

1. Contexto. Por qué el sitio web elegido proporciona dicha información.

Este dataset es una contiene información sobre las cotizaciones en bolsa de: apertura, máximo, mínimo, cierre y cierre ajustado (todas en Euros) de la empresa Telefónica, desde el 2015-03-30 hasta la fecha actual 2020-03-28, utilizando las librerías: pandas_datareader, argparse y datetime.

El sitio web elegido es yahoo por su confiabilidad y por facilitar el campo cierre ajustado.

2. Definir un título para el dataset

EVOLUCIÓN DE LAS COTIZACIONES DE TELEFÓNICA DURANTE LOS ÚLTIMOS 5 AÑOS

3. Descripción del dataset.

El dataset contiene datos de cotizaciones de la empresa Telefónica desde el 30 de marzo de 2015 hasta el 28 de marzo de 2020 (fecha actual del proyecto). El periodo abarca un periodo temporal de aproximadamente 5 años. Hay que tener en cuenta sin embargo, que el número de registros (días) del dataset no se corresponde con 1.825 (el correspondiente a aproximadamente 5 años), dado que la bolsa no está activa durante los fines de semana y algunos festivos.

4. Representación gráfica. Presentar una imagen o esquema que identifique el dataset visualmente**5. Contenido. Explicar los campos que incluye el dataset, el periodo de tiempo de los datos y cómo se ha recogido.**

El dataset está compuesto por 1280 registros (filas) con los siguientes campos (columnas):

- Date: La fecha en la que fueron recogidos los demás datos.
- High: La cotización máxima que alcanzaron las acciones de Telefónica en la fecha indicada.
- Low: La cotización mínima que alcanzaron las acciones de Telefónica en la fecha indicada.
- Open: La cotización de apertura (inicial) de las acciones de Telefónica en la fecha indicada.
- Close: La cotización de cierre (final) de las acciones de Telefónica en la fecha indicada.

Autores:

- Enrique Javier Andrés Orera
- Rubén Silva Marín

- Volume: El volumen de operaciones que se llevaron a cabo con las acciones de Telefónica en la fecha indicada.

- Adj Close: La cotización final de las acciones de Telefónica en la fecha indicada. El cierre ajustado es el precio de cierre después de los ajustes para todas las distribuciones de splits y dividendos aplicables. Los datos se ajustan utilizando los multiplicadores de splits y dividendos correspondientes, de conformidad con las Center for Research in Security Prices (CRSP).

(Fuente: <https://es.ayuda.yahoo.com/kb/%C2%BFQu%C3%A9-es-el-cierre-ajustado-sln28256.html> visitada 29/03/2020)

El periodo recogido incluye aproximadamente los 5 años anteriores a la fecha actual: desde el 30 de marzo de 2015 hasta el 28 de marzo de 2020.

Durante el trabajo de extracción se ha optado por el uso de la librería `pandas_datareader` dado que ofrece claras ventajas como: sencillez del código y, por tanto, mejor legibilidad, además de una extracción más precisa a través del uso de parámetros concretos de símbolo (empresa en bolsa) y fechas específicas, así como parámetros por defecto.

También se ha hecho uso de la librería `argparse`, con el fin de posibilitar la interacción con personas para que estas puedan extraer los datos requeridos en cada momento.

Finalmente se ha incorporado además la librería `datetime` para el manejo de las fechas de manera adecuada.

6. Agradecimientos

Los datos han sido recogidos de la base de datos online de yahoo finance. Utilizando para ello el lenguaje de programación Python y la librería: `pandas_datareader`.

7. Inspiración

Dada la situación actual en la que se ve sumergida el mundo entero por el COVID-19, se ha dispuesto el siguiente dataset para evaluar el impacto económico que ha tenido en la empresa Telefónica. Esto se ha llevado a cabo seleccionando el periodo de los 5 años anteriores.

Este dataset puede ser utilizado como base para elaborar modelos predictivos, con la intención de predecir cotizaciones futuras.

8. Licencia. Seleccione una de estas licencias para su dataset y explique el motivo de su selección:

Se han de tener en cuenta las siguientes consideraciones:

- La información extraída de la página de Yahoo Finance u otras entidades asociadas, no puede utilizarse con fines comerciales por ninguno de sus usuarios.
- Zenodo, la página de destino para el dataset de esta práctica está certificada por OPENAire, lo que lo convierte en Open Data Source.
- Teniendo esto en cuenta, la publicación del dataset en Zenodo no se establecería como una actividad comercial.
- En base a lo anterior se establece que la posible licencia para aplicar es: CC-BY-NC-SA 4.0, que permite la reproducción y distribución del dataset, pero prohíbe su comercialización, a la vez que obliga a replicar esta licencia en todas sus redistribuciones.

Autores:

- Enrique Javier Andrés Orera
- Rubén Silva Marín

9. Código. Adjuntar el código con el que se ha generado el dataset, preferiblemente en Python o, alternativamente, en R.

```
# Import libraries
import pandas_datareader
import datetime
import argparse
import pandas as pd
import plotly.graph_objects as go

# Parse command line arguments
parser = argparse.ArgumentParser()
parser.add_argument("--ticker", help="Enter corporation ticker symbol",
                    default = 'TEF.MC')
parser.add_argument("--startDate", help="Enter start date of interval YY-MM-DD",
                    default = str(datetime.date.today()-datetime.timedelta(days = 365)))
parser.add_argument("--endDate", help="Enter end date of interval YY-MM-DD",
                    default = str(datetime.date.today()))
args = parser.parse_args()

# Variables
ticker = args.ticker
startDate = datetime.datetime.strptime(str(args.startDate), "%Y-%m-%d")
endDate = datetime.datetime.strptime(str(args.endDate), "%Y-%m-%d")

# Read data from yahoo
df = pandas_datareader.DataReader(ticker, 'yahoo', startDate, endDate)

# Get Date into a column value
df.reset_index(inplace=True, drop=False)

# Print dataframe head
print(df.head)
```

Autores:

- Enrique Javier Andrés Orera
- Rubén Silva Marín

```
# Export dataframe to csv
df.to_csv(ticker + '.csv')

# This code is adapted from https://plot.ly/python/ohlc-charts/
# OHLC Chart without Rangeslider
fig = go.Figure(data=go.Ohlc(x=df['Date'],
                             open=df['Open'],
                             high=df['High'],
                             low=df['Low'],
                             close=df['Close'])))
fig.update(layout_xaxis_rangeslider_visible=False)

# Adding customized text
fig.update_layout(
    title=ticker,
    yaxis_title='Cotización')

# Show OHLC Chart
fig.show()

# Save OHLC Chart html
fig.write_html("file.html")
```

10. Dataset. Publicación del dataset en formato CSV en Zenodo con una pequeña descripción.

Autores:

- Enrique Javier Andrés Orera
- Rubén Silva Marín