Effect of COVID-19 on Economy

Badrinath Reddy Gujjula
New York University Tandon School
of Engineering
Brooklyn, New York
bg2310@nyu.edu

Lokesh Munagala
New York University Tandon School
of Engineering
Brooklyn, New York
lm4508@nyu.edu

Susmita Karyakarte
New York University Tandon School
of Engineering
Brooklyn, New York
sk8663@nyu.edu

ABSTRACT

The coronavirus pandemic has left national economies and businesses counting the costs with governments struggling to come up with measures and policies to tackle the spread of the virus. In this project we try to unfold the effects caused by the pandemic mainly on the US economy and try to explain a few trends and patterns observed on certain economic factors alongside the patterns observed in countries like India, China, Italy and Spain on the same factors. We try to explicate these trends with real life events such as strategies and policies that the corresponding governments came up with at the time and how these events affected society and in turn helped to shape the pandemic.

KEYWORDS

datasets, COVID-19, visualizations, CPS data, analysis

1 GROUP NAME:

Lazy Coders

2 ESSENTIAL DATA AND IT'S ACQUISITION

As we navigate through this pandemic and its long-lasting impact it is necessary to analyze trends and patterns observed throughout the world in the past two years in order to understand how the society will thrive post pandemic and what could be done in order to reduce the repercussions of the failures in handling the pandemic. In order to show how the pandemic and its impact grew with time we decided to choose five countries which were the most affected during the first half of the year 2020, with majority of the cases out of the world population being from these countries.

As we started searching for the data required for the project we came across many datasets all providing different dimensions and volumes of data for **COVID-19** which we got from**[5]** for all the countries in the world. Hence we had to reduce the dimensions as we only wanted to show how the number of cases increased with time for the selected five countries and how that increase resulted in economic losses to these nations.

In order to survey the economic impact, we needed data of various economic indicators such as **Share Market** indexes of these five countries, **Gold and Crude Oil** prices for the years 2020 and 2021 which we got from **Yahoo Finance [4]** and **International trade data from WTO [1]**that includes monthly imports and exports along with data for certain obviously affected sectors like international travel, restaurants and small businesses. We also required quarterly **Gross Domestic Product (GDP) from [3]** for various sectors of the US to examine how these individual sectors were affected by the pandemic. As USA was severely hit by the pandemic in the year 2020 and continues to be a severely affected

country we decided to analyze the effects mainly for the US economy and try to forecast what trends we can expect to be observed in the future with the help of the trends that were observed in countries like India, Italy, China and Spain. We also used **Flight Data from [2]** to show how Pandemic effected Air Travel

The data set which proved to be of crucial importance and helped us to dive deeper into the analysis was the **Current Population Survey Dataset[CPS]**maintained and collected by the **United States Census Bureau [6]** . There are very few national data sets that provide information on the general population with a lot of additional and accurate information on demographic characteristics of the population. Surveys are conducted monthly by the US Bureau of the Census and the US BLS, are representative of the entire US population and contain observations for more than 130,000 people. Using microdata from the basic monthly files of the CPSs, we were able to gather information about population that owns small businesses, restaurants, other local businesses such as shops, bookstores that were hit hard economically due to lockdowns during the pandemic.

3 DATA CLEANING AND PREPROCESSING

Data gathered from public sources is often riddled with data quality issues such as missing values, missing metadata, duplicate data, inconsistent and invalid entries and so on. Detecting and repairing dirty data is one of the perennial challenges in data analysis, and failure to do so can result in inaccurate analytics and unreliable decisions. As data is gathered from different sources, when multiple data sources need to be integrated, the need for data cleaning increases.

After **dimensionality reduction** of the COVID-19 dataset, next we moved to cleaning of the share market data. Since we are considering the data of the following five countries – USA, India, Italy, China and Spain we had obtained data for the NASDAQ, NIFTY, NSEBANK, FTSE MIB, Hang Seng, Spain IBEX 35 which are few of the major composite indexes for these countries and are representative of the stock market. We performed dimensionality reduction for these datasets along with the datasets we obtained for crude oil and gold prices and trade data as well for these countries as we only wanted to observe and analyze this data side by side with the COVID-19 data along the year 2020.

Acquiring data for the gross domestic product (GDP) for various sectors of the US was a bit challenging as the data acquired was in the form of statistics more than straightforward data entries and required encoding and some manual extraction and its conversion to a csv file. Lastly, we moved to the cleaning of the **CPS** data that consisted of 12 csv files for 12 months of the year 2020 and had more than 388 columns and 100,000 records per file with column headers and row data for few of those columns encoded in the form

of predefined codes that were stored separately in a data dictionary provided by the United States Census Bureau for our reference. **We had to go through the entire dictionary to understand what each of those 388 columns represent**. It was the most challenging part of the project.

For all the 12 files we performed dimensionality reduction and removed records with missing values and checked for invalid entries and inconsistent data after the number of columns were reduced to make sure the data for the analysis was accurate. The data dictionary also had predefined invalid entries that made the process less time consuming however navigating through the dictionary itself and understanding the particular encodings for each value took up most of the time due to the size of the data. The datasets we used for this project consisted of more than 25 csv files amounting to around 1.8 GB of data.

4 PROJECT TASKS

We divided the datasets and the visualizations we had to come up for further analysis into individual tasks and made observations from the respective plots.

4.1 Task 1

In Task 1, we plotted the total cases of COVID-19 since February 2020 to March 2021 across all the 5 countries using a line plot.
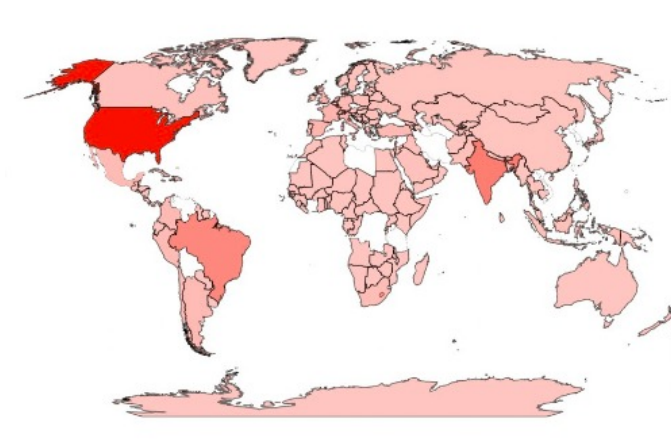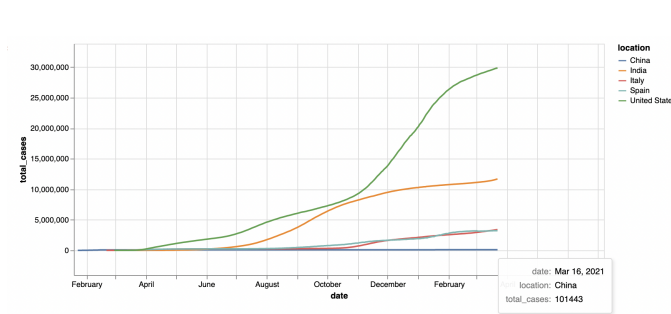

Figure 1: The Total Cases of COVID-19

Figure 2: The Total Cases of COVID-19

We observed a sharp rise in the total cases in the US, followed by India with the US reaching the 30 million mark in total cases in March 2021. Hence the fact that US had the most number of COVID-19 cases until date was proven with our collected data. Initially Italy followed by Spain were the most affected countries and the outbreak hadn't reached India until that time hence we can see, for the months of April, May and June the number of cases in India are less compared to Italy and Spain. After the first and second wave of the outbreak in Italy and Spain, the two countries managed to contain the outbreak up to some extent in comparison to the US.

Even if the actual figures reported for China are considered valid, the empirical integrity of China's numbers has been repeatedly questioned. According to a few news outlets the estimates for confirmed cases for China would be around 232,000 just until April 2020.There was growing concern that China is not being entirely honest about the extent of its infections and deaths and can be seen from the given data. According to our dataset the total cases for China had reached only the 101443 mark by the month of March 2021 which explains the straight line shown by the plot for China.
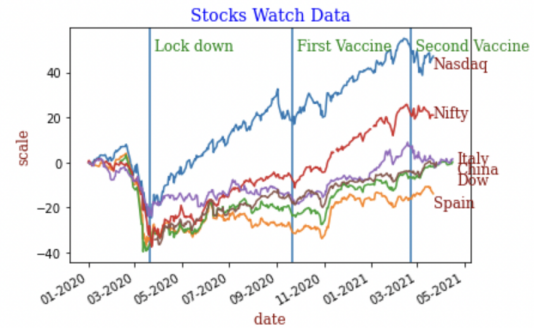
4.2 Task 2

Figure 3: Stock Market Data

For Task 2, we plotted the stock market data of all the indexes of all the five countries we chose. As we expected, initially with the rise in the number of COVID-19 cases and governments implementing lock-downs in all the countries there was a sharp decline in the indexes that can be seen in the plot.

It was reported that global stocks surged to record highs after the drug companies Pfizer and BioNTech announced that their coronavirus vaccine was 90% effective in trials, giving hope that economies around the world can soon return to normal. The Dow Jones industrial average and the tech-focused Nasdaq Composite index hit new records before dropping back. The tech heavy Nasdaq ended the day down 1.5% as investors sold Netflix, Peloton and Zoom and other companies that have benefitted from a pandemic boost. The Dow ended the day up 2.95%.Many of the most pronounced stock market gainers were the companies most affected by restrictions on movement brought in to fight the pandemic, while companies that have most benefited from lockdowns retreated from some of their spectacular gains.These patterns are also observed in

our plot where we have shown the temporary rise in market when the first and second vaccines were released.

4.3 Task 3

We started the plotting of Oil prices that certainly took a huge hit during the pandemic and can also be seen with the plot. In the months from March to June when most countries were in a lockdown, travel and transportation industry reduced it's operations upto a great extent and hence there was a less need for oil that made the oil prices go down. As countries started lifting lockdowns in phases we can see the graph slowly moving upwards with time and it came back to normal after year in the month of February 2021.
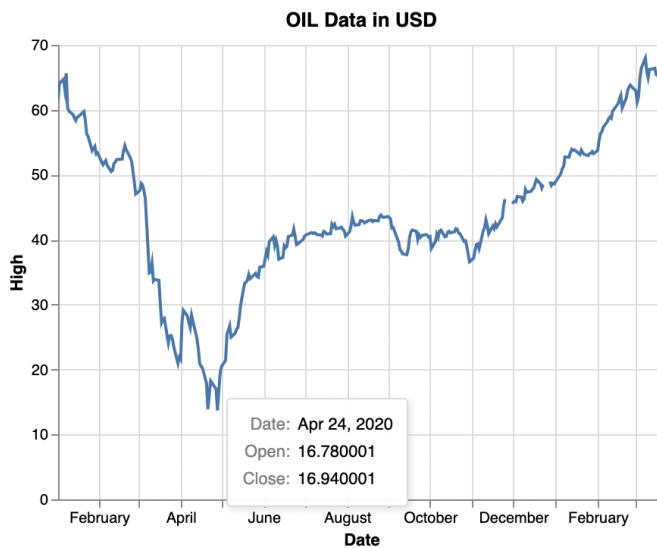
GOLD Data in USD

Figure 5: Gold Prices Data

OIL Data in USD

Figure 4: Oil Prices Data

4.5 Task 5

We noticed the travel and tourism industry was one of the most affected industries due to the pandemic as most countries had banned travel to other countries and tourist activities leading to a decrease in the number of flights that were available for the general public.Hence we decided to collect and plot the data for the number of flights for just the US as the same pattern was going to be observed for data that was present for the other countries.On plotting we observed a sharp decline in the total number of flights that were operated during the initial months as expected, with the plot climbing upwards with time. An important observation here to note was that the number of flights plotted for the month of January 2020 is more compared to February 2021, that is even after a year of the declaration of the pandemic airports and flight services are still not operating at their full capacity and that may happen only when things start going back to normal until then the flight and air industry will be facing a substantial loss economically compared to the profits it would have brought in to the US Economy in normal times.

4.4 Task 4

Next, we went to plotting the gold prices which are seen to follow a up and down pattern in the prices. We can clearly see a decrease in the prices for the months of March and April and a steady graph from there onwards until a hike in August which was when lockdowns had been lifted by most countries. Lastly the plot ends with an increase in the gold prices in comparison to the prices in February 2020.
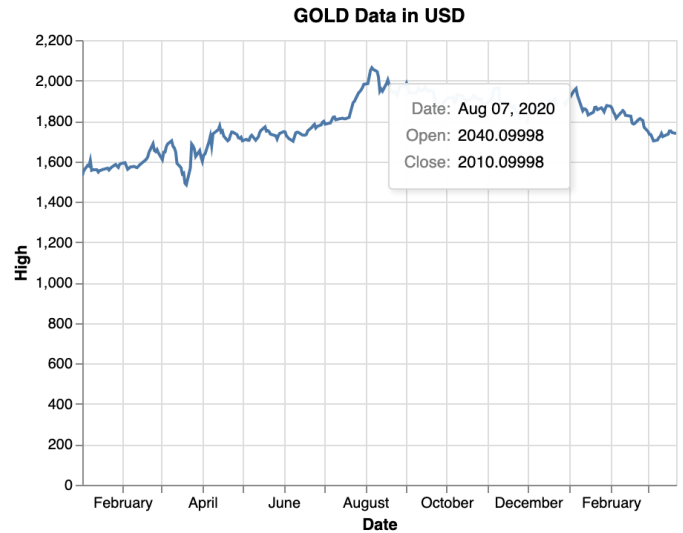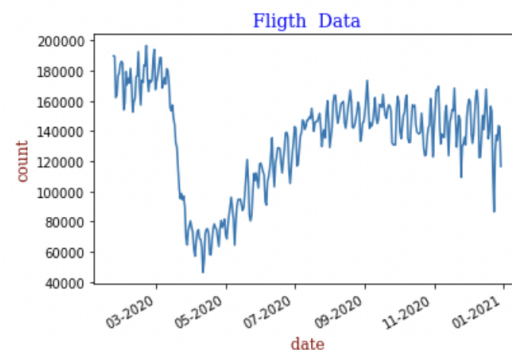
Fligth Data

Figure 6: Flights Data

4.6 Task 6

For Task 6 we started plotting the trading data for all of the five countries.The fields of the data we took into consideration were total merchandise imports and exports of these countries with the world.
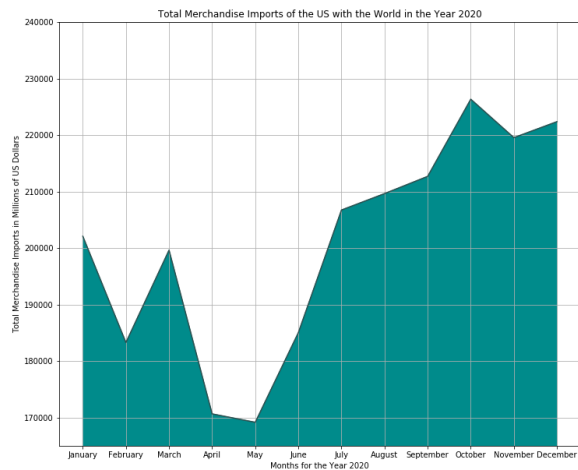

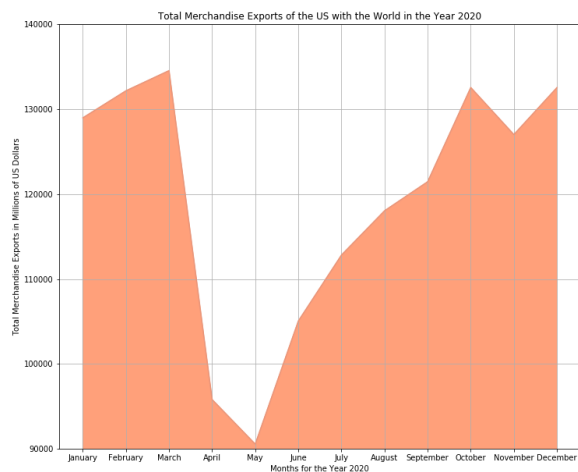**Figure 7: Total Merchandise Imports Data for the US**

Figure 8: Total Merchandise Exports Data for the US

5 ANALYSIS

We started by examining how many people in the CPS dataset had work for pay or profit for both types of businesses; incorporated and unincorporated individually and together for every month

of the year 2020. Then we realized that the count of these people was distributed over various occupational sectors within the dataset and we sorted these counts for each mentioned sector in the data dictionary for the CPS data. We compare the number of unemployed people for each month by computing the percentage of active business owners from the employment characteristics given in the dataset for each individual, sector-wise. Based on this computation we observed a dramatic drop in the number of active business owners in April 2020 and the partial rebound in May and a continuing rebound in June.

The number of working business owners dropped from 15.0 million in February 2020 to 11.7 million in April 2020 because of COVID-19. March 2020 only shows a small drop in business owners likely because of the limited effect from shelter-in-place restrictions. May 2020 shows a partial rebound from April 2020 adding back 1.1 million active business owners (7 percentage points relative to February levels). The losses due to COVID-19 from February remain high at 15%, but the rebound suggests that not all of the losses of active business owners in April 2020 were permanent closures. June experienced a further rebound with business activity being down 8% from February levels.

As mentioned in a study by Robert Fairlie published in the J Econ Manag Strategy a loss of 3.3 million businesses (or 22%) was observed from February to April 2020 and was the largest drop on record and hence explains the observed pattern in our analysis.

The CPS data also provides detailed information on gender, race, and immigrant status.

5.1 GDP:

Gross Domestic Product in Q1,Q2,Q3,Q4 for 2019 is 21115309, 21329877, 21540325, 21747394 Millions of Dollars. And for 2020 in Q1,Q2,Q3,Q4 is 21561139, 19520114, 21170252, 21494731 Millions of dollars. There has been a huge drop in Q2 of 2020 when compared to previous year quarter because of the pandemic. In the **Figure 9** we can see the contribution towards GDP from various sector and we can see all the sectors which have huge impact of the pandemic have decreased in their contribution towards GDP even when Total GDP has decreased showing some sectors got much more effected than the others. For example from **Figure 10** we can see in Q2 the contribution of Finance and Real Estate increased from 0.2121% to 0.2338% where as Transportation decreased from 0.0324% to 0.0266%. Showing Transportation got much more worstly effected than the Finance and Real Estate. Just looking at the contributions we can not get the full picture of how all the sectors effected in terms of business owners and Unemployment so we use CPS data to dig deeper.
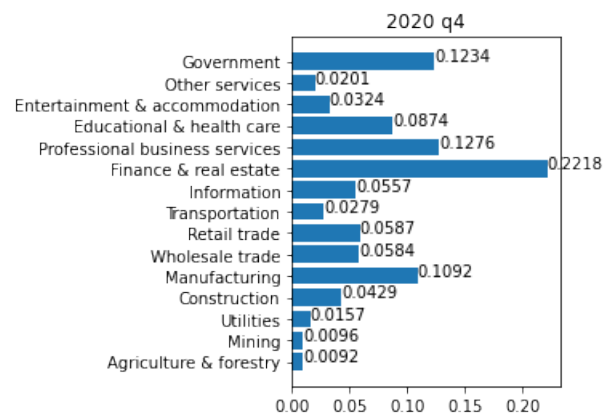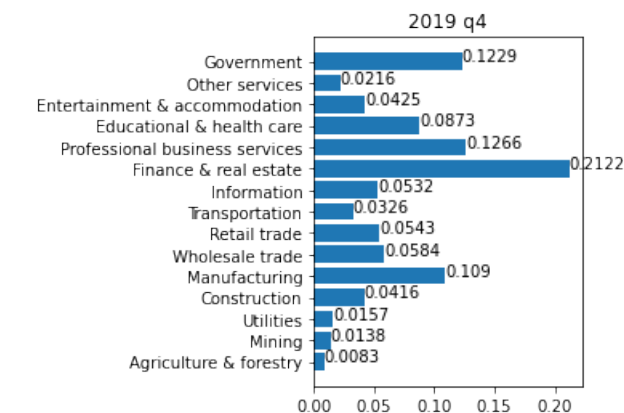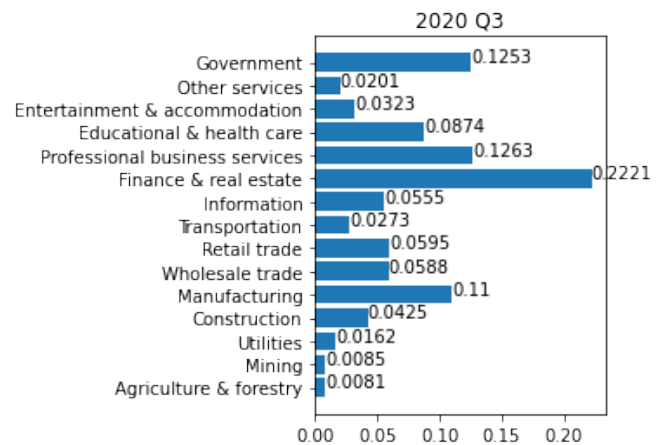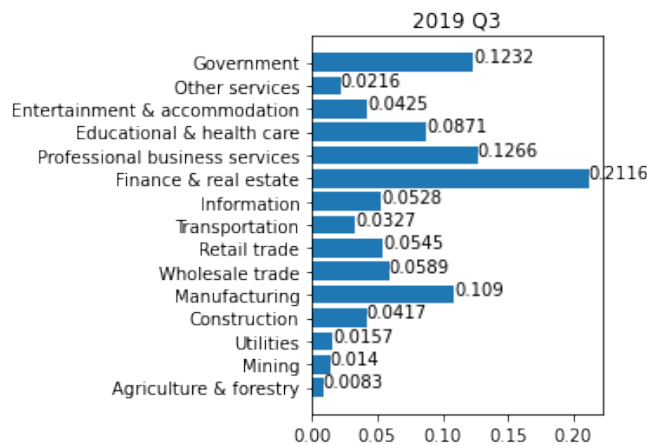
2019 Q1

Sector	Value
Government	0.1234
Other services	0.0214
Entertainment & accommodation	0.0421
Educational & health care	0.0869
Professional business services	0.1264
Finance & real estate	0.2117
Information	0.0518
Transportation	0.0323
Retail trade	0.054
Wholesale trade	0.0595
Manufacturing	0.1101
Construction	0.0417
Utilities	0.0156
Mining	0.015
Agriculture & forestry	0.0081

2020 Q1

Sector	Value
Government	0.1244
Other services	0.0213
Entertainment & accommodation	0.04
Educational & health care	0.0872
Professional business services	0.1276
Finance & real estate	0.2137
Information	0.0537
Transportation	0.0318
Retail trade	0.0548
Wholesale trade	0.0584
Manufacturing	0.1088
Construction	0.0424
Utilities	0.0155
Mining	0.0115
Agriculture & forestry	0.0088

2019 Q2

Sector	Value
Government	0.1231
Other services	0.0215
Entertainment & accommodation	0.0422
Educational & health care	0.0869
Professional business services	0.1263
Finance & real estate	0.2121
Information	0.0526
Transportation	0.0324
Retail trade	0.0542
Wholesale trade	0.0588
Manufacturing	0.1097
Construction	0.0416
Utilities	0.0156
Mining	0.015
Agriculture & forestry	0.0081

2020 Q2

Sector	Value
Government	0.1329
Other services	0.019
Entertainment & accommodation	0.0244
Educational & health care	0.083
Professional business services	0.1294
Finance & real estate	0.2338
Information	0.0572
Transportation	0.0266
Retail trade	0.0565
Wholesale trade	0.057
Manufacturing	0.1052
Construction	0.0437
Utilities	0.0171
Mining	0.007
Agriculture & forestry	0.0073

2019 Q3

Sector	Value
Government	0.1232
Other services	0.0216
Entertainment & accommodation	0.0425
Educational & health care	0.0871
Professional business services	0.1266
Finance & real estate	0.2116
Information	0.0528
Transportation	0.0327
Retail trade	0.0545
Wholesale trade	0.0589
Manufacturing	0.109
Construction	0.0417
Utilities	0.0157
Mining	0.014
Agriculture & forestry	0.0083

2020 Q3

Sector	Value
Government	0.1253
Other services	0.0201
Entertainment & accommodation	0.0323
Educational & health care	0.0874
Professional business services	0.1263
Finance & real estate	0.2221
Information	0.0555
Transportation	0.0273
Retail trade	0.0595
Wholesale trade	0.0588
Manufacturing	0.11
Construction	0.0425
Utilities	0.0162
Mining	0.0085
Agriculture & forestry	0.0081

2019 q4

Sector	Value
Government	0.1229
Other services	0.0216
Entertainment & accommodation	0.0425
Educational & health care	0.0873
Professional business services	0.1266
Finance & real estate	0.2122
Information	0.0532
Transportation	0.0326
Retail trade	0.0543
Wholesale trade	0.0584
Manufacturing	0.109
Construction	0.0416
Utilities	0.0157
Mining	0.0138
Agriculture & forestry	0.0083

2020 q4

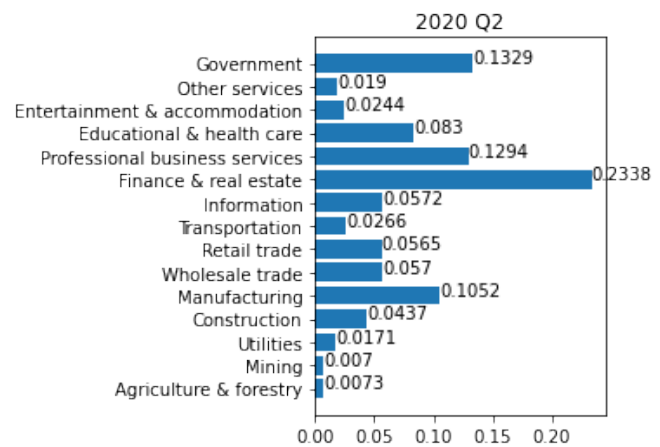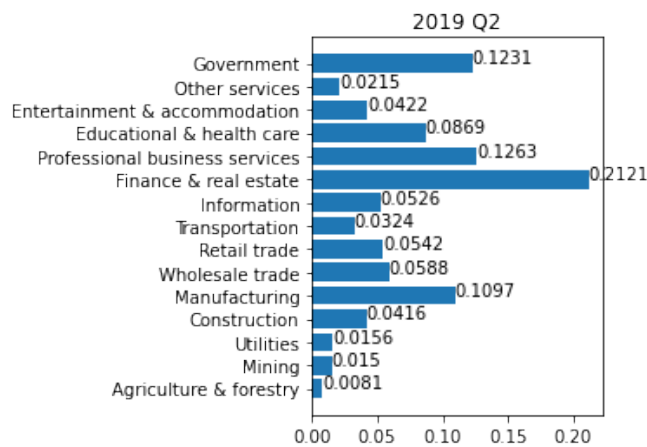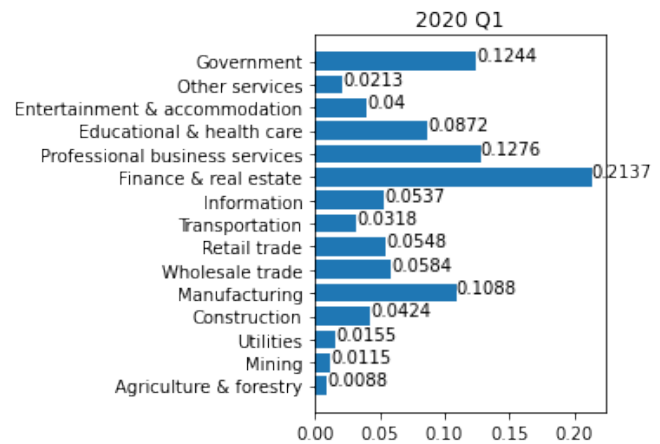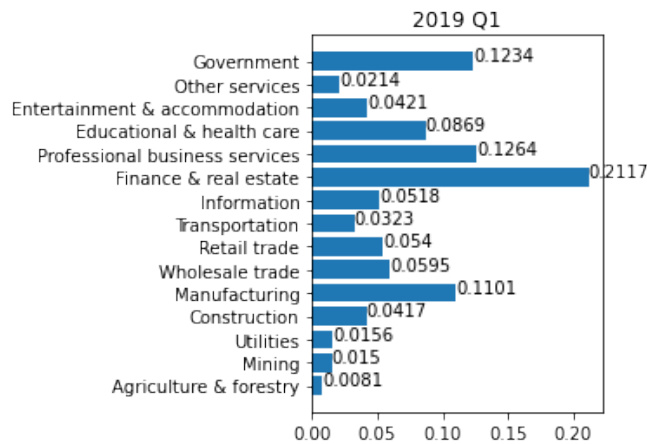| Sector	Value
Government	0.1234
Other services	0.0201
Entertainment & accommodation	0.0324
Educational & health care	0.0874
Professional business services	0.1276
Finance & real estate	0.2218
Information	0.0557
Transportation	0.0279
Retail trade	0.0587
Wholesale trade	0.0584
Manufacturing	0.1092
Construction	0.0429
Utilities	0.0157
Mining	0.0096
Agriculture & forestry	0.0092

Figure 9: GDP of 2019 in Millions of dollars   **Figure 10: GDP of 2020 in Millions of dollars**

5.2 Unemployment:

The unemployment rate peaked at an unprecedented level touching **all time high of 14.8%** in April 2020 as we can see from **Figure 11** but it does not paint the full picture as we don't know whether the jobs lost are permanent or temporary and which sectors lost more number of jobs. So in the next section we see the duration of jobs lost to get better understanding of effect of pandemic on employmnet.
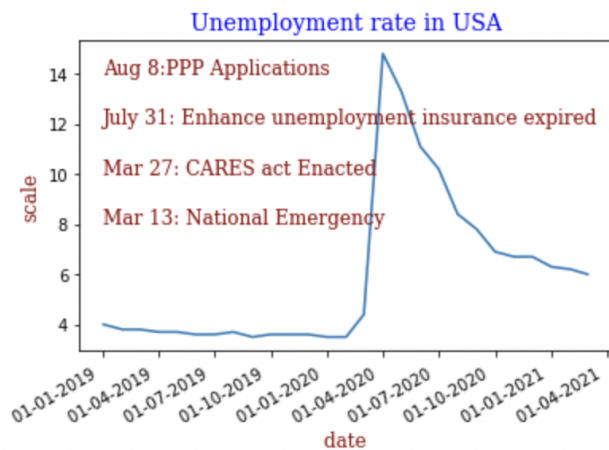

Figure 11: Unemployment Rate

5.3 Unemployment Time:

When the public health crisis began, many workers were laid off on temporary furloughs, but since then many of those temporary job losses have become permanent. Assuming jobs return eventually after the pandemic subsides, this increase in permanent layoffs would be considered an increase in cyclical unemployment—unemployment that is a result of the business cycle. However, if the pandemic results in permanent changes to and job losses in some industries, the level of structural unemployment relatively long-lasting unemployment as a result of shifts in the economy could increase. **Figure 12** explains this phenomenon. In April 2020, due to the sudden closure of many businesses, the percentage of individuals unemployed for less than five weeks increased to more than 50%. Since the shock, the duration of unemployment has been increasing, with those unemployed for more than 14 weeks accounting for over half of all unemployed individuals in October 2020.By December, the percentage of unemployed individuals who had been unemployed for 27 or more weeks was a seasonally adjusted 35%, up from 22% in February, before the pandemic began showing that the some jobs lost in April are **permanent losses** not just temporary because of pandemic.
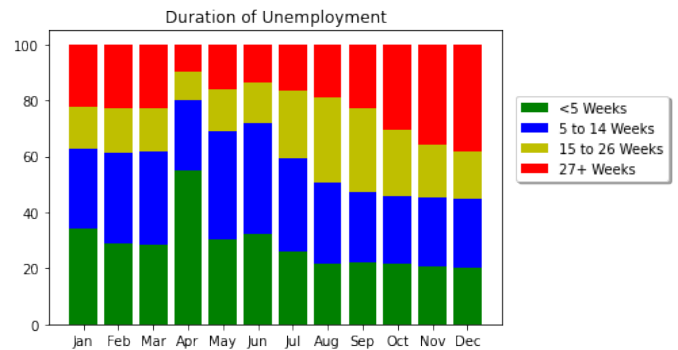

Figure 12: Duration of Unemployment

5.4 Small Scale Active Business Owners:

We need to note The CPS survey reference period is generally the calendar week that contains the 12th day of the month. To estimate business ownership in the CPS data, I identify all individuals who own a business as their main job in the survey month (based on the class of worker question and monthly labor force re code). The main job is defined as the one with the most hours worked during the survey week. Thus, individuals who start side businesses will not be counted if they are working more hours on a wage and salary job. The CPS captures the current work activity of the business owner, and whether that business owner is currently operating the business. Thus, the number of active business owners can be captured in the data, but there is no way of telling whether these are temporary or permanent business closures. But, inactive business owners regardless of whether the business is temporary or permanently closed are suffering losses in business income during those months of non operation.The measure of business ownership in the CPS captures all business owners including those who own incorporated or unincorporated businesses, and those who are employers or non employers
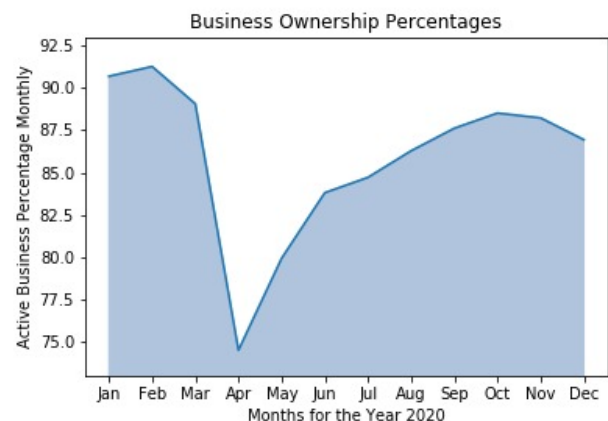

Figure 13: Percentage of Active Small Scale Business

As we can see from above graph Active small scale business owners dropped from around 89% in march to 75% in April and

and the partial rebound in May and a continuing rebound in June. March 2020 only shows a small drop in business owners likely because of the limited effect from shelter-in-place restrictions.May 2020 shows a partial rebound from April 2020.

5.5 Active Business owners by Ethnicity:

Turning to racial patterns. The **Figure 14** displays the number of active business owners by major racial groups. The findings are alarming. The percentage of active business owners for White is 91.63% in Feb which decreased to 76.14% in April whereas for black's it reduced from 85.38% to 59% and for Asian people it reduced from 87.27% to 59.45% even though all the government policies to revive the economy are applicable to people irrespective of the ethnicity White's are in a better position compared to back's and Asian business people. This might be because of one of the following reasons.

- major black-owned firms are more likely to be located in **COVID-19 hot spots**, whereas white-owned firms are less likely to be in the most heavily affected areas
- Black-owned businesses also had **fewer resources** to fall back on when the pandemic struck.
- there might be a disparity in government aid through policies like **Paycheck Protection Program(PPP)**.

5.6 Active Business owners by Immigration Status:

As we can see from **Figure 15** Immigrants were the most effected compared to native it may be because native have a strong financial standing and Compared to Immigrants Native's get more help from government policies.

5.7 Employment Percentage by industry:

Almost every industry experienced sizeable drops in the number of employments from February to April as shown in **Figure 17**. The only exception was **Agriculture** for which the employment increased slightly.
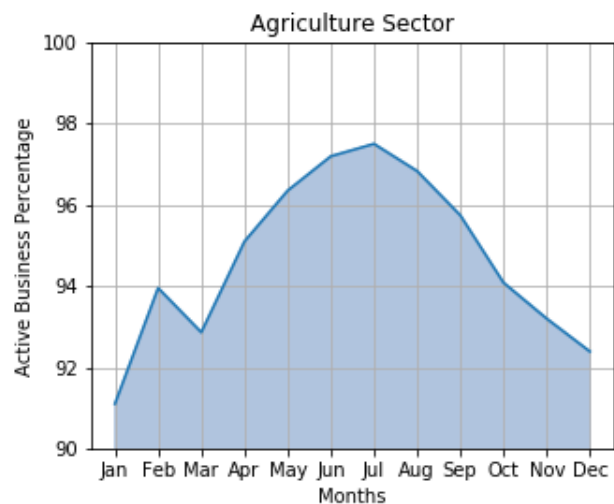
Figure 16: **Employment Percentage for the Agriculture Sector**

Personal and Laundry Services were especially hard hit with losses of 67% of business owner activity in April and continuing losses of 44% in May and 22% in June
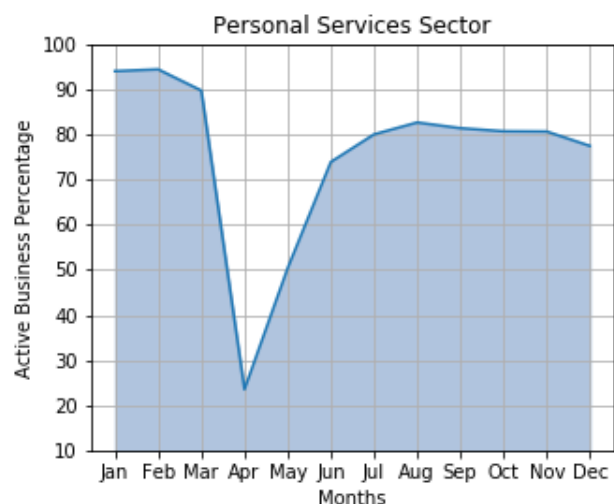
Figure 18: **Active businesses for the Personal Services Sector**

6 COMMENT ON GOVERNMENT STEPS:

Based on the analysis we have done we will comment on two of the steps taken by the government to revive the economy and decrease the unemployment.

6.1 Employment Visa Ban:

On June 24, 2020 the then President **Trump** issued a proclamation that will suspend the entry of foreign nationals of various work
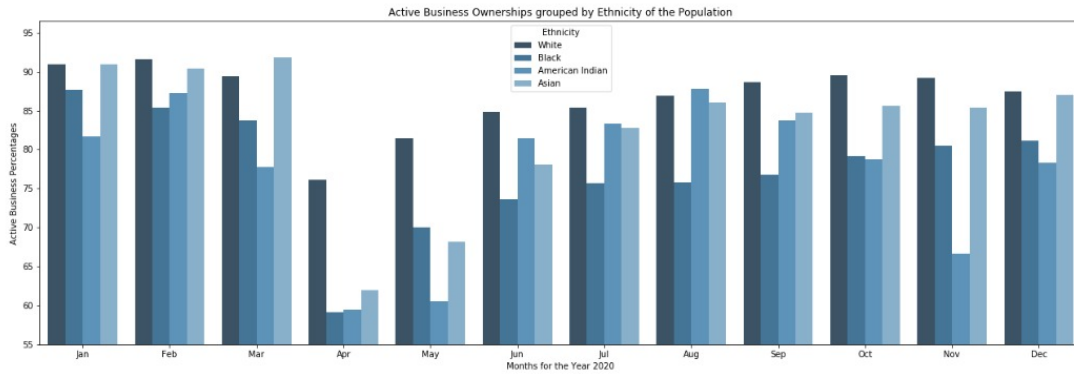
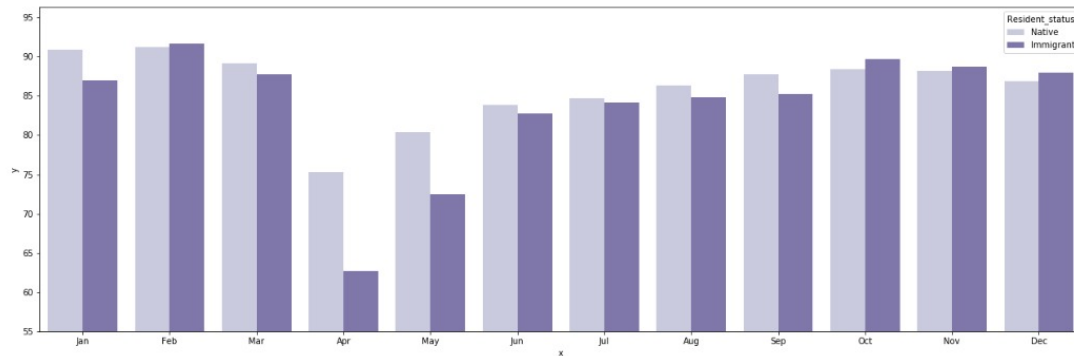**Figure 14: Percentage of Active Business owners by Etnicity**

Figure 15: Percentage of Active Business owners by Immigration Status

visa's like H-1B, L-1, H-2B, J-1 categories, and related categories for dependents, with some exceptions and the reason sated for this proclamation is that "since Unemployment is at all time high we need to put more emphasis on hiring Americans over foreign nationals". Though Unemployment is at all time high this alone does not paint the whole picture as we can see from **Figure 17** the percentage drop in active business in **Financial Activities and Professional and Business** is a little over 10% in April from March. where as other sectors such as **Personal Services and Entertainment sectors** are much more worstly effected than the former sectors so the increase in Unemployment is mostly likely because of the other sectors rather than **Financial Activities and Professional and Business** and majority of the high skilled labour under these banned visa's come under this sector so banning these visa is not going to help in increasing the job growth.

On October 02,2020 An **Us court stopped the proclamation** stating that while the proclamation cited unemployment figures for the early months of the pandemic, the Judge observed that, "The statistics regarding pandemic-related unemployment actually indicate that unemployment is concentrated in service occupations and that

large number of job vacancies remain in the area most affected by the ban, computer operations which require high-skilled workers." which is **consistent with our analysis**.

6.2 Paycheck Protection Program:

The Paycheck Protection Program (PPP) is a $953-billion business loan program established by the United States federal government in April 3rd 2020 through the Coronavirus Aid, Relief, and Economic Security Act (CARES Act) to help certain businesses, self-employed workers, sole proprietors, certain nonprofit organizations, and tribal businesses continue paying their workers.

As we can see from **Figure 17** Personal Services sector like laundry and Saloons are the worst effected as its employment percentage dropped from 94.318% in February to 23.55% in April. But it increased to 50% followed by 73% in May and June. There are two possible explanations for this change.

- The Pandemic took the world and US by surprise and people's first reaction to pandemic is much more intense which decreased over the next few months.

Industry	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Agriculture	91.101695	93.945720	92.857143	95.089286	96.338673	97.183099	97.488584	96.819085
Construction	87.412587	90.466733	87.431092	73.525872	80.221130	83.592814	83.476132	84.529148
Manufacturing	92.653061	92.477876	93.301435	84.158416	80.219780	89.119171	88.500000	87.939698
Wholesale_trade	87.962963	88.990826	94.047619	87.951807	87.500000	86.363636	89.473684	86.734694
Retail_trade	93.456033	91.125541	90.222222	73.568282	80.465116	84.848485	88.324873	88.341969
Transportation	89.169675	88.926174	85.992218	70.037453	71.851852	77.816901	77.536232	82.061069
Information	89.887640	90.697674	80.681818	76.543210	77.777778	74.666667	84.615385	80.459770
Financial_Activities	90.804598	92.090395	90.293454	80.338266	84.700665	86.652542	88.800000	91.287129
Professional & Buisiness	90.742158	89.655172	88.701518	77.956989	83.621474	84.673605	84.642234	85.374150
Educational Services	90.476190	91.970803	91.071429	64.035088	73.809524	73.333333	76.146789	85.964912
Health Services	94.082840	95.223881	93.092105	79.617834	83.214286	89.230769	88.811189	89.354839
Entertainment	90.733591	90.441176	81.896552	60.851064	64.285714	69.268293	65.384615	67.452830
Restaurants	92.156863	92.857143	90.277778	73.376623	78.048780	79.562044	80.291971	80.891720
Repair Maintenance	90.212766	88.392857	90.055249	79.545455	82.911392	86.111111	86.330935	88.079470
Personal Services	93.920973	94.318182	89.632107	23.550725	50.326797	73.825503	79.935275	82.568807

Figure 17: Employment Percentage by industry

- Paycheck Protection Program Started by the Government to help small scale business in pandemic was successful in its implementation.

The second explanation is seems more likely than the first as the increase in percentage in other sectors is not that high. So Personal Services Sector which is an Ideal candidate for PPP has much increase when compared to other sectors.

7 CONCLUSION

The Pandemic has resulted in an all time high unemployment rate in USA. Even though it got better over the next few months some of the jobs lost in April are permanent losses. Government needs to to take steps in restoring the economy. PPF was a success according to the data It was a good step to implement PPf for the second time as it gives boost to small scale business. But even though policies are not biased based on Colour we have seen black business owners are much more worstly effected than white owners. Government needs to try to find the reason for this and take necessary steps. Travel data is also one sector which got effected mostly as it can bee seen from flight data. Now as more and more people are getting vaccinated once the situation stabilizes government should be more encouraging towards travel. One such way is incentivize people towards travel.

8 CONTRIBUTIONS

During the data acquisition process, we came across lots of datasets containing data for COVID-19 of different dimensions. As most datasets contained time series information and a lot of columns were unnecessary considering the purpose of our project and as we only wanted to plot its effect on the economy the three of us agreed to plot only the total number of COVID-19 cases against the months for the year of 2020 and concentrate on the economic factors that would help us achieve our objective.We mainly wanted to focus on the US Economy and how it has been affected in the past year and also try to predict in what direction things are headed to considering what has been happening in the world, in other countries, what trends were observed in their economies, what changes were brought into place by the governments and how that affected them economically.

Badrinath suggested we make use of the data that was available for the countries – India, Italy, Spain and China as other than US these economies play a substantial role in the world market and were the most affected economies in the past year. As we were collecting data we had to make sure to preprocess it well and develop visualizations initially for our own understanding of the data before we proceeded with the analysis and hence Lokesh suggested we make use of Python libraries numpy, pandas, and matplotlib alongwith a notebook system – Jupyter

Once we were done with all of data collection required for our analysis we realized the need for using big data infrastructure for preprocessing and the data and also the subsequent analysis that followed and hence Badrinath suggested the use of Google Colab. Colab is a free Jupyter notebook environment that runs entirely in the cloud.

As we moved to the analysis part, Susmita noticed there was a data dictionary along with the CPS data that decoded the variable names of the columns present in the CPS microdata and also that the dataset could be used for finding out more information on how small businesses and restaurants have been affected due to COVID-19 in the US. Badrinath went through the microdata explaining the 388 columns in CPS data noticed it provided data about the sectors to which the people who were interviewed belonged and could be segregated to find active business percentages. On analysis Badrinath noticed that sectors like Arts/Entertainment and Personal Services were most affected and remain affected until now upto some extent hence we collectively agreed to plot the data for these sectors.

As we were visualizing the data, Badrinath realized that we could explain the trends observed in the data with the help of real life events and policies that were brought into place by the US government during that time such as the Paycheck Protection Program and many more and hence we agreed to show how the plots are affected before and after these programs were put into action. Susmita suggested to use The CPS data which contains demographic information about people such as their race and heritage. Many news outlets reported that the Black and Asian community were being affected more compared to the general population during the pandemic. Susmita suggested verifying whether we could observe such a trend in our data. Lokesh researched about the different policies and actions that were carried out by various governments around the world that helped us to predict the future trends of the US Economy.

9 SUBMISSION

You can find the code on the GitHub repository
https://github.com/badri449/Bigdata_Project

REFERENCES
[1] Trade Data From World Trade Organization

[2] Flight DataFrom flightradar24.com

[3] GDP of USA sector wise from fred.stlouisfed.org

[4] Share Market,Gold and Oil taken From Yahoo Finance

[5] Covid Data

[6] CPS Data Taken from census.gov

[7] Department of Economics, University of California, Santa Cruz California, Robert Fairlie, Department of Economics, University of California, Santa Cruz, CA, 95064; Stanford University (Visiting Scholar), and NBER. on Small Scale Business

[8] US Court Blocks ban on H1b by Priyanka Sangani,ET Bureau

[9] Paycheck Protection Program Details from US small Business Administration

[10] Coronavirus: How the pandemic has changed the world economy by Lora Jones, Daniele Palumbo and David Brown