FINAL PROJECT PHYTON FOR FINANCE

PROJECT TITLE

Exploring the Impact of Covid-19 on Financial Performance of Health and Industrial Firms:
Evidence from Cross-Sectional Study of US Market

GROUP 1

1. Amin Ilyas
2. Hélène Do
3. Logba Elvire Ursule Vera Yapi
4. Nico Benedikt Horstmann
5. Pritam Ritu Raj

Contents

1. Project Overview
2. Data and Sample
3. Data Preparation
4. Python (brief coding) and Results
5. Discussion (all interpretations: hypotheses)
6. Conclusion
7. Recommendation
8. List of References

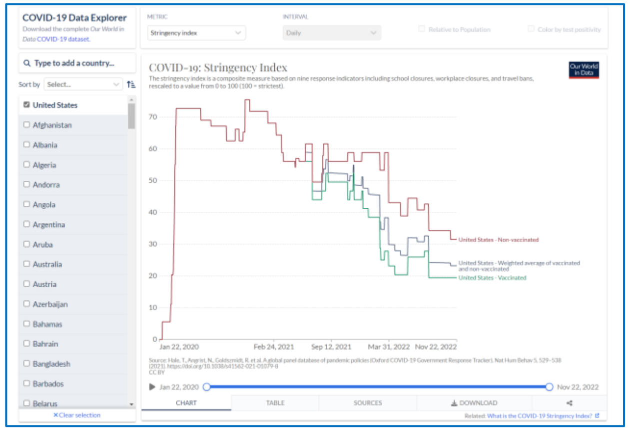
Project Overview

- ***Project Topic***:
  - Analysing Impact of Covid-19 on quarterly firm performance in US
- ***Time Horizon***:
  - COVID-19 and quarterly (continuous)
  - 2020 - 2022: Q1 – Q4
- ***Research Method***:
  - CROSS SECTIONAL
- ***Aim***:
  - To analyze the relationship between:
    - Firms' financial performance (Altman Z-score) in a sector in US
    - COVID-19 data in US
    - Macroeconomic variable: Federal Funds Effective Rate in US
  - With analysis on heterogeneity between 2 (two) industries/ sectors in US: **HEALTH** and **INDUSTRIAL**
- ***Variables***:
  - **Dependent variable (DV)**
    - Altman Z-Score : Z_Score
  - **Independent variable**
    - Stringency Index : *SI*
    - Number of Cases : *NC*
    - Federal funds Effective Rate : *FF*
    - Size of firms : *size*

- ***Data Processing***
- **PART 0 Data Preparation**
  - 0.1. IV1 and IV2 csv in quarterly data
  - 0.2. Split WRDS csv data into two sectors
  - 0.3. For each sector, convert WRDS csv into DV csv (i.e. compute z-score, using excel)
- **PART 1 Descriptive Statistics**
  - 1.1. Preliminary Data Visualization (# samples, etc.)
  - 1.2. Univariate Descriptive Statistics
  - 1.3. Histograms
  - 1.4. Winsorizing (if necessary)
  - 1.5. Correlation coefficients:
    - 1.5.1. Pearson
    - 1.5.2. Spearman's rank
- **PART 2 Inferential Statistics**
  - 2.1 Classic assumption test
    - 2.1.1 Jarque Bera
    - 2.1.2 Kolmogorov-Smirnov
  - 2.2. Hypothesis Test:
    - 2.2.1. One Sample t-Test
    - 2.2.2. Two Sample t-Test
    - 2.2.3. Simple Regression
    - 2.2.4 Multiple Linear Regression
    - 2.2.5. Quantile Regression
- **PART 3 Heterogeneity**

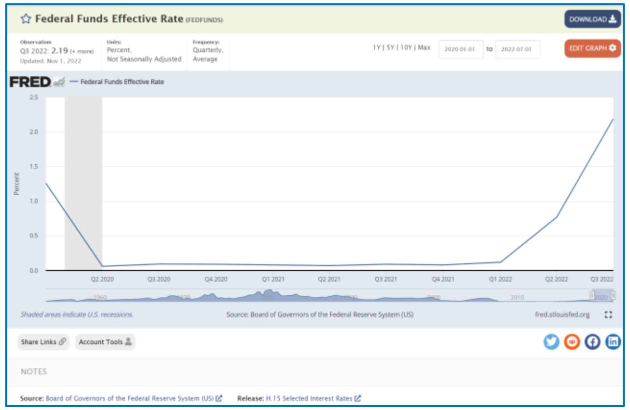
Data and Sample (1)

COVID 19 Data: Stringency Index (DAILY)



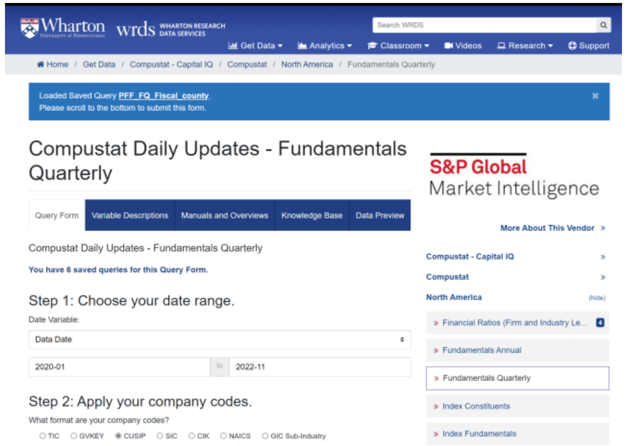
https://ourworldindata.org/explorers/coronavirus-data-explorer?uniformYAxis=0&hideControls=true&Metric=Stringency+index&Interval=7-day+rolling+average&Relative+to+Population=true&Color+by+test+positivity=false&country=~USA

Macroeconomic Data: Federal Funds ER (QUARTERLY)



https://fred.stlouisfed.org/series/FEDFUNDS#0

Financial Data – WRDS Fundamentals (QUARTERLY)



https://wrds-www.wharton.upenn.edu/pages/get-data/compustat-capital-iq-standard-poors/compustat/north-america-daily/fundamentals-quarterly/?saved_query=2316916

Data and Sample (2)

gvkey	datadate	fyearq	fqtr	indfmt	consol	popsrc	datafmt	tic	cusip	conm	curcdq	datacqtr	datafqtr	cogsq	niq	req	revtq	txtq	wcapq	xintq	xoprq	costat	mkvaltq	busdesc	gsector
1004	20200229	2019	3	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2020Q1	2019Q3	461.9	2.3	685.2	562.9	0.2	672.8	2.4	518.7	A	1212.74	AAR Corp.	20
1004	20200531	2019	4	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2020Q2	2019Q4	353.7	-16.5	661.4	424	-4	1055.6	2.8	397.7	A	707.9065	AAR Corp.	20
1004	20200831	2020	1	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2020Q3	2020Q1	341.3	-14.5	648.3	402.7	-3.8	679.9	1.7	383.7	A	712.3742	AAR Corp.	20
1004	20201130	2020	2	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2020Q4	2020Q2	332.2	8.2	656.6	401.7	5.2	633.5	1.3	376.4	A	1001.234	AAR Corp.	20
1004	20210228	2020	3	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2021Q1	2020Q3	336.8	28.1	685.3	411.8	12	640.7	1.1	383.4	A	1404.99	AAR Corp.	20
1004	20210531	2020	4	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2021Q2	2020Q4	354.3	14	723.4	435.2	4.8	600.2	0.9	406.1	A	1476.906	AAR Corp.	20
1004	20210831	2021	1	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2021Q3	2021Q1	373.7	11.5	734.6	456.1	3.9	620.7	0.7	422	A	1200.693	AAR Corp.	20
1004	20211130	2021	2	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2021Q4	2021Q2	352.5	20.8	751.8	434.1	7.9	626.5	0.5	398.3	A	1158.418	AAR Corp.	20
1004	20220228	2021	3	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2022Q1	2021Q3	365.3	22.5	776.4	452	8.2	634.1	0.6	412.7	A	1581.753	AAR Corp.	20
1004	20220531	2021	4	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2022Q2	2021Q4	378.8	23.9	800.8	474.9	6.6	659	0.6	434.8	A	1706.554	AAR Corp.	20
1004	20220831	2022	1	INDL	C	D	STD	AIR	361105	AAR CORP	USD	2022Q3	2022Q1	358.7	22.7	820.4	446.4	8.1	677.3	1.1	408.1	A	1504.359	AAR Corp.	20
1019	20200331	2020	1	INDL	C	D	STD	AFAP	1038108	AFA PROT	USD	2020Q1	2020Q1									A		AFA Prote	20
1019	20200630	2020	2	INDL	C	D	STD	AFAP	1038108	AFA PROT	USD	2020Q2	2020Q2									A		AFA Prote	20
1019	20200930	2020	3	INDL	C	D	STD	AFAP	1038108	AFA PROT	USD	2020Q3	2020Q3									A		AFA Prote	20
1019	20201231	2020	4	INDL	C	D	STD	AFAP	1038108	AFA PROT	USD	2020Q4	2020Q4	55.673	2.603	10.262	82.679	1.2	15.525	0.109	76.909	A	28.188	AFA Prote	20
1045	20200331	2020	1	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2020Q1	2020Q1	7541	-2241	-6501	8515	-649	-12038	257	9196	A	5154.993	American	20
1045	20200630	2020	2	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2020Q2	2020Q2	4538	-2067	-8551	1622	-592	-4211	255	5195	A	6646.827	American	20
1045	20200930	2020	3	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2020Q3	2020Q3	5191	-2399	-10963	3172	-696	-4244	339	5986	A	6250.743	American	20
1045	20201231	2020	4	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2020Q4	2020Q4	5237	-2178	-13767	4028	-631	-5474	376	5982	A	9800.74	American	20
1045	20210331	2021	1	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2021Q1	2021Q1	5820	-1250	-14931	4008	-323	756	371	6687	A	15328.84	American	20
1045	20210630	2021	2	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2021Q2	2021Q2	6697	19	-14873	7478	-10	1126	485	7932	A	13732.33	American	20
1045	20210930	2021	3	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2021Q3	2021Q3	7448	170	-14664	8969	37	257	476	8875	A	13286.23	American	20
1045	20211231	2021	4	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2021Q4	2021Q4	8004	-932	-14580	9427	-259	-1670	468	9567	A	11633.19	American	20
1045	20220331	2022	1	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2022Q1	2022Q1	8276	-1635	-16189	8899	-451	-4104	463	9893	A	11853.34	American	20
1045	20220630	2022	2	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2022Q2	2022Q2	9933	476	-15687	13422	127	-4245	469	11827	A	8239.198	American	20
1045	20220930	2022	3	INDL	C	D	STD	AAL	02376R10	AMERICA	USD	2022Q3	2022Q3	10064	483	-15176	13462	176	-4593	498	11921	A	7824.351	American	20
1050	20200331	2020	1	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2020Q1	2020Q1	51.723	3.412	-59.705	80.486	0.779	109.905	1.023	73.699	A	164.384	CECO Envi	20
1050	20200630	2020	2	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2020Q2	2020Q2	48.688	3.258	-55.28	75.17	0.564	73.997	0.944	67.095	A	232.996	CECO Envi	20
1050	20200930	2020	3	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2020Q3	2020Q3	52.087	-0.239	-54.155	77.425	0.206	74.055	0.772	71.07	A	257.8254	CECO Envi	20
1050	20201231	2020	4	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2020Q4	2020Q4	56.073	1.78	-52.637	82.93	2.123	74.129	0.796	73.633	A	246.1543	CECO Envi	20
1050	20210331	2021	1	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2021Q1	2021Q1	46.841	1.181	-51.401	71.892	0.551	75.434	0.725	66.295	A	280.9599	CECO Envi	20
1050	20210630	2021	2	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2021Q2	2021Q2	53.235	0.293	-50.818	78.68	0.199	75.305	0.705	73.746	A	254.9819	CECO Envi	20
1050	20210930	2021	3	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2021Q3	2021Q3	56.564	-1.25	-52.317	79.979	0.063	71.299	0.722	77.493	A	247.456	CECO Envi	20
1050	20211231	2021	4	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2021Q4	2021Q4	64.514	1.202	-48.785	93.589	1.878	72.326	0.8	85.417	A	218.2244	CECO Envi	20
1050	20220331	2022	1	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2022Q1	2022Q1	65.172	2.792	-46.524	92.436	1.112	79.248	0.822	86.324	A	192.5672	CECO Envi	20
1050	20220630	2022	2	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2022Q2	2022Q2	72.768	4.388	-45.105	105.375	1.86	74.69	1.098	95.756	A	206.5133	CECO Envi	20
1050	20220930	2022	3	INDL	C	D	STD	CECO	1.25E+08	CECO ENV	USD	2022Q3	2022Q3	75.086	1.942	-49.052	108.414	0.315	82.867	1.569	100.252	A	303.8205	CECO Envi	20
1062	20200229	2020	1	INDL	C	D	STD	ASA	G3156P10	ASA GOLD	USD	2020Q1	2020Q1	0.886	-8.931	256.285	0.157	0		0	0.886	A	237.8457	ASA Gold	40
1062	20200531	2020	2	INDL	C	D	STD	ASA	G3156P10	ASA GOLD	USD	2020Q2	2020Q2	0.842	87.373	343.466	0.503	0		0	0.842	A	295.137	ASA Gold	40
1062	20200831	2020	3	INDL	C	D	STD	ASA	G3156P10	ASA GOLD	USD	2020Q3	2020Q3	1.201	158.462	501.927	0.243	0		0	1.201	A	449.0712	ASA Gold	40
1062	20201130	2020	4	INDL	C	D	STD	ASA	G3156P10	ASA GOLD	USD	2020Q4	2020Q4	1.045	-58.461	443.274	0.43	0		0	1.045	A	384.0639	ASA Gold	40

WRDS FundQuart_NorthAmerica_Q1- | GICS

10	Energy	3694	
15	Materials	4346	
20	Industrials	7773	
25	Consumer Discretionary	7099	
30	Consumer Staples	2806	
35	Health Care	14558	
40	Financials	10833	
45	IT	8696	
50	Communication Services	3305	
55	Utilities	2492	
60	Real estates	2677	
	Total	68279	TRUE

GCSI Structure

https://www.msci.com/documents/1296102/11185224/GICS+Methodology+2022.pdf/f9910041-6127-17d2-1246-4052926adaf7?t=1645738126436

Data Preparation

1. Altman Z-Score computation
2. Creation of COVID- and FederalFunds Quarterly Data
3. Duplication of Stringency Index Quarterly Data
4. Duplication of Covid and Federal Funds Quarterly Data
5. Merging data and using the common Sample Period

Sample Period:

2020Q1 to 2022Q3 (Covid period)

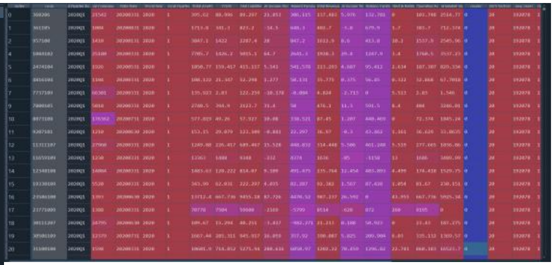
$$Z\ score = 1.2 \times \frac{WC}{TA} + 1.4 \times \frac{RE}{TA} + 3.3 \times \frac{EBIT}{TA} + 0.6 \times \frac{ME}{TL} + 1.0 \times \frac{SALE}{TA}$$

```
#EBIT calculation
FundQ_sector["EBIT"] = FundQ_sector["Total Revenue"] - FundQ_sector["COGS"] - FundQ_sector["Total Operating Expense"]
```

The Z-Score is then calculated the same way like we did in Class

Index	il Quarter by	new cases	ringency indi
0	2020Q1	192078	17.0476
1	2020Q2	2.45658e+06	72.0304
2	2020Q3	4.59754e+06	66.5305
3	2020Q4	1.29711e+07	68.5199
4	2021Q1	1.03601e+07	67.6377
5	2021Q2	3.20289e+06	56.6393

Index	DATE	FEDFUNDS	il Quarter by
0	1/1/2020	0.0126	2020Q1
1	4/1/2020	0.0006	2020Q2
2	7/1/2020	0.000933333	2020Q3
3	10/1/2020	0.0009	2020Q4
4	1/1/2021	0.0008	2021Q1
5	4/1/2021	0.0007	2021Q2
6	7/1/2021	0.0009	2021Q3
7	10/1/2021	0.0008	2021Q4

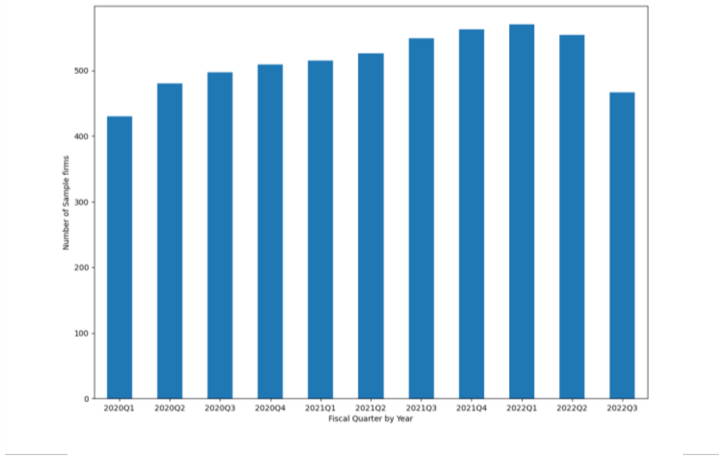


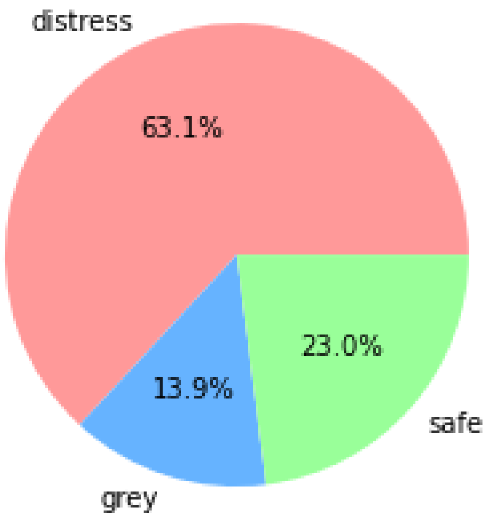
- Only 11 values for the COVID Data ➔ We needed to merge them with the firm data such that the values are duplicated and matched to the respective quarter
- Selection of the columns we needed for the statistical analysis.

Index	cusip	il Quarter by	new cases	ringency inde	FEDFUNDS	al Market Va	Z Score	size	[illegible]	logNC	logSI	logFF
0	360206	2020Q1	192078	17.0476	0.0126	2514.77	18.2697	7.82994	safe	12.1657	2.89301	0.0125213
1	361105	2020Q1	192078	17.0476	0.0126	712.374	1.13925	6.5686	distress	12.1657	2.89301	0.0125213
2	957100	2020Q1	192078	17.0476	0.0126	2545.96	0.36901	7.84226	distress	12.1657	2.89301	0.0125213
4	1084102	2020Q1	192078	17.0476	0.0126	3537.23	0.808733	8.1711	distress	12.1657	2.89301	0.0125213
5	2474104	2020Q1	192078	17.0476	0.0126	829.334	1.81317	6.72062	grey	12.1657	2.89301	0.0125213
6	4816104	2020Q1	192078	17.0476	0.0126	67.7018	1.82042	4.21511	grey	12.1657	2.89301	0.0125213
7	7737109	2020Q1	192078	17.0476	0.0126	1.546	0.0219159	0.435671	distress	12.1657	2.89301	0.0125213
9	7800105	2020Q1	192078	17.0476	0.0126	3246.01	0.986652	8.08518	distress	12.1657	2.89301	0.0125213
10	8073108	2020Q1	192078	17.0476	0.0126	1845.24	20.8038	7.52036	safe	12.1657	2.89301	0.0125213
11	9207101	2020Q1	192078	17.0476	0.0126	33.8635	0.331576	3.52234	distress	12.1657	2.89301	0.0125213
12	11311107	2020Q1	192078	17.0476	0.0126	1036.86	1.5988	6.94395	distress	12.1657	2.89301	0.0125213
13	11659109	2020Q1	192078	17.0476	0.0126	3489.99	0.320884	8.15766	distress	12.1657	2.89301	0.0125213
14	12348108	2020Q1	192078	17.0476	0.0126	1529.75	1.9928	7.33286	grey	12.1657	2.89301	0.0125213
15	19330109	2020Q1	192078	17.0476	0.0126	230.151	1.03733	5.43873	distress	12.1657	2.89301	0.0125213
16	23586100	2020Q1	192078	17.0476	0.0126	5925.34	0.820626	8.68699	distress	12.1657	2.89301	0.0125213
18	30111207	2020Q1	192078	17.0476	0.0126	187.275	-9.52394	5.23258	distress	12.1657	2.89301	0.0125213
19	30506109	2020Q1	192078	17.0476	0.0126	1369.57	1.09835	7.22225	distress	12.1657	2.89301	0.0125213
20	31100100	2020Q1	192078	17.0476	0.0126	16523.7	2.82231	9.71255	grey	12.1657	2.89301	0.0125213
21	37598109	2020Q1	192078	17.0476	0.0126	544.52	0.629387	6.2999	distress	12.1657	2.89301	0.0125213

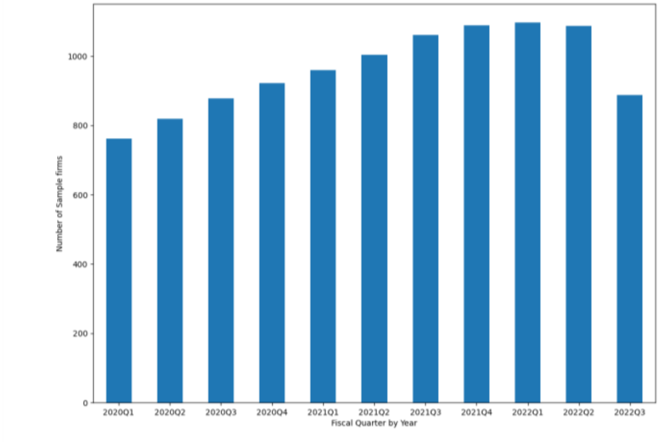
1.1. Preliminary Data Visualization (Number of .)

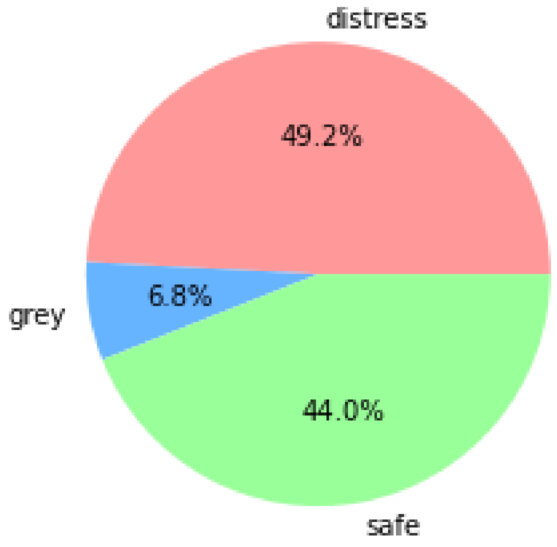
SECTOR 1: INDUSTRIAL





SECTOR 2: HEALTH





1.2. Univariate Descriptive Statistics

SECTOR 1: INDUSTRIAL

	new_cases	stringency_index	FEDFUNDS	Z_Score
count	5.660000e+03	5660.000000	5660.000000	5660.000000
mean	9.108190e+06	50.307819	0.004157	-114.186900
std	6.607299e+06	17.299089	0.006382	5821.615094
min	1.920780e+05	17.047597	0.000600	-430411.821234
25%	3.202890e+06	29.152842	0.000800	0.228439
50%	8.686982e+06	51.214645	0.000900	1.189588
75%	1.136729e+07	67.637665	0.007700	2.776037
max	2.533413e+07	72.030444	0.021900	2492.176716

	logNC	logSI	logFF
count	5660.000000	5660.000000	5660.000000
mean	15.614944	3.864351	0.004129
std	1.174522	0.411697	0.006319
min	12.165662	2.893013	0.000600
25%	14.979564	3.406279	0.000800
50%	15.977336	3.955363	0.000900
75%	16.246250	4.228841	0.007671
max	17.047663	4.290876	0.021664

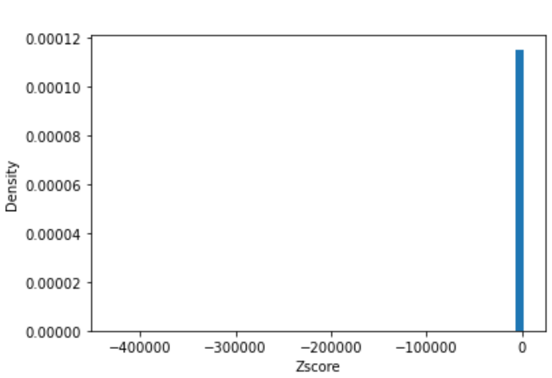
SECTOR 2: HEALTH

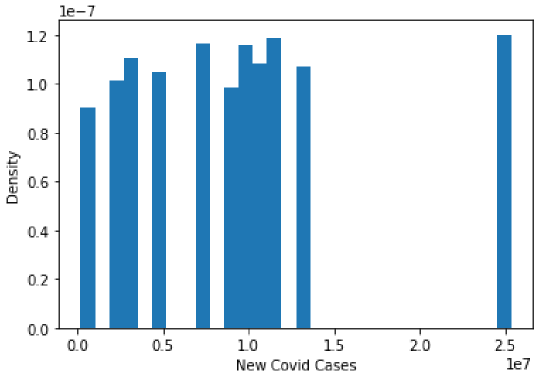
	new_cases	stringency_index	FEDFUNDS	Z_Score
count	1.056200e+04	10562.000000	10562.000000	10562.000000
mean	9.240568e+06	50.008811	0.004179	-174.501036
std	6.617038e+06	17.104406	0.006398	4089.558398
min	1.920780e+05	17.047597	0.000600	-177398.552000
25%	4.597537e+06	29.152842	0.000800	-2.007426
50%	8.686982e+06	51.214645	0.000900	1.808619
75%	1.136729e+07	67.637665	0.007700	9.108121
max	2.533413e+07	72.030444	0.021900	2300.534229

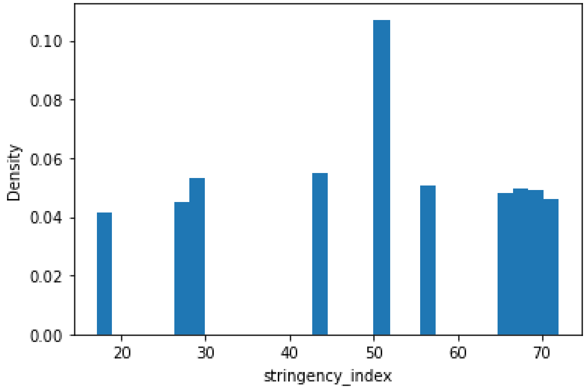
	logNC	logSI	logFF
count	10562.000000	10562.000000	10562.000000
mean	15.642457	3.860041	0.004150
std	1.155428	0.406362	0.006334
min	12.165662	2.893013	0.000600
25%	15.341032	3.406279	0.000800
50%	15.977336	3.955363	0.000900
75%	16.246250	4.228841	0.007671
max	17.047663	4.290876	0.021664

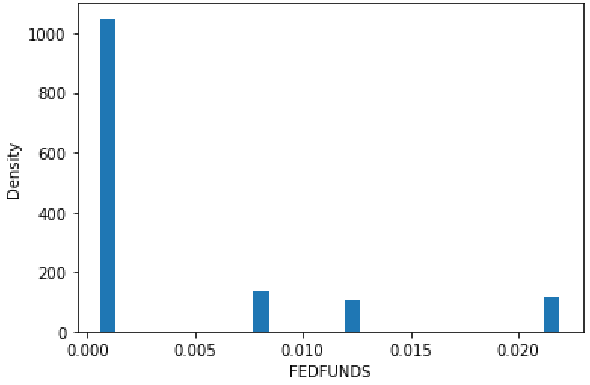
1.3. Histograms

SECTOR 1: INDUSTRIAL

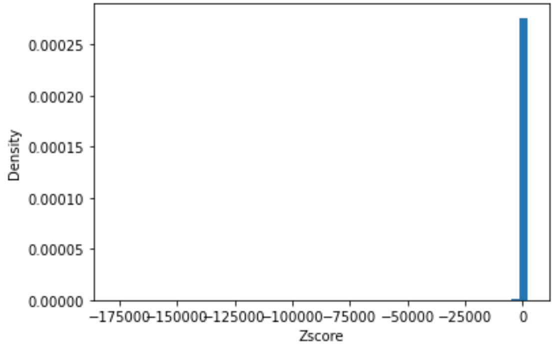


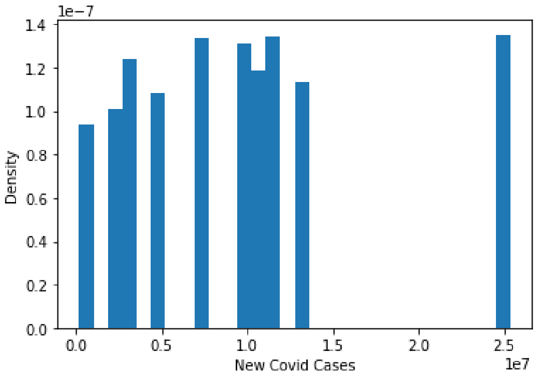


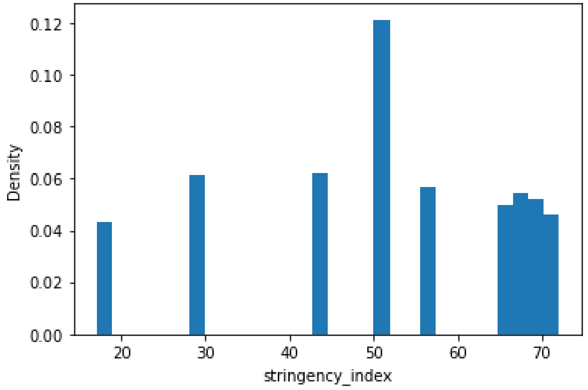


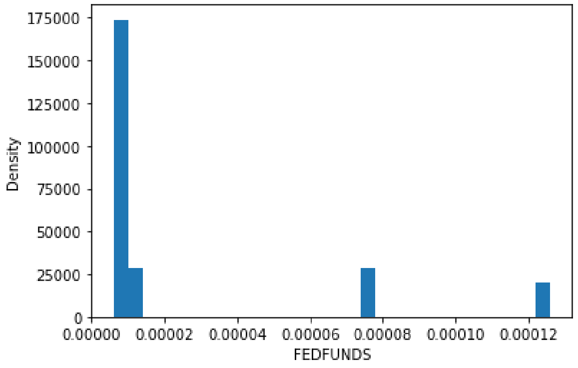


SECTOR 2: HEALTH



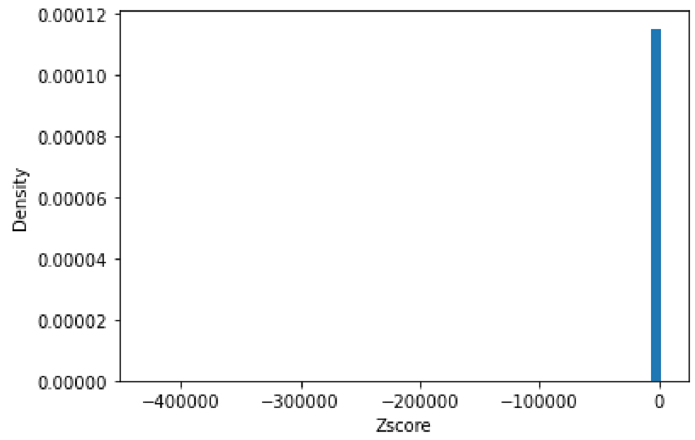


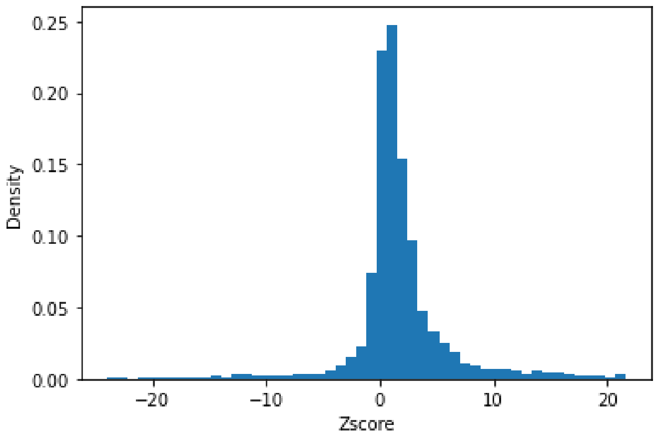




1.4. Winsorizing of the Z-Score

SECTOR 1: INDUSTRIAL





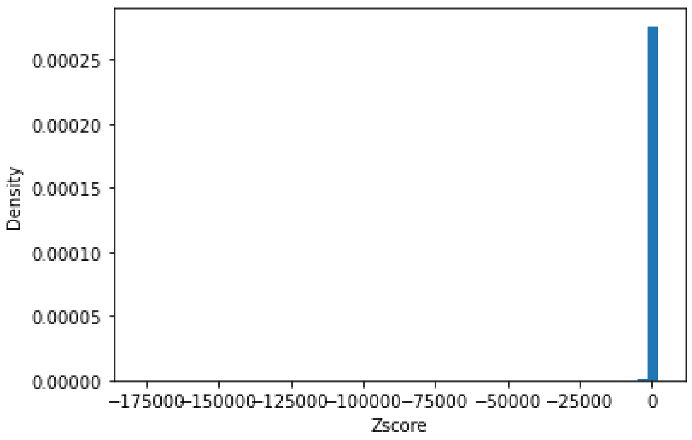
- Values that exceeded the 2.5% or 97.5% percentile were replaced with the percentile
- Makes the distribution more normalized
- And the mean has a value that is in the range of the actual Z-Score

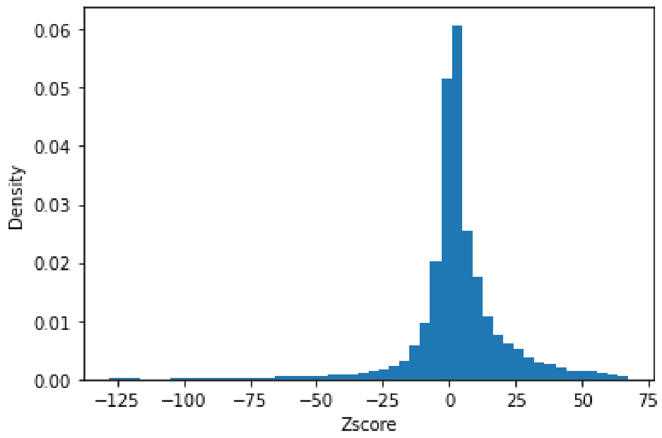
For example for the Industrial sector we get the following Statistics:

Before: After:

	Z_Score (Before)	Z_Score (After)
count	5660.000000	5376.000000
mean	-114.186900	1.699561
std	5821.615094	4.347951
min	-430411.821234	-24.034667
25%	0.228439	0.285224
50%	1.189588	1.189588
75%	2.776037	2.654219
max	2492.176716	21.648497

SECTOR 2: HEALTH

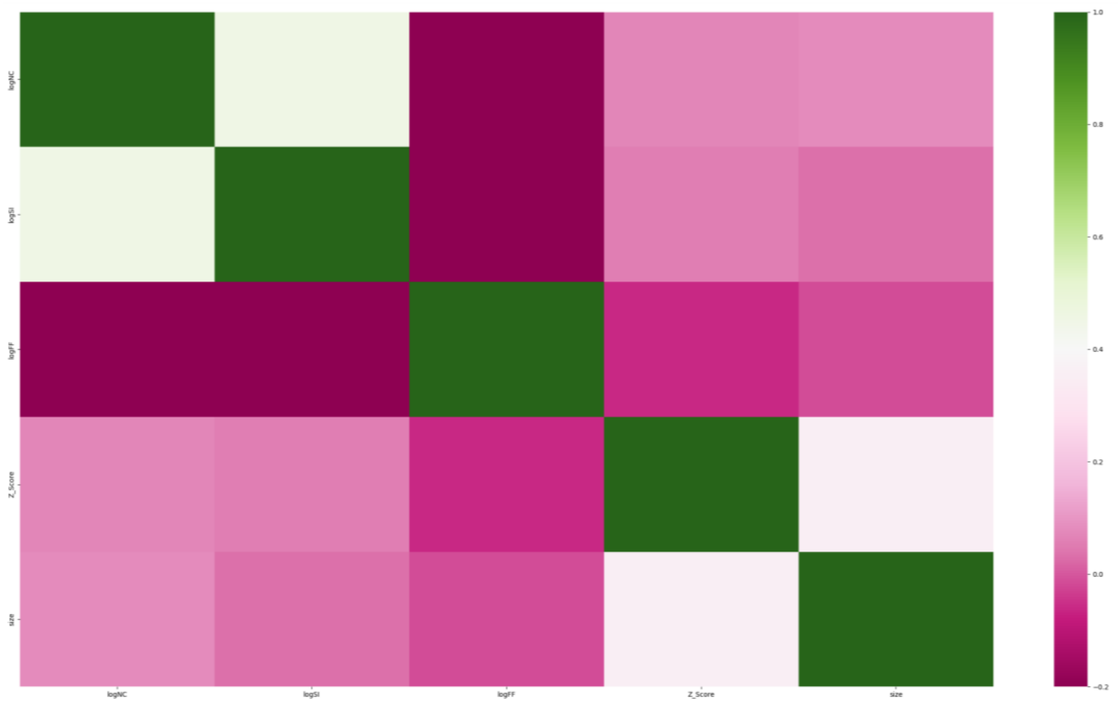




1.5. Correlation Coefficients: Pearson

SECTOR 1: INDUSTRIAL

	logNC	logSI	logFF	Z_Score	size
logNC	1.000000	0.451678	-0.295122	0.069025	0.078571
logSI	0.451678	1.000000	-0.783331	0.056297	0.030001
logFF	-0.295122	-0.783331	1.000000	-0.059580	-0.013437
Z_Score	0.069025	0.056297	-0.059580	1.000000	0.349764
size	0.078571	0.030001	-0.013437	0.349764	1.000000



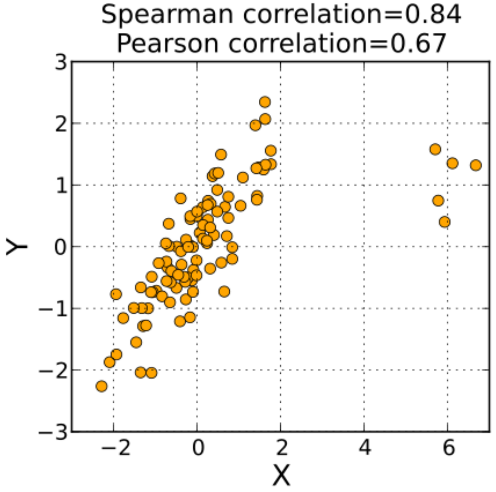
SECTOR 2: HEALTH

	logNC	logSI	logFF	Z_Score	size
logNC	1.000000	0.510591	-0.671029	0.051455	0.043267
logSI	0.510591	1.000000	-0.928709	0.091064	0.096937
logFF	-0.671029	-0.928709	1.000000	-0.098879	-0.098745
Z_Score	0.051455	0.091064	-0.098879	1.000000	0.400855
size	0.043267	0.096937	-0.098745	0.400855	1.000000

1.5. Correlation Coefficients: Spearman's Rank

SECTOR 1: INDUSTRIAL

	logNC	logSI	logFF	Z_Score	size
logNC	1.000000	0.049512	0.043613	0.089391	0.060635
logSI	0.049512	1.000000	-0.789485	0.034115	0.004085
logFF	0.043613	-0.789485	1.000000	-0.071404	-0.026252
Z_Score	0.089391	0.034115	-0.071404	1.000000	0.404897
size	0.060635	0.004085	-0.026252	0.404897	1.000000

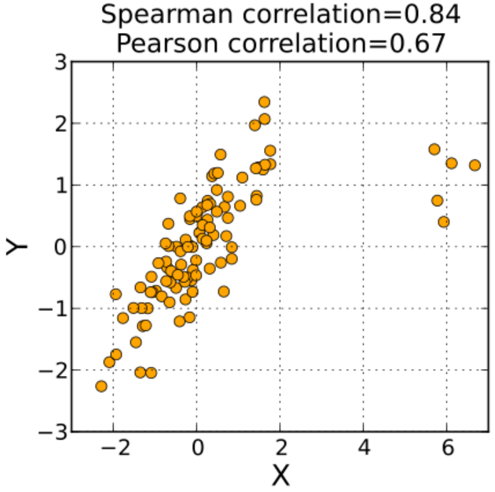


Predicted graphs

$$r_s = \rho_{\mathrm{R}(X),\mathrm{R}(Y)} = \frac{\mathrm{cov}(\mathrm{R}(X),\mathrm{R}(Y))}{\sigma_{\mathrm{R}(X)}\sigma_{\mathrm{R}(Y)}},$$

SECTOR 2: HEALTH

	logNC	logSI	logFF	Z_Score	size
logNC	1.000000	-0.012001	0.084017	0.038203	0.023209
logSI	-0.012001	1.000000	-0.731064	0.096633	0.091844
logFF	0.084017	-0.731064	1.000000	-0.119743	-0.101599
Z_Score	0.038203	0.096633	-0.119743	1.000000	0.492085
size	0.023209	0.091844	-0.101599	0.492085	1.000000



Predicted graphs

2.1.1 Classic assumption test: Jarque Bera

SECTOR 1: INDUSTRIAL

```
The Jarque-Bera test stat of New Cases is 4600.01
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of Stringency-index is 845.67
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of FEDFUNDS is 3933.99
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of Z-score is 13620.47
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of size is 109.41
The p value of Jarque-Bera test is 0.0
```

Reject H_0
Non-normal distribution

SECTOR 2: HEALTH

```
The Jarque-Bera test stat of New Cases is 6789.91
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of Stringency-index is 2882.59
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of FEDFUNDS is 6806.68
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of Z-score is 35725.88
The p value of Jarque-Bera test is 0.0

The Jarque-Bera test stat of size is 156.81
The p value of Jarque-Bera test is 0.0
```

Reject H_0
Non-normal distribution

2.1.2. Classic assumption test: Kolmogorov-Smirnov

SECTOR 1: INDUSTRIAL

```
The KS test stat of New Cases is 1.0
The p value of KS test is 0.0

The KS test stat of Stringency is 1.0
The p value of KS test is 0.0

The KS test stat of FEDFUNDS is 0.5
The p value of KS test is 0.0

The KS test stat of Z_Score is 0.39
The p value of KS test is 0.0

The KS test stat of size is 0.96
The p value of KS test is 0.0
```

Reject H_0
Non-normal distribution

SECTOR 2: HEALTH

```
The KS test stat of New Cases is 1.0
The p value of KS test is 0.0

The KS test stat of Stringency is 1.0
The p value of KS test is 0.0

The KS test stat of FEDFUNDS is 0.5
The p value of KS test is 0.0

The KS test stat of Z_Score is 0.48
The p value of KS test is 0.0

The KS test stat of size is 0.96
The p value of KS test is 0.0
```

Reject H_0
Non-normal distribution

2.1.3 Justification on Parametric Test on Non-Normal

PARAMETRIC TESTS CAN PERFORM WELL WITH SKEWED AND NONNORMAL DISTRIBUTIONS

Parametric tests can perform well with continuous data that are nonnormal if you satisfy the sample size guidelines in the table below. These guidelines are based on simulation studies conducted by statisticians here at Minitab.

Parametric analyses	Sample size guidelines for nonnormal data
1-sample t test	Greater than 20
2-sample t test	Each group should be greater than 15
One-Way ANOVA	• If you have 2-9 groups, each group should be greater than 15. • If you have 10-12 groups, each group should be greater than 20.

https://blog.minitab.com/en/adventures-in-statistics-2/choosing-between-a-nonparametric-test-and-a-parametric-test

2.2.1. One Sample and 2.2.2. Two Sample t-Test

SECTOR 1: INDUSTRIAL

One Sample t-test:

```
The t-test stat of Z_Score against 1.81 is -1.86
The p value of t-test is 0.063
```

Two Sample t-test:

```
The t-test stat of the diff between two groups is -6.63
The p value of t-test is 0.0

The t-test stat of the diff between two groups is -5.88
The p value of t-test is 0.0

The t-test stat of the diff between two groups is 6.41
The p value of t-test is 0.0

The t-test stat of the diff between two groups is -18.09
The p value of t-test is 0.0
```

SECTOR 2: HEALTH

One Sample t-test:

```
The t-test stat of Z_Score against 1.81 is 7.92
The p value of t-test is 0.0
```

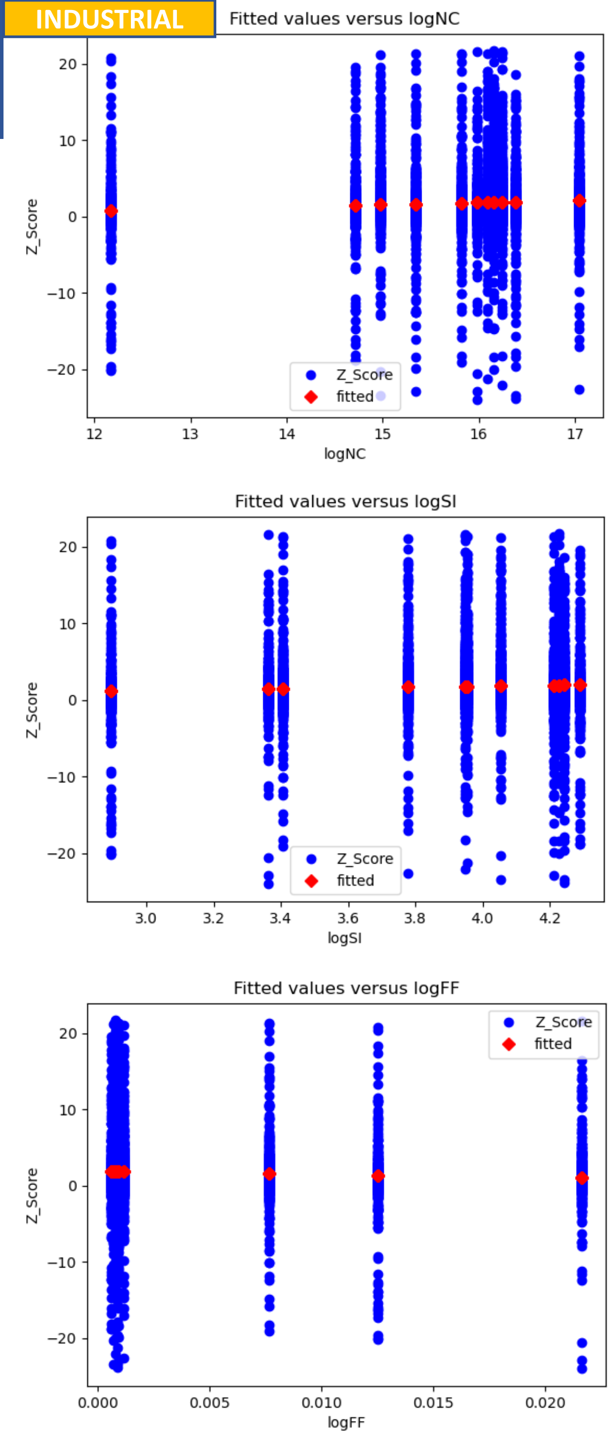
Two Sample t-test:

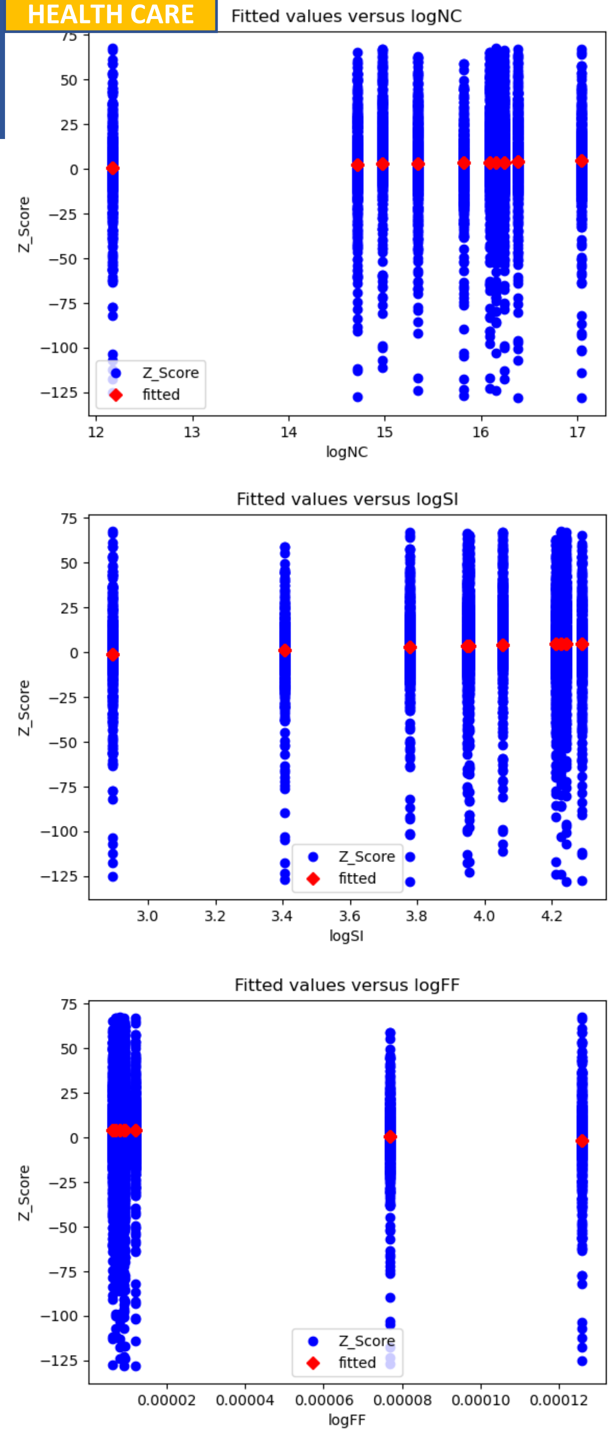
```
The t-test stat of the diff between two groups is -5.33
The p value of t-test is 0.0

The t-test stat of the diff between two groups is -10.24
The p value of t-test is 0.0

The t-test stat of the diff between two groups is 10.99
The p value of t-test is 0.0

The t-test stat of the diff between two groups is -39.61
The p value of t-test is 0.0
```

2.2.3. Simple Regression

```python
NC_zscore_reg = smf.ols(formula='Z_Score ~ 1 + logNC', data = Total_df).fit()
print(NC_zscore_reg.summary())

SI_zscore_reg = smf.ols(formula='Z_Score ~ 1 + logSI', data = Total_df).fit()
print(SI_zscore_reg.summary())

FF_zscore_reg = smf.ols(formula='Z_Score ~ 1 + logFF', data = Total_df).fit()
print(FF_zscore_reg.summary())

Size_zscore_reg = smf.ols(formula='Z_Score ~ 1 + size', data = Total_df).fit()
print(Size_zscore_reg.summary())

# visualizing simple linear regression
# binscatter plot with fitted line



#model = sm.OLS(Y, X, missing='drop')
#model_result = model.fit()
sm.graphics.plot_fit(NC_zscore_reg,1, vlines=False)
sm.graphics.plot_fit(SI_zscore_reg,1, vlines=False)
sm.graphics.plot_fit(FF_zscore_reg,1, vlines=False)
```

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                Z_Score   R-squared:                       0.008
Model:                            OLS   Adj. R-squared:                  0.008
Method:                 Least Squares   F-statistic:                     31.70
Date:                Tue, 29 Nov 2022   Prob (F-statistic):           1.92e-08
Time:                        21:07:15   Log-Likelihood:                -11654.
No. Observations:                3865   AIC:                         2.331e+04
Df Residuals:                    3863   BIC:                         2.332e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -3.8067      0.981     -3.879      0.000      -5.731      -1.882
logNC          0.3591      0.064      5.631      0.000       0.234       0.484
==============================================================================
Omnibus:                      888.944   Durbin-Watson:                   1.949
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            16610.218
Skew:                          -0.604   Prob(JB):                         0.00
Kurtosis:                      13.084   Cond. No.                         191.
==============================================================================
```

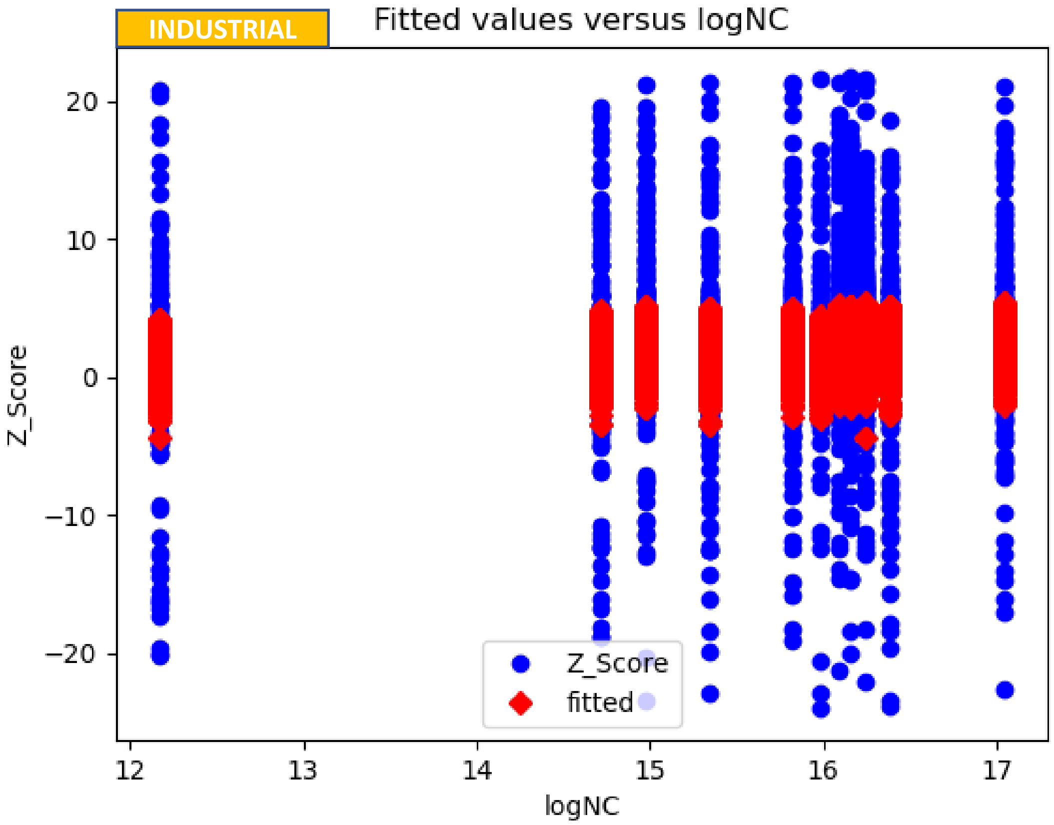


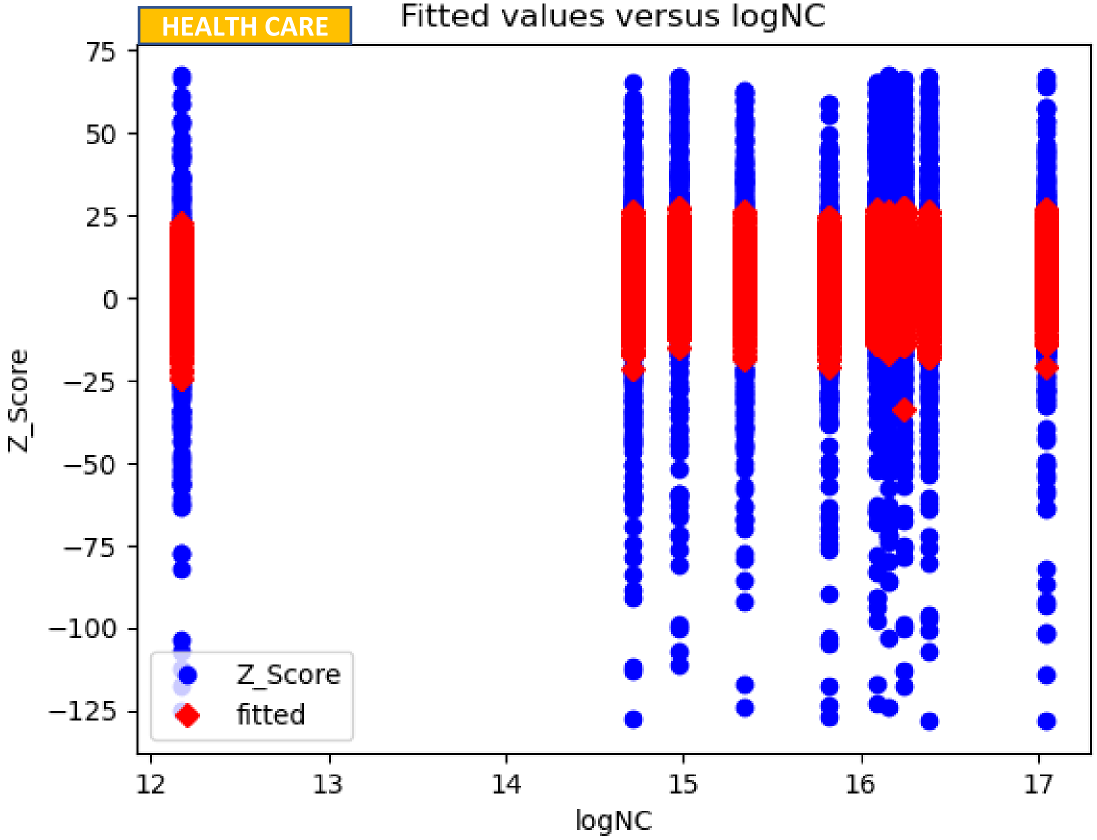


2.2.4. Multiple Linear Regression

```python
Covid_zscore_reg = smf.ols(formula='Z_Score ~ 1 + logNC + logSI + logFF + size', data = Total_df).fit()
print(Covid_zscore_reg.summary())

sm.graphics.plot_fit(Covid_zscore_reg,1, vlines=False)
```

```
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
                            OLS Regression Results
==============================================================================
Dep. Variable:                Z_Score   R-squared:                       0.130
Model:                            OLS   Adj. R-squared:                  0.129
Method:                 Least Squares   F-statistic:                     143.8
Date:                Tue, 29 Nov 2022   Prob (F-statistic):          9.57e-115
Time:                        21:07:15   Log-Likelihood:                -11401.
No. Observations:                3865   AIC:                         2.281e+04
Df Residuals:                    3860   BIC:                         2.284e+04
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      0.2776      3.686      0.075      0.940      -6.949       7.504
logNC          0.1332      0.127      1.049      0.294      -0.116       0.382
logSI         -1.3325      0.616     -2.164      0.030      -2.539      -0.125
logFF      -1.723e+08   8.51e+07     -2.024      0.043   -3.39e+08   -5.37e+06
size           0.7377      0.032     22.943      0.000       0.675       0.801
==============================================================================
Omnibus:                      650.696   Durbin-Watson:                   1.949
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            13555.067
Skew:                           0.033   Prob(JB):                         0.00
Kurtosis:                      12.174   Cond. No.                     1.99e+10
==============================================================================
```

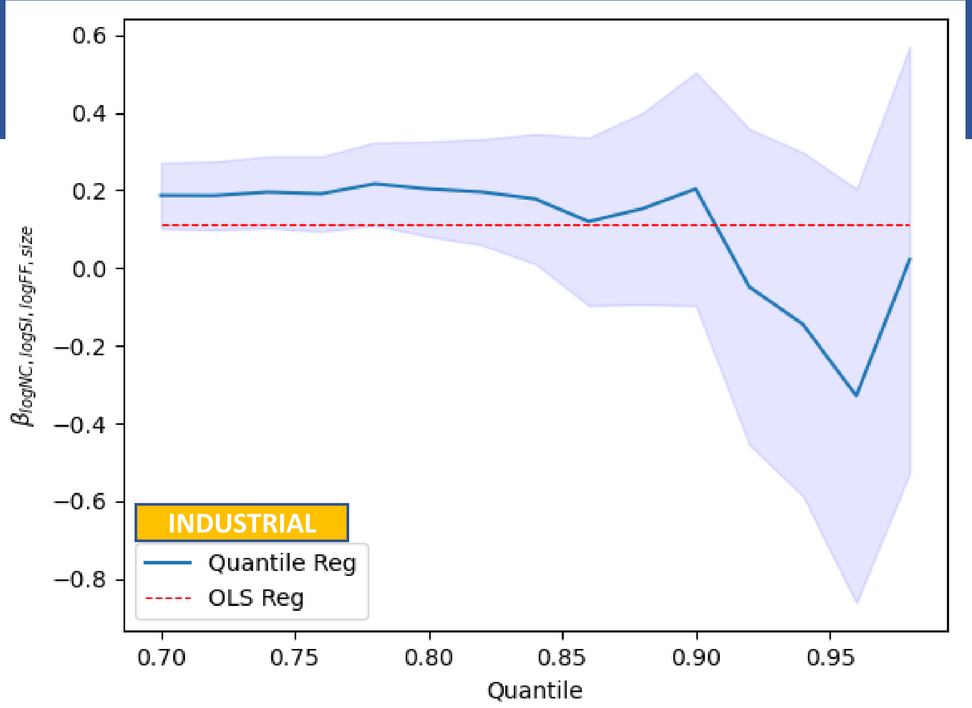


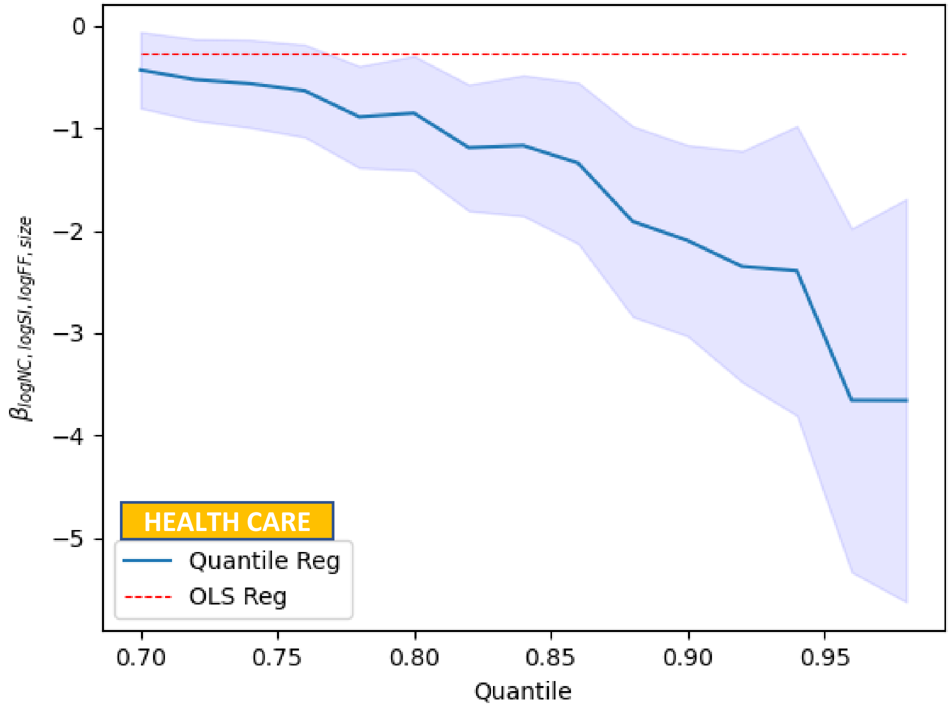


2.2.5. Quantile Regression

```python
# the 60% unconditional quantiles of Z_score
Z_Score_p60=Total_df['Z_Score'].quantile(0.6)
print("The 60% (and below) percentile of Z_SCore is 1.61.")

# quantile regression for 80% quantile
e_qreg = smf.quantreg(formula='Z_Score ~ 1 + logNC + logSI + logFF + size', data = Total_df).fit(q=0.8)
print(e_qreg.summary())

# plot the quantile regression coeffcient and confidence intervals
q_start=0.7
q_end=0.98
q_inc=0.02
n=(q_end-q_start)/q_inc+1
n=int(n)
q_forplot=np.linspace(q_start,q_end,n)
beta_forplot=np.zeros(n)
se_forplot=np.zeros(n)
for i in range(n):
    e_qreg = smf.quantreg(formula='Z_Score ~ 1 + logNC + logSI + logFF + size', data = Total_df).fit(q=q_forplot[i])
    beta=e_qreg.params
    beta_forplot[i]=beta[1]
    se=e_qreg.bse
    se_forplot[i]=se[1]

e_reg_control = smf.ols(formula='Z_Score ~ 1 + logNC + logSI + logFF + size', data = Total_df).fit()
beta=e_reg_control.params
beta_ols=beta[1]

fig = plt.figure()
ax = fig.add_subplot(111)
ax.plot(q_forplot,beta_forplot,label='Quantile Reg')
ax.fill_between(q_forplot,beta_forplot-1.96*se_forplot,
                beta_forplot+1.96*se_forplot,color='b',
                alpha=0.1)
ax.plot(q_forplot,beta_ols*np.ones(n),color='r',
        linestyle='dashed',linewidth=0.8,
        label='OLS Reg')
ax.legend(loc='lower left')
ax.set_xlabel('Quantile')
ax.set_ylabel(r'$\beta_{logNC,logSI, logFF, size}$')
plt.show()
```

```
Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The smallest eigenvalue is 2.95e-15. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
The 60% (and below) percentile of Z_SCore is 1.61.
                         QuantReg Regression Results
==============================================================================
Dep. Variable:                Z_Score   Pseudo R-squared:              0.02713
Model:                       QuantReg   Bandwidth:                      0.4887
Method:                 Least Squares   Sparsity:                        13.23
Date:                Tue, 29 Nov 2022   No. Observations:                 3865
Time:                        21:07:15   Df Residuals:                     3860
                                        Df Model:                            4
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -4.8938      1.049     -4.666      0.000      -6.950      -2.837
logNC          0.5436      0.112      4.853      0.000       0.324       0.763
logSI         -0.7514      0.344     -2.181      0.029      -1.427      -0.076
logFF       -1.96e-07   4.41e-08     -4.446      0.000   -2.82e-07    -1.1e-07
size           0.3984      0.036     11.163      0.000       0.328       0.468
==============================================================================

The smallest eigenvalue is 6.48e-15. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```





Discussion: Hypotheses Test, Results and Interpretation

No.	Hpotheses	Test	Sector 1: Industrial	Sector 2: Health
H1	Z is significantly different from 1.81	One-sample t-test	Do no reject H_0	Reject H_0
H2	$SI_{distressed}$ is significantly different from $SI_{non\text{-}distressed}$	Two-sample t-test	Reject H_0	Reject H_0
H3	$FF_{distressed}$ is significantly different from $FF_{non\text{-}distressed}$	Two-sample t-test	Reject H_0	Reject H_0
H4	$Size_{distressed}$ is significantly different from $Size_{non\text{-}distressed}$	Two-sample t-test	Reject H_0	Reject H_0
H5	There is a significant relationship between Z and logNC	SRA (p-value)	Reject H_0	Reject H_0
H6	There is a significant relationship between Z and logSI	SRA (p-value)	Reject H_0	Reject H_0
H7	There is a significant relationship between Z and logFF	SRA (p-value)	Reject H_0	Reject H_0
H8	There is a significant relationship between Z and size	SRA (p-value)	Reject H_0	Reject H_0
H9	At least one of the independent variables is related to Z	MRA (p of F)	Reject H_0	Reject H_0
H10	At least one of the independent variables is related to Z	QRA (p-value)	Reject H_0	Reject H_0

Conclusion

Observations based on our analysis : -

- Industrial sector performed better during the Covid period. During pandemic FED kept on reducing the interest rate thereby, the debt burden on industrial sector diminished.

- Health sector performed even better than industrial sector during the same tenure. With the reduced debt burden and in anticipation of vaccine development the health sector performed as per the expectation.

- Industrial sector and Health sector both are in positive co-relation with increase in number of covid cases. Both these sector were inversely propotional to the FED rate.

Limitation

- This analysis measured the impact of COVID-19 using the growth rate of all confirmed cases and the stringency index. However, other measures, such as the net growth rate (i.e., numbers subtracting recovered cases from confirmed cases), may capture net declines in active confirmed cases which may help avoid unnecessarily inflating the negative impact of global pandemic.

- It should be also noted that there is much uncertainty in calculating recovered cases due to questionable data source reporting recovered cases in many countries, including U.S.

List of References

- Cho, B., & Saki, Z. (2021). Firm performance under the COVID-19 pandemic: The case of the U.S. textile and apparel industry. *The Journal of the Textile Institute*, 1–11. https://doi.org/10.1080/00405000.2021.1943258
- *COVID-19 Data Explorer*. (n.d.). Our World in Data. Retrieved November 29, 2022, from https://ourworldindata.org/explorers/coronavirus-data-explorer?uniformYAxis=0&hideControls=true&Metric=Stringency+index&Interval=7-day+rolling+average&Relative+to+Population=true&Color+by+test+positivity=false&country=~USA
- Editor, M. B. (2015, February 19). *Choosing Between a Nonparametric Test and a Parametric Test*. Blog.minitab.com. https://blog.minitab.com/en/adventures-in-statistics-2/choosing-between-a-nonparametric-test-and-a-parametric-test
- *Effective Federal Funds Rate*. (2019). Stlouisfed.org. https://fred.stlouisfed.org/series/FEDFUNDS#0
- *GLOBAL INDUSTRY CLASSIFICATION STANDARD (GICS®) METHODOLOGY Guiding Principles and Methodology for GICS*. (2022). https://www.msci.com/documents/1296102/11185224/GICS+Methodology+2022.pdf/f9910041-6127-17d2-1246-4052926adaf7?t=1645738126436
- *Wharton WRDS Login*. (n.d.). Wrds-Www.wharton.upenn.edu. Retrieved November 29, 2022, from https://wrds-www.wharton.upenn.edu/pages/get-data/compustat-capital-iq-standard-poors/compustat/north-america-daily/fundamentals-quarterly/?saved_query=2316916
- Xu, J., & Jin, Z. (2022). Exploring the Impact of the COVID-19 Pandemic on Firms' Financial Performance and Cash Holding: New Evidence from China's Agri-food Sector. *Agronomy*, *12*(8), 1951. https://doi.org/10.3390/agronomy12081951