

Ciclo 1 Fundamentos de Programación

Reto 5: Análisis Financiero TRM



Para realizar negocios a nivel nacional, es importante monitorear el valor del dólar o Tasa Representativa del Mercado (TRM – Peso colombiano por dólar de los Estados Unidos), el valor de las acciones de Ecopetrol, dado que el petróleo representa el principal ingreso del país, y en el contexto de pandemia, el número de contagios también puede constituir un elemento importante para tener en cuenta. La junta directiva de una asociación, necesita observar el comportamiento de la TRM los días en que se ha superado la media de contagios, desde el primero de marzo de 2021 hasta los primeros días de junio de 2021.

El indicador debe construirse obteniendo el valor promedio de variación absoluta del dólar, la cual se calcula de la siguiente manera:

$$\text{Variación} = \frac{| \text{Precio anterior} - \text{Precio actual} |}{\text{Precio anterior}} * 100$$

Los directivos necesitan conocer el promedio de variación del dólar, junto con los registros (días) donde el número de contagios es igual o mayor a la media, para analizar el impacto que los casos de COVID-19 pueden tener en los convenios comerciales que están en proceso.

Para lo anterior, se requiere una función que procese una base de datos en formato CSV (del inglés, *comma-separated values*), que contiene fecha, la TRM, el número de casos presentados el día registrado, y el valor de la acción de Ecopetrol. A continuación se especifica el tratamiento que se debe realizar de la información:

- 1) Validar que el origen de los datos tenga formato (extensión) CSV antes de realizar cualquier cómputo, retornando el mensaje “Extensión inválida” en caso de que no cumpla.
- 2) Intentar realizar la carga del archivo CSV a un dataframe de pandas con separación de comas “,”. En caso de que falle el intento, retornar el mensaje “Error al leer el archivo de datos”.

- 3) Cambiar el campo que contiene la fecha en formato string a tipo fecha de pandas, y convertirla en el índice del dataframe, dado que este especifica el día de cada registro, pero no se realizarán cálculos con esta información.
- 4) Calcular el valor promedio de la variación absoluta del dólar.
- 5) Obtener un dataframe que contenga únicamente los registros donde el número de contagios es mayor o igual a la media de casos de COVID-19.

A continuación se especifican los detalles técnicos del requerimiento. El esqueleto de la función, indicado en la Figura 1, recibe la ruta del archivo (que podría ser también una dirección de Internet), y retorna un diccionario con el promedio de variación, y el dataframe resultante de la manipulación indicada en los numerales anteriores.

```
def analisisTRM(rutaArchivo: str) -> dict:
    """Desarrollar aquí el requerimiento"""
    pass
```

Figura 1. Prototipo de la función

Especificación de la entrada

Entrada	Tipo de Dato	Descripción
rutaArchivo	string	Ruta del archivo, la cual podría ser local (mismo directorio) o podría ser una dirección web. Tener esto en cuenta para la validación de formato.

El archivo o base de datos que se requiere procesar, se encuentra en el siguiente enlace:
<https://github.com/luismescobarf/clasesCiclo1/blob/master/BaseDeDatosReto5.csv>

Se recomienda para la realización de las pruebas y observación de la base de datos, descargar el archivo del enlace anterior, en la misma carpeta donde se desarrollará la función del requerimiento.

Estructura de la Salida

Tipo de Retorno	Descripción
dict	Diccionario con el promedio de variación del dólar (float) y dataframe tipo pandas con los registros que son iguales o superiores al promedio de casos.

Ejemplo Salida:



{'Promedio Variación': 0.43869307986456585, 'Registros Mayores':

TRM Casos

Ecopetrol

Fecha

2021-05-23 3738.19 21669 2317.81

2021-05-22 3738.19 18737 2325.25

.

.

.

2021-04-08 3639.62 12464 2281.00

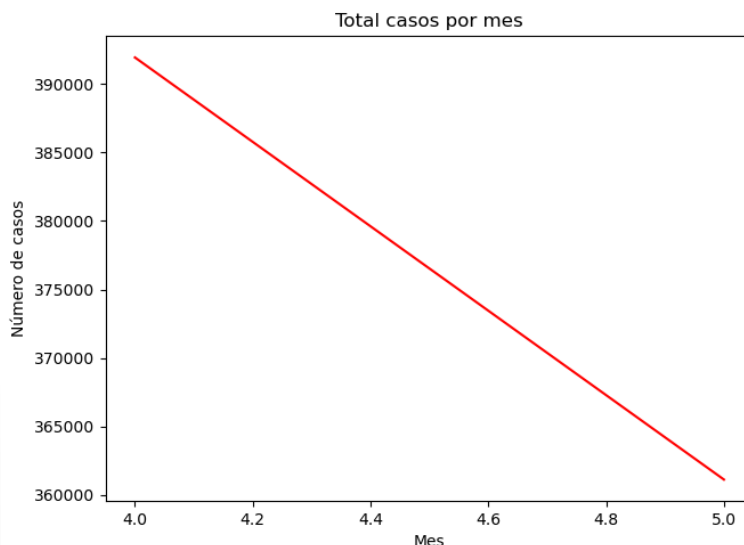
2021-04-06 3645.79 11827 2289.00}

Visualizaciones Complementarias:

A continuación se presentan las gráficas que deben obtenerse si el tratamiento de los datos fue el adecuado. Es importante resaltar que estas gráficas son opcionales y **no deben ser incluidas en la función que se subirá a la plataforma iMaster**. Suponiendo que en el dataframe denominado *data*, se ha almacenado el dataframe denominado en el diccionario de retorno con la llave 'Registros Mayores', las siguientes instrucciones de graficado deben producir la visualización correspondiente:

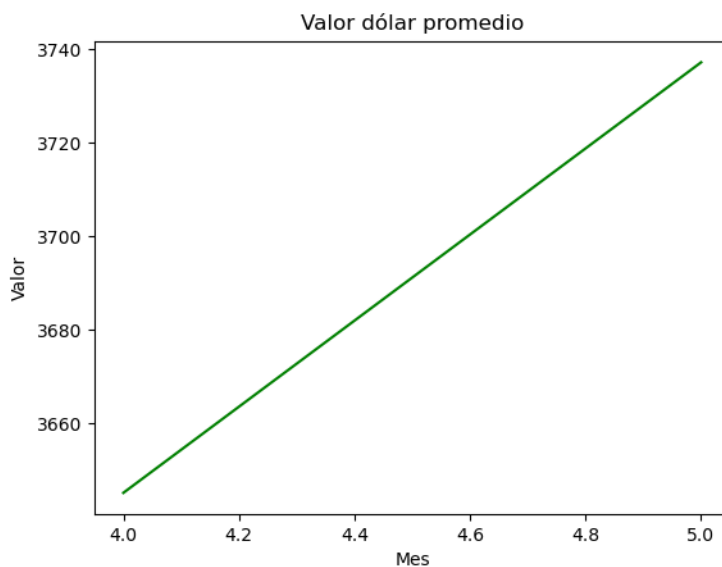
Visualización 1

```
#Graficado opcional de los casos acumulados en cada mes después de la  
manipulación  
import matplotlib.pyplot as plt  
dataCasosMes = data["Casos"].groupby(data.index.month).sum()  
dataCasosMes.plot(color="red")  
plt.title("Total casos por mes")  
plt.xlabel("Mes")  
plt.ylabel("Número de casos")  
plt.show()
```



Visualización 2

```
#Graficado opcional del TRM medio de cada mes después de la manipulación
import matplotlib.pyplot as plt
dataMediaPrecioMes = data['TRM'].groupby(data.index.month).mean()
dataMediaPrecioMes.plot(color = "green")
plt.title("Valor dólar promedio")
plt.xlabel("Mes")
plt.ylabel("Valor")
plt.show()
```



Partiendo de lo anterior, los casos de prueba son los siguientes:

Caso de Prueba 1:

Llamado:

```
print(analisisTRM('BaseDeDatos.csv'))
```

Resultado esperado:

```
Error al leer el archivo de datos
```

Caso de Prueba 2:

Llamado:

```
print(analisisTRM('BaseDeDatosReto5.ppt'))
```



Resultado esperado:

Extensión inválida

Caso de Prueba 3:

Llamado:

```
print( analisisTRM('https://raw.githubusercontent.com/luismescobarf/clasesCiclo1/master/BaseDeDatosReto5.csv'))
```

Resultado esperado:

```
{'Promedio Variación': 0.4386930798645658, 'Registros Mayores':
```

```
TRM  Casos  Ecopetrol  
Fecha  
2021-05-23  3738.19  21669  2317.81  
2021-05-22  3738.19  18737  2325.25  
2021-05-21  3721.57  14838  2285.00  
2021-05-20  3682.66  16086  2314.00  
2021-05-19  3655.74  16579  2347.00  
2021-05-18  3682.84  13137  2355.00  
2021-05-17  3682.84  12984  2349.56  
2021-05-16  3682.84  15093  2343.80  
2021-05-15  3682.84  18873  2346.00  
2021-05-14  3728.09  16581  2387.00  
2021-05-13  3734.09  19160  2301.00  
2021-05-12  3703.20  16993  2370.00  
2021-05-11  3714.94  16425  2315.00  
2021-05-10  3765.33  12543  2357.00  
2021-05-09  3765.33  17222  2270.84  
2021-05-08  3765.33  16910  2262.20  
2021-05-07  3800.33  17525  2288.00  
2021-05-06  3846.28  16490  2291.00  
2021-05-05  3831.35  14806  2288.00  
2021-05-04  3816.65  14551  2225.00  
2021-05-02  3740.14  15909  2284.16  
2021-05-01  3740.14  18022  2281.80  
2021-04-30  3712.89  17790  2210.00  
2021-04-29  3699.74  17308  2310.00  
2021-04-28  3717.46  19745  2332.00  
2021-04-27  3659.62  17578  2287.00  
2021-04-26  3640.07  12839  2270.00  
2021-04-25  3640.07  17190  2257.04  
2021-04-24  3640.07  16730  2264.20  
2021-04-23  3630.81  19925  2228.00
```



2021-04-22	3639.12	19306	2250.00
2021-04-21	3636.26	17212	2262.00
2021-04-20	3606.42	16965	2281.00
2021-04-19	3595.57	14189	2300.00
2021-04-18	3595.57	16871	2310.04
2021-04-17	3595.57	16654	2306.20
2021-04-16	3620.40	16703	2301.00
2021-04-15	3665.49	16918	2319.00
2021-04-14	3666.17	16487	2344.00
2021-04-13	3653.57	16377	2280.00
2021-04-12	3650.23	16739	2287.00
2021-04-11	3650.23	17483	2293.08
2021-04-10	3650.23	14509	2298.40
2021-04-09	3634.07	12125	2260.00
2021-04-08	3639.62	12464	2281.00
2021-04-06	3645.79	11827	2289.00}

Recordar: En la plataforma debe subirse una función con **el mismo nombre**, la **misma cantidad de argumentos**, y el **retorno debe tener exactamente la estructura de la salida** que se presenta en este documento.