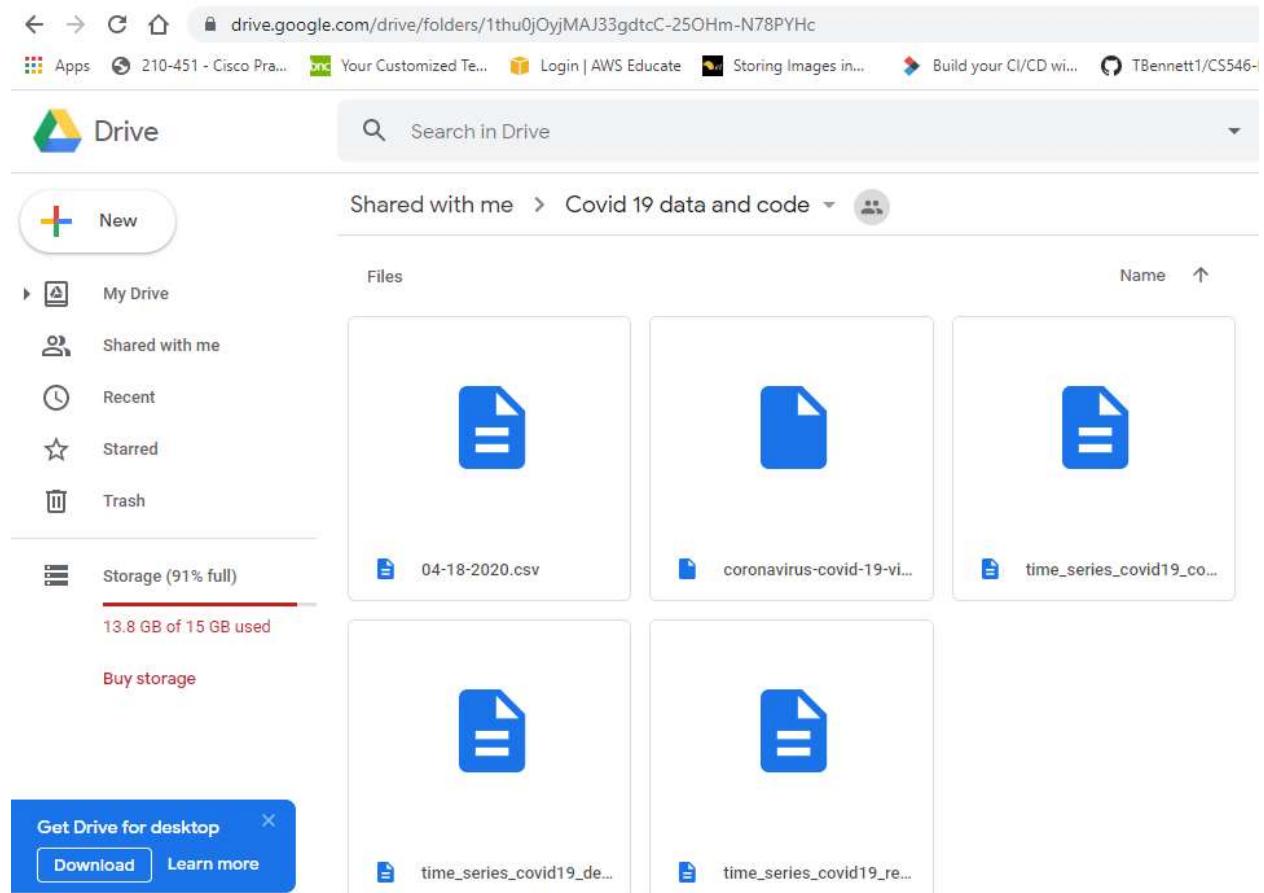


Lab # 5

A.

Step 1: Download files from: <https://drive.google.com/drive/folders/1thu0jOyjMAJ33gdtcC-25OHm-N78PYHc>



The screenshot shows a Google Drive interface. On the left, there's a sidebar with options like 'New', 'My Drive', 'Shared with me', 'Recent', 'Starred', and 'Trash'. Below that is a storage section showing 'Storage (91% full)' and '13.8 GB of 15 GB used'. A 'Buy storage' button is also present. At the top, the URL is 'drive.google.com/drive/folders/1thu0jOyjMAJ33gdtcC-25OHm-N78PYHc'. The main area shows five files under the folder 'Covid 19 data and code': '04-18-2020.csv', 'coronavirus-covid-19-vi...', 'time_series_covid19_co...', 'time_series_covid19_de...', and 'time_series_covid19_re...'. Each file has a blue document icon. A 'Get Drive for desktop' button is visible at the bottom left.

The five files are:

- 1) 04-18-2020.csv (data)
- 2) time_series_covid19_confirmed_global.csv (data)
- 3) time_series_covid19_confirmed_global.csv (data)
- 4) time_series_covid19_recovered_global.csv (data)
- 5) coronavirus-covid-19-visualization-prediction.ipynb (code)

B.

Create an Amazon S3 Bucket and upload the above data into that bucket:

Step 1: Login to your AWS account.

From the console, Services dropdown, select S3 under storage category:

The screenshot shows the AWS Services dropdown menu. The 'Storage' category is highlighted in blue. Under 'Storage', the 'S3' service is listed. Other services shown include Compute (EC2, Lightsail, Lambda, Batch, Elastic Beanstalk, Serverless Application Repository, AWS Outposts, EC2 Image Builder), Blockchain (Amazon Managed Blockchain), Analytics (Athena, EMR, CloudSearch, Elasticsearch Service, Kinesis, QuickSight), End User Computing (WorkSpaces, AppStream 2.0, WorkDocs, WorkLink), Internet Of Things (IoT Core, FreeRTOS, IoT 1-Click, IoT Analytics, IoT Device Defender, IoT Device Management, IoT Events, IoT Greengrass, IoT SiteWise, IoT Things Graph), Management & Governance (AWS Organizations, CloudWatch, AWS Auto Scaling, CloudFormation, CloudTrail, Config), Security, Identity, & Compliance (IAM, Resource Access Manager, Cognito), and Storage (S3, EFS, FSx, S3 Glacier, Storage Gateway, AWS Backup).

On the next page, select Create Bucket option:

The screenshot shows the Amazon S3 console. A blue banner at the top states: "We're gradually updating the design of the Amazon S3 console. You will notice some updated screens as we improve the performance and user interface. To help us improve the experience, give feedback on the recent updates." On the left sidebar, there are links for Buckets (Batch Operations, Access analyzer for S3), Block public access (account settings), and Feature spotlight. The main content area is titled "Amazon S3" and shows a table for "Buckets (0)". The table has columns for Name, Region, Access, and Bucket created. A message below the table says "No buckets" and "You don't have any buckets." At the bottom right of the table area is a large orange "Create bucket" button.

Give the bucket a suitable name:

The screenshot shows the 'Create bucket' page in the AWS S3 console. The 'General configuration' section is visible, containing fields for 'Bucket name' (set to 'mynewcovidbucket') and 'Region' (set to 'US East (N. Virginia) us-east-1'). A note below the bucket name field states: 'Bucket name must be unique and must not contain spaces or uppercase letters. See rules for bucket naming'.

For now, all public accesses are blocked to this bucket:

The screenshot shows the 'Advanced settings' section of the 'Create bucket' page. It includes a note about blocking public access and four checkboxes under the heading 'Block all public access':

- Block public access to buckets and objects granted through new access control lists (ACLs)**
S3 will block public access permissions applied to newly added buckets or objects, and prevent the creation of new public access ACLs for existing buckets and objects. This setting doesn't change any existing permissions that allow public access to S3 resources using ACLs.
- Block public access to buckets and objects granted through any access control lists (ACLs)**
S3 will ignore all ACLs that grant public access to buckets and objects.
- Block public access to buckets and objects granted through new public bucket or access point policies**
S3 will block new bucket and access point policies that grant public access to buckets and objects. This setting doesn't change any existing policies that allow public access to S3 resources.
- Block public and cross-account access to buckets and objects through any public bucket or access point policies**
S3 will ignore public and cross-account access for buckets or access points with policies that grant public access to buckets and objects.

At the bottom right, there are 'Cancel' and 'Create bucket' buttons.

Click on Create Bucket.

s3.console.aws.amazon.com/s3/home?region=us-east-1

Apps 210-451 - Cisco Pra... Your Customized Te... Login | AWS Educate Storing Images in... Build your CI/CD wi... T8Bennett1/CS546-F... Website | Trello

AWS Services Resource Groups

Amazon S3

Buckets Batch Operations Access analyzer for S3

Block public access (account settings)

Feature spotlight

Successfully created bucket "mynewcovidbucket"
To upload files and folders, or to configure additional bucket settings such as Bucket Versioning, tags, and default encryption, choose Go to bucket details.

Go to bucket details

Amazon S3

Buckets (1)

Copied ARN Empty Delete Create bucket

Find bucket by name

Name	Region	Access	Bucket created
mynewcovidbucket	US East (N. Virginia) us-east-1	Not public	2020-04-22T19:03:36.000Z

Bucket has been created successfully.

Click on the name of the bucket and add files by upload button on the Next page:

s3.console.aws.amazon.com/s3/buckets/mynewcovidbucket/?region=us-east-1

Apps 210-451 - Cisco Pra... Your Customized Te... Login | AWS Educate Storing Images in... Build your CI/CD wi... T8Bennett1/CS546-F... Website | Trello

AWS Services Resource Groups

Amazon S3 > mynewcovidbucket

mynewcovidbucket

Overview Properties Permissions Management Access points

Upload Create folder Download Actions US East (N. Virginia)

This bucket is empty. Upload new objects to get started.



Select the five files downloaded in A.

Click Next.

And Upload these files.

Files can be seen in Overview Tab:

The screenshot shows the AWS S3 console interface. At the top, there's a navigation bar with links like 'Services', 'Resource Groups', 'Global', and 'Support'. Below the navigation bar, the bucket name 'mynewcovidbucket' is displayed. A tab bar at the top of the main content area includes 'Overview', 'Properties', 'Permissions', 'Management', and 'Access points', with 'Management' being the active tab. A search bar below the tabs contains the placeholder text 'Type a prefix and press Enter to search. Press ESC to clear.' Under the search bar are buttons for 'Upload', '+ Create folder', 'Download', and 'Actions'. To the right of these buttons, it says 'US East (N. Virginia)' with a refresh icon. A table follows, showing the details of the five files in the bucket. The columns are 'Name', 'Last modified', 'Size', and 'Storage class'. The files are listed as follows:

Name	Last modified	Size	Storage class
04-18-2020.csv	Apr 22, 2020 3:07:12 PM GMT-0400	308.5 KB	Standard
time_series_covid19_confirmed_global.csv	Apr 22, 2020 3:07:12 PM GMT-0400	73.2 KB	Standard
time_series_covid19_deaths_global.csv	Apr 22, 2020 3:07:12 PM GMT-0400	57.9 KB	Standard
time_series_covid19_recovered_global.csv	Apr 22, 2020 3:07:12 PM GMT-0400	61.4 KB	Standard

At the bottom of the table, there's a summary bar with the text 'Operations 0 In progress 1 Success 0 Error'.

C.

Open Amazon Sage Maker.

Go to the console, select Services dropdown and select Amazon SageMaker from Machine Learning division:

The screenshot shows the AWS Services dropdown menu open. On the left, a sidebar lists services grouped by category: S3, EC2, CloudFormation, CodePipeline, Elastic Beanstalk, Developer Tools, Robotics, and others. The 'Machine Learning' section is expanded, showing the following services:

- Machine Learning
- Amazon SageMaker
- Amazon Comprehend
- Amazon Forecast
- Amazon Fraud Detector
- Amazon Kendra
- Amazon Lex
- Amazon Machine Learning
- Amazon Personalize
- Amazon Polly
- Amazon Rekognition
- Amazon Transcribe
- Amazon Translate
- AWS DeepLens
- AWS DeepRacer
- Amazon Augmented AI

Other collapsed sections include Application Integration, AWS Cost Management, and Customer Engagement. At the bottom of the dropdown, there are 'Group' and 'A-Z' buttons.

The page would look like:

D. Launch a Notebook instance (and at the same time create the new IAM permissions for the Amazon S3 bucket you had created)

An Amazon SageMaker notebook instance is a fully managed machine learning (ML) Amazon Elastic Compute Cloud (Amazon EC2) compute instance that runs the Jupyter Notebook App. You use the notebook instance to create and manage Jupyter notebooks that you can use to prepare and process data and to train and deploy machine learning models.

Step 1: On the left hand panel in the Amazon SageMaker page, click on Notebook Instances.

Click on Create notebook instance.

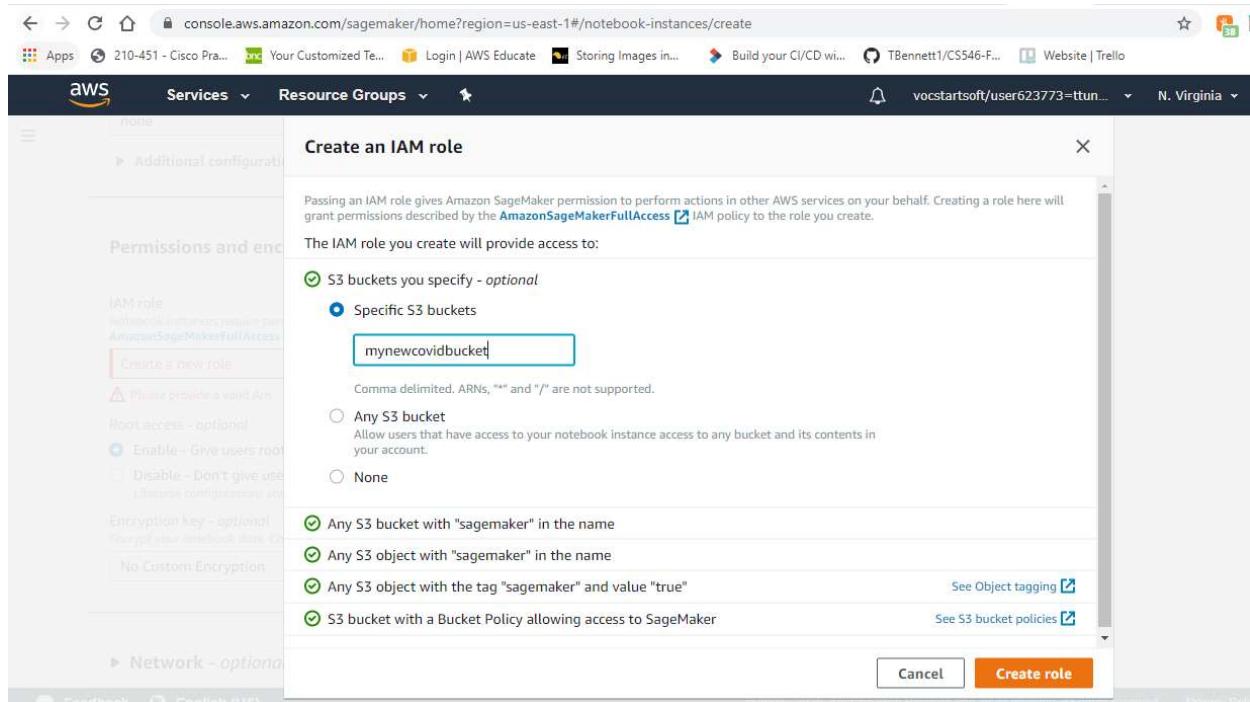
The screenshot shows the 'Create notebook instance' page in the AWS SageMaker console. The 'Notebook instance settings' section is visible, containing fields for the notebook instance name ('mycovidnotebook'), instance type ('ml.t2.medium'), and elastic inference configuration ('none'). A link to 'Additional configuration' is also present.

Provide a name for your notebook, for instance type, select ml.t2.medium as this is the least expensive one.

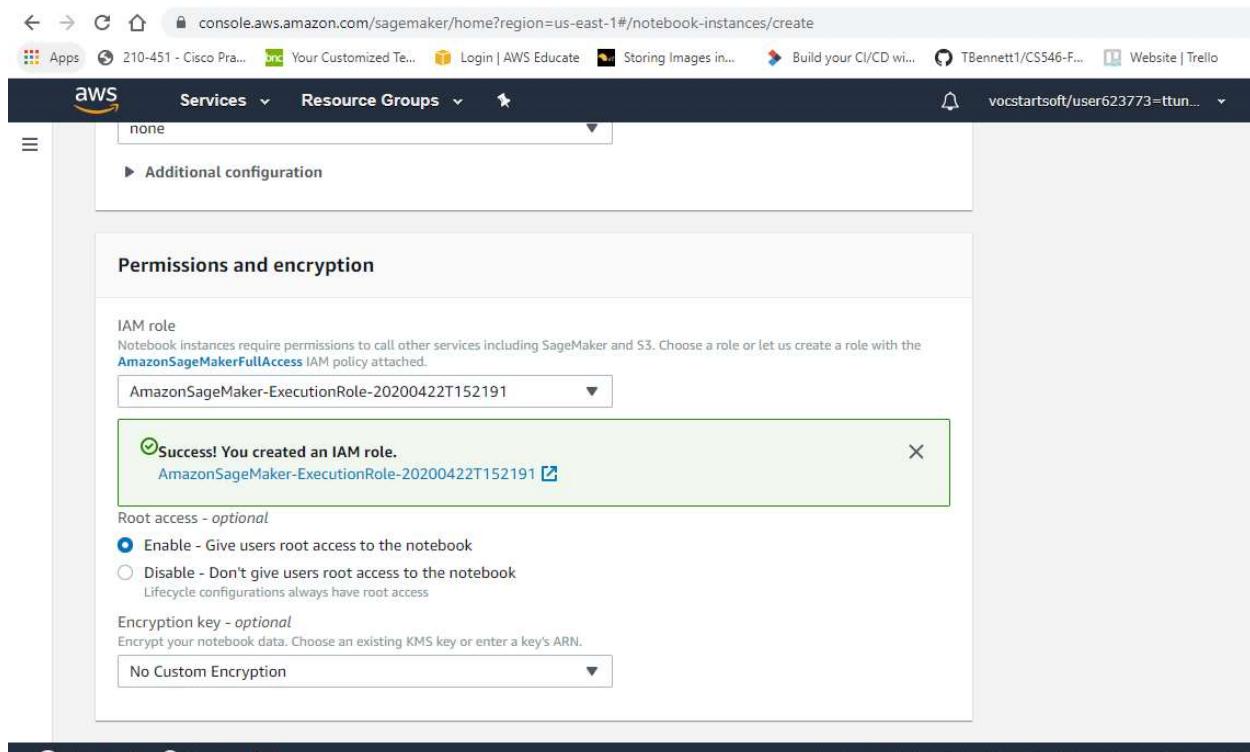
The screenshot shows the 'Permissions and encryption' configuration page. Under the 'IAM role' section, the 'Create a new role' button is highlighted. The 'Encryption key - optional' section shows 'No Custom Encryption' selected.

For IAM role, click on Create a new role.

Enter the name of your bucket:



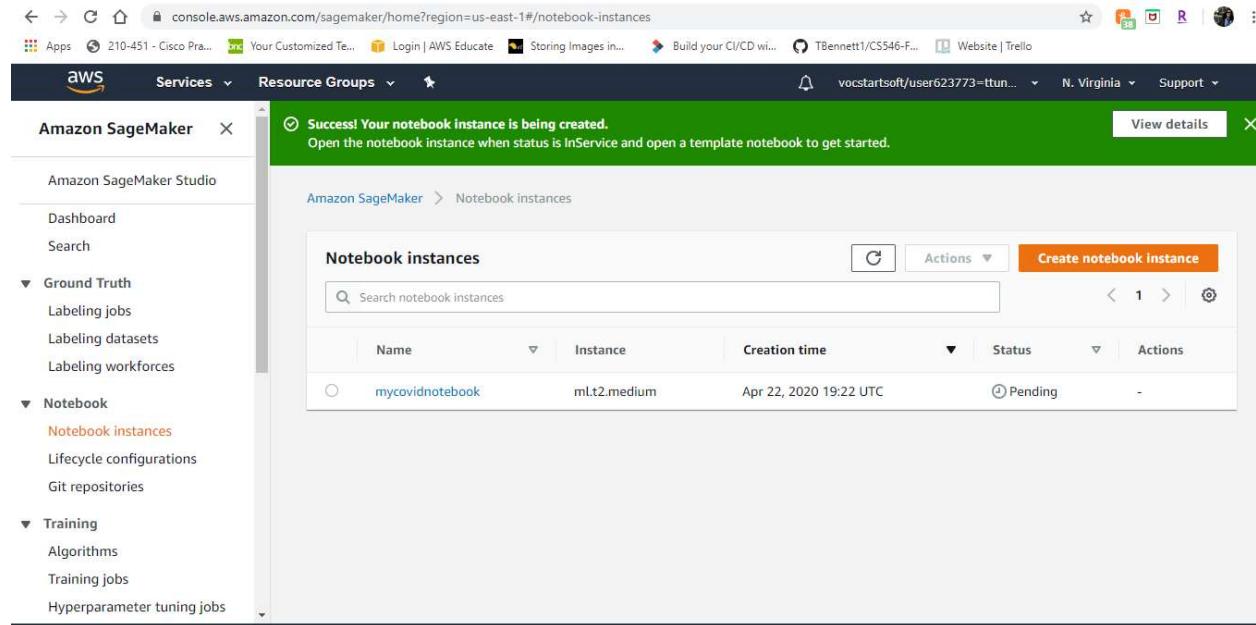
Click on Create Role.



The new IAM role has been created successfully. Also, make sure the root access to the notebook is enabled.

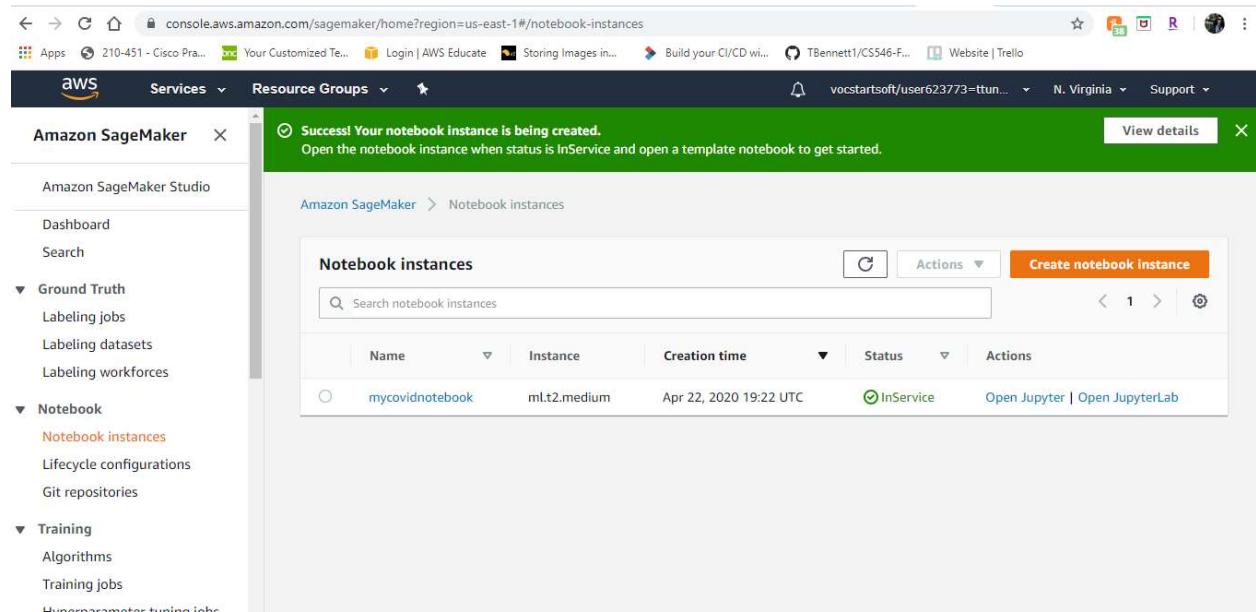
Click on Create Notebook Instance.

The notebook instance is in pending state, wait for it to be created completely.



A screenshot of the AWS SageMaker console. The left sidebar shows navigation options like Amazon SageMaker Studio, Dashboard, Search, Ground Truth, Notebook, Training, and more. The main content area displays a success message: "Success! Your notebook instance is being created. Open the notebook instance when status is InService and open a template notebook to get started." Below this, a table lists the "Notebook instances". There is one entry: "mycovidnotebook" (ml.t2.medium), created on Apr 22, 2020 19:22 UTC, with a status of "Pending".

The Notebook instance has been launched.



A screenshot of the AWS SageMaker console, identical to the previous one but with a key difference: the status of the notebook instance "mycovidnotebook" has changed to "InService". The table now shows: "mycovidnotebook" (ml.t2.medium), created on Apr 22, 2020 19:22 UTC, with a status of "InService". The "Actions" column for this row contains "Open Jupyter" and "Open JupyterLab".

The notebook instance has a preconfigured Jupyter notebook server and a set of Anaconda libraries.

E. Load the Notebook:

Click on Open Jupyter and Using Upload button upload the .pynb file as:

The screenshot shows a Jupyter Notebook interface. At the top, there's a navigation bar with links like 'mycovidnotebook.notebook.us-east-1.sagemaker.aws/tree', 'Open JupyterLab', and 'Quit'. Below the navigation bar is a header with tabs: 'Files', 'Running', 'Clusters', 'SageMaker Examples', and 'Conda'. A sub-header says 'Select items to perform actions on them.' There's a file list with one item: 'coronavirus-covid-19-visualization-prediction.ipynb'. To the right of the file list are buttons for 'Upload', 'New', and a trash icon. Below the file list are filters for 'Name', 'Last Modified', and 'File size'.

Click on this file.

The screenshot shows a Jupyter Notebook interface with a modal dialog box titled 'Kernel not found'. The dialog message says 'Could not find a kernel matching Python 3. Please select a kernel.' with a dropdown menu currently set to 'conda_python3'. In the background, the notebook cell contains text about COVID-19, a link to GitHub, and a CDC link. It also includes a warning about running the notebook to see graphs better. At the bottom of the cell is a large image of a COVID-19 virus particle.

Select python 3 and run this code.

This notebook tracks the spread of the novel coronavirus, also known as SARS-CoV-2. It is a contagious respiratory virus that first started in Wuhan in December 2019. On 2/11/2020, the disease is officially named COVID-19 by the World Health Organization.

Data: <https://github.com/CSSEGISandData/COVID-19>
Learn more from the [WHO](#)
Learn more from the [CDC](#)

Make sure you run the notebook to see the graphs better. Some diagrams are hard to see in the default view.

F.

Edit name of the bucket:

After the first block of code is executed, edit the name of bucket:

```
import operator
plt.style.use('fivethirtyeight')
%matplotlib inline
```

Connect to your bucket.

```
In [1]: role = get_execution_role()
bucket='mynewcovidbucket'
key1 = '04-18-2020.csv'
key2 = 'time_series_covid19_confirmed_global.csv'
key3 = 'time_series_covid19_deaths_global.csv'
key4 = 'time_series_covid19_recovered_global.csv'
data_location1 = 's3://{}{}'.format(bucket, key1)
data_location2 = 's3://{}{}'.format(bucket, key2)
data_location3 = 's3://{}{}'.format(bucket, key3)
data_location4 = 's3://{}{}'.format(bucket, key4)
```

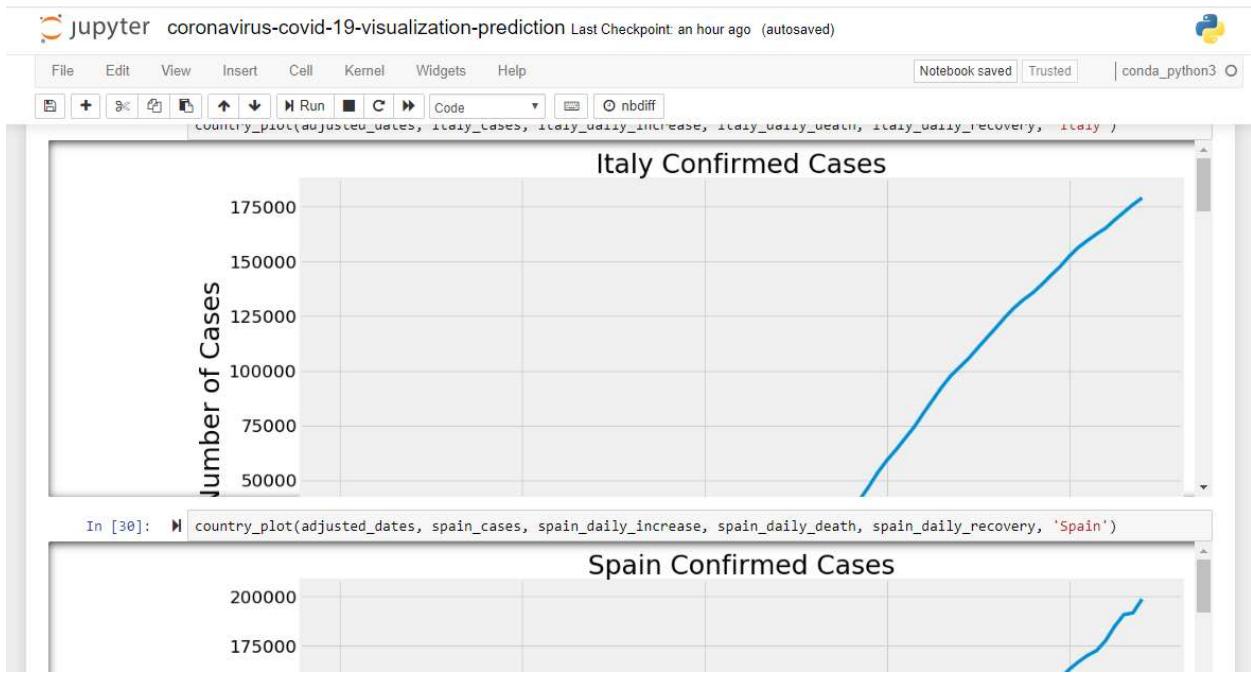
```
NameError: name 'get_execution_role' is not defined
Traceback (most recent call last)
<ipython-input-1-31daa24a9363> in <module>()
----> 1 role = get_execution_role()
      2 bucket=<bucket_name>
      3 key1 = '04-18-2020.csv'
      4 key2 = 'time_series_covid19_confirmed_global.csv'
      5 key3= 'time_series_covid19_deaths_global.csv'
```

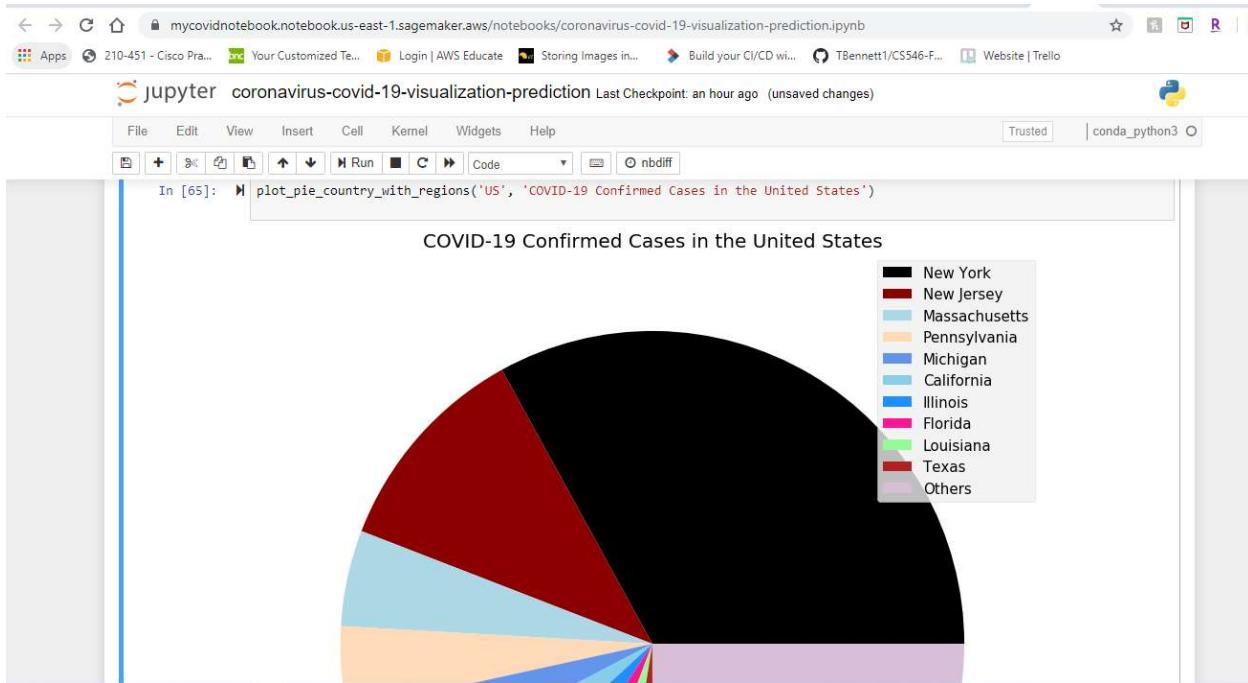
Click on Restart and Run all option from Kernel menu.

Screenshot of a Jupyter Notebook interface showing a code cell with a context menu open over it. The menu includes options like 'Interrupt', 'Restart', 'Reconnect', 'Shutdown', 'Change kernel', 'Conda Packages', and 'Visit anaconda.org'. The code cell contains Python code related to COVID-19 visualization.

```
for i in range(len(regions)):
    confirmed = regions[i]['confirmed']
    # additional logic here
    if len(regions) > 10:
        regions = regions[:10]
        confirmed = confirmed[:10]
        plot_pie_charts(regions, confirmed, title)
    else:
        img_data.seek(0)
        s3 = boto3.resource('s3')
        bucket1 = s3.Bucket(bucket)
        bucket1.put_object(Body=img_data, ContentType='image/png', Key=f"Output/{title}.png")
```

After sometime, these graphs are generated:





G. To check the results, go to your S3 bucket

Click on the name of your bucket and you can see a new Output folder has been created:

Name	Last modified	Size	Storage class
04-18-2020.csv	Apr 22, 2020 3:07:12 PM GMT-0400	308.5 KB	Standard
time_series_covid19_confirmed_global.csv	Apr 22, 2020 3:07:12 PM GMT-0400	73.2 KB	Standard

Click on this file and you can see many output files have been generated:

s3.console.aws.amazon.com/s3/buckets/mynewcovidbucket/Output/?region=us-east-1			
AWS Services ▾ Resource Groups ▾			
<input type="text"/> Type a prefix and press Enter to search. Press ESC to clear.			
Actions	Upload	Create folder	Download
			US East (N. Virginia)
Name	Last modified	Size	Storage class
Active cases over time	Apr 22, 2020 4:47:57 PM GMT-0400	57.5 KB	Standard
Bayesian	Apr 22, 2020 4:47:55 PM GMT-0400	25.3 KB	Standard
Bayesian Ridge Regression Predictions	Apr 22, 2020 4:48:15 PM GMT-0400	65.3 KB	Standard
COVID-19 Confirmed Cases in the United States	Apr 22, 2020 4:51:34 PM GMT-0400	69.3 KB	Standard
Corona_Cases_World	Apr 22, 2020 4:48:11 PM GMT-0400	72.8 KB	Standard
Corona_Death_World	Apr 22, 2020 4:48:12 PM GMT-0400	79.9 KB	Standard
Corona_Recoveries_World	Apr 22, 2020 4:48:12 PM GMT-0400	86.3 KB	Standard
Covid-19 Confirmed Cases per Country	Apr 22, 2020 4:48:24 PM GMT-0400	65.8 KB	Standard

s3.console.aws.amazon.com/s3/buckets/mynewcovidbucket/Output/?region=us-east-1			
AWS Services ▾ Resource Groups ▾			
<input type="text"/> Type a prefix and press Enter to search. Press ESC to clear.			
Actions	Upload	Create folder	Download
			US East (N. Virginia)
Name	Last modified	Size	Storage class
Covid-19 Confirmed Cases per State	Apr 22, 2020 4:48:25 PM GMT-0400	63.3 KB	Standard
Death over time	Apr 22, 2020 4:47:56 PM GMT-0400	48.3 KB	Standard
Death_vs_Recoveries	Apr 22, 2020 4:48:17 PM GMT-0400	57.9 KB	Standard
Log of Common Log Number of Coronavirus Confirmed Cases in Countries	Apr 22, 2020 4:48:22 PM GMT-0400	32.9 KB	Standard
Log of Number of Coronavirus Confirmed Cases in Provinces	Apr 22, 2020 4:48:23 PM GMT-0400	34.1 KB	Standard
Mortality_Rate	Apr 22, 2020 4:48:16 PM GMT-0400	63.6 KB	Standard
Number of Coronavirus Confirmed Cases in Provinces	Apr 22, 2020 4:48:23 PM GMT-0400	35.1 KB	Standard
Number of Covid-19 Confirmed Cases in Countries	Apr 22, 2020 4:48:21 PM GMT-0400	32.5 KB	Standard
Polynomial Regression Predictions	Apr 22, 2020 4:48:14 PM GMT-0400	63.6 KB	Standard

s3.console.aws.amazon.com/s3/buckets/mynewcovidbucket/Output/?region=us-east-1

Apps 210-451 - Cisco Pra... Your Customized Te... Login | AWS Educate Storing Images in... Build your CI/CD wi... TBennett1/CS546-F... Website | Trello

Services Resource Groups

aws Services Resource Groups

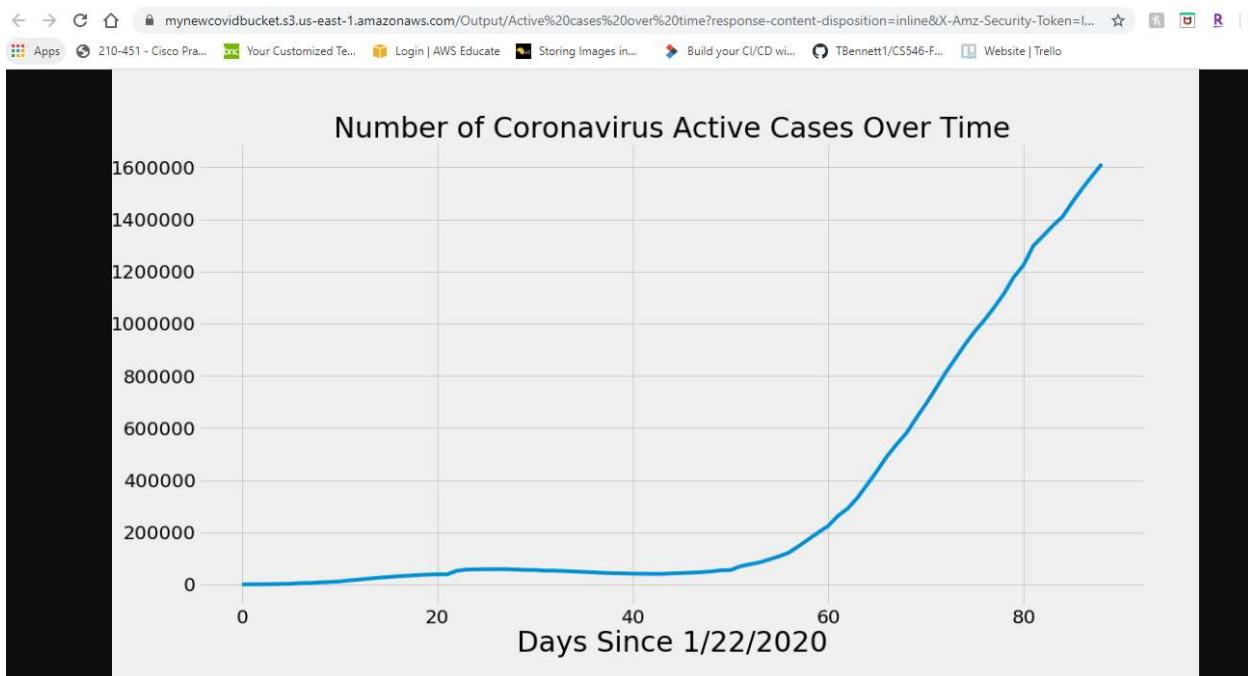
Type a prefix and press Enter to search. Press ESC to clear.

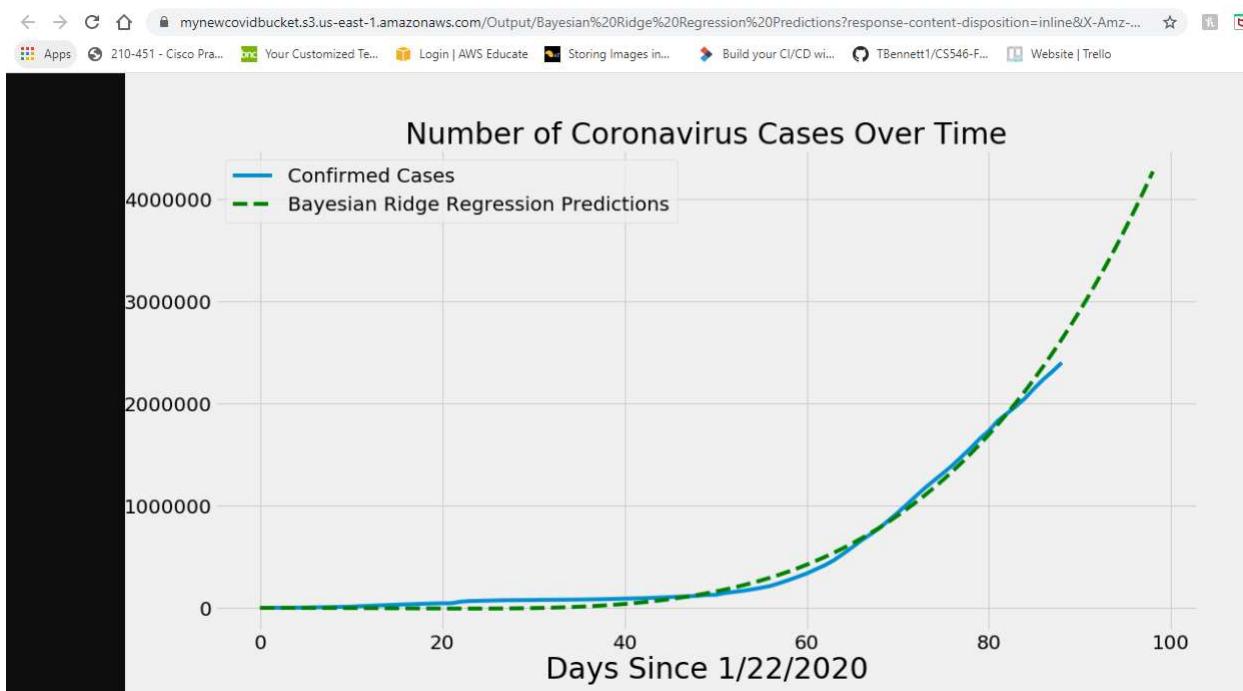
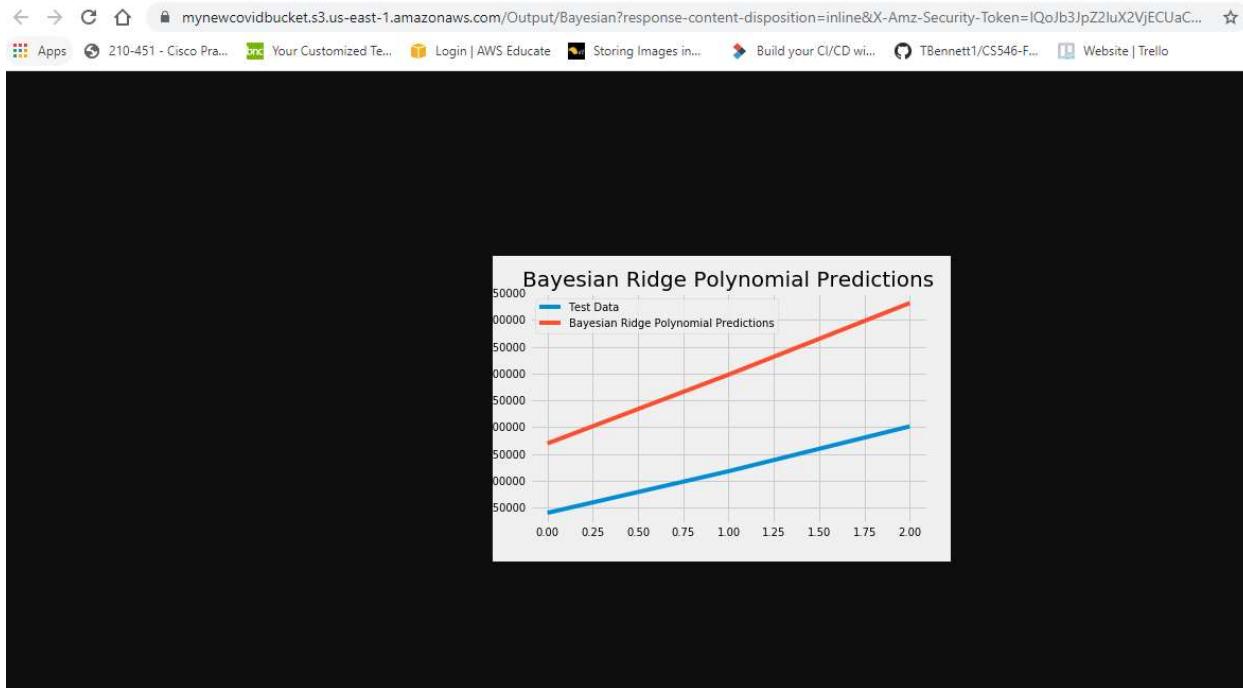
Upload Create folder Download Actions US East (N. Virginia)

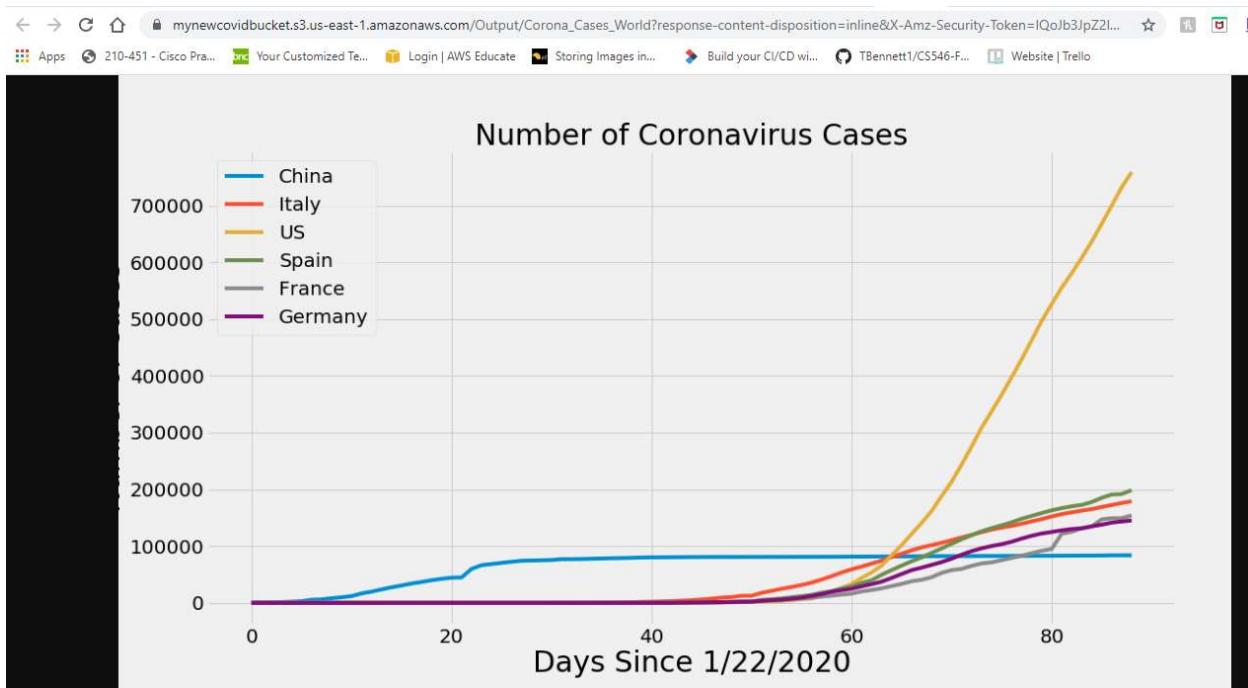
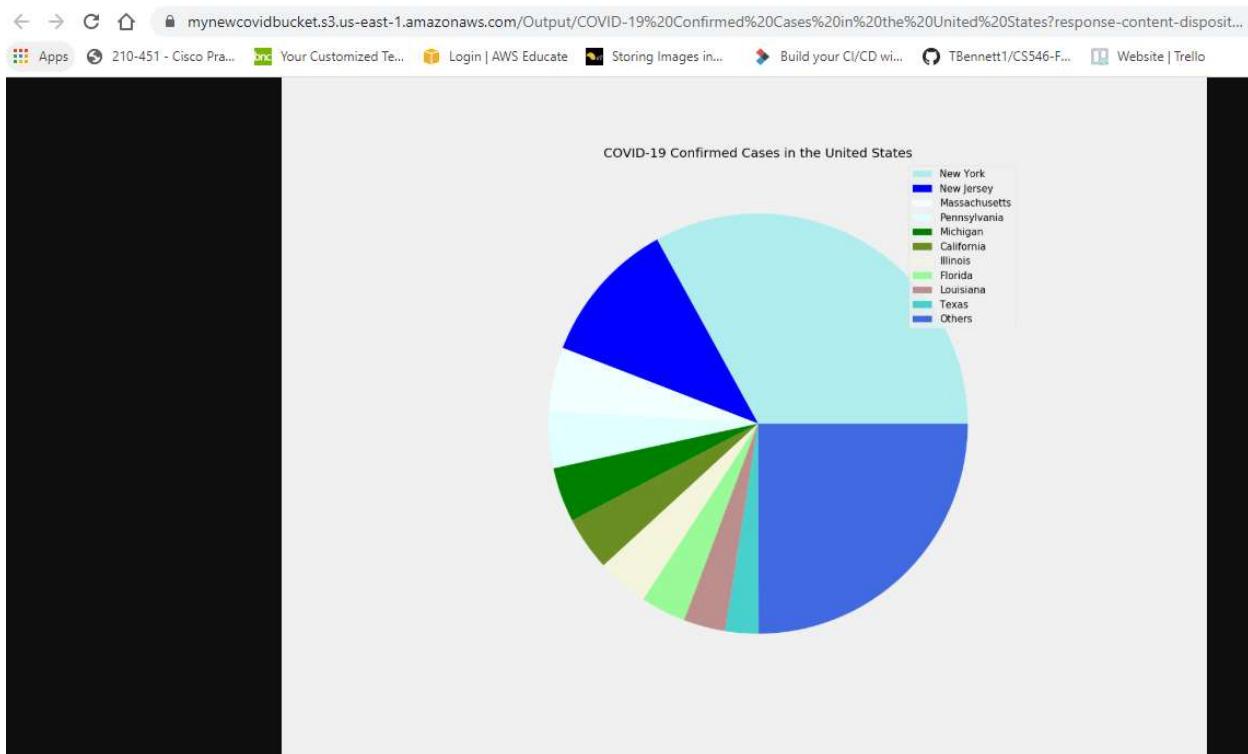
<input type="checkbox"/>	Polynomial_Regression	Apr 22, 2020 4:47:53 PM GMT-0400	24.8 KB
<input type="checkbox"/>	Recoveries over time	Apr 22, 2020 4:47:56 PM GMT-0400	52.4 KB
<input type="checkbox"/>	Recovery_Rate	Apr 22, 2020 4:48:16 PM GMT-0400	64.6 KB
<input type="checkbox"/>	SVM	Apr 22, 2020 4:47:52 PM GMT-0400	23.7 KB
<input type="checkbox"/>	SVM Predictions	Apr 22, 2020 4:48:14 PM GMT-0400	68.0 KB
<input type="checkbox"/>	cases over time	Apr 22, 2020 4:47:56 PM GMT-0400	49.7 KB

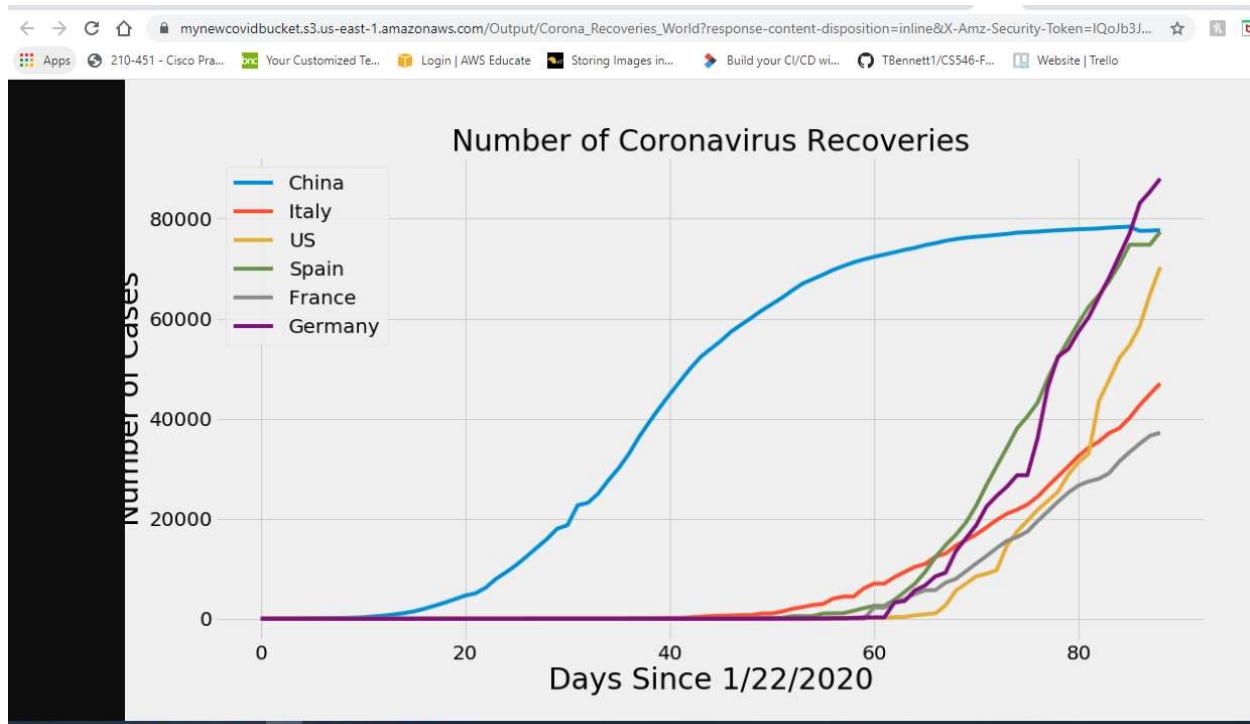
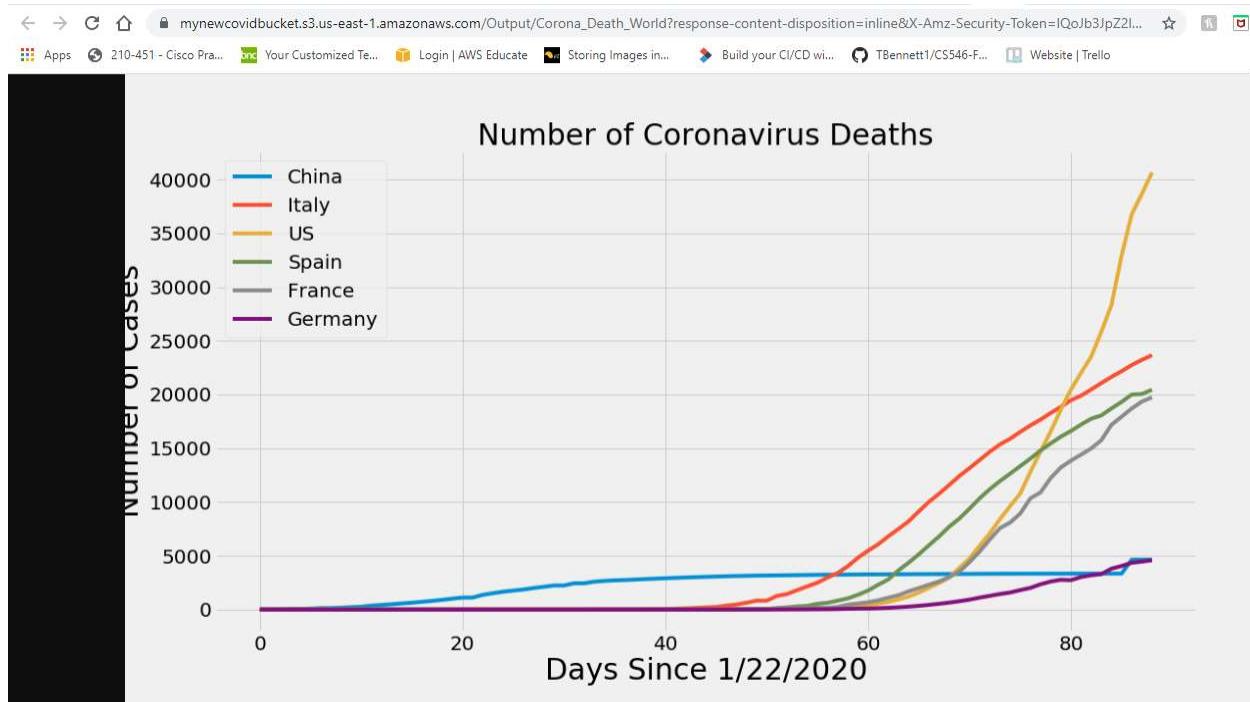
Viewing 1 to 23

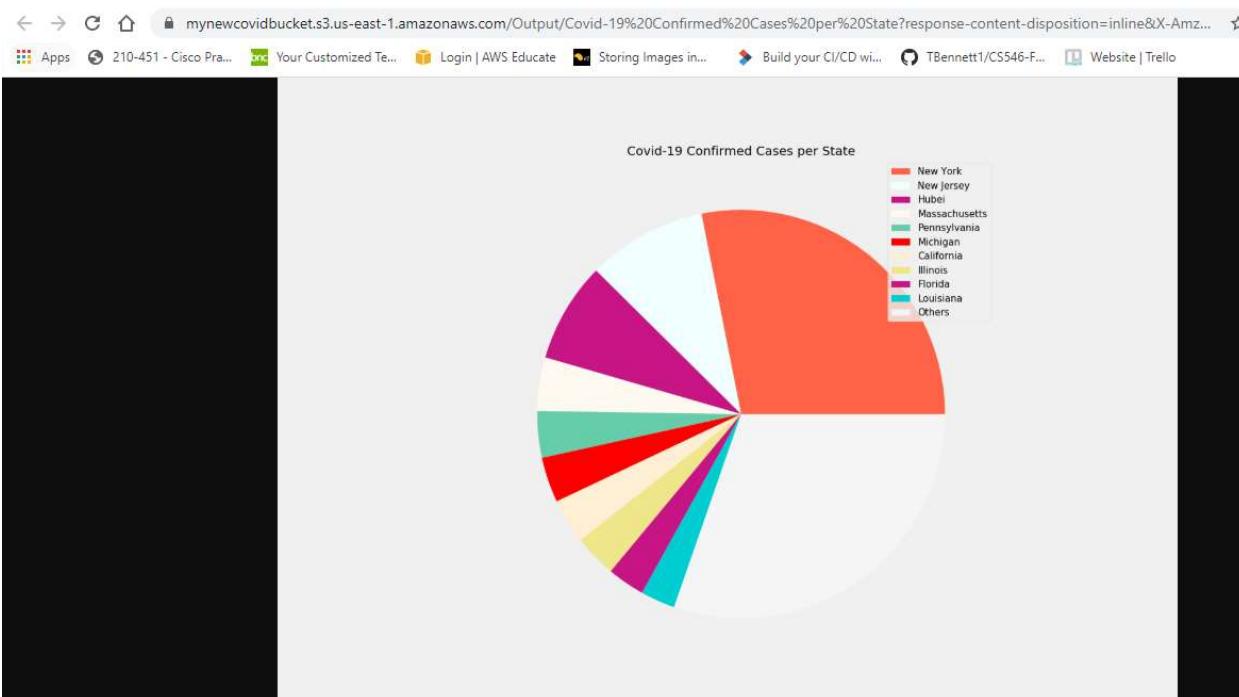
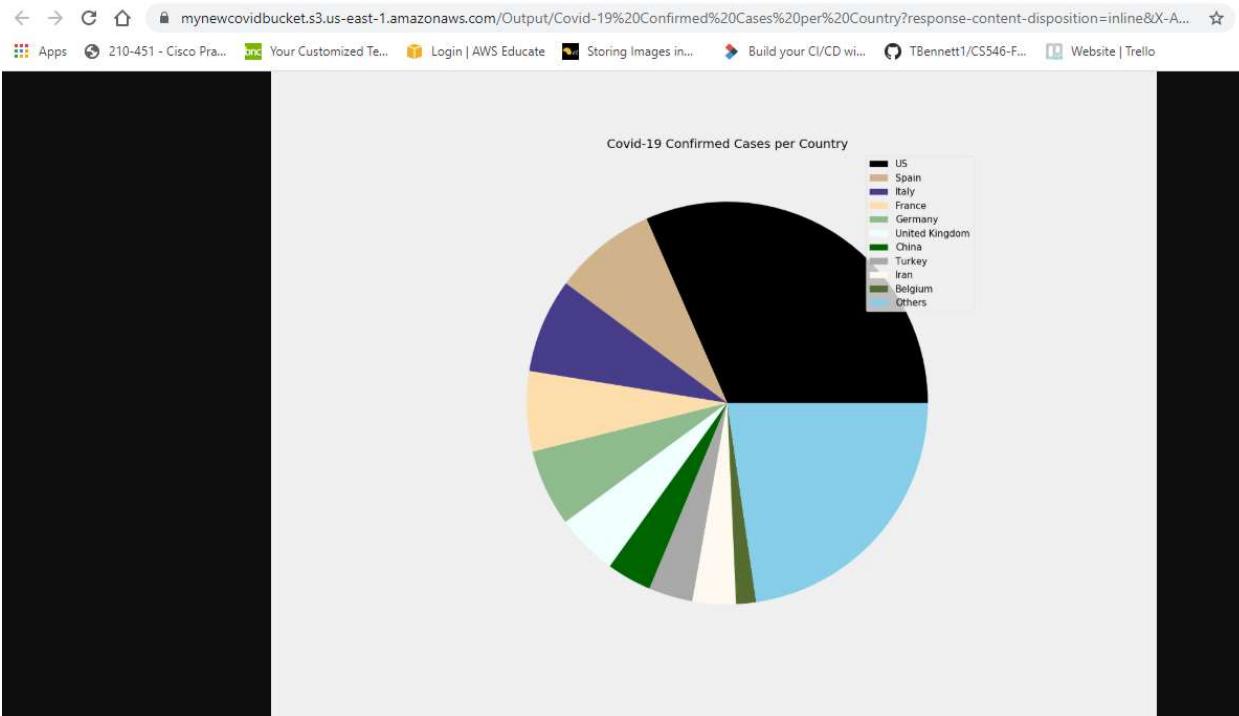
The following graphs are generated:

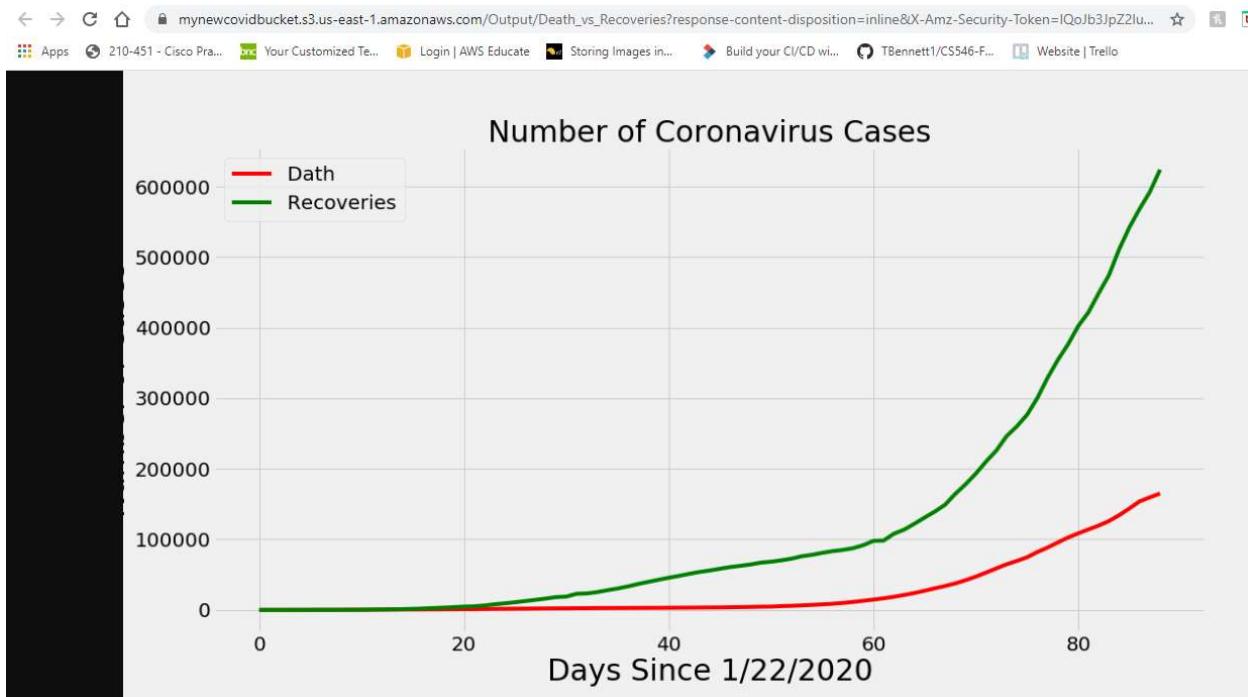
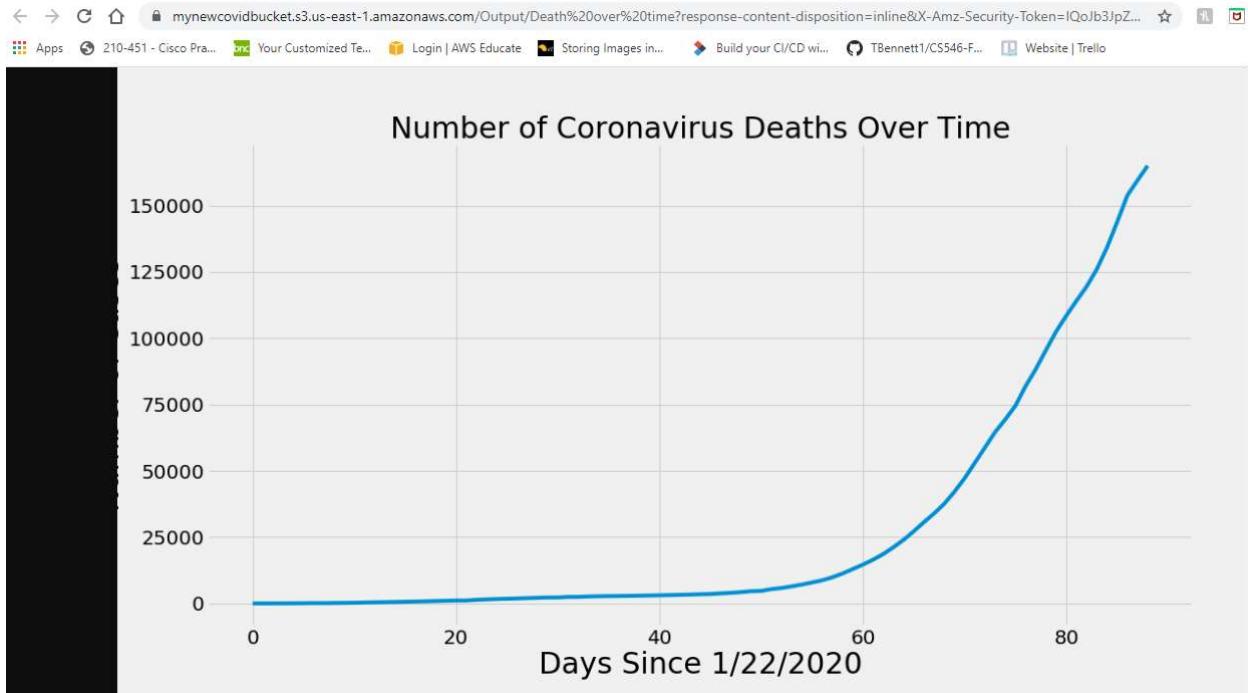


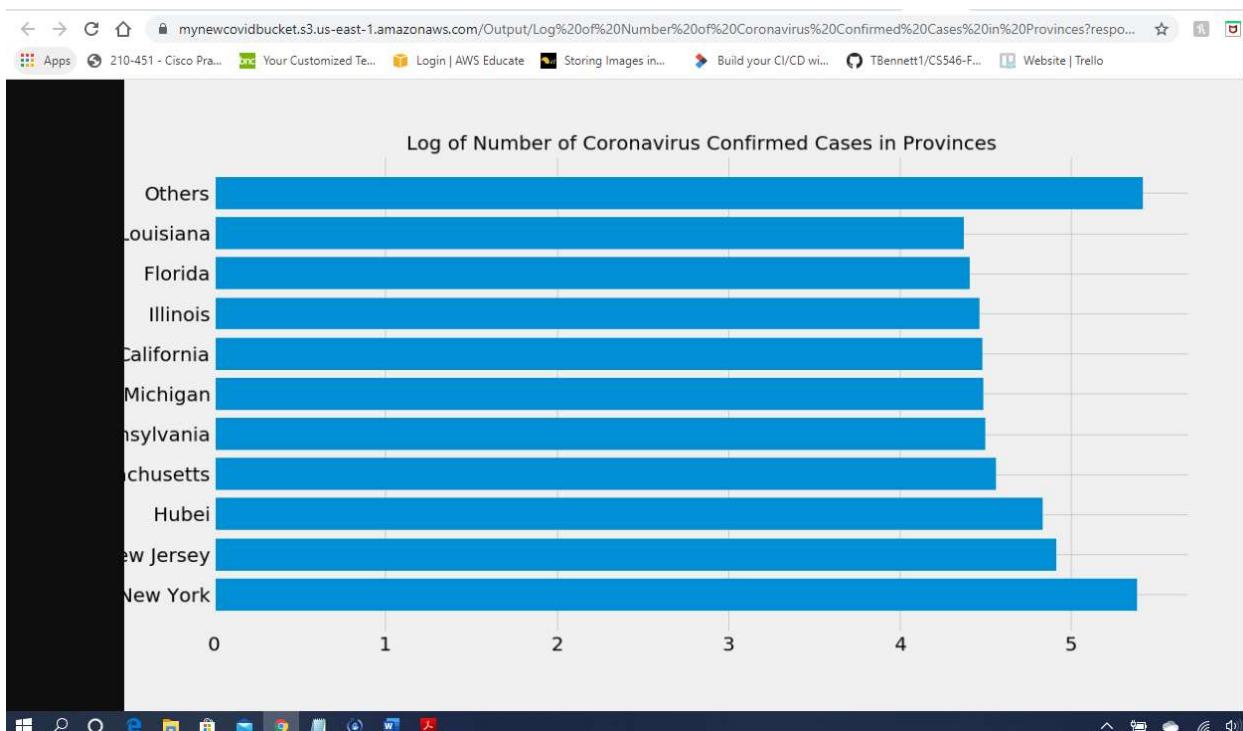
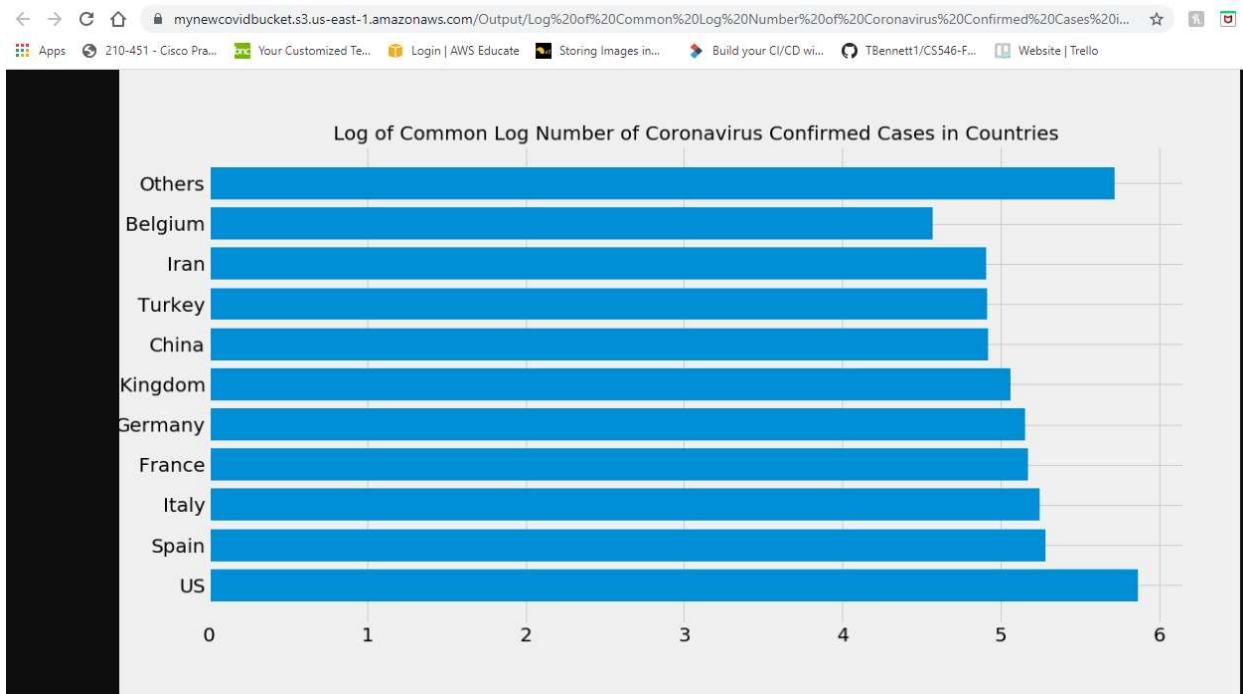


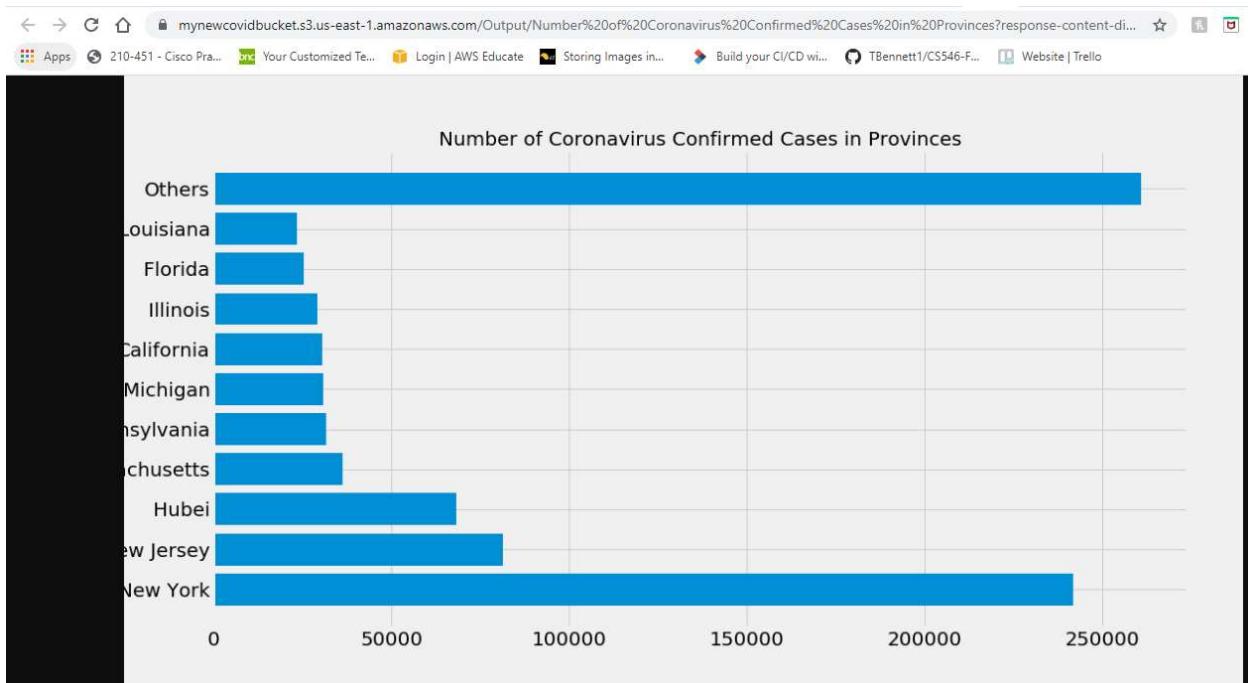
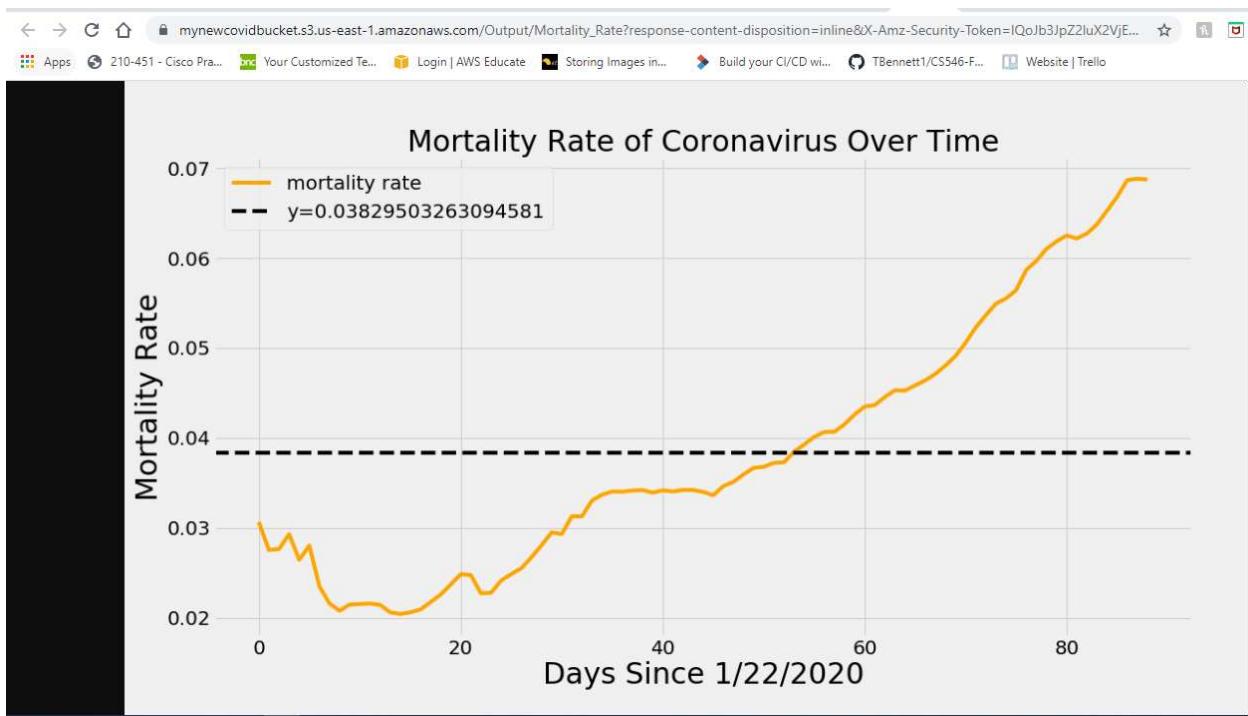


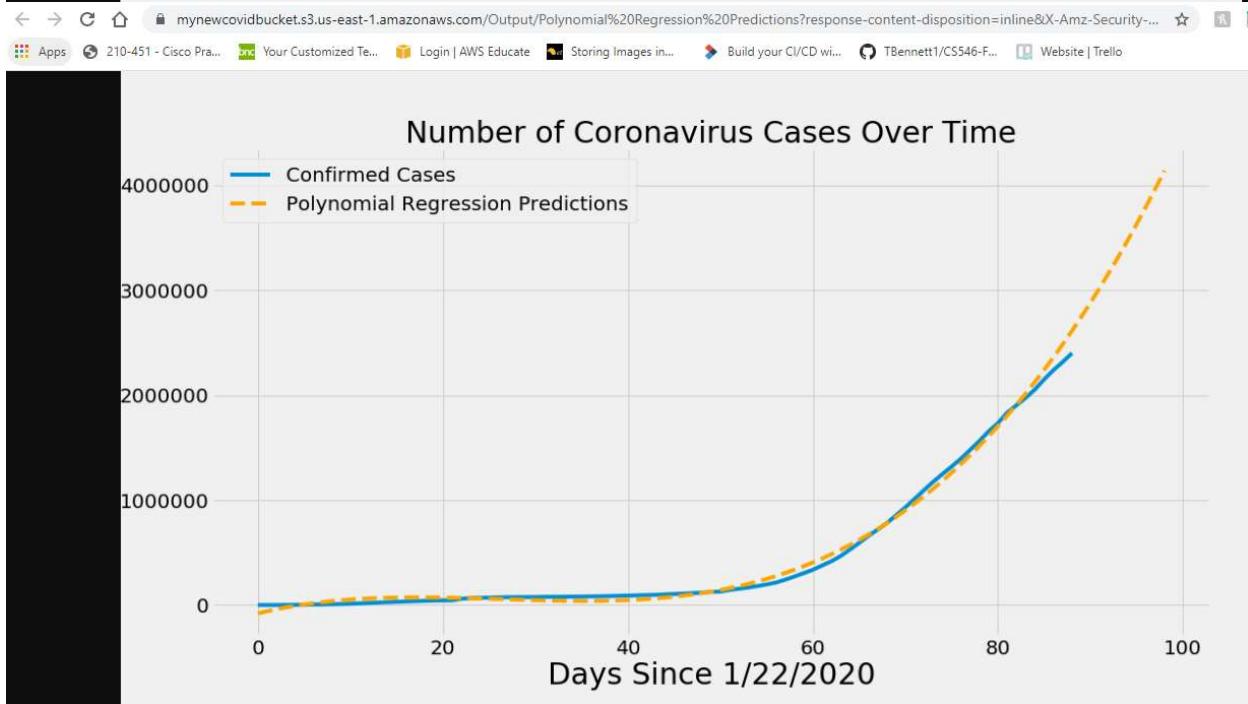
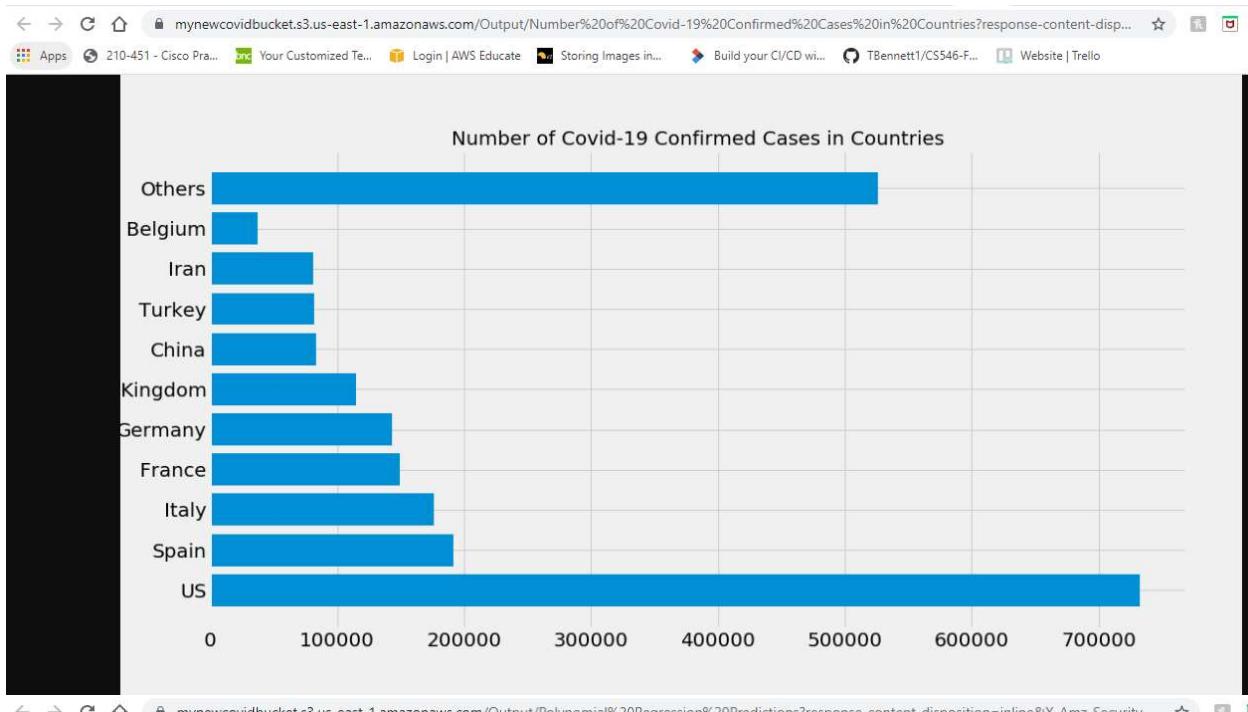


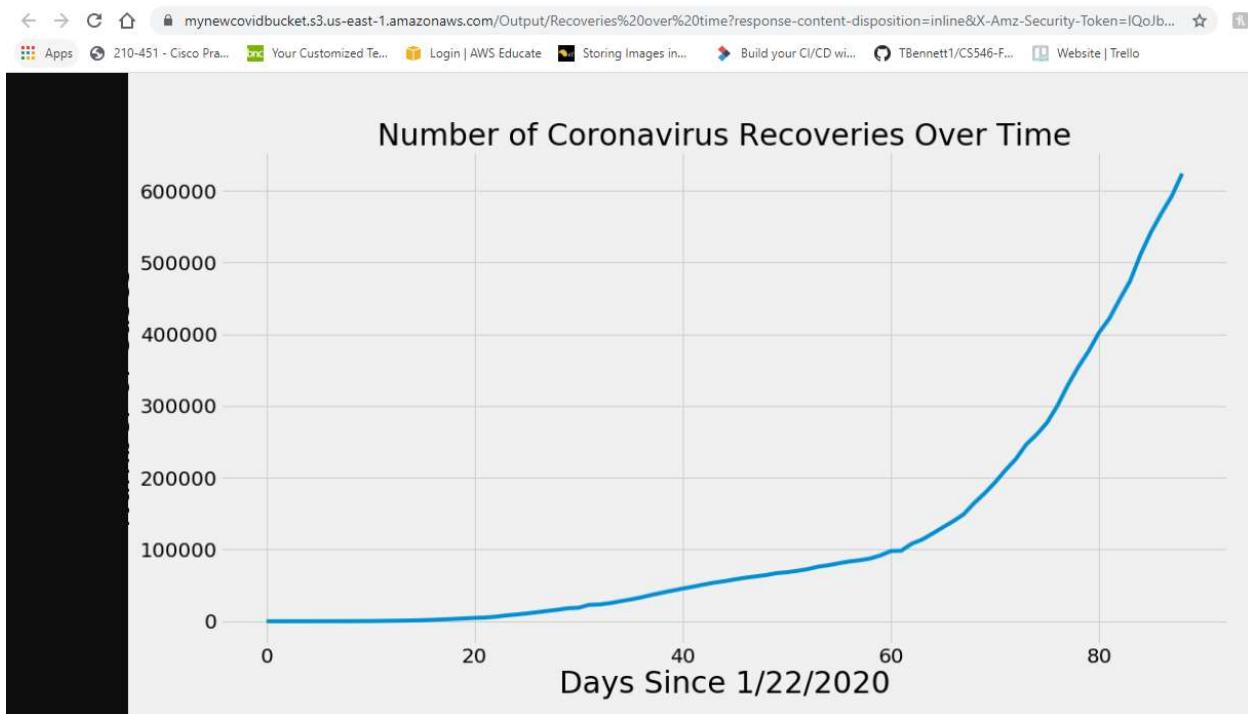
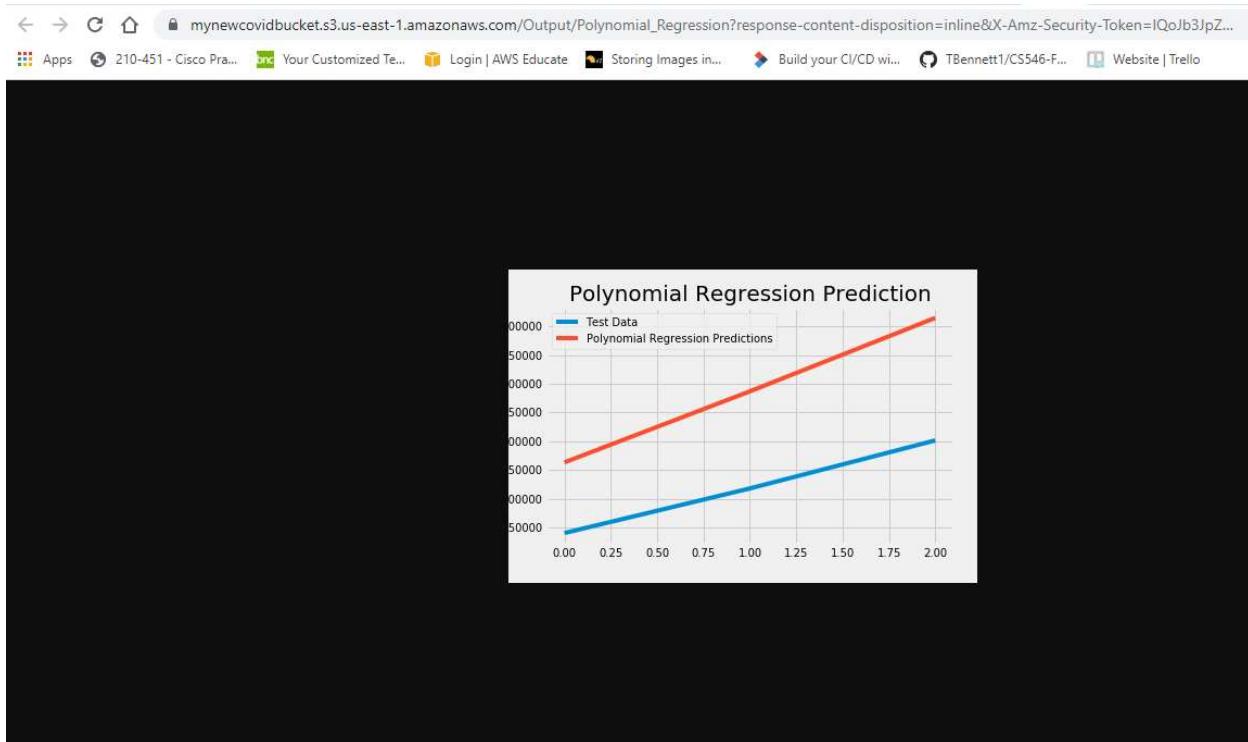


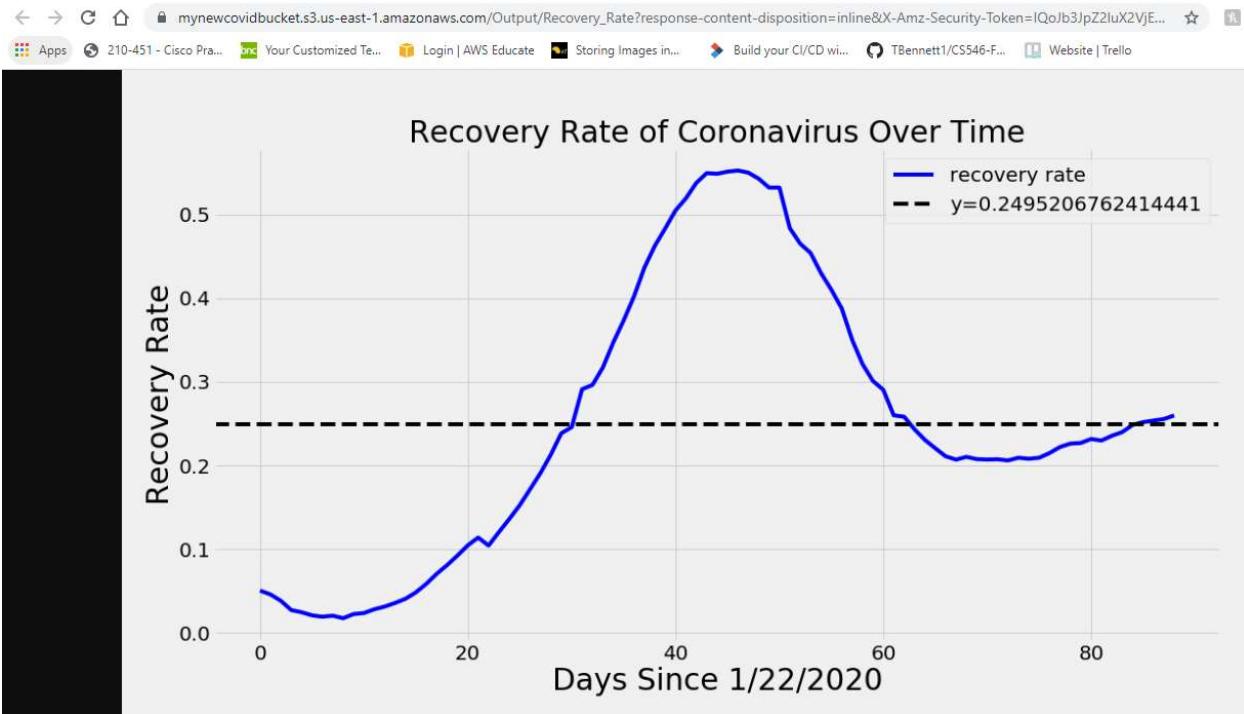


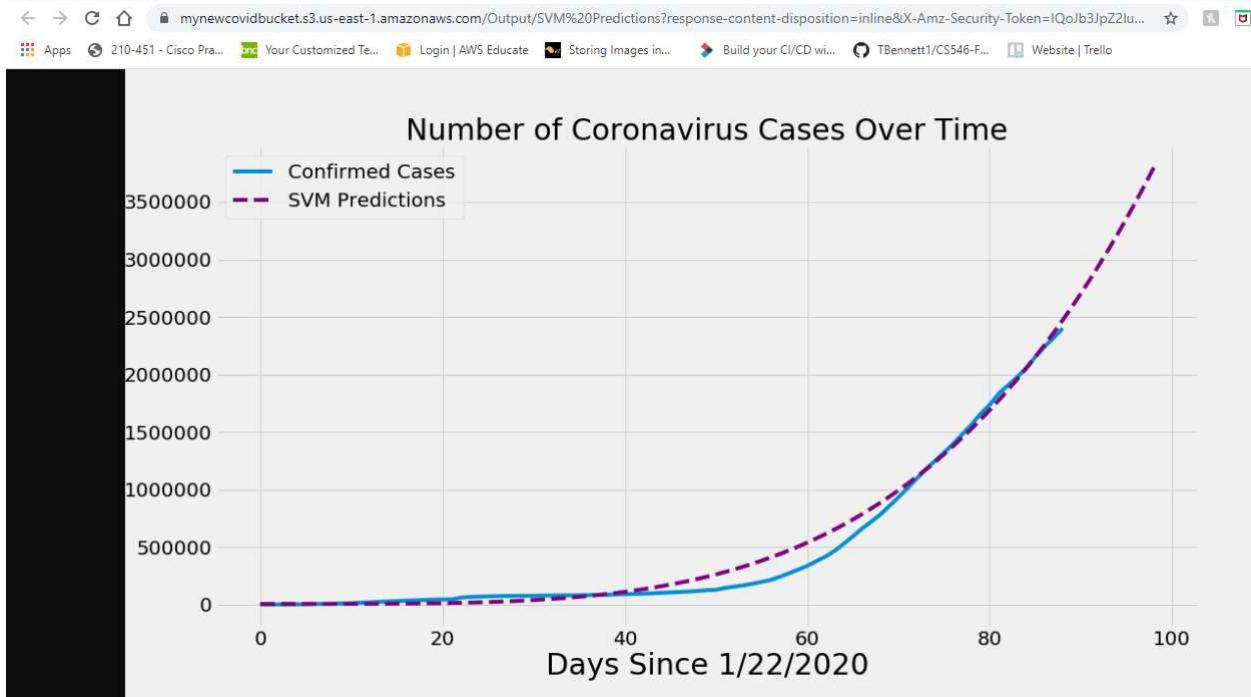
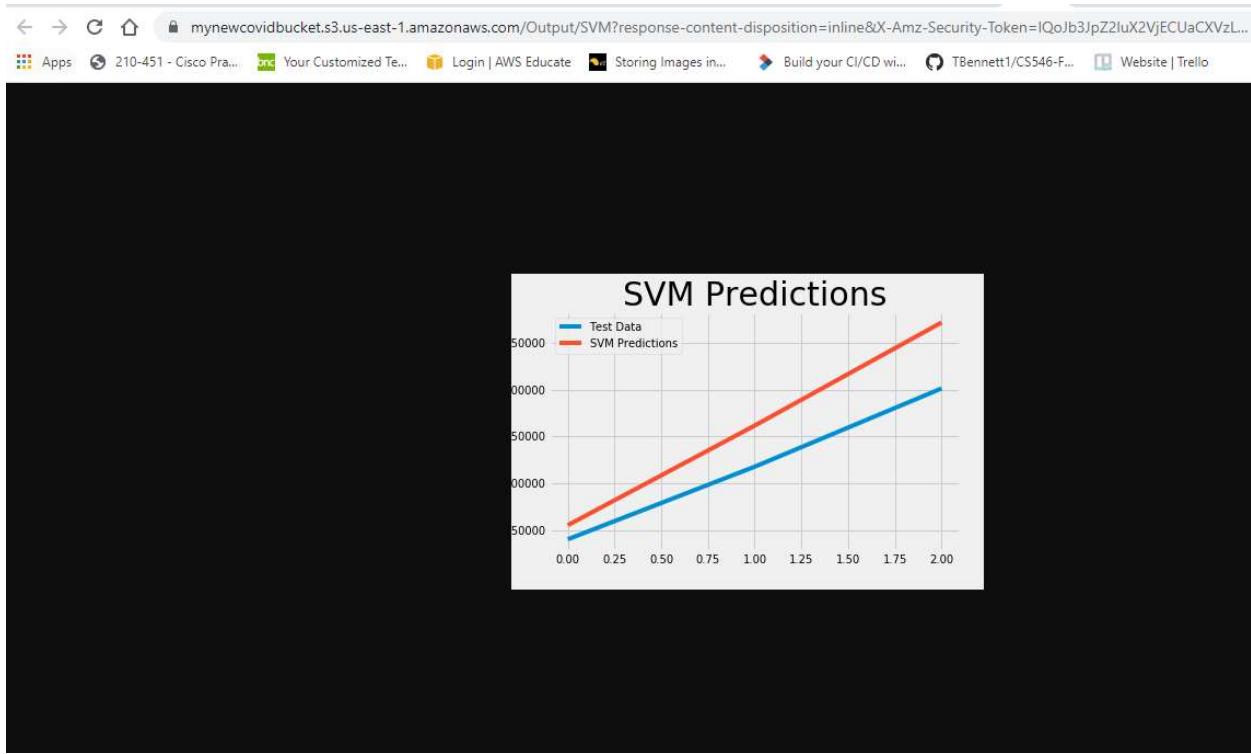


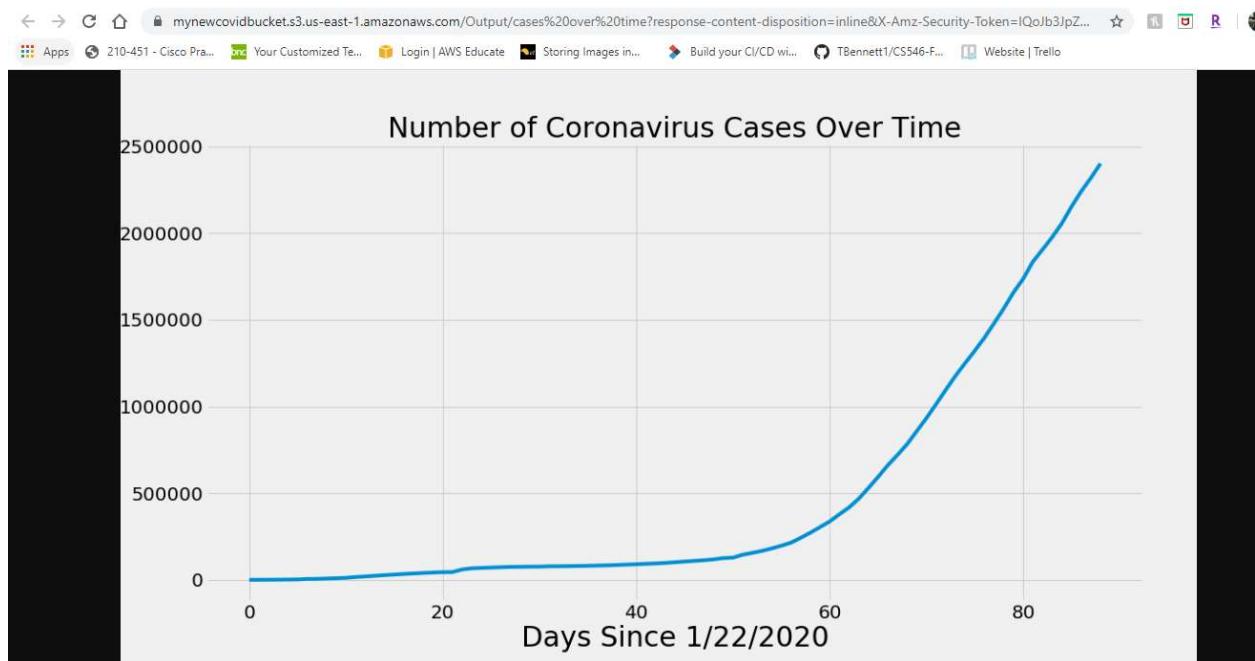
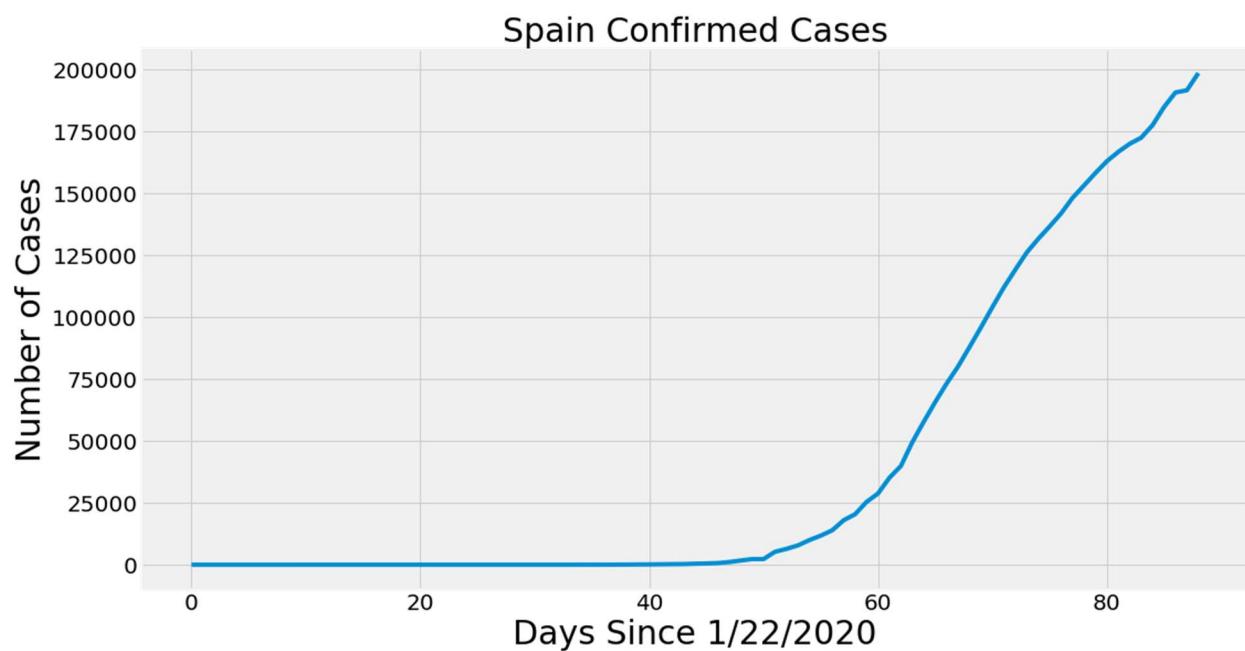


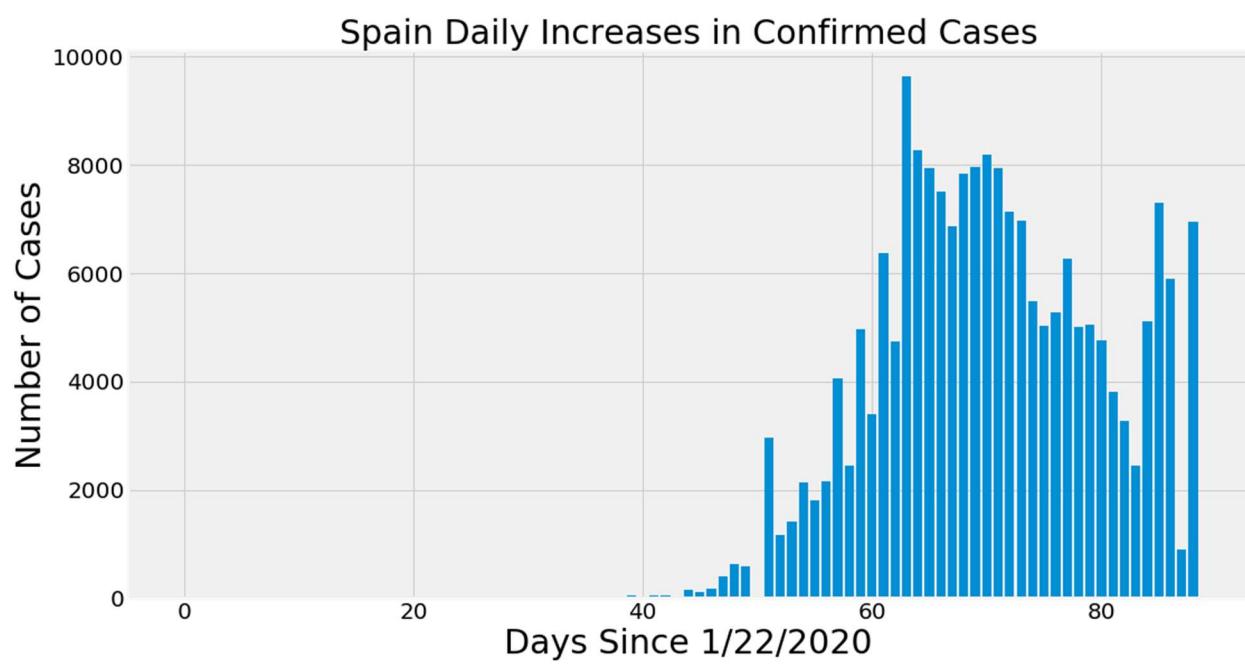
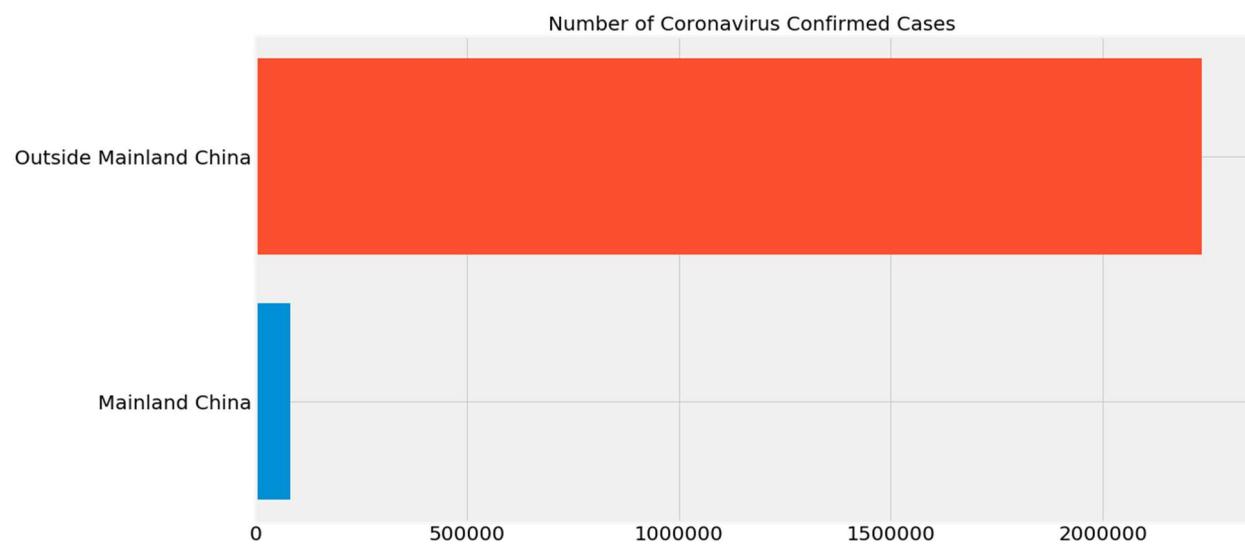


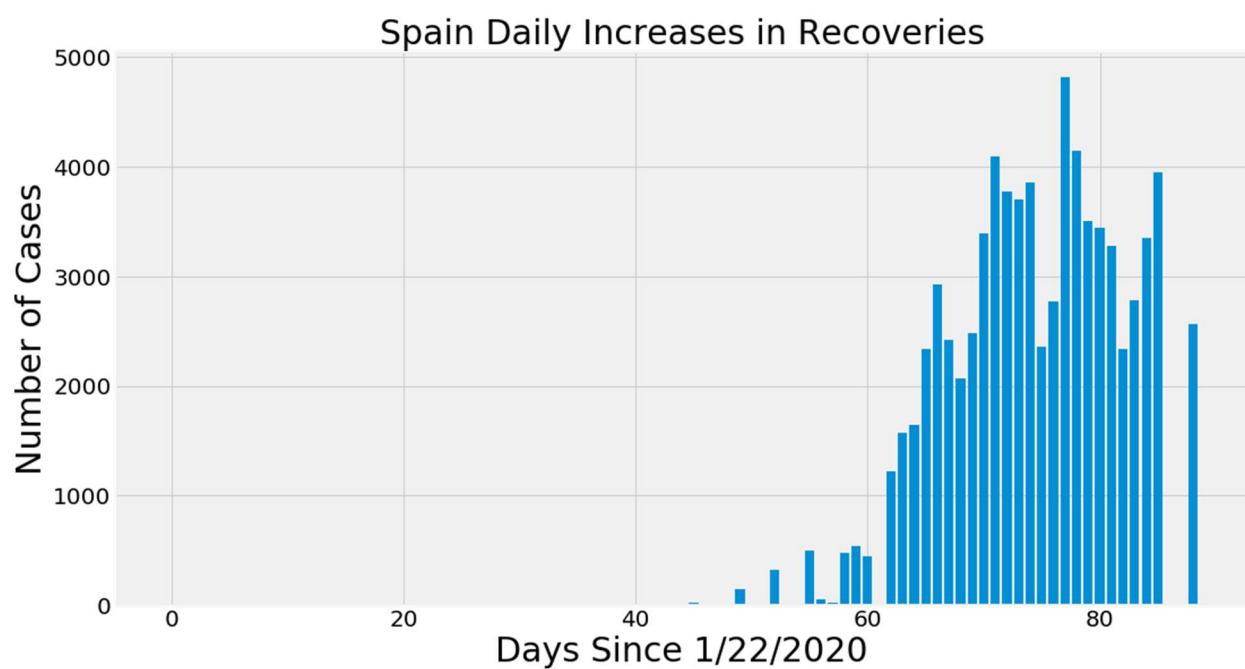
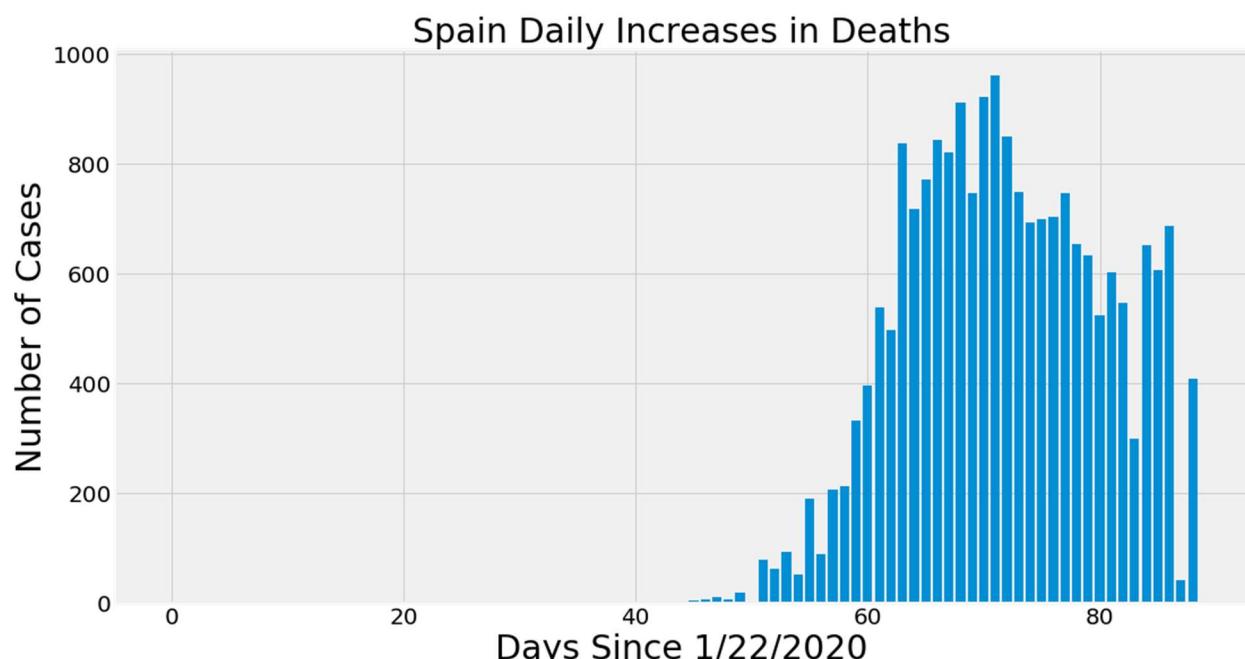


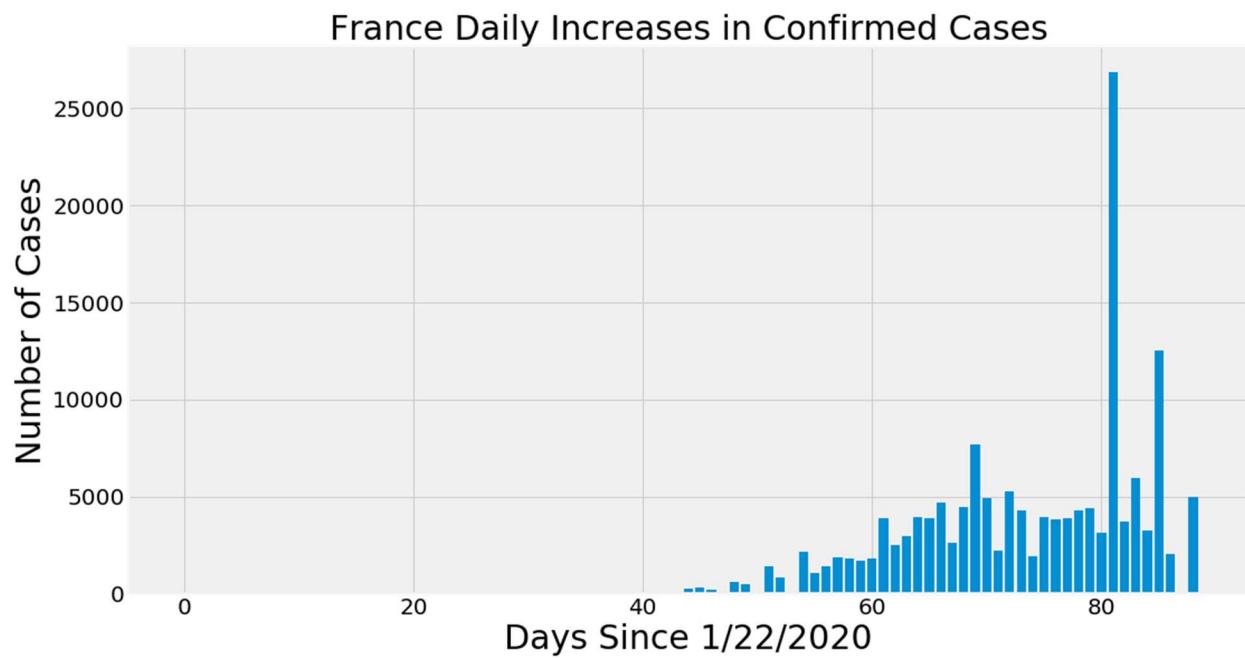
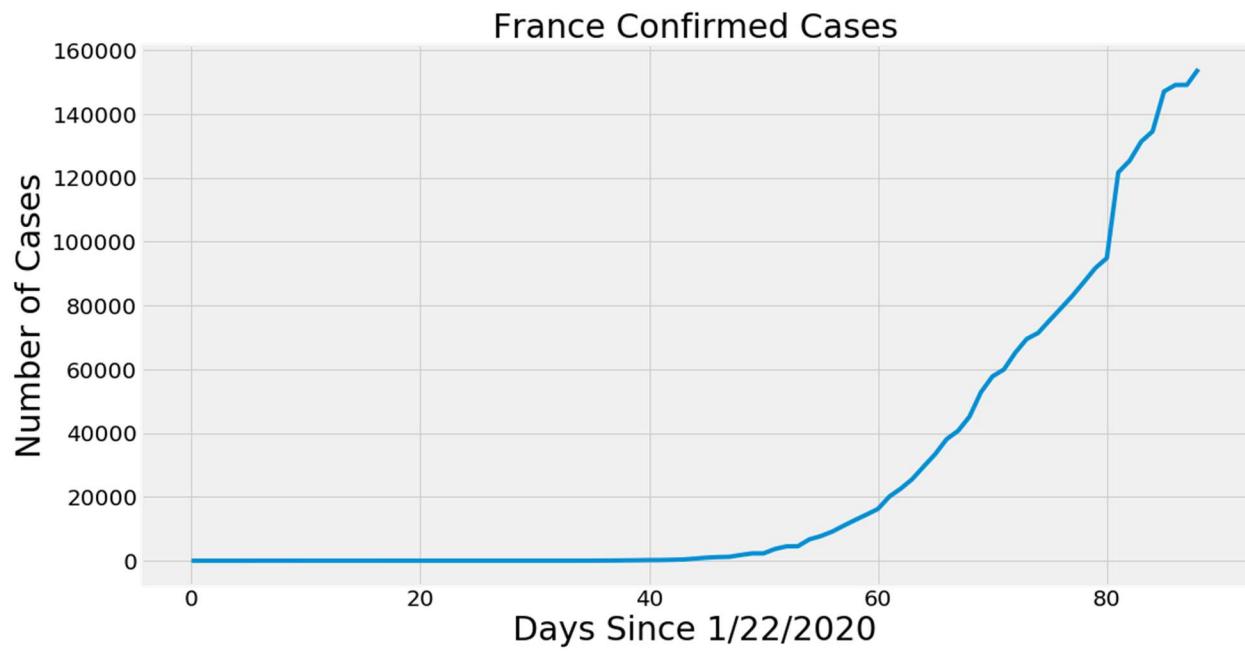




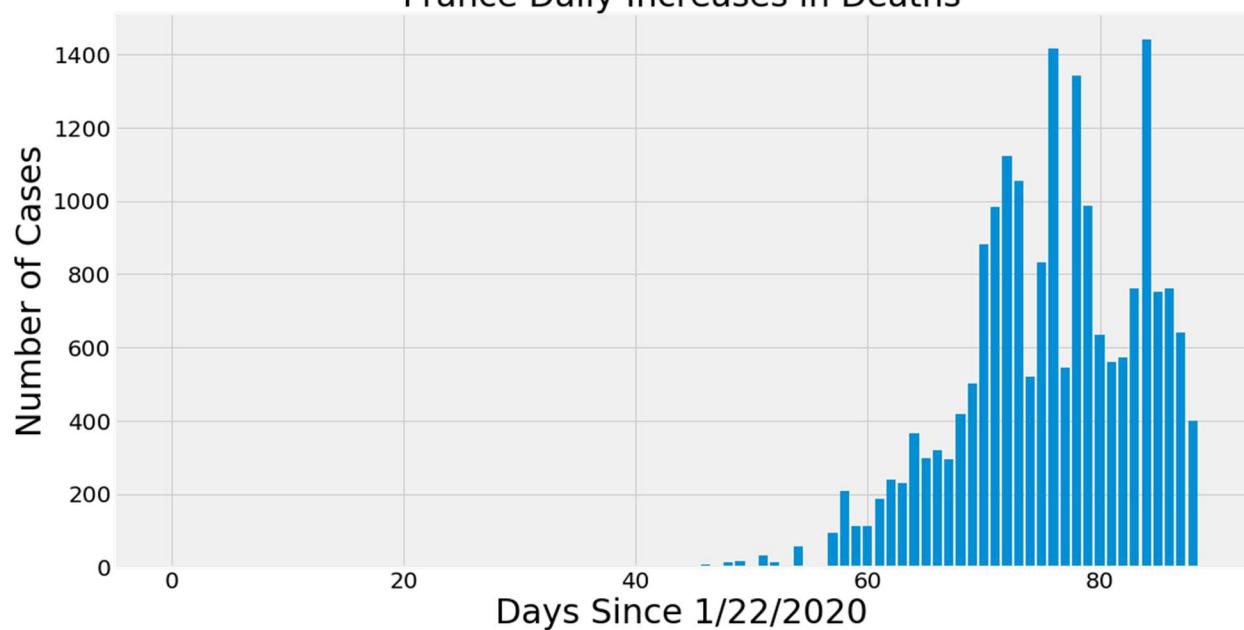




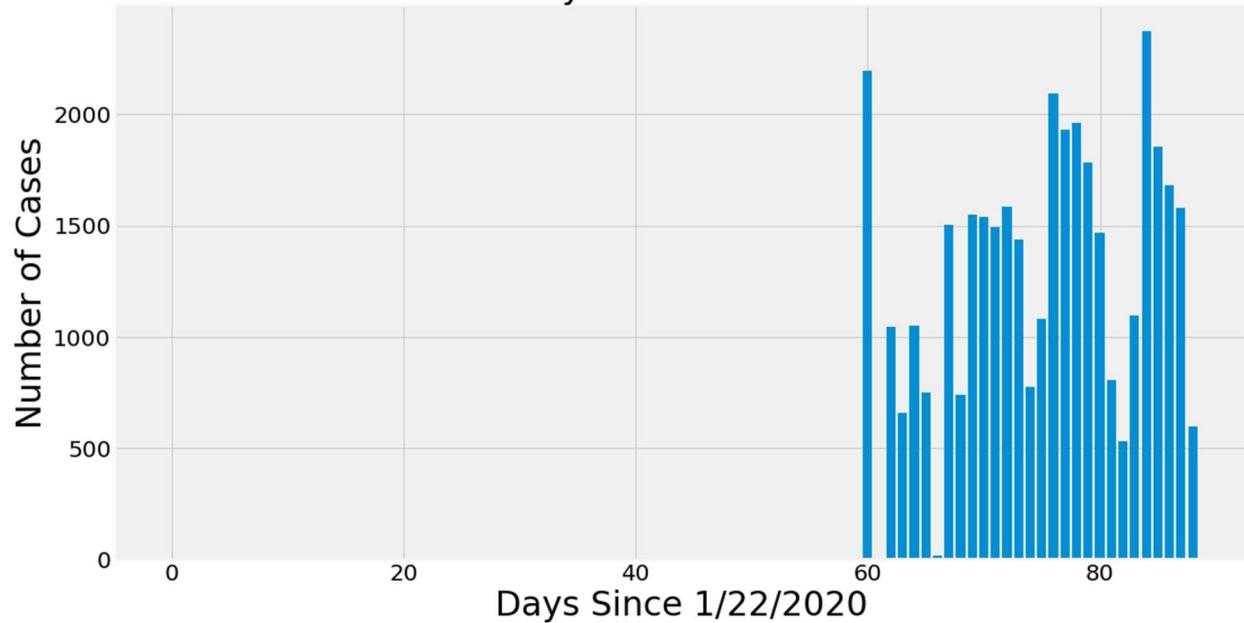


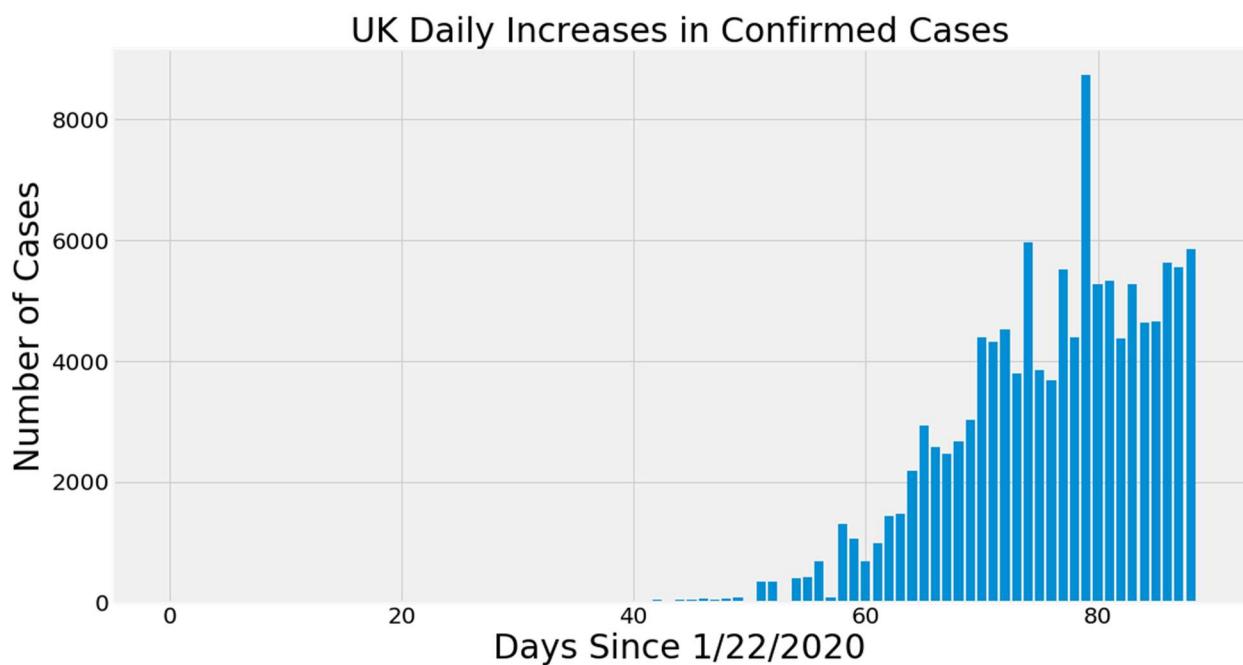
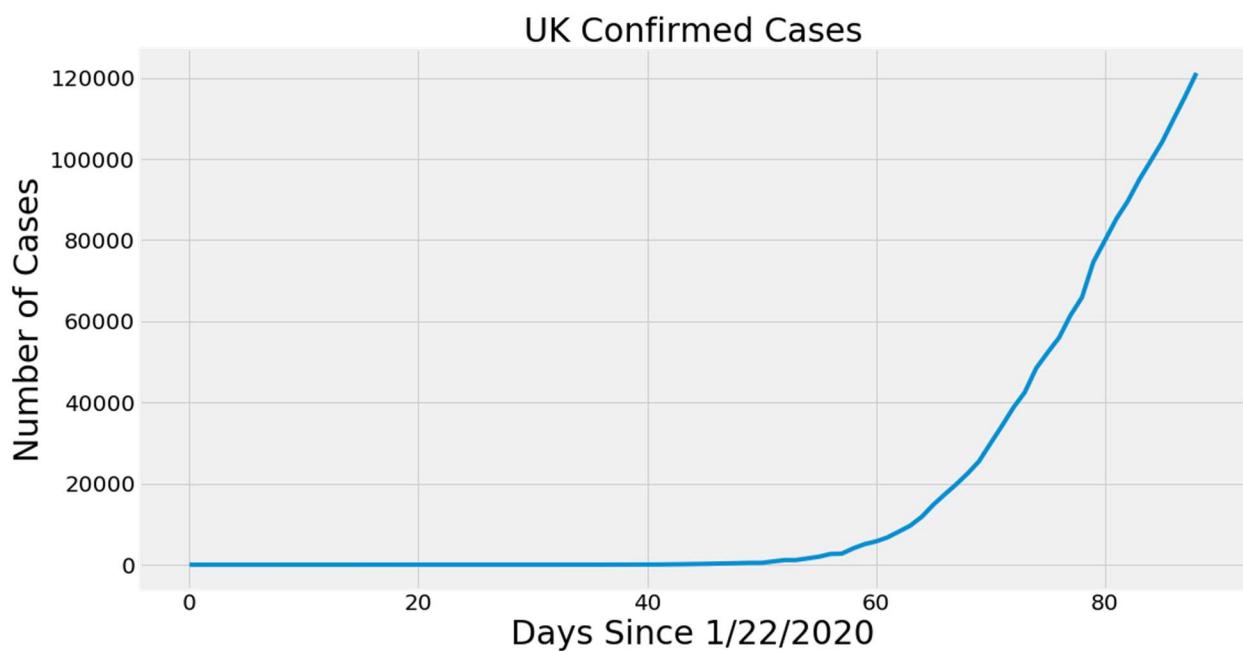


France Daily Increases in Deaths

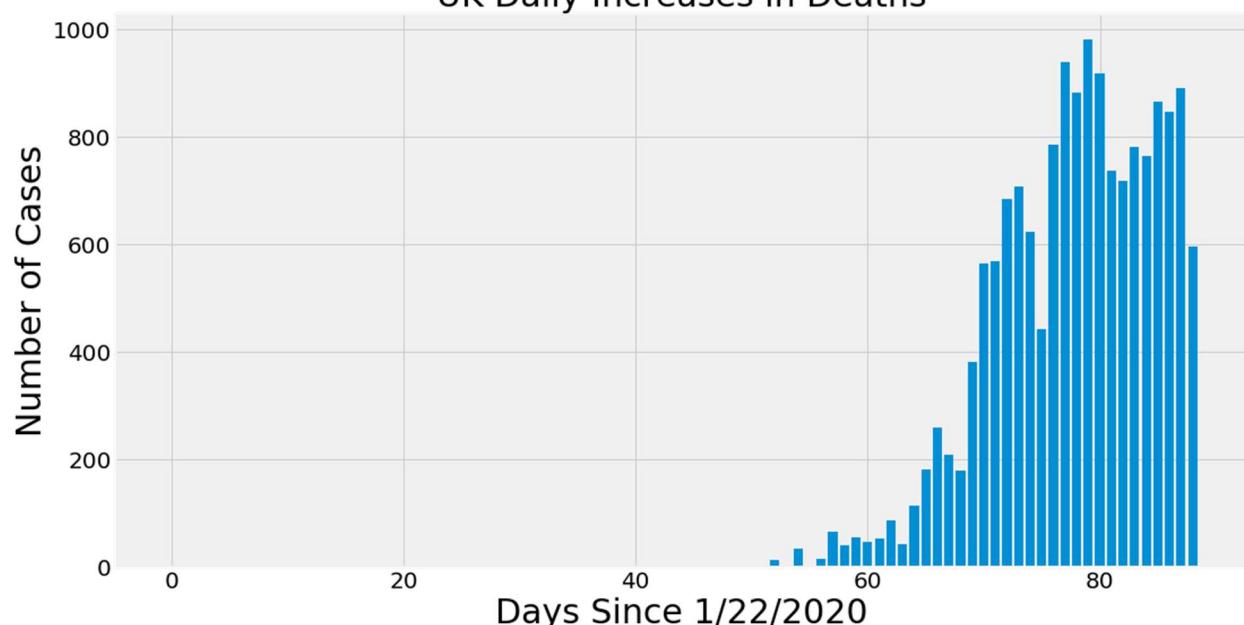


France Daily Increases in Recoveries

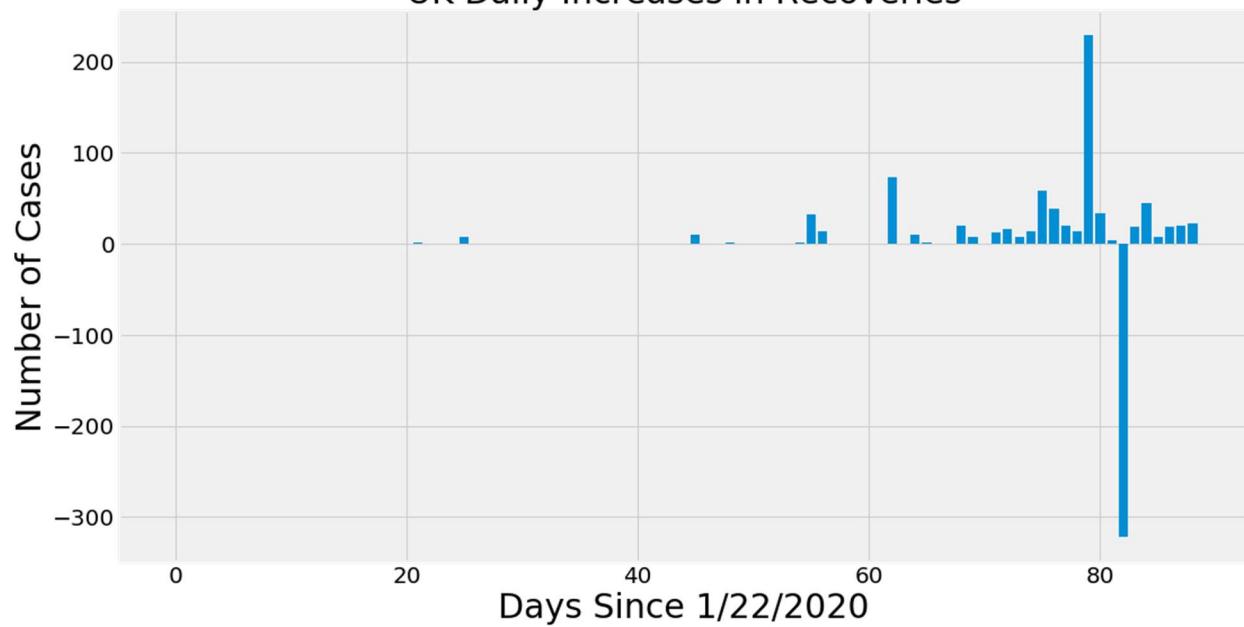




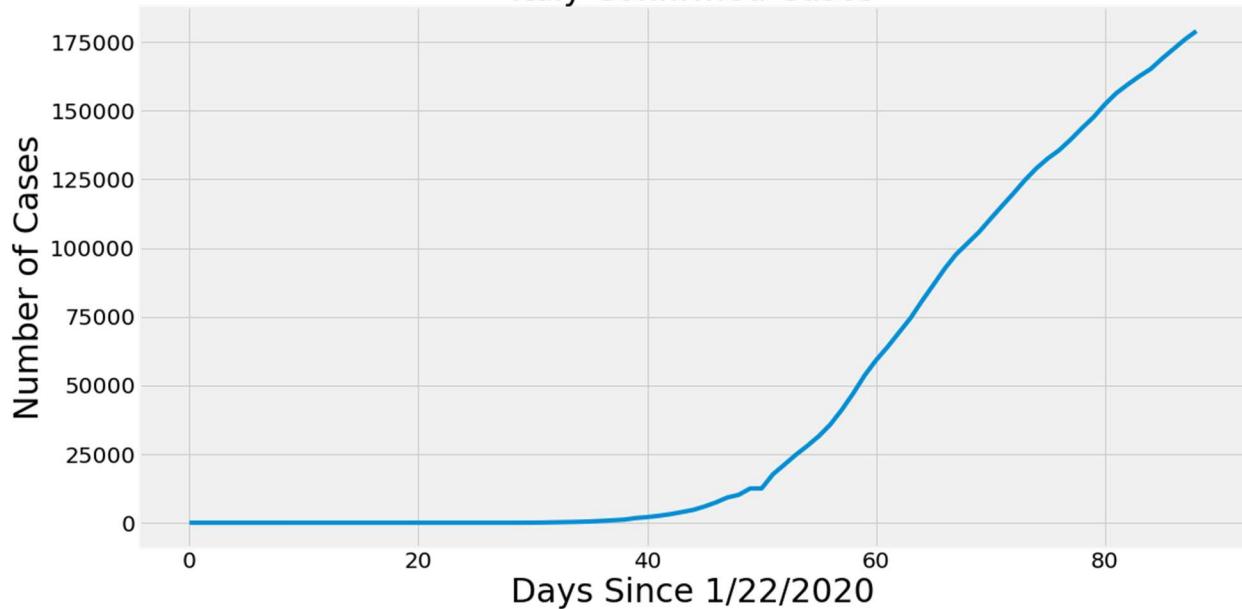
UK Daily Increases in Deaths



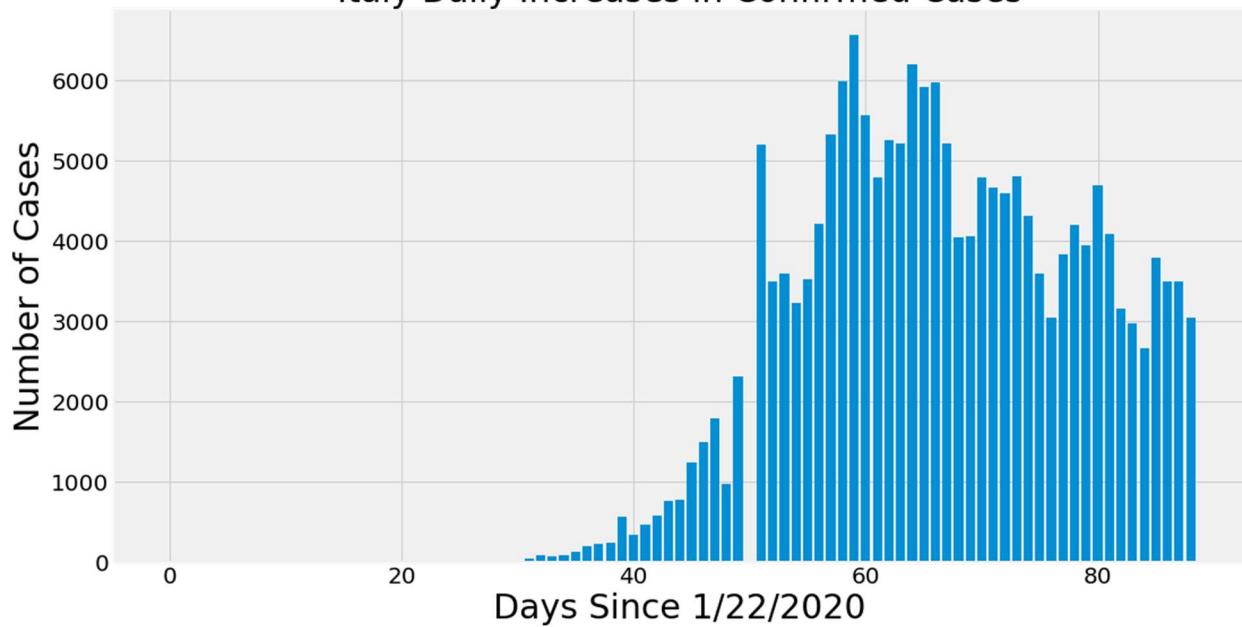
UK Daily Increases in Recoveries



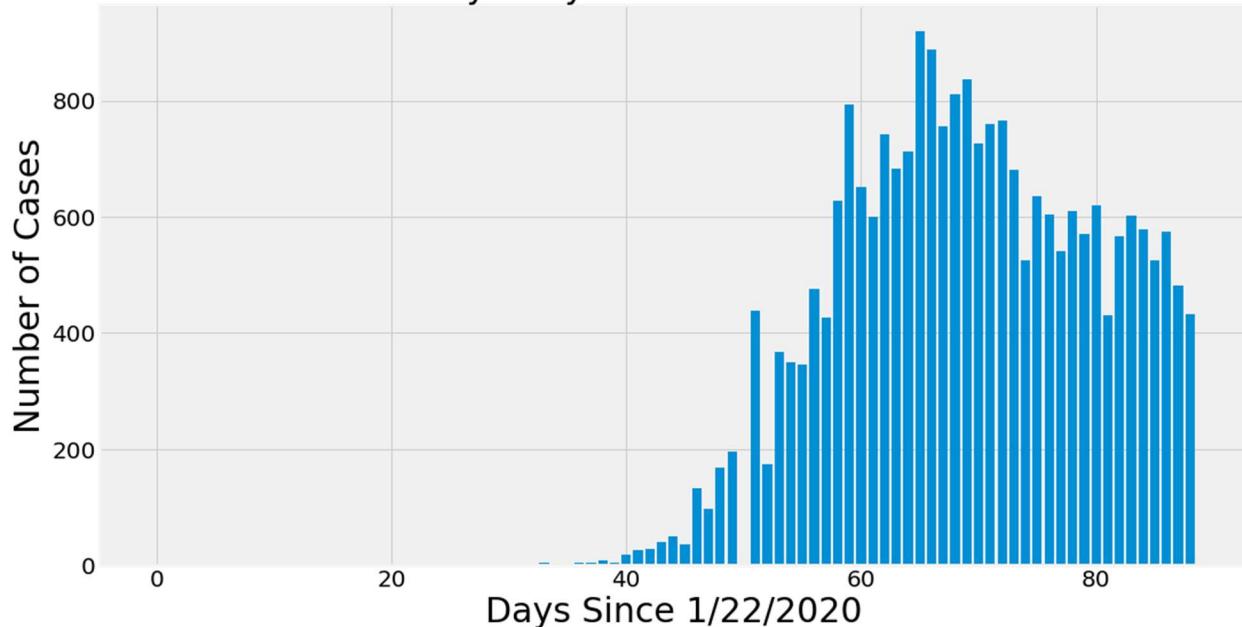
Italy Confirmed Cases



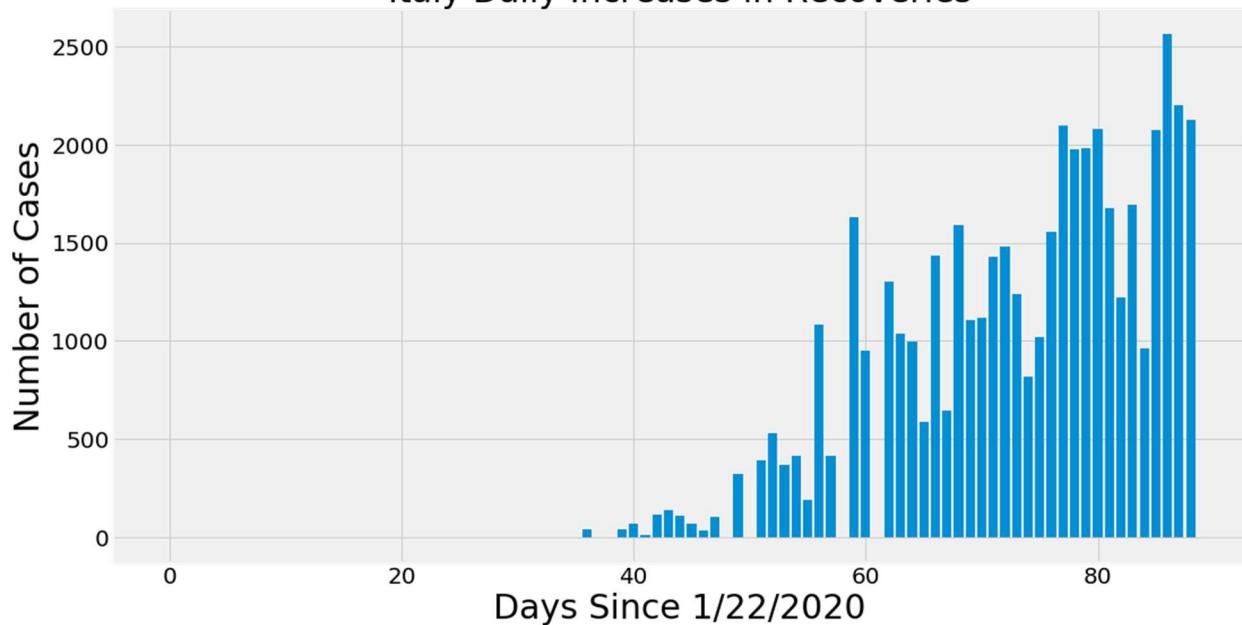
Italy Daily Increases in Confirmed Cases



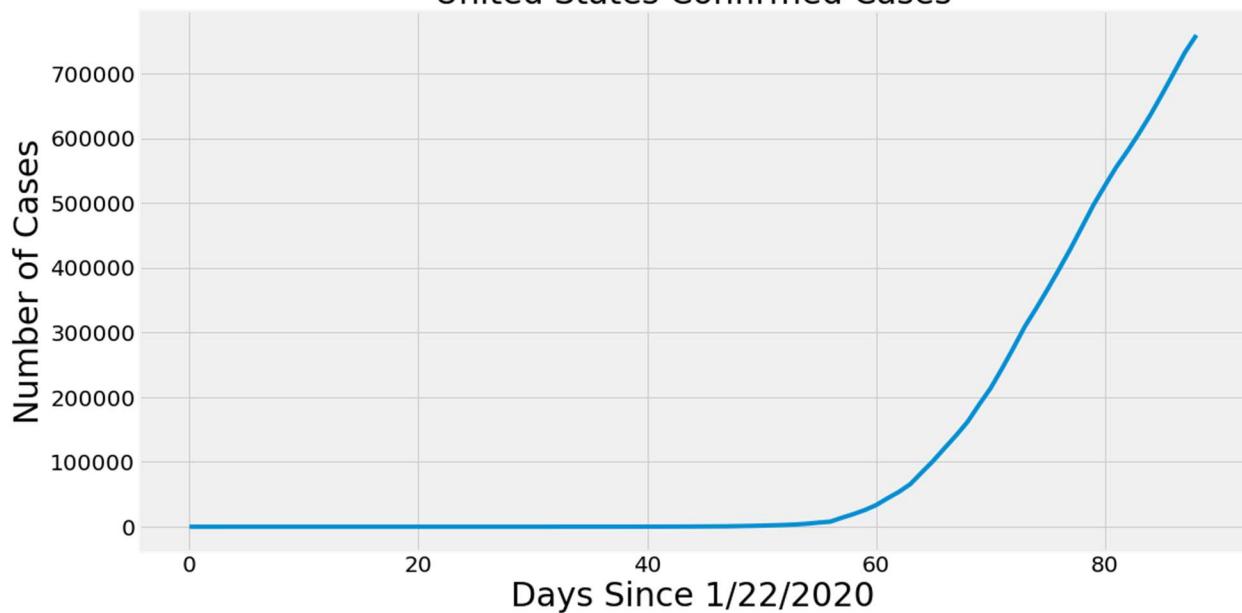
Italy Daily Increases in Deaths



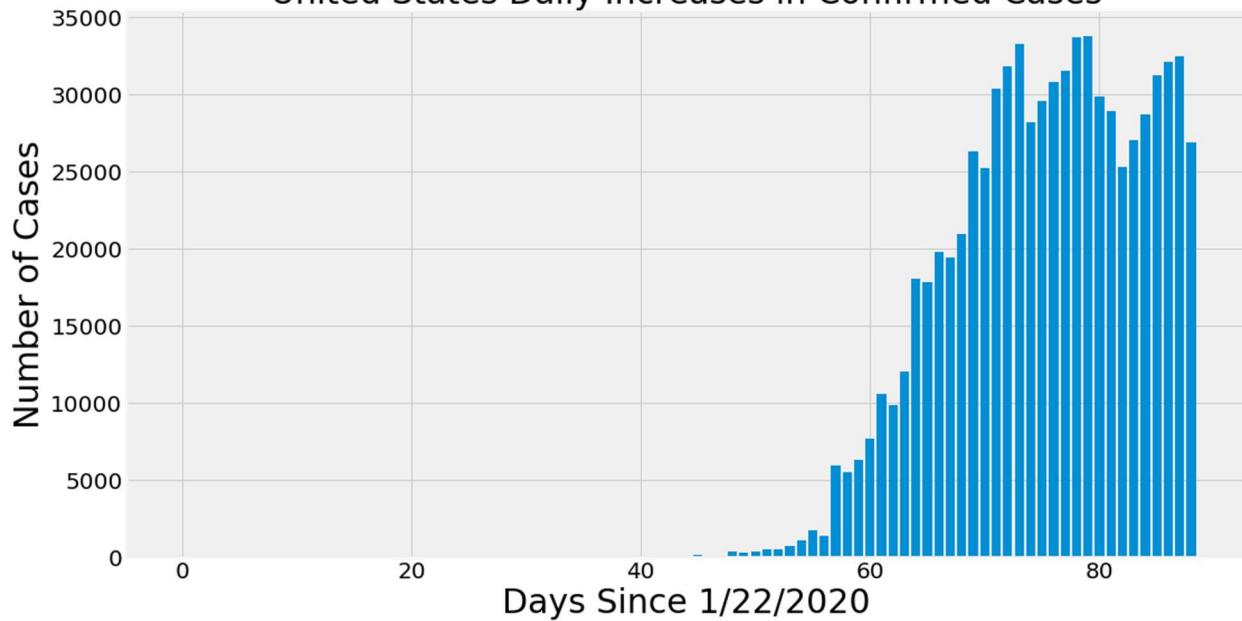
Italy Daily Increases in Recoveries



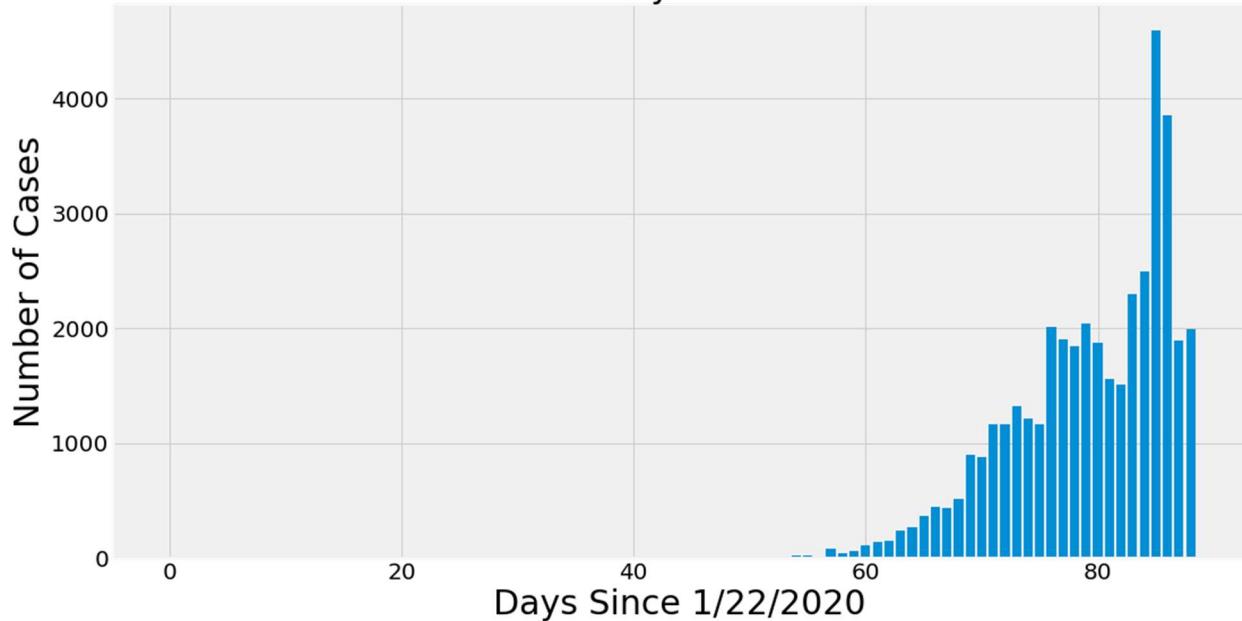
United States Confirmed Cases



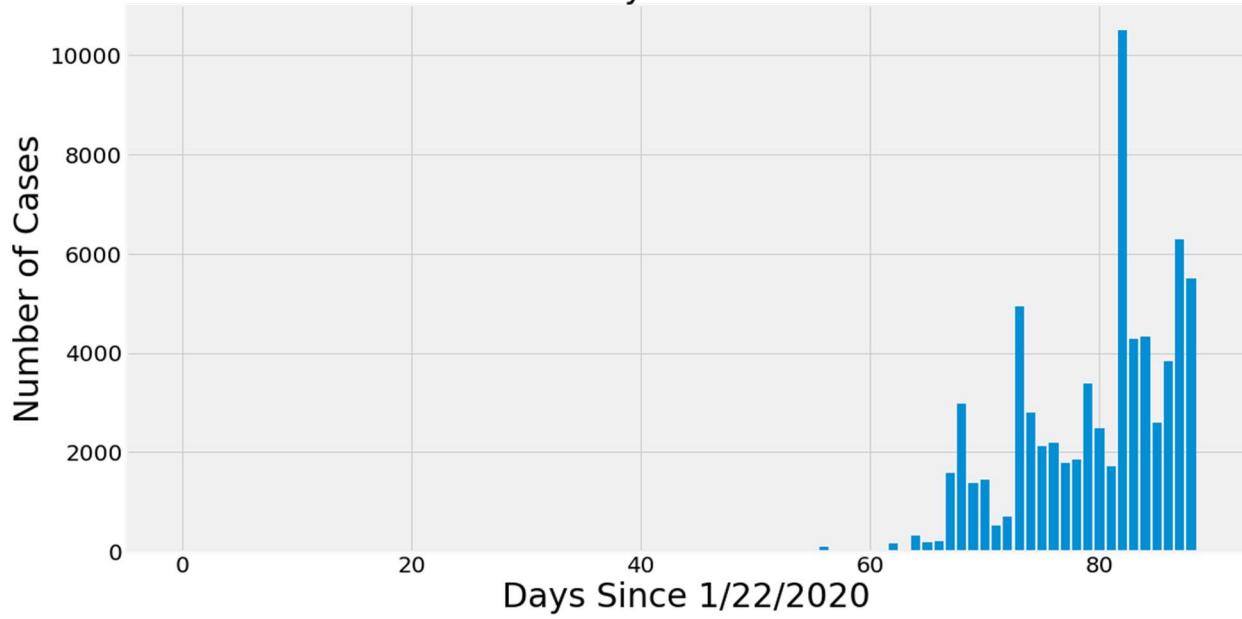
United States Daily Increases in Confirmed Cases



United States Daily Increases in Deaths



United States Daily Increases in Recoveries



Stop the notebook instance and delete your S3 bucket:

The screenshot shows the AWS SageMaker console with the URL console.aws.amazon.com/sagemaker/home?region=us-east-1#/notebook-instances. The left sidebar has sections for Amazon SageMaker Studio, Dashboard, Search, Ground Truth (Labeling jobs, Labeling datasets, Labeling workforces), Notebook (Notebook instances, Lifecycle configurations, Git repositories), Training (Algorithms, Training jobs, Hyperparameter tuning jobs), and ML Models. The main content area shows a 'Notebook instances' table with one row:

Name	Instance	Creation time	Status	Actions
mycovidnotebook	ml.t2.medium	Apr 22, 2020 19:22 UTC	Stopping	-

A modal window titled 'Amazon Elastic Inference' is open, stating: 'Amazon Elastic Inference adds GPU acceleration to any Amazon SageMaker or EC2 instance for faster inference at much lower cost, with up to 75% savings. Find out if Elastic Inference is right for you.' It includes a 'Learn more' button.

Conclusions:

- Machine Learning Models can be deployed to Amazon SageMaker very quickly.
- It is pretty simple to upload massive data on S3 buckets or GitHub and then use SageMaker by running a not so complex script and massive analysis can be made.
- It can be a complex task if the traditional Machine Learning approach is followed, moreover, it can be very expensive too. Making multiple tools work together can be very challenging, all these hurdles are addressed smoothly by Amazon SageMaker, so that the developer can just focus on the script rather than the tool complexity.
- The above study shows how Amazon SageMaker can be effectively utilized for machine learning. All the statistics, whether it is in terms of country, province, state, city or the entire world, can be easily analyzed and it can prove to be very useful for scientists and analysts.