

A dark blue vertical bar runs along the left edge of the page. A blue arrow-shaped banner points to the right from this bar, containing the date '23-3-2022'. In the bottom-left corner, there are several thin, curved, light blue lines that sweep upwards and to the right.

23-3-2022

Manual de estadísticas descriptivas

Edgar Asael Martínez

Tania Sayuri Guizado Hernández | A01640092

Ricardo Jiménez Ureña | A01636825

Carlos David Blanco | A01633323

Rafael Jiménez | A01637850

Carlos Estrada Ceballos A01638214

HERRAMIENTAS COMPUTACIONALES: EL ARTE DE LA
ANALÍTICA - TC1002S.100

Para este ejercicio se eligió las estadísticas del radar sobre las pruebas covid

Con la librería pandas se lee el archivo csv y convertimos los datos de Federal a float

```
In [1]: import pandas as pd
        #import numpy as np
        from matplotlib import pyplot as plt
        #import random

In [2]: df = pd.read_csv("01.01.2022_da_radar_pruebas_covid19.csv")
        del df['Unnamed: 6']

In [3]: df['Federal'] = df['Federal'].str.replace(',','')

In [4]: df["Federal"] = df.Federal.astype(float)
```

En la siguiente línea visualizamos con head los primeros 10 datos

```
In [5]: df.head(10)
```

Out[5]:

	Fecha	Resultado	Federal	U de G	Privado	Rápidas
0	17/04/2020	Confirmados	190.0	0	18	Na
1	17/04/2020	Descartados	2601.0	110	271	Na
2	17/04/2020	Sospechosos	573.0	18	Na	Na
3	18/04/2020	Confirmados	196.0	0	18	Na
4	18/04/2020	Descartados	2666.0	110	271	Na
5	18/04/2020	Sospechosos	547.0	18	Na	Na
6	19/04/2020	Confirmados	198.0	0	18	Na
7	19/04/2020	Descartados	2805.0	110	271	Na
8	19/04/2020	Sospechosos	372.0	0	Na	Na
9	20/04/2020	Confirmados	216.0	0	18	Na

Observando la tabla que arrojo el código las variables que vamos a considerar son:

- **Fecha - 1875 rows × 1 column**
 - Dia/Mes/Año
- **Resultado - 1875 rows × 2 column**
 - Confirmados: Pruebas covid con resultado positivo
 - Descartados: Pruebas covid con resultado negativo
 - Sospechosos: Pruebas covid con probabilidad de positivo o negativo
- **Federal - 1875 rows × 3 column**
 - Número de pruebas
- **U de G - 1875 rows × 4 column**
 - Número de pruebas
- **Privado - 1822 rows × 5 column**
 - Número de pruebas
- **Rápidas 1822 rows × 6 column**

- Número de pruebas

Con la siguiente línea identificamos que hay un total de 1875 entradas

```
In [6]: len(df.index)
```

```
Out[6]: 1875
```

Pero con la línea donde requerimos la información por columna nos percatamos que las variables privado y rápidas tienen 1822 datos y las demás 1875

```
In [7]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1875 entries, 0 to 1874
Data columns (total 6 columns):
#   Column      Non-Null Count  Dtype
---  ---
0   Fecha       1875 non-null   object
1   Resultado    1875 non-null   object
2   Federal     1875 non-null   float64
3   U de G      1875 non-null   object
4   Privado     1822 non-null   object
5   Rápidas     1822 non-null   object
dtypes: float64(1), object(5)
memory usage: 88.0+ KB
```

Media:

Media

```
In [8]: print('Media: ', df["Federal"].mean())
```

```
Media: 60256.4832
```

Mediana:

Mediana

```
In [9]: print('Mediana:', df["Federal"].median())
```

```
Mediana: 29800.0
```

Desviación estándar:

Desviación estándar

```
In [10]: print('Desviación estándar: ', df["Federal"].std())
```

```
Desviación estándar: 59558.50207360816
```

Conclusiones:

En la observación de los resultados, que nos dio el análisis de los datos, pudimos observar que existe una media general de 60,256 de pruebas entre las tres variables (confirmadas, descartadas y sospechosas), además de esto se puede notar que existen muchos valores atípicos en los datos lo que nos hace entender que hay una gran variación en la cantidad de pruebas aplicadas al día, dependiendo de la fase de emergencia y la disponibilidad de estas, así como también los resultados positivos nos dan una idea general de cuando fue la fase mas

crítica de la pandemia y su vez las pruebas descartadas nos dan un panorama general de que periodo de la pandemia tuvo una reducción en la cantidad de contagios.

Git: <https://github.com/carlosdblancoretoAnalitica>