1.950096	0.000000	0.166453	0.431943
<b>cash_total_liab</b>	1729.0	3.726009e-01	0.825282	0.000000	0.046967	0.138495
<b>cash_interest</b>	1729.0	inf	NaN	0.000000	0.041921	0.164594
<b>fixed_total</b>	1729.0	2.674180e-01	0.239589	0.000000	0.089561	0.170764
<b>fixed_employees</b>	1729.0	1.170752e+03	9915.986013	0.000000	41.221940	74.695200
<b>working_coverage</b>	1729.0	4.630823e+01	1107.557764	-14227.319149	-3.046942	10.787431

```

In [322]: numeric_col = list(metrics.describe(include=[np.number]).columns)
metrics_numeric = metrics[numeric_col]
fig, axes = plt.subplots(len(numeric_col)//11,11, figsize=(15, 12))
i = 0
for triaxis in axes:
    for axis in triaxis:
        metrics_numeric.boxplot(column = numeric_col[i], ax=axis)
        i = i+1

```





After looking at the box and whisker plot many variables have outliers which need to be dealt with. I will be winsorizing the outliers to the 99.5 percentile to still maintain their large quantity but not throw off the models.

## **Metrics that need Winsorizing**

```

In [323]: def mstats_winsorize(s):

            return mstats.winsorize(s, limits=[0.005, 0.005])

def winsorize(df):

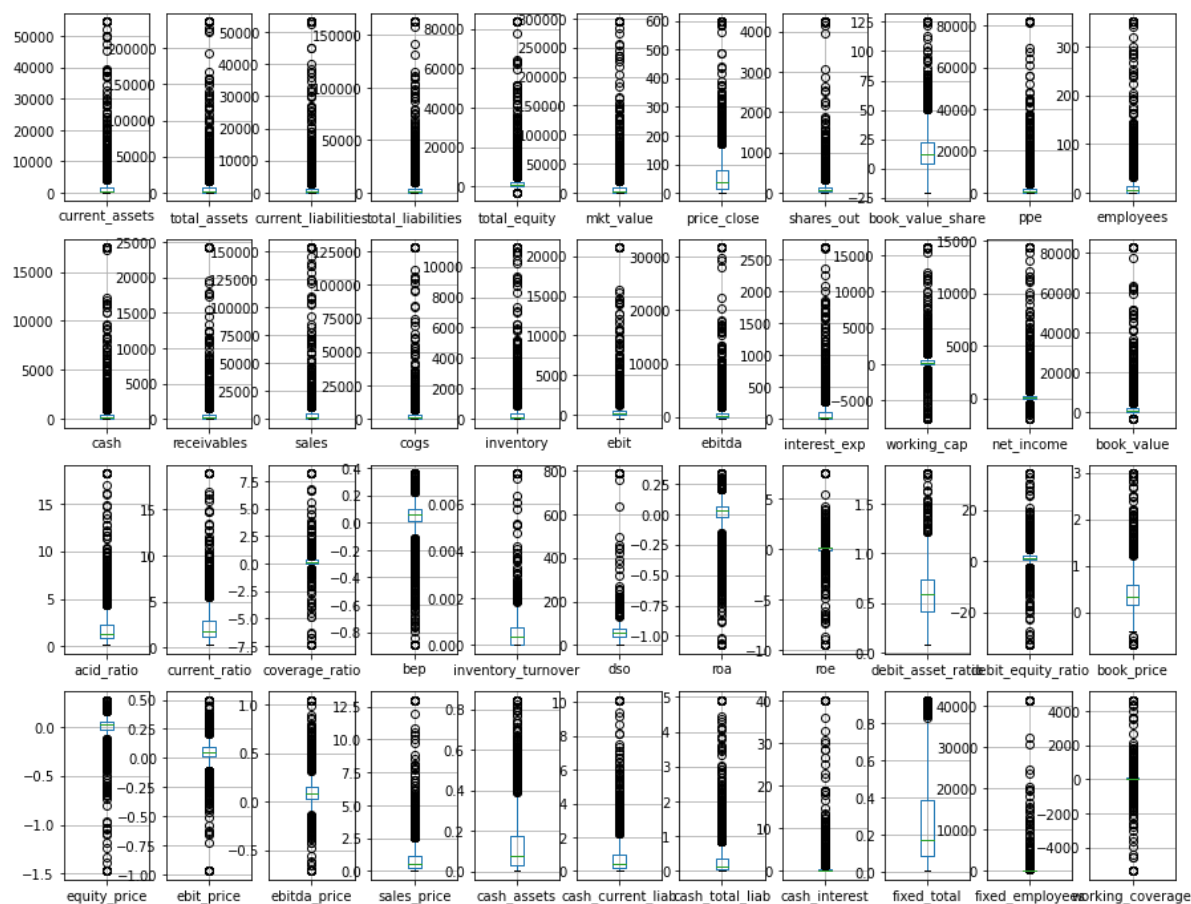
    temp = df.copy()

    return temp.apply(mstats_winsorize, axis=0)

metrics_win = winsorize(metrics)

# infinity and nonnumeric values not able to be plotted
numeric_col = list(metrics_win.describe(include=[np.number]).columns)
metrics_win_numeric = metrics_win[numeric_col]
fig, axes = plt.subplots(len(numeric_col)//11,11, figsize=(15, 12))
i = 0
for triaxis in axes:
    for axis in triaxis:
        metrics_win_numeric.boxplot(column = numeric_col[i], ax=axis)
        i = i+1

```



The box plots look significantly better outliers still persist but not at the level of before and most values have centralized.

# Correlation Matrix

```
In [324]: stock_metrics = pd.merge(returns2020, metrics_win, on="tic", how="inner")
stock_metrics.to_csv(r'stock_metrics.csv', index = False, header=True)

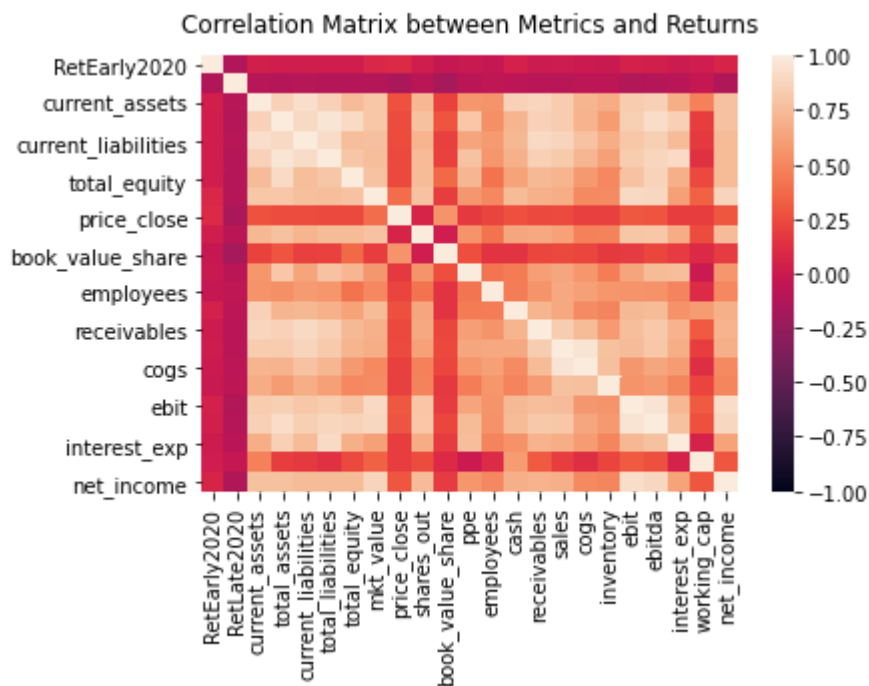
corr_metrics = stock_metrics[['RetEarly2020', 'RetLate2020', 'current_assets',
'total_assets',
    'current_liabilities', 'total_liabilities', 'total_equity', 'mkt_value'
,
    'price_close', 'shares_out', 'book_value_share', 'ppe', 'employees',
    'cash', 'receivables', 'sales', 'cogs', 'inventory', 'ebit', 'ebitda',
    'interest_exp', 'working_cap', 'net_income',]]

corr_ratios = stock_metrics[['RetEarly2020', 'RetLate2020', 'book_value', 'aci
d_ratio',
    'current_ratio', 'coverage_ratio', 'bep', 'inventory_turnover', 'dso',
    'roa', 'roe', 'debt_asset_ratio', 'debt_equity_ratio', 'book_price',
    'equity_price', 'ebit_price', 'ebitda_price', 'sales_price',
    'cash_assets', 'cash_current_liab', 'cash_total_liab', 'cash_interest',
    'fixed_total', 'fixed_employees', 'working_coverage']]
```

```
In [325]: len(corr_ratios.columns)
```

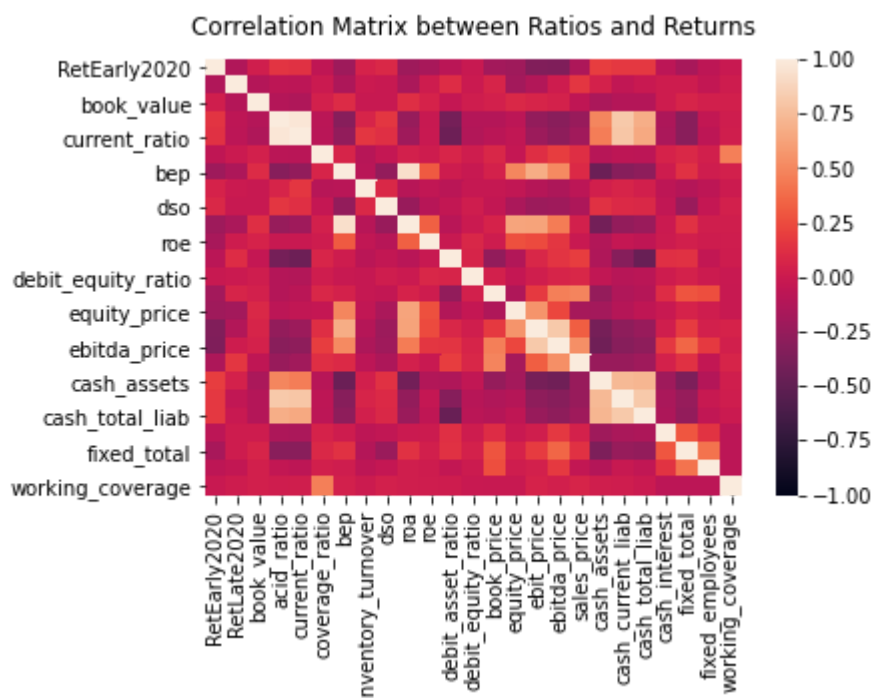
Out[325]: 25

```
In [326]: corr_matrix = sns.heatmap(corr_metrics.corr(), vmin=-1, vmax=1, annot=False)
# Give a title to the heatmap. Pad defines the distance of the title from the
# top of the heatmap.
corr_matrix.set_title('Correlation Matrix between Metrics and Returns', fontdi
ct={'fontsize':12}, pad=12);
# plt.figure(figsize=(15, 12))
```



Many of the metrics are based off values of the balance sheet which as the name insuates must ultimately balance. Therefore, many of the values are highly correlated.

```
In [327]: corr_matrix = sns.heatmap(corr_ratios.corr(), vmin=-1, vmax=1, annot=False)
# Give a title to the heatmap. Pad defines the distance of the title from the
# top of the heatmap.
corr_matrix.set_title('Correlation Matrix between Ratios and Returns', fontdic
t={'fontsize':12}, pad=12);
# plt.figure(figsize=(15, 12))
```



Many of the ratios share compents or explain similar things such as the acid ratio\_and current\_ratio. It is comforting to see there is a mix of highly correlated values and uncorrelated values.

## Linear Model on Metrics

### Early 2020 Returns

```
In [328]: # Dependent Variables
X = stock_metrics.iloc[:,3:]
X = sm.add_constant(X)

# Independent Variables
Y = stock_metrics.iloc[:,1:2]

# Fit Model
early_metrics = sm.OLS(Y, X).fit()

# Predict Model
# predictions = model.predict(X)

# Summary Statistics
print(early_metrics.summary(title="Early 2020 Returns with Metrics"))
early_metrics_results = pd.read_html(early_metrics.summary().tables[1].as_html(
),header=0,index_col=0)[0]
```

## Early 2020 Returns with Metrics

```

=====
=
Dep. Variable:          RetEarly2020    R-squared:                0.20
3
Model:                  OLS             Adj. R-squared:           0.18
2
Method:                Least Squares    F-statistic:              9.61
3
Date:                  Thu, 08 Jul 2021  Prob (F-statistic):       2.43e-5
5
Time:                  14:36:36          Log-Likelihood:           -458.6
8
No. Observations:      1709             AIC:                      100
7.
Df Residuals:          1664             BIC:                      125
2.
Df Model:              44
Covariance Type:       nonrobust
=====
=====
                                coef      std err          t      P>|t|      [0.025
0.975]
-----
-----
const                    -0.1363      0.045      -3.037      0.002      -0.224
-0.048
current_assets          -3.303e-06   7.79e-06     -0.424      0.672     -1.86e-05
1.2e-05
total_assets            4.776e-07   3.91e-06      0.122      0.903     -7.19e-06
8.14e-06
current_liabilities    -2.795e-06   8.44e-06     -0.331      0.740     -1.93e-05
1.38e-05
total_liabilities       3.878e-06   4.08e-06      0.949      0.343     -4.13e-06
1.19e-05
total_equity            -2.016e-05   1.42e-05     -1.418      0.156     -4.81e-05
7.73e-06
mkt_value               9.876e-08   6.9e-07      0.143      0.886     -1.25e-06
1.45e-06
price_close             0.0003      0.000      2.007      0.045      7.39e-06
0.001
shares_out              -1.616e-05   3.72e-05     -0.435      0.664     -8.91e-05
5.68e-05
book_value_share        0.0004      0.001      0.592      0.554      -0.001
0.002
ppe                     1.146e-06   1.9e-06      0.602      0.547     -2.59e-06
4.88e-06
employees               -0.0005      0.000     -1.743      0.082      -0.001
5.7e-05
cash                   9.397e-06   1.08e-05      0.873      0.383     -1.17e-05
3.05e-05
receivables             -1.137e-05   8.9e-06     -1.278      0.202     -2.88e-05
6.08e-06
sales                   3.262e-06   3.58e-06      0.912      0.362     -3.76e-06
1.03e-05
cogs                   -2.832e-06   3.58e-06     -0.791      0.429     -9.86e-06
4.19e-06

```

inventory 2.08e-05	6.917e-07	1.03e-05	0.067	0.946	-1.95e-05
ebit 3.59e-05	2.205e-06	1.72e-05	0.128	0.898	-3.15e-05
ebitda -2e-06	-2.924e-05	1.39e-05	-2.105	0.035	-5.65e-05
interest_exp 0.000	2.057e-05	7.59e-05	0.271	0.786	-0.000
working_cap 1.7e-05	-1.964e-06	9.67e-06	-0.203	0.839	-2.09e-05
net_income 7.62e-05	4.894e-05	1.39e-05	3.517	0.000	2.16e-05
book_value 4.72e-05	1.93e-05	1.42e-05	1.355	0.176	-8.64e-06
acid_ratio 0.036	-0.0013	0.019	-0.070	0.944	-0.039
current_ratio 0.034	-0.0015	0.018	-0.085	0.932	-0.037
coverage_ratio 0.016	-0.0021	0.009	-0.232	0.816	-0.020
bep 0.559	0.2288	0.169	1.357	0.175	-0.102
inventory_turnover 18.006	-3.0309	10.726	-0.283	0.778	-24.068
dso 0.000	6.17e-05	0.000	0.508	0.612	-0.000
roa 0.143	-0.1726	0.161	-1.073	0.283	-0.488
roe -0.017	-0.0308	0.007	-4.377	0.000	-0.045
debit_asset_ratio -0.055	-0.1491	0.048	-3.117	0.002	-0.243
debit_equity_ratio 0.004	0.0008	0.001	0.553	0.580	-0.002
book_price -0.068	-0.1226	0.028	-4.370	0.000	-0.178
equity_price -0.065	-0.2258	0.082	-2.759	0.006	-0.386
ebit_price -0.150	-0.4673	0.162	-2.888	0.004	-0.785
ebitda_price -0.002	-0.2307	0.117	-1.975	0.048	-0.460
sales_price 0.035	0.0205	0.008	2.698	0.007	0.006
cash_assets 0.099	-0.1034	0.103	-1.002	0.316	-0.306
cash_current_liab 0.040	0.0090	0.016	0.576	0.565	-0.022
cash_total_liab 0.080	0.0310	0.025	1.233	0.218	-0.018
cash_interest 0.003	-0.0011	0.002	-0.517	0.605	-0.005
fixed_total 0.007	-0.0867	0.048	-1.815	0.070	-0.180
fixed_employees 7.97e-06	3.042e-06	2.51e-06	1.212	0.226	-1.88e-06
working_coverage	4.799e-06	1.57e-05	0.307	0.759	-2.59e-05



3.55e-05

=====

=

Omnibus: 2600.741 Durbin-Watson: 1.94

2  
Prob(Omnibus): 0.000 Jarque-Bera (JB): 2213503.983  
Skew: 8.846 Prob(JB): 0.00  
Kurtosis: 178.419 Cond. No. 7.71e+0

7

=====

=

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 7.71e+07. This might indicate that there are strong multicollinearity or other numerical problems.

## Determining the most predictive Metrics for Early 2020 Return

In [329]: `early_metrics_results.loc[early_metrics_results['P>|t|']<0.05,].sort_values("coef", ascending=False)`

Out[329]:

	coef	std err	t	P> t	[0.025	0.975]
<b>sales_price</b>	0.020500	0.008000	2.698	0.007	0.006000	0.035000
<b>price_close</b>	0.000300	0.000000	2.007	0.045	0.000007	0.001000
<b>net_income</b>	0.000049	0.000014	3.517	0.000	0.000022	0.000076
<b>ebitda</b>	-0.000029	0.000014	-2.105	0.035	-0.000056	-0.000002
<b>roe</b>	-0.030800	0.007000	-4.377	0.000	-0.045000	-0.017000
<b>book_price</b>	-0.122600	0.028000	-4.370	0.000	-0.178000	-0.068000
<b>const</b>	-0.136300	0.045000	-3.037	0.002	-0.224000	-0.048000
<b>debit_asset_ratio</b>	-0.149100	0.048000	-3.117	0.002	-0.243000	-0.055000
<b>equity_price</b>	-0.225800	0.082000	-2.759	0.006	-0.386000	-0.065000
<b>ebitda_price</b>	-0.230700	0.117000	-1.975	0.048	-0.460000	-0.002000
<b>ebit_price</b>	-0.467300	0.162000	-2.888	0.004	-0.785000	-0.150000

## Early 2020 Return Metric Explaintion

**Adj R<sup>2</sup>: 0.182**

The top positive metrics at the 95% statisitcal signifiant level were:

- sales\_price
- price\_close
- net\_income

The top negative metrics at the 95% statisitcal signifiant level were:

- ebit\_price
- ebitda\_price
- equity\_price
- debit\_asset\_ratio
- const
- book\_price
- roe
- ebitda

## Late 2020 Returns

```
In [330]: # Dependent Variables
X = stock_metrics.iloc[:,3:]
X = sm.add_constant(X)

# Independent Variables
Y = stock_metrics.iloc[:,2:3]

# Fit Model
late_metrics = sm.OLS(Y, X).fit()

# # Predict Model
# predictions = model.predict(X)

# Summary Statistics
print(late_metrics.summary(title="Late 2020 Returns with Metrics"))
late_metrics_results = pd.read_html(late_metrics.summary().tables[1].as_html
(),header=0,index_col=0)[0]
```

## Late 2020 Returns with Metrics

```

=====
=
Dep. Variable:          RetLate2020    R-squared:                0.13
0
Model:                  OLS    Adj. R-squared:            0.10
7
Method:                 Least Squares    F-statistic:              5.66
0
Date:                   Thu, 08 Jul 2021    Prob (F-statistic):       6.69e-2
8
Time:                   14:36:36    Log-Likelihood:           -2606.
0
No. Observations:      1709    AIC:                      530
2.
Df Residuals:          1664    BIC:                      554
7.
Df Model:               44
Covariance Type:       nonrobust
=====
=====
                                coef    std err          t      P>|t|      [0.025
0.975]
-----
const                0.9297      0.158      5.899      0.000      0.621
1.239
current_assets      -7.699e-06    2.74e-05    -0.281      0.779    -6.14e-05
4.6e-05
total_assets       -3.411e-06    1.37e-05    -0.248      0.804    -3.03e-05
2.35e-05
current_liabilities  2.32e-05    2.96e-05     0.783      0.434    -3.49e-05
8.13e-05
total_liabilities  -1.098e-05    1.44e-05    -0.765      0.444    -3.91e-05
1.72e-05
total_equity        6.453e-05     5e-05     1.292      0.197    -3.35e-05
0.000
mkt_value           1.79e-07    2.42e-06     0.074      0.941    -4.57e-06
4.93e-06
price_close         -0.0001      0.001     -0.243      0.808     -0.001
0.001
shares_out          -0.0002      0.000     -1.366      0.172     -0.000
7.78e-05
book_value_share    -0.0085      0.002     -3.492      0.000     -0.013
-0.004
ppe                 9.851e-07    6.69e-06     0.147      0.883    -1.21e-05
1.41e-05
employees           0.0004      0.001     0.459      0.646     -0.001
0.002
cash                2.819e-05    3.78e-05     0.745      0.456     -4.6e-05
0.000
receivables         5.938e-06    3.13e-05     0.190      0.849    -5.54e-05
6.73e-05
sales               -4.801e-06    1.26e-05    -0.382      0.703    -2.95e-05
1.99e-05
cogs                -1.805e-06    1.26e-05    -0.143      0.886    -2.65e-05
2.29e-05

```

inventory 6.37e-05	-7.131e-06	3.61e-05	-0.198	0.843	-7.79e-05
ebit 6.25e-05	-5.586e-05	6.04e-05	-0.925	0.355	-0.000
ebitda 0.000	6.232e-05	4.88e-05	1.277	0.202	-3.34e-05
interest_exp 0.001	0.0002	0.000	0.842	0.400	-0.000
working_cap 7.32e-05	6.58e-06	3.4e-05	0.194	0.846	-6e-05
net_income 5.49e-05	-4.101e-05	4.89e-05	-0.839	0.402	-0.000
book_value 4.6e-05	-5.212e-05	5e-05	-1.042	0.298	-0.000
acid_ratio 0.070	-0.0613	0.067	-0.919	0.358	-0.192
current_ratio 0.152	0.0276	0.064	0.434	0.664	-0.097
coverage_ratio 0.013	-0.0496	0.032	-1.544	0.123	-0.113
bep 0.250	-0.9110	0.592	-1.538	0.124	-2.072
inventory_turnover 84.235	10.3308	37.680	0.274	0.784	-63.573
dso 0.000	-0.0006	0.000	-1.453	0.146	-0.001
roa 1.526	0.4178	0.565	0.740	0.460	-0.690
roe 0.144	0.0951	0.025	3.843	0.000	0.047
debit_asset_ratio 0.636	0.3060	0.168	1.820	0.069	-0.024
debit_equity_ratio 0.001	-0.0094	0.005	-1.796	0.073	-0.020
book_price 0.207	0.0138	0.099	0.140	0.888	-0.179
equity_price 0.353	-0.2109	0.288	-0.733	0.463	-0.775
ebit_price -1.292	-2.4069	0.569	-4.233	0.000	-3.522
ebitda_price 2.256	1.4511	0.410	3.537	0.000	0.646
sales_price 0.139	0.0867	0.027	3.250	0.001	0.034
cash_assets 1.066	0.3544	0.363	0.978	0.328	-0.357
cash_current_liab -0.015	-0.1231	0.055	-2.241	0.025	-0.231
cash_total_liab 0.369	0.1959	0.088	2.216	0.027	0.022
cash_interest 0.009	-0.0051	0.007	-0.684	0.494	-0.020
fixed_total 0.291	-0.0385	0.168	-0.230	0.818	-0.368
fixed_employees 2.54e-06	-1.475e-05	8.82e-06	-1.673	0.095	-3.2e-05
working_coverage	6.124e-05	5.5e-05	1.113	0.266	-4.66e-05

0.000

=====

=

```

Omnibus:                1800.212    Durbin-Watson:                1.99
1
Prob(Omnibus):           0.000    Jarque-Bera (JB):            207925.95
4
Skew:                    4.909    Prob(JB):                    0.0
0
Kurtosis:                56.137    Cond. No.                    7.71e+0
7

```

=====

=

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 7.71e+07. This might indicate that there are strong multicollinearity or other numerical problems.

## Determining the most predictive Metrics for Late 2020 Return

```
In [331]: late_metrics_results.loc[late_metrics_results['P>|t|']<0.05,].sort_values("coef", ascending=False)
```

Out[331]:

	coef	std err	t	P> t	[0.025	0.975]
<b>ebitda_price</b>	1.4511	0.410	3.537	0.000	0.646	2.256
<b>const</b>	0.9297	0.158	5.899	0.000	0.621	1.239
<b>cash_total_liab</b>	0.1959	0.088	2.216	0.027	0.022	0.369
<b>roe</b>	0.0951	0.025	3.843	0.000	0.047	0.144
<b>sales_price</b>	0.0867	0.027	3.250	0.001	0.034	0.139
<b>book_value_share</b>	-0.0085	0.002	-3.492	0.000	-0.013	-0.004
<b>cash_current_liab</b>	-0.1231	0.055	-2.241	0.025	-0.231	-0.015
<b>ebit_price</b>	-2.4069	0.569	-4.233	0.000	-3.522	-1.292

## Late 2020 Return Metric Explaintion

Adj R<sup>2</sup>: 0.107

The top positive metrics at the 95% statisitcal signifiant level were:

- ebitda\_price
- debit\_asset\_ratio
- cash\_assets
- cash\_total\_liab
- book\_price
- roe
- sales\_price

The top negative metrics at the 95% statisitcal signifiant level were:

- ebit\_price
- cash\_current\_liab
- book\_value\_share

## 2020 Metric Returns Explained

The top positive metric at the 95% statisitcal significant level is: **sales\_price**

- Price to Sale was a postive statistcally significant variable for both Early and Late 2020 Returns. This makes sense because Sales is a the top of the balance sheet. High top line health on the balance sheet means money is coming in the company and connected with a high price, is a great signals for investors and subsequently postive stock returns.

The top negative metrics at the 95% statisitcal signifiant level is: **ebit\_price**

- Price to EBIT was a negative statistcally significant variable for both Early and Late 2020 Returns. This makes sense because EBIT is low on the balance sheet just above interest and taxes. Low bottom line health on the balnce sheet means there isn't much to pay off Interest Expenses, Taxes, or Working Capital. This would mean low new income and stock holders equity, and combined with a low price, subsequently negative stock retutn

## Linear Model on Top Metrics and Industry Classifier

```
In [332]: stock = pd.merge(stock_class, stock_metrics.drop(["RetEarly2020", "RetLate2020"], axis =1), on='tic', how="inner")
stock.to_csv(r'stock.csv', index = False, header=True)
stock.sample(3)
```

Out[332]:

	tic	RetEarly2020	RetLate2020	gic_sub_industry	gic_sub_industry_name	Aerospace & Defense	A
803	VXRT	4.049931	2.225995	35201010	Biotechnology	0	
1585	ADT	-0.450820	0.841093	20201080	Security & Alarm Services	0	
577	CELH	-0.128365	10.950120	30201030	Soft Drinks	0	

3 rows × 195 columns

## Early 2020 Returns



```
In [333]: # Dependent Variables
X = stock.iloc[:,5:]
X = sm.add_constant(X)

# Independent Variables
Y = stock.iloc[:,1:2]

# Fit Model
early_stock = sm.OLS(Y, X).fit()

# Predict Model
# predictions = model.predict(X)

# Summary Statistics
print(early_stock.summary(title="Early 2020 Returns with Metrics and Classification"))
early_stock_results = pd.read_html(early_stock.summary().tables[1].as_html(), header=0, index_col=0)[0]
```

## Early 2020 Returns with Metrics and Classifier

```

=====
=
Dep. Variable:          RetEarly2020    R-squared:                0.29
9
Model:                  OLS             Adj. R-squared:           0.21
8
Method:                 Least Squares   F-statistic:              3.67
2
Date:                   Thu, 08 Jul 2021 Prob (F-statistic):       7.20e-4
2
Time:                   14:36:37        Log-Likelihood:          -351.5
2
No. Observations:      1653            AIC:                      104
9.
Df Residuals:          1480            BIC:                      198
5.
Df Model:               172
Covariance Type:       nonrobust
=====
=====

```

```

=====
                                coef    std err          t
P>|t|    [0.025    0.975]
-----
const                                -0.2152    0.054    -3.960
0.000    -0.322    -0.109
Aerospace & Defense                   -0.0182    0.065    -0.281
0.779    -0.145    0.109
Agricultural & Farm Machinery          0.0683    0.182    0.375
0.708    -0.290    0.426
Agricultural Products                 0.0752    0.122    0.616
0.538    -0.164    0.315
Air Freight & Logistics                0.1422    0.116    1.225
0.221    -0.085    0.370
Airlines                             -0.1338    0.108    -1.243
0.214    -0.345    0.077
Airport Services                     -0.0372    0.315    -0.118
0.906    -0.656    0.581
Alternative Carriers                 0.1991    0.163    1.223
0.221    -0.120    0.518
Aluminum                            -0.2180    0.224    -0.975
0.330    -0.657    0.221
Apparel Retail                      -0.1635    0.080    -2.041
0.041    -0.321    -0.006
Apparel, Accessories & Luxury Goods   -0.1723    0.086    -1.996
0.046    -0.342    -0.003
Application Software                 0.0666    0.042    1.571
0.116    -0.017    0.150
Asset Management & Custody Banks       -0.1305    0.143    -0.910
0.363    -0.412    0.151
Auto Parts & Equipment                -0.0936    0.075    -1.248
0.212    -0.241    0.054
Automobile Manufacturers              0.0719    0.167    0.431
0.667    -0.256    0.399
Automotive Retail                   -0.1209    0.094    -1.289
0.198    -0.305    0.063
=====

```

Biotechnology			0.0708	0.044	1.600
0.110	-0.016	0.158			
Brewers			0.1238	0.162	0.763
0.446	-0.194	0.442			
Building Products			0.0635	0.064	0.985
0.325	-0.063	0.190			
Casinos & Gaming			-0.1848	0.085	-2.178
0.030	-0.351	-0.018			
Coal & Consumable Fuels			-0.1388	0.146	-0.954
0.340	-0.424	0.147			
Commercial Printing			-0.0394	0.143	-0.276
0.782	-0.320	0.241			
Commodity Chemicals			-0.0651	0.083	-0.780
0.436	-0.229	0.099			
Communications Equipment			0.0206	0.063	0.326
0.745	-0.103	0.144			
Computer & Electronics Retail			-0.0356	0.190	-0.188
0.851	-0.409	0.337			
Construction & Engineering			0.0234	0.071	0.329
0.742	-0.116	0.163			
Construction Machinery & Heavy Trucks			-0.0202	0.082	-0.247
0.805	-0.181	0.141			
Construction Materials			-0.0747	0.121	-0.618
0.537	-0.312	0.162			
Consumer Electronics			-0.0990	0.143	-0.690
0.490	-0.380	0.182			
Consumer Finance			0.1236	0.159	0.776
0.438	-0.189	0.436			
Copper			-0.0708	0.235	-0.301
0.764	-0.532	0.391			
Data Processing & Outsourced Services			-0.0065	0.066	-0.098
0.922	-0.136	0.123			
Department Stores			-0.2453	0.163	-1.507
0.132	-0.564	0.074			
Distillers & Vintners			0.0163	0.146	0.111
0.911	-0.271	0.303			
Distributors			-0.0423	0.143	-0.296
0.767	-0.323	0.238			
Diversified Banks			7.185e-14	1.07e-13	0.673
0.501	-1.38e-13	2.81e-13			
Diversified Chemicals			-0.0557	0.182	-0.305
0.760	-0.413	0.302			
Diversified Metals & Mining			-0.0922	0.223	-0.413
0.679	-0.530	0.345			
Diversified REITs			-1.465e-14	2.19e-14	-0.668
0.504	-5.77e-14	2.84e-14			
Diversified Real Estate Activities			-0.0753	0.319	-0.236
0.813	-0.700	0.550			
Diversified Support Services			-0.0735	0.129	-0.568
0.570	-0.328	0.180			
Drug Retail			0.2233	0.237	0.942
0.346	-0.242	0.688			
Education Services			-0.0257	0.098	-0.263
0.792	-0.217	0.166			
Electric Utilities			0.1751	0.085	2.051
0.040	0.008	0.343			
Electrical Components & Equipment			-0.0479	0.071	-0.673

0.501	-0.187	0.092			
Electronic Components			-0.0528	0.101	-0.520
0.603	-0.252	0.146			
Electronic Equipment & Instruments			-0.0496	0.079	-0.625
0.532	-0.205	0.106			
Electronic Manufacturing Services			0.0499	0.107	0.468
0.640	-0.159	0.259			
Environmental & Facilities Services			-0.0090	0.081	-0.112
0.911	-0.167	0.149			
Fertilizers & Agricultural Chemicals			0.0252	0.121	0.209
0.835	-0.212	0.262			
Financial Exchanges & Data			0.1176	0.118	0.998
0.318	-0.114	0.349			
Food Distributors			-0.0552	0.117	-0.473
0.637	-0.284	0.174			
Food Retail			0.2742	0.116	2.358
0.019	0.046	0.502			
Footwear			-0.1080	0.131	-0.823
0.410	-0.365	0.149			
Forest Products			-0.1704	0.315	-0.541
0.589	-0.788	0.448			
Gas Utilities			0.1869	0.105	1.782
0.075	-0.019	0.393			
General Merchandise Stores			-0.0359	0.149	-0.241
0.809	-0.327	0.256			
Gold			-0.0200	0.237	-0.084
0.933	-0.485	0.445			
Health Care Services			0.1035	0.070	1.476
0.140	-0.034	0.241			
Health Care Distributors			0.3056	0.088	3.490
0.000	0.134	0.477			
Health Care Equipment			0.1392	0.049	2.853
0.004	0.043	0.235			
Health Care Facilities			-0.0078	0.088	-0.088
0.930	-0.181	0.166			
Health Care REITs			8.61e-15	1.25e-14	0.690
0.490	-1.59e-14	3.31e-14			
Health Care Supplies			0.0327	0.076	0.433
0.665	-0.115	0.181			
Health Care Technology			0.1478	0.079	1.878
0.061	-0.007	0.302			
Heavy Electrical Equipment			-0.0140	0.182	-0.077
0.939	-0.372	0.344			
Home Furnishings			-0.0862	0.130	-0.665
0.506	-0.341	0.168			
Home Improvement Retail			-0.0413	0.172	-0.240
0.810	-0.378	0.296			
Homebuilding			-0.0878	0.159	-0.554
0.580	-0.399	0.223			
Homefurnishing Retail			-0.1653	0.108	-1.527
0.127	-0.378	0.047			
Hotel & Resort REITs			-1.221e-14	1.8e-14	-0.679
0.497	-4.75e-14	2.31e-14			
Hotels, Resorts & Cruise Lines			-0.2777	0.109	-2.557
0.011	-0.491	-0.065			
Household Appliances			-0.0380	0.159	-0.239
0.811	-0.349	0.273			

Household Products			0.1775	0.115	1.547
0.122	-0.048	0.403			
Housewares & Specialties			0.0221	0.184	0.120
0.904	-0.338	0.382			
Human Resource & Employment Services			-0.0871	0.101	-0.866
0.387	-0.284	0.110			
Hypermarkets & Super Centers			0.1600	0.192	0.835
0.404	-0.216	0.536			
IT Consulting & Other Services			0.0555	0.098	0.565
0.572	-0.137	0.248			
Independent Power Producers & Energy Traders			0.2667	0.144	1.848
0.065	-0.016	0.550			
Industrial Conglomerates			-0.0679	0.161	-0.422
0.673	-0.384	0.248			
Industrial Gases			0.0477	0.230	0.207
0.836	-0.403	0.498			
Industrial Machinery			-0.0485	0.045	-1.080
0.280	-0.136	0.040			
Industrial REITs			-2.097e-15	3.29e-15	-0.637
0.524	-8.55e-15	4.36e-15			
Insurance Brokers			0.1733	0.126	1.376
0.169	-0.074	0.420			
Integrated Oil & Gas			-0.0144	0.307	-0.047
0.962	-0.616	0.587			
Integrated Telecommunication Services			0.5253	0.161	3.255
0.001	0.209	0.842			
Internet & Direct Marketing Retail			0.0863	0.078	1.110
0.267	-0.066	0.239			
Investment Banking & Brokerage			-0.0063	0.315	-0.020
0.984	-0.625	0.612			
Leisure Facilities			-0.1736	0.143	-1.217
0.224	-0.453	0.106			
Leisure Products			-0.0146	0.082	-0.178
0.859	-0.175	0.146			
Life & Health Insurance			-0.0517	0.315	-0.164
0.870	-0.669	0.566			
Life Sciences Tools & Services			0.0582	0.068	0.852
0.394	-0.076	0.192			
Managed Health Care			0.0543	0.166	0.326
0.745	-0.272	0.381			
Marine			-0.0745	0.152	-0.491
0.623	-0.372	0.223			
Metal & Glass Containers			0.1531	0.093	1.653
0.099	-0.029	0.335			
Mortgage REITs			-1.419e-16	2.04e-16	-0.695
0.487	-5.42e-16	2.59e-16			
Motorcycle Manufacturers			-0.0615	0.315	-0.195
0.845	-0.680	0.557			
Multi-Sector Holdings			0.1478	0.316	0.468
0.640	-0.472	0.767			
Multi-Utilities			0.1921	0.093	2.072
0.038	0.010	0.374			
Multi-line Insurance			-1.902e-17	1.86e-16	-0.103
0.918	-3.83e-16	3.45e-16			
Office REITs			1.782e-17	2.22e-16	0.080
0.936	-4.18e-16	4.54e-16			
Office Services & Supplies			-0.0046	0.097	-0.048

0.962	-0.194	0.185			
Oil & Gas Drilling			-0.4661	0.170	-2.748
0.006	-0.799	-0.133			
Oil & Gas Equipment & Services			-0.3117	0.070	-4.445
0.000	-0.449	-0.174			
Oil & Gas Exploration & Production			-0.3474	0.076	-4.569
0.000	-0.497	-0.198			
Oil & Gas Refining & Marketing			-0.1858	0.096	-1.945
0.052	-0.373	0.002			
Oil & Gas Storage & Transportation			-0.0716	0.111	-0.643
0.520	-0.290	0.147			
Other Diversified Financial Services			-1.454e-16	1.64e-16	-0.888
0.375	-4.67e-16	1.76e-16			
Packaged Foods & Meats			0.1075	0.058	1.845
0.065	-0.007	0.222			
Paper Packaging			0.0845	0.101	0.838
0.402	-0.113	0.282			
Paper Products			0.0578	0.130	0.444
0.657	-0.198	0.313			
Personal Products			-0.0273	0.094	-0.291
0.771	-0.211	0.156			
Pharmaceuticals			-0.0116	0.057	-0.204
0.839	-0.123	0.100			
Property & Casualty Insurance			-5.302e-17	1.41e-16	-0.376
0.707	-3.29e-16	2.23e-16			
Railroads			0.0075	0.168	0.044
0.965	-0.323	0.338			
Real Estate Development			0.1196	0.318	0.377
0.707	-0.503	0.743			
Real Estate Operating Companies			-4.412e-17	1.8e-16	-0.244
0.807	-3.98e-16	3.1e-16			
Real Estate Services			-0.1503	0.132	-1.141
0.254	-0.408	0.108			
Regional Banks			-3.527e-17	1.41e-16	-0.251
0.802	-3.11e-16	2.4e-16			
Reinsurance			-1.329e-16	1.9e-16	-0.700
0.484	-5.05e-16	2.4e-16			
Renewable Electricity			0.1789	0.229	0.780
0.436	-0.271	0.629			
Research & Consulting Services			0.0681	0.071	0.959
0.338	-0.071	0.208			
Residential REITs			1.103e-16	1.7e-16	0.649
0.516	-2.23e-16	4.43e-16			
Restaurants			-0.1203	0.065	-1.857
0.063	-0.247	0.007			
Retail REITs			-1.543e-16	1.26e-16	-1.224
0.221	-4.02e-16	9.29e-17			
Security & Alarm Services			0.0194	0.184	0.106
0.916	-0.341	0.380			
Semiconductor Equipment			0.0369	0.073	0.503
0.615	-0.107	0.181			
Semiconductors			-0.0011	0.057	-0.019
0.984	-0.113	0.111			
Silver			-0.1837	0.318	-0.578
0.563	-0.807	0.440			
Soft Drinks			0.0198	0.125	0.158
0.874	-0.226	0.265			

Specialized Consumer Services	-0.1007	0.122	-0.825
0.409 -0.340 0.139			
Specialized Finance	-2.713e-17	2.19e-16	-0.124
0.902 -4.57e-16 4.03e-16			
Specialized REITs	0.0998	0.100	1.000
0.317 -0.096 0.296			
Specialty Chemicals	-0.0368	0.060	-0.614
0.539 -0.154 0.081			
Specialty Stores	-0.0332	0.098	-0.338
0.735 -0.225 0.159			
Steel	-0.1003	0.089	-1.129
0.259 -0.275 0.074			
Systems Software	0.1660	0.072	2.311
0.021 0.025 0.307			
Technology Distributors	-0.0071	0.133	-0.053
0.957 -0.269 0.254			
Technology Hardware, Storage & Peripherals	0.0260	0.094	0.278
0.781 -0.158 0.210			
Textiles	-0.2499	0.316	-0.791
0.429 -0.869 0.369			
Thrifts & Mortgage Finance	1.107e-16	1.28e-16	0.862
0.389 -1.41e-16 3.63e-16			
Tires & Rubber	-0.0769	0.224	-0.344
0.731 -0.515 0.362			
Tobacco	0.1263	0.152	0.831
0.406 -0.172 0.424			
Trading Companies & Distributors	-0.0398	0.066	-0.605
0.546 -0.169 0.089			
Trucking	0.0889	0.089	0.995
0.320 -0.086 0.264			
Water Utilities	0.2495	0.111	2.243
0.025 0.031 0.468			
Wireless Telecommunication Services	0.1884	0.122	1.549
0.122 -0.050 0.427			
current_assets	-8.304e-06	9.38e-06	-0.886
0.376 -2.67e-05 1.01e-05			
total_assets	5.923e-06	5.92e-06	1.001
0.317 -5.68e-06 1.75e-05			
current_liabilities	5.459e-06	9.77e-06	0.559
0.576 -1.37e-05 2.46e-05			
total_liabilities	-4.321e-06	6.44e-06	-0.671
0.502 -1.7e-05 8.31e-06			
total_equity	-1.085e-05	1.67e-05	-0.648
0.517 -4.37e-05 2.2e-05			
mkt_value	-1.728e-07	7.63e-07	-0.227
0.821 -1.67e-06 1.32e-06			
price_close	0.0003	0.000	1.588
0.112 -6.53e-05 0.001			
shares_out	-3.063e-05	4.74e-05	-0.647
0.518 -0.000 6.23e-05			
book_value_share	0.0002	0.001	0.209
0.834 -0.001 0.002			
ppe	-3.83e-07	2.37e-06	-0.162
0.872 -5.03e-06 4.26e-06			
employees	-4.984e-05	0.000	-0.165
0.869 -0.001 0.001			
cash	1.063e-05	1.23e-05	0.867

0.386	-1.34e-05	3.47e-05			
receivables			-8.548e-06	1.01e-05	-0.850
0.395	-2.83e-05	1.12e-05			
sales			-5.211e-07	4.07e-06	-0.128
0.898	-8.5e-06	7.46e-06			
cogs			-3.439e-07	4.05e-06	-0.085
0.932	-8.28e-06	7.59e-06			
inventory			8.511e-06	1.21e-05	0.701
0.483	-1.53e-05	3.23e-05			
ebit			4.631e-06	2.47e-05	0.187
0.851	-4.38e-05	5.31e-05			
ebitda			-2.138e-05	2.03e-05	-1.055
0.292	-6.12e-05	1.84e-05			
interest_exp			3.684e-06	8.9e-05	0.041
0.967	-0.000	0.000			
working_cap			4.774e-06	1.09e-05	0.438
0.661	-1.66e-05	2.62e-05			
net_income			5.252e-05	1.69e-05	3.114
0.002	1.94e-05	8.56e-05			
book_value			6.346e-06	1.63e-05	0.390
0.696	-2.55e-05	3.82e-05			
acid_ratio			-0.0227	0.023	-1.004
0.316	-0.067	0.022			
current_ratio			0.0191	0.022	0.889
0.374	-0.023	0.061			
coverage_ratio			-0.0087	0.010	-0.916
0.360	-0.027	0.010			
bep			0.4771	0.179	2.670
0.008	0.127	0.828			
inventory_turnover			-7.8962	11.831	-0.667
0.505	-31.104	15.311			
dso			4.768e-05	0.000	0.382
0.702	-0.000	0.000			
roa			-0.3658	0.168	-2.181
0.029	-0.695	-0.037			
roe			-0.0293	0.007	-4.057
0.000	-0.043	-0.015			
debit_asset_ratio			-0.1164	0.053	-2.181
0.029	-0.221	-0.012			
debit_equity_ratio			0.0005	0.002	0.327
0.744	-0.003	0.004			
book_price			-0.0713	0.031	-2.323
0.020	-0.132	-0.011			
equity_price			-0.1396	0.086	-1.622
0.105	-0.308	0.029			
ebit_price			-0.6716	0.176	-3.806
0.000	-1.018	-0.325			
ebitda_price			-0.1022	0.134	-0.762
0.446	-0.365	0.161			
sales_price			0.0163	0.009	1.864
0.062	-0.001	0.033			
cash_assets			-0.0851	0.111	-0.765
0.445	-0.304	0.133			
cash_current_liab			0.0089	0.017	0.524
0.600	-0.024	0.042			
cash_total_liab			0.0328	0.026	1.269
0.205	-0.018	0.084			



cash_interest			-0.0030	0.002	-1.357
0.175	-0.007	0.001			
fixed_total			0.0157	0.068	0.233
0.816	-0.117	0.148			
fixed_employees			5.146e-06	3.41e-06	1.508
0.132	-1.55e-06	1.18e-05			
working_coverage			1.123e-05	1.67e-05	0.672
0.501	-2.15e-05	4.4e-05			

=====

=

Omnibus:	2734.935	Durbin-Watson:	1.97
5			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	3334110.27
3			
Skew:	10.402	Prob(JB):	0.0
0			
Kurtosis:	222.033	Cond. No.	1.17e+1
8			

=====

=

**Notes:**

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The smallest eigenvalue is 3.42e-24. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

## Late 2020 Returns

```
In [334]: # Dependent Variables
X = stock.iloc[:,5:]
X = sm.add_constant(X)

# Independent Variables
Y = stock.iloc[:,2:3]

# Fit Model
late_stock = sm.OLS(Y, X).fit()

# Predict Model
# predictions = model.predict(X)

# Summary Statistics
print(late_stock.summary(title="Late 2020 Returns with Metrics and Classifier"
))
late_stock_results = pd.read_html(late_stock.summary().tables[1].as_html(),header=0,index_col=0)[0]
```

## Late 2020 Returns with Metrics and Classifier

```

=====
=
Dep. Variable:          RetLate2020    R-squared:                0.33
5
Model:                  OLS            Adj. R-squared:           0.25
8
Method:                 Least Squares   F-statistic:              4.33
3
Date:                   Thu, 08 Jul 2021 Prob (F-statistic):        1.42e-5
4
Time:                   14:36:37        Log-Likelihood:           -2310.
0
No. Observations:      1653            AIC:                      496
6.
Df Residuals:           1480            BIC:                      590
2.
Df Model:                172
Covariance Type:        nonrobust
=====
=====

```

```

=====
                                coef    std err          t
P>|t|    [0.025    0.975]
-----
const                                1.1719      0.178      6.594
0.000      0.823      1.521
Aerospace & Defense                   -0.2664      0.212     -1.258
0.209     -0.682      0.149
Agricultural & Farm Machinery          -0.1105      0.597     -0.185
0.853     -1.281      1.060
Agricultural Products                 -0.4557      0.399     -1.141
0.254     -1.239      0.328
Air Freight & Logistics                -0.3745      0.380     -0.987
0.324     -1.119      0.370
Airlines                             -0.1534      0.352     -0.436
0.663     -0.844      0.537
Airport Services                      -0.3830      1.031     -0.371
0.710     -2.406      1.640
Alternative Carriers                  -0.3164      0.532     -0.595
0.552     -1.360      0.727
Aluminum                             0.2418      0.731      0.331
0.741     -1.192      1.676
Apparel Retail                       0.5077      0.262      1.938
0.053     -0.006      1.022
Apparel, Accessories & Luxury Goods    0.3441      0.282      1.219
0.223     -0.210      0.898
Application Software                  0.0294      0.139      0.212
0.832     -0.243      0.301
Asset Management & Custody Banks       -0.5568      0.469     -1.188
0.235     -1.477      0.363
Auto Parts & Equipment                 0.4227      0.245      1.723
0.085     -0.059      0.904
Automobile Manufacturers              3.4729      0.546      6.362
0.000      2.402      4.544
Automotive Retail                     0.6928      0.307      2.258
0.024      0.091      1.295

```

Biotechnology			-0.5558	0.145	-3.840
0.000	-0.840	-0.272			
Brewers			-0.4975	0.531	-0.938
0.349	-1.538	0.543			
Building Products			0.0743	0.211	0.352
0.725	-0.339	0.488			
Casinos & Gaming			0.7688	0.277	2.771
0.006	0.225	1.313			
Coal & Consumable Fuels			-0.7231	0.476	-1.520
0.129	-1.656	0.210			
Commercial Printing			-0.9262	0.467	-1.984
0.047	-1.842	-0.010			
Commodity Chemicals			0.0591	0.273	0.216
0.829	-0.476	0.595			
Communications Equipment			-0.2318	0.206	-1.123
0.262	-0.637	0.173			
Computer & Electronics Retail			1.0852	0.622	1.746
0.081	-0.134	2.305			
Construction & Engineering			-0.1845	0.233	-0.793
0.428	-0.641	0.272			
Construction Machinery & Heavy Trucks			-0.1290	0.268	-0.481
0.631	-0.655	0.397			
Construction Materials			0.0365	0.395	0.092
0.926	-0.739	0.812			
Consumer Electronics			0.4125	0.469	0.880
0.379	-0.507	1.332			
Consumer Finance			-0.3563	0.520	-0.685
0.494	-1.377	0.665			
Copper			1.3312	0.769	1.731
0.084	-0.178	2.840			
Data Processing & Outsourced Services			-0.3177	0.215	-1.474
0.141	-0.740	0.105			
Department Stores			0.3496	0.532	0.657
0.511	-0.694	1.393			
Distillers & Vintners			0.0037	0.479	0.008
0.994	-0.935	0.943			
Distributors			-0.4156	0.468	-0.889
0.374	-1.333	0.501			
Diversified Banks			-3.042e-13	3.49e-13	-0.872
0.384	-9.89e-13	3.8e-13			
Diversified Chemicals			0.3830	0.596	0.642
0.521	-0.787	1.553			
Diversified Metals & Mining			-0.0631	0.730	-0.086
0.931	-1.494	1.368			
Diversified REITs			5.724e-14	7.18e-14	0.798
0.425	-8.35e-14	1.98e-13			
Diversified Real Estate Activities			-0.0007	1.042	-0.001
0.999	-2.044	2.043			
Diversified Support Services			-0.1046	0.423	-0.247
0.805	-0.935	0.726			
Drug Retail			-1.0863	0.775	-1.402
0.161	-2.607	0.434			
Education Services			-0.3726	0.319	-1.167
0.243	-0.999	0.254			
Electric Utilities			-0.9183	0.279	-3.288
0.001	-1.466	-0.371			
Electrical Components & Equipment			0.2852	0.233	1.225

0.221	-0.171	0.742			
Electronic Components			-0.0310	0.332	-0.094
0.925	-0.682	0.620			
Electronic Equipment & Instruments			-0.0934	0.260	-0.360
0.719	-0.603	0.416			
Electronic Manufacturing Services			-0.5574	0.349	-1.599
0.110	-1.241	0.126			
Environmental & Facilities Services			-0.5496	0.264	-2.084
0.037	-1.067	-0.032			
Fertilizers & Agricultural Chemicals			-0.1593	0.395	-0.403
0.687	-0.935	0.616			
Financial Exchanges & Data			-0.1182	0.385	-0.307
0.759	-0.874	0.638			
Food Distributors			-0.6866	0.382	-1.797
0.072	-1.436	0.063			
Food Retail			-0.7862	0.380	-2.067
0.039	-1.532	-0.040			
Footwear			0.3511	0.429	0.819
0.413	-0.490	1.192			
Forest Products			0.2246	1.030	0.218
0.827	-1.796	2.246			
Gas Utilities			-0.9668	0.343	-2.818
0.005	-1.640	-0.294			
General Merchandise Stores			0.6847	0.486	1.408
0.159	-0.269	1.638			
Gold			-0.2351	0.775	-0.304
0.762	-1.755	1.285			
Health Care Services			-0.0643	0.229	-0.280
0.779	-0.514	0.385			
Health Care Distributors			-0.4311	0.286	-1.505
0.132	-0.993	0.131			
Health Care Equipment			-0.3475	0.160	-2.178
0.030	-0.660	-0.035			
Health Care Facilities			0.1753	0.289	0.606
0.545	-0.392	0.743			
Health Care REITs			-3.74e-14	4.08e-14	-0.917
0.359	-1.17e-13	4.26e-14			
Health Care Supplies			-0.4294	0.247	-1.739
0.082	-0.914	0.055			
Health Care Technology			-0.2301	0.257	-0.894
0.371	-0.735	0.275			
Heavy Electrical Equipment			1.1727	0.597	1.965
0.050	0.002	2.343			
Home Furnishings			0.7258	0.424	1.711
0.087	-0.106	1.558			
Home Improvement Retail			3.0662	0.562	5.457
0.000	1.964	4.168			
Homebuilding			0.2621	0.518	0.506
0.613	-0.755	1.279			
Homefurnishing Retail			1.4691	0.354	4.150
0.000	0.775	2.163			
Hotel & Resort REITs			5.051e-14	5.88e-14	0.859
0.390	-6.48e-14	1.66e-13			
Hotels, Resorts & Cruise Lines			0.2021	0.355	0.569
0.569	-0.495	0.899			
Household Appliances			0.0993	0.519	0.191
0.848	-0.918	1.117			

Household Products			-0.3890	0.375	-1.037
0.300	-1.125	0.347			
Housewares & Specialties			5.9281	0.601	9.872
0.000	4.750	7.106			
Human Resource & Employment Services			0.1362	0.329	0.414
0.679	-0.509	0.781			
Hypermarkets & Super Centers			-0.0551	0.626	-0.088
0.930	-1.284	1.173			
IT Consulting & Other Services			-0.3684	0.321	-1.147
0.252	-0.999	0.262			
Independent Power Producers & Energy Traders			-0.4318	0.472	-0.915
0.360	-1.357	0.494			
Industrial Conglomerates			-0.0180	0.526	-0.034
0.973	-1.050	1.014			
Industrial Gases			-0.4093	0.752	-0.545
0.586	-1.883	1.065			
Industrial Machinery			-0.0346	0.147	-0.236
0.814	-0.322	0.253			
Industrial REITs			9.515e-15	1.08e-14	0.884
0.377	-1.16e-14	3.06e-14			
Insurance Brokers			-0.1801	0.412	-0.437
0.662	-0.988	0.627			
Integrated Oil & Gas			-0.5492	1.002	-0.548
0.584	-2.515	1.417			
Integrated Telecommunication Services			-1.5726	0.528	-2.980
0.003	-2.608	-0.537			
Internet & Direct Marketing Retail			1.0856	0.254	4.268
0.000	0.587	1.585			
Investment Banking & Brokerage			0.7032	1.031	0.682
0.495	-1.319	2.725			
Leisure Facilities			0.6691	0.467	1.434
0.152	-0.246	1.584			
Leisure Products			0.5708	0.268	2.131
0.033	0.045	1.096			
Life & Health Insurance			2.3709	1.030	2.302
0.021	0.351	4.391			
Life Sciences Tools & Services			0.2054	0.223	0.920
0.358	-0.233	0.643			
Managed Health Care			-0.2993	0.544	-0.550
0.583	-1.367	0.769			
Marine			-0.6320	0.496	-1.274
0.203	-1.605	0.341			
Metal & Glass Containers			-0.4619	0.303	-1.525
0.128	-1.056	0.132			
Mortgage REITs			-6.753e-16	6.68e-16	-1.011
0.312	-1.99e-15	6.35e-16			
Motorcycle Manufacturers			-0.0671	1.032	-0.065
0.948	-2.091	1.956			
Multi-Sector Holdings			-0.7563	1.033	-0.732
0.464	-2.782	1.269			
Multi-Utilities			-0.8967	0.303	-2.958
0.003	-1.492	-0.302			
Multi-line Insurance			1.532e-15	6.07e-16	2.524
0.012	3.41e-16	2.72e-15			
Office REITs			-1.229e-15	7.27e-16	-1.691
0.091	-2.65e-15	1.97e-16			
Office Services & Supplies			-0.4382	0.316	-1.386

0.166	-1.058	0.182			
Oil & Gas Drilling			-0.7828	0.555	-1.411
0.158	-1.871	0.305			
Oil & Gas Equipment & Services			0.0488	0.229	0.213
0.831	-0.401	0.499			
Oil & Gas Exploration & Production			-0.0354	0.249	-0.142
0.887	-0.523	0.452			
Oil & Gas Refining & Marketing			-0.1854	0.312	-0.594
0.553	-0.798	0.427			
Oil & Gas Storage & Transportation			-0.6453	0.364	-1.772
0.077	-1.360	0.069			
Other Diversified Financial Services			2.914e-16	5.35e-16	0.544
0.586	-7.59e-16	1.34e-15			
Packaged Foods & Meats			-0.5545	0.191	-2.910
0.004	-0.928	-0.181			
Paper Packaging			-0.2788	0.330	-0.845
0.398	-0.926	0.368			
Paper Products			-0.6392	0.426	-1.501
0.134	-1.475	0.196			
Personal Products			-0.3424	0.306	-1.118
0.264	-0.943	0.258			
Pharmaceuticals			-0.8778	0.186	-4.725
0.000	-1.242	-0.513			
Property & Casualty Insurance			-2.214e-16	4.61e-16	-0.480
0.631	-1.13e-15	6.83e-16			
Railroads			-0.1517	0.550	-0.276
0.783	-1.231	0.928			
Real Estate Development			-0.8950	1.039	-0.862
0.389	-2.932	1.142			
Real Estate Operating Companies			-6.121e-17	5.9e-16	-0.104
0.917	-1.22e-15	1.1e-15			
Real Estate Services			-0.2403	0.430	-0.558
0.577	-1.085	0.604			
Regional Banks			-3.85e-16	4.6e-16	-0.837
0.402	-1.29e-15	5.17e-16			
Reinsurance			4.972e-16	6.21e-16	0.801
0.423	-7.21e-16	1.72e-15			
Renewable Electricity			0.6620	0.750	0.882
0.378	-0.810	2.134			
Research & Consulting Services			-0.5081	0.232	-2.187
0.029	-0.964	-0.052			
Residential REITs			-1.03e-15	5.55e-16	-1.855
0.064	-2.12e-15	5.91e-17			
Restaurants			0.0968	0.212	0.457
0.648	-0.319	0.512			
Retail REITs			3.987e-16	4.12e-16	0.967
0.334	-4.1e-16	1.21e-15			
Security & Alarm Services			-0.9680	0.601	-1.611
0.107	-2.146	0.210			
Semiconductor Equipment			0.3557	0.240	1.481
0.139	-0.115	0.827			
Semiconductors			0.2991	0.187	1.599
0.110	-0.068	0.666			
Silver			1.3642	1.039	1.312
0.190	-0.675	3.403			
Soft Drinks			1.1416	0.409	2.792
0.005	0.339	1.944			

Specialized Consumer Services	-0.4209	0.399	-1.055
0.292 -1.204 0.362			
Specialized Finance	-5.936e-16	7.17e-16	-0.827
0.408 -2e-15 8.14e-16			
Specialized REITs	-0.6855	0.326	-2.101
0.036 -1.326 -0.045			
Specialty Chemicals	-0.0488	0.196	-0.249
0.803 -0.433 0.335			
Specialty Stores	0.6986	0.320	2.180
0.029 0.070 1.327			
Steel	0.0224	0.291	0.077
0.939 -0.548 0.592			
Systems Software	-0.1992	0.235	-0.848
0.397 -0.660 0.262			
Technology Distributors	-0.2342	0.436	-0.537
0.591 -1.090 0.621			
Technology Hardware, Storage & Peripherals	-0.1255	0.306	-0.410
0.682 -0.726 0.475			
Textiles	-0.3822	1.032	-0.370
0.711 -2.407 1.643			
Thrifts & Mortgage Finance	-1.749e-16	4.2e-16	-0.416
0.677 -9.98e-16 6.49e-16			
Tires & Rubber	0.1232	0.731	0.169
0.866 -1.311 1.557			
Tobacco	-0.0687	0.497	-0.138
0.890 -1.044 0.906			
Trading Companies & Distributors	-0.0246	0.215	-0.114
0.909 -0.447 0.398			
Trucking	-0.5445	0.292	-1.864
0.063 -1.118 0.029			
Water Utilities	-0.9199	0.364	-2.530
0.012 -1.633 -0.207			
Wireless Telecommunication Services	-0.7390	0.398	-1.858
0.063 -1.519 0.041			
current_assets	2.383e-05	3.07e-05	0.777
0.437 -3.63e-05 8.4e-05			
total_assets	-4.225e-05	1.93e-05	-2.184
0.029 -8.02e-05 -4.3e-06			
current_liabilities	-2.512e-05	3.19e-05	-0.786
0.432 -8.78e-05 3.75e-05			
total_liabilities	3.682e-05	2.11e-05	1.748
0.081 -4.49e-06 7.81e-05			
total_equity	7.86e-05	5.48e-05	1.436
0.151 -2.88e-05 0.000			
mkt_value	2.831e-06	2.49e-06	1.135
0.257 -2.06e-06 7.72e-06			
price_close	-0.0008	0.001	-1.457
0.145 -0.002 0.000			
shares_out	-0.0002	0.000	-1.440
0.150 -0.001 8.07e-05			
book_value_share	-0.0053	0.002	-2.167
0.030 -0.010 -0.000			
ppe	3.181e-06	7.74e-06	0.411
0.681 -1.2e-05 1.84e-05			
employees	-0.0014	0.001	-1.455
0.146 -0.003 0.000			
cash	-3.414e-05	4.01e-05	-0.852



0.394	-0.000	4.45e-05			
receivables			2.805e-05	3.29e-05	0.854
0.393	-3.64e-05	9.25e-05			
sales			6.359e-06	1.33e-05	0.478
0.633	-1.97e-05	3.24e-05			
cogs			-6.648e-06	1.32e-05	-0.503
0.615	-3.26e-05	1.93e-05			
inventory			-7.797e-05	3.97e-05	-1.965
0.050	-0.000	-1.17e-07			
ebit			-5.84e-05	8.08e-05	-0.723
0.470	-0.000	0.000			
ebitda			3.299e-05	6.63e-05	0.498
0.619	-9.71e-05	0.000			
interest_exp			8.72e-05	0.000	0.300
0.765	-0.000	0.001			
working_cap			-5.147e-07	3.56e-05	-0.014
0.988	-7.04e-05	6.94e-05			
net_income			-4.022e-05	5.52e-05	-0.729
0.466	-0.000	6.8e-05			
book_value			-2.502e-05	5.32e-05	-0.471
0.638	-0.000	7.93e-05			
acid_ratio			0.1026	0.074	1.387
0.166	-0.043	0.248			
current_ratio			-0.1269	0.070	-1.803
0.072	-0.265	0.011			
coverage_ratio			-0.0272	0.031	-0.874
0.382	-0.088	0.034			
bep			-1.5506	0.584	-2.654
0.008	-2.697	-0.404			
inventory_turnover			33.7459	38.689	0.872
0.383	-42.145	109.637			
dso			-0.0003	0.000	-0.715
0.475	-0.001	0.001			
roa			0.6904	0.548	1.259
0.208	-0.386	1.766			
roe			0.0953	0.024	4.033
0.000	0.049	0.142			
debit_asset_ratio			0.2444	0.174	1.401
0.162	-0.098	0.587			
debit_equity_ratio			-0.0085	0.005	-1.679
0.093	-0.018	0.001			
book_price			-0.0702	0.100	-0.698
0.485	-0.267	0.127			
equity_price			-0.2962	0.281	-1.052
0.293	-0.848	0.256			
ebit_price			-2.5756	0.577	-4.464
0.000	-3.707	-1.444			
ebitda_price			1.7403	0.439	3.969
0.000	0.880	2.600			
sales_price			0.0630	0.029	2.205
0.028	0.007	0.119			
cash_assets			0.0817	0.364	0.224
0.822	-0.632	0.796			
cash_current_liab			-0.0863	0.055	-1.563
0.118	-0.195	0.022			
cash_total_liab			0.1693	0.085	2.001
0.046	0.003	0.335			

cash_interest			0.0019	0.007	0.255
0.799	-0.013	0.016			
fixed_total			-0.0366	0.221	-0.166
0.868	-0.470	0.396			
fixed_employees			-8.859e-06	1.12e-05	-0.794
0.427	-3.07e-05	1.3e-05			
working_coverage			-1.144e-05	5.46e-05	-0.209
0.834	-0.000	9.57e-05			
=====					
=					
Omnibus:	1311.683	Durbin-Watson:	2.13		
0					
Prob(Omnibus):	0.000	Jarque-Bera (JB):	71710.85		
2					
Skew:	3.247	Prob(JB):	0.0		
0					
Kurtosis:	34.607	Cond. No.	1.17e+1		
8					
=====					
=					

#### Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The smallest eigenvalue is 3.42e-24. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

## Final Results

```
In [335]: early_stock_results.loc[early_stock_results['P>|t|']<0.05,].sort_values("coef", ascending=False)
```

Out[335]:

	coef	std err	t	P> t	[0.025	0.975]
<b>Integrated Telecommunication Services</b>	0.525300	0.161000	3.255	0.001	0.209000	0.842000
<b>bep</b>	0.477100	0.179000	2.670	0.008	0.127000	0.828000
<b>Health Care Distributors</b>	0.305600	0.088000	3.490	0.000	0.134000	0.477000
<b>Food Retail</b>	0.274200	0.116000	2.358	0.019	0.046000	0.502000
<b>Water Utilities</b>	0.249500	0.111000	2.243	0.025	0.031000	0.468000
<b>Multi-Utilities</b>	0.192100	0.093000	2.072	0.038	0.010000	0.374000
<b>Electric Utilities</b>	0.175100	0.085000	2.051	0.040	0.008000	0.343000
<b>Systems Software</b>	0.166000	0.072000	2.311	0.021	0.025000	0.307000
<b>Health Care Equipment</b>	0.139200	0.049000	2.853	0.004	0.043000	0.235000
<b>net_income</b>	0.000053	0.000017	3.114	0.002	0.000019	0.000086
<b>roe</b>	-0.029300	0.007000	-4.057	0.000	-0.043000	-0.015000
<b>book_price</b>	-0.071300	0.031000	-2.323	0.020	-0.132000	-0.011000
<b>debit_asset_ratio</b>	-0.116400	0.053000	-2.181	0.029	-0.221000	-0.012000
<b>Apparel Retail</b>	-0.163500	0.080000	-2.041	0.041	-0.321000	-0.006000
<b>Apparel, Accessories &amp; Luxury Goods</b>	-0.172300	0.086000	-1.996	0.046	-0.342000	-0.003000
<b>Casinos &amp; Gaming</b>	-0.184800	0.085000	-2.178	0.030	-0.351000	-0.018000
<b>const</b>	-0.215200	0.054000	-3.960	0.000	-0.322000	-0.109000
<b>Hotels, Resorts &amp; Cruise Lines</b>	-0.277700	0.109000	-2.557	0.011	-0.491000	-0.065000
<b>Oil &amp; Gas Equipment &amp; Services</b>	-0.311700	0.070000	-4.445	0.000	-0.449000	-0.174000
<b>Oil &amp; Gas Exploration &amp; Production</b>	-0.347400	0.076000	-4.569	0.000	-0.497000	-0.198000
<b>roa</b>	-0.365800	0.168000	-2.181	0.029	-0.695000	-0.037000
<b>Oil &amp; Gas Drilling</b>	-0.466100	0.170000	-2.748	0.006	-0.799000	-0.133000
<b>ebit_price</b>	-0.671600	0.176000	-3.806	0.000	-1.018000	-0.325000

## Early 2020 Return Explaintion

**Adj R<sup>2</sup>: 0.218**

The top positive or least negative variables at the 95% statisitcal signifiant level were:

- Integrated Telecommunication Services
- bep
- Health Care Distributors
- Food Retail
- Water Utilities

Telecommunications, Health Care Distributors, and Water Utilites makes sense. I am surprised Food Retail scored so high.

The top negative variables at the 95% statisitcal signifiant level were:

- ebit\_price
- Oil & Gas Drilling
- roa
- Oil & Gas Exploration & Production
- Oil & Gas Equipment & Services

All of these variables were the top statisitcal signifiant negative variables in their respective models (metric and classification). Intersting to see more clasification than metrics. Additionally I must point out EBIT to Price persisted in magnitude and statisitcal signifiant unlike Sales to Price. Also iteresting to see Return on Assets (roa) showing up.

```
In [336]: late_stock_results.loc[late_stock_results['P>|t|']<0.05,].sort_values("coef",  
ascending=False)
```

Out[336]:

	coef	std err	t	P> t	[0.025	0.975]
<b>Housewares &amp; Specialties</b>	5.928100e+00	6.010000e-01	9.872	0.000	4.750000e+00	7.106000e+00
<b>Automobile Manufacturers</b>	3.472900e+00	5.460000e-01	6.362	0.000	2.402000e+00	4.544000e+00
<b>Home Improvement Retail</b>	3.066200e+00	5.620000e-01	5.457	0.000	1.964000e+00	4.168000e+00
<b>Life &amp; Health Insurance</b>	2.370900e+00	1.030000e+00	2.302	0.021	3.510000e-01	4.391000e+00
<b>ebitda_price</b>	1.740300e+00	4.390000e-01	3.969	0.000	8.800000e-01	2.600000e+00
<b>Homefurnishing Retail</b>	1.469100e+00	3.540000e-01	4.150	0.000	7.750000e-01	2.163000e+00
<b>const</b>	1.171900e+00	1.780000e-01	6.594	0.000	8.230000e-01	1.521000e+00
<b>Soft Drinks</b>	1.141600e+00	4.090000e-01	2.792	0.005	3.390000e-01	1.944000e+00
<b>Internet &amp; Direct Marketing Retail</b>	1.085600e+00	2.540000e-01	4.268	0.000	5.870000e-01	1.585000e+00
<b>Casinos &amp; Gaming</b>	7.688000e-01	2.770000e-01	2.771	0.006	2.250000e-01	1.313000e+00
<b>Specialty Stores</b>	6.986000e-01	3.200000e-01	2.180	0.029	7.000000e-02	1.327000e+00
<b>Automotive Retail</b>	6.928000e-01	3.070000e-01	2.258	0.024	9.100000e-02	1.295000e+00
<b>Leisure Products</b>	5.708000e-01	2.680000e-01	2.131	0.033	4.500000e-02	1.096000e+00
<b>cash_total_liab</b>	1.693000e-01	8.500000e-02	2.001	0.046	3.000000e-03	3.350000e-01
<b>roe</b>	9.530000e-02	2.400000e-02	4.033	0.000	4.900000e-02	1.420000e-01
<b>sales_price</b>	6.300000e-02	2.900000e-02	2.205	0.028	7.000000e-03	1.190000e-01
<b>Multi-line Insurance</b>	1.532000e-15	6.070000e-16	2.524	0.012	3.410000e-16	2.720000e-15
<b>total_assets</b>	-4.225000e-05	1.930000e-05	-2.184	0.029	-8.020000e-05	-4.300000e-06
<b>book_value_share</b>	-5.300000e-03	2.000000e-03	-2.167	0.030	-1.000000e-02	-0.000000e+00
<b>Health Care Equipment</b>	-3.475000e-01	1.600000e-01	-2.178	0.030	-6.600000e-01	-3.500000e-02
<b>Research &amp; Consulting Services</b>	-5.081000e-01	2.320000e-01	-2.187	0.029	-9.640000e-01	-5.200000e-02
<b>Environmental &amp; Facilities Services</b>	-5.496000e-01	2.640000e-01	-2.084	0.037	-1.067000e+00	-3.200000e-02
<b>Packaged Foods &amp; Meats</b>	-5.545000e-01	1.910000e-01	-2.910	0.004	-9.280000e-01	-1.810000e-01
<b>Biotechnology</b>	-5.558000e-01	1.450000e-01	-3.840	0.000	-8.400000e-01	-2.720000e-01
<b>Specialized REITs</b>	-6.855000e-01	3.260000e-01	-2.101	0.036	-1.326000e+00	-4.500000e-02
<b>Food Retail</b>	-7.862000e-01	3.800000e-01	-2.067	0.039	-1.532000e+00	-4.000000e-02
<b>Pharmaceuticals</b>	-8.778000e-01	1.860000e-01	-4.725	0.000	-1.242000e+00	-5.130000e-01
<b>Multi-Utilities</b>	-8.967000e-01	3.030000e-01	-2.958	0.003	-1.492000e+00	-3.020000e-01
<b>Electric Utilities</b>	-9.183000e-01	2.790000e-01	-3.288	0.001	-1.466000e+00	-3.710000e-01
<b>Water Utilities</b>	-9.199000e-01	3.640000e-01	-2.530	0.012	-1.633000e+00	-2.070000e-01

	coef	std err	t	P> t	[0.025	0.975]
<b>Commercial Printing</b>	-9.262000e-01	4.670000e-01	-1.984	0.047	-1.842000e+00	-1.000000e-02
<b>Gas Utilities</b>	-9.668000e-01	3.430000e-01	-2.818	0.005	-1.640000e+00	-2.940000e-01
<b>bep</b>	-1.550600e+00	5.840000e-01	-2.654	0.008	-2.697000e+00	-4.040000e-01
<b>Integrated Telecommunication Services</b>	-1.572600e+00	5.280000e-01	-2.980	0.003	-2.608000e+00	-5.370000e-01
<b>ebit_price</b>	-2.575600e+00	5.770000e-01	-4.464	0.000	-3.707000e+00	-1.444000e+00

## Late 2020 Return Explaintion

**Adj R<sup>2</sup>: 0.258**

The top positive variables at the 95% statisitcal signifiant level were:

- Housewares & Specialties
- Automobile Manufacturers
- Home Improvement Retail
- Life & Health Insurance
- ebitda\_price

Intersting to see all top positive variables were classification. This means a big indicator of success in late 2020 was the industry.

The top negative variables or least positive at the 95% statisitcal signifiant level were:

- ebit\_price
- Integrated Telecommunication Services
- bep
- Gas Utilities
- Commerical Printing

Intersting to see all top negative variables were metrics. EBIT to Price remains an important indicator even for late 2020 returns. Basic Earning Power was a new predictor of low returns for late 2020. Telecommunications and Gas Utilities makes sense but im not so sure about Commerical Printing