

Primera Entrega Proyecto COVID 19 - COLOMBIA

Daniel Julian Siachoque Peralta
Escuela de Ciencias Exactas e Ingeniería
Universidad Sergio Arboleda
Bogotá, Colombia
daniel.siachoque01@correo.usa.edu.co

Juan Guillermo Torres Delgado
Escuela de Ciencias Exactas e Ingeniería
Universidad Sergio Arboleda
Bogotá, Colombia
juan.torres01@correo.usa.edu.co

Carlos Antonio Plaza Amado
Escuela de Ciencias Exactas e Ingeniería
Universidad Sergio Arboleda
Bogotá, Colombia
carlos.plaza01@correo.usa.edu.co

Resumen

Durante el avance del proyecto se habló sobre la actual pandemia “Coronavirus” o “covid19”, enfocándonos en los datos actualizados para Colombia. Para esto fue necesario una pagina que nos brindó en tiempo real los datos de la pandemia, así mismo, se utilizó el Python como lenguaje de programación, sql como método de extracción de información a través de una base de datos y Látex como método de escritura para presentar los datos. El objetivo principal del proyecto se basa en plasmar mediante gráficas de barras, tortas y 2 dimensiones las diferentes estadísticas de la enfermedad.

Palabras clave:

Bases de datos sql, Coronavirus, Python.

1. Marco teórico

1.1. Covid 19

Es una enfermedad que se descubre por primera vez en Wuhan-China, la cual se ha esparcido por todas las áreas del mundo, llegando así a identificarse casos en cada uno de los continentes. El primer caso confirmado en Colombia se dio a conocer el 6 de marzo del presente 2020.

Esta enfermedad produce una infección Respiratoria Aguda o lo que conocemos como una gripe, pero variando desde ser grave, moderada o simplemente leve. Algunos de sus síntomas parten de la fiebre, tos seca y cansancio, sin dejar de lado dolores o molestias generales. Cabe recalcar que un 80 % de las personas infectadas por el virus se logran recuperar sin un tratamiento hospitalario, teniendo una mortalidad baja respecto a otras enfermedades.

Algunas recomendaciones generales para evitar el contagio de esta enfermedad parte del lavado constante de las manos y el uso constante de la mascarilla junto con gel antibacterial.

En el momento de este proyecto, se tiene conocimiento de 33 Millones de casos a nivel mundial, de los cuales Estados Unidos, India, Brasil, Rusia y Colombia se encuentran en el top 5 de países con más casos registrados (7 Millones, 6 Millones, 4 Millones, 1 Millón y 800 Mil) respectivamente.

En Colombia, como ya se comentó anteriormente, se tiene conocimiento de más de 800 Mil casos confirmados, un número de muertes aproximadas a 25 mil y un aproximado de 5 Mil casos diarios. Datos con los cuales se trabajó durante este proyecto para graficar posteriormente.

Dicho virus tiene al mundo actual limitado en sus numerosos campos de acción, desde el campo social hasta el económico, pausando las economías a nivel mundial y produciendo nuevas alternativas a las acciones de la vida cotidiana. [1].

1.2. Bases de datos

Se entiende como el conjunto de datos los cuales se almacenan en un mismo lugar con los cuales se trabajara a corto o largo plazo. Compuestas de columnas y filas que forman múltiples registros.

Algunas ventajas notables a la hora de usar dichas bases, son la posibilidad de guardar cantidades inmensas de información, encontrarla y utilizarla de una manera práctica, ordenada y eficaz.

Con el paso del tiempo, como la mayoría de cosas en nuestra actualidad, se busca una optimización de nuestros recursos con el uso de estas bases, dando seguridad, compatibilidad máxima, la menor cantidad de redundancia y respaldo a nuestros datos. [2].

Se encuentran diferentes tipos de bases de datos como lo son:

1. SQL: También llamada como bases de datos relacionales. Su lenguaje se basa en consultas estructuradas de los datos. Dichas bases son escalables verticalmente a través de mejoras de hardware en su servidor y en cuanto a su estructura, se encuentra que están basadas en tablas.

2. NOSQL: Son bases conocidas por tener un esquema dinámico para datos no estructurados. Tiene una escalabilidad horizontal, lo cual se entiende como un mayor tráfico de fragmentación de datos. Usualmente su estructura está basada en pares clave-valor, gráficas y columnas.

1.3. Python

Es un lenguaje de programación de código abierto, orientado a objetos y con múltiples bibliotecas de herramientas al uso de todos. Python permite ejecutar el código sin necesidad de compilar nuestro código, dando agilidad durante su uso. Python posee la característica principal de multiplataforma, así mismo, posee una notación indentada, la cual trabaja con tabulación en el código de sus respectivas funciones o bucles, permitiendo una mayor organización y optimización del proyecto. Por otro lado encontramos las librerías, que son un código prediseñado que nos ofrece un sin número de funciones a nuestros proyectos. Para el desarrollo de este proyecto fue necesario el uso de librerías como: , Requests, Json.

1. Matplotlib: Se encarga del trazado 2D y 3D, produciendo diferentes gráficos, histogramas, espectros de potencia, gráficos de barras, gráficos de error, gráficos de dispersión, etc., con solo unas pocas líneas de código. [3].
2. Requests: Permite enviar solicitudes HTTP via python, entendiéndose como “Web Scraping” o simplemente como la recolección de información web. [4].
3. Json: Se encarga de parsear el JSON de archivos o strings, convirtiendo los datos y listas de python en cadenas JSON. [5].

Para el uso de estas librerías, es necesario realizar su instalación correspondiente por medio de las siguientes líneas de comando. [6].

```
pp3 install requests
pp3 install json
pp3 install matplotlib
```

Figura 1: Librerías Python

2. Resultados

Recordamos que el lenguaje que se utilizó durante este proyecto fue python y que se deben instalar las librerías previamente mencionadas en la sección 1.3. Inicialmente se procede a extraer la información de la página web, por medio de sentencias sql para su debido almacenamiento.

```
1 import requests
2 import matplotlib.pyplot as plt
3 import numpy as np
4
5 ListaValor=[]
6 ListaLocalidad=[]
7 Localidad=[]
8 Casos=[]
9 Tiempo=[]
10
11 urlDatos = 'https://datosabiertos.bogota.gov.co/api/3/action/datastore_search_sql?'
12 urlDatosSQL = 'sql=SELECT "Localidad de residencia", count(*) as Cantidad from ...
13               "b64ba3c4-9e41-41b8-b3fd-2da21d627558" GROUP BY "Localidad de residencia" ORDER BY ...
14               "Localidad de residencia"'
15
16 req = requests.get(url=urlDatos+urlDatosSQL)
17 reqJson = req.json() #Organizar datos
18 Datos = reqJson['result']['records']
19 i=0
```

Listing 1: Código de obtención datos COVID-19

La primera gráfica que se realizó, fue la de barras de forma horizontal, teniendo el número de casos en el eje x y la localidad en el eje y.

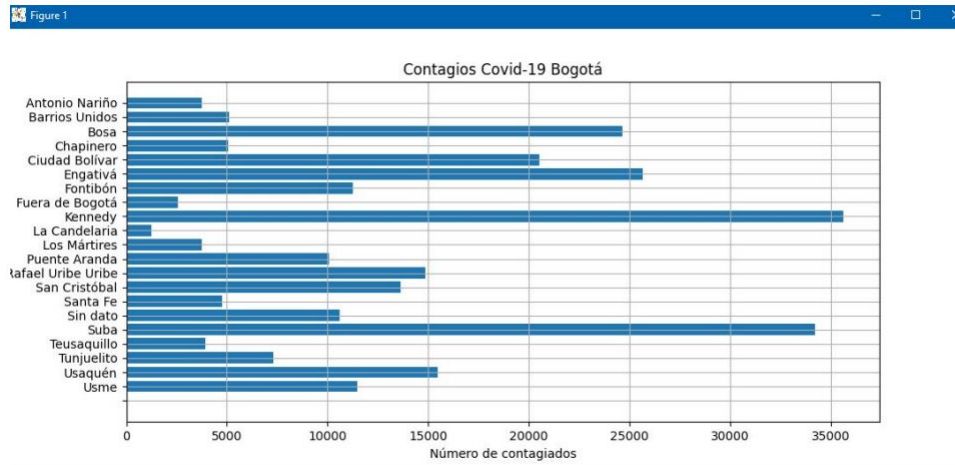


Figura 2: Gráfica de Barras Contagio Por Localidades

```

1 #Grafica de barras
2
3 for fila in Datos:
4     ListaValor.append(int(fila["cantidad"]))
5     ListaLocalidad.append(fila["Localidad de residencia"])
6     i = i + 1
7
8 plt.rcParams()
9 fig, ax = plt.subplots(figsize=(11, 5))
10
11 y_pos = np.arange(len(ListaLocalidad))
12
13 ax.barh(y_pos, ListaValor, align='center')
14 ax.set_yticks(y_pos)
15 ax.set_yticklabels(ListaLocalidad)
16 ax.invert_yaxis()
17 ax.set_xlabel('Número de contagiados')
18 ax.set_ylabel('Localidades')
19 ax.set_title('Contagios Covid-19 Bogotá')
20
21 plt.grid()

```

Listing 2: Gráfica de barras

A continuación se procedió con realizar dos gráficas de torta y mostrar los códigos correspondientes.

```

1 #Graficas tortas
2
3 labels = ListaLocalidad
4 sizes = ListaValor
5
6 fig1, ax1 = plt.subplots(figsize=(10,7))
7 ax1.pie(sizes, labels=labels, autopct='%1.1f%%',
8         shadow=True, startangle=90)
9 ax1.axis('equal')
10 ax1.set_title("Contagios Covid-19 Bogotá")
11
12 fig, ax = plt.subplots(figsize=(15, 10), subplot_kw=dict(aspect="equal"))
13 recipe = ListaLocalidad
14 data = ListaValor
15 wedges, texts = ax.pie(data, wedgeprops=dict(width=0.5), startangle=-40)
16
17 bbox_props = dict(boxstyle="square,pad=0.3", fc="w", ec="k", lw=0.72)
18 kw = dict(arrowprops=dict(arrowstyle="-"),
19         bbox=bbox_props, zorder=0, va="center")
20

```

```

21 for i, p in enumerate(wedges):
22     ang = (p.theta2 - p.theta1)/2. + p.theta1
23     y = np.sin(np.deg2rad(ang))
24     x = np.cos(np.deg2rad(ang))
25     horizontalalignment = {-1: "right", 1: "left"}[int(np.sign(x))]
26     connectionstyle = "angle,angleA=0,angleB={}".format(ang)
27     kw["arrowprops"].update({"connectionstyle": connectionstyle})
28     ax.annotate(recipe[i], xy=(x, y), xytext=(1.35*np.sign(x), 1.4*y),
29                 horizontalalignment=horizontalalignment, **kw)
30
31 ax.set_title("Contagios Covid-19 Bogot ")

```

Listing 3: Gráficas de torta

Por último, se graficó en dos dimensiones.

```

1
2 #Grafica 2 dimensiones
3
4 t = np.arange(0, 22)
5 s = np.array(ListaValor)
6
7 fig, ax = plt.subplots()
8 ax.plot(t, s)
9
10 ax.grid(True, linestyle='dashed')
11 ax.tick_params(labelcolor='r', labelsize='medium', width=3)
12 ax.set_title("Contagios Covid-19 Bogot ")
13 ax.set_xlabel('Numero de Localidades')
14 ax.set_ylabel('Contagiados')
15
16 plt.show()

```

Listing 4: Código Gráfica en dos dimensiones

Para encontrar este documento, junto con el código y video correspondiente por favor visite nuestro repositorio <https://github.com/CarlosPA90-666/ProyectoSeniales.git>

3. Conclusiones

- Se evidencia que el 2020 ha sido un año complicado para la humanidad, trayendo con él diferentes problemas a nivel mundial, en este caso el virus “Covid19, del cual es necesario tener una información actualizada de sus síntomas, tratamientos, prevención y expansión. Para esto, encontramos diferentes páginas web que nos brindan dicha información. Evidenciando la necesidad de extraer la información con recursos que se tienen a la mano durante nuestra formación académica como python y sql. Este proyecto nos servirá para presentar de forma gráfica el día a día de esta enfermedad para así tomar decisiones, controlar y actuar a tiempo.

Referencias

- [1] M. Colombia, “Coronavirus (covid-19).” Colombia, 2020.
- [2] A. Oppel, “Fundamentos de bases de datos.” México D.F: McGraw-Hill, 2011.
- [3] J. Hunter, “Python plotting — matplotlib 3.3.2 documentation.” Matplotlib, 2012.
- [4] Unypython, “Solicitudes http en python con requests.” Unypython, 2018.
- [5] A. Ahmed, “Cómo trabajar con datos json utilizando python.” Ahmed, 2016.
- [6] T. Donaldson, “Python.” Berkeley, Calif.: Peachpit Press, 2014.

4. Anexos

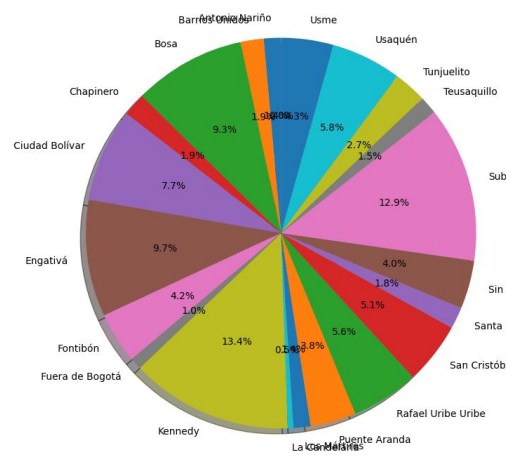


Figura 3: Gráfico de Torta por Porcentajes

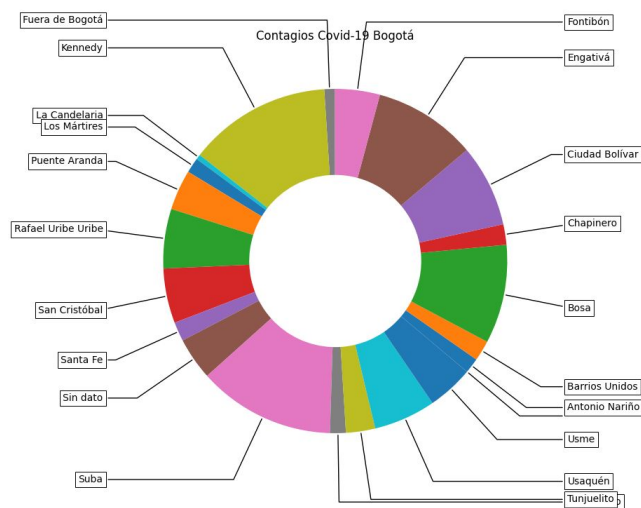


Figura 4: Gráfico de Torta por Nombres

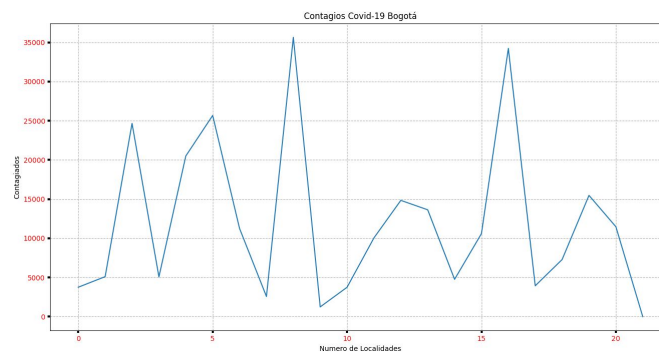


Figura 5: Gráfico de Dos Dimensiones