

# WEB SCRAPING

Martín Cordón – Erick Ruiz

# ¿QUÉ ES WEB SCRAPING?

Proceso para extraer data de sitios web utilizando Python.

El objetivo es simular la navegación desde la perspectiva de cualquier internauta.



# EXPLICACIÓN



## ¿cómo funciona?

Se utiliza un driver, es decir un archivo ejecutable de Chrome, al cuál se le asignan comandos y que posteriormente execute el programa.



## ¿qué puede lograr?

- Imprimir títulos
- Hacer Click's
- Extraer data útil



## tipos de herramientas

- beautiful soup
- selenium
- query

## DEFINICIÓN

Extraer datos a través de la web.  
Especificamente datos de activos  
financieros de Yahoo Finance.

## DELIMITACIÓN

Recopilación de datos y generar un  
pronósticos de series de tiempo  
univariadas + machine learning.

# EXPERIENCIA

*trabajando con Yahoo Finance*

Hacer un código  
utilizando la  
librería de  
selenium.



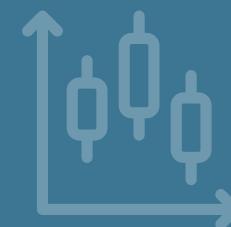
Yahoo Finance no  
permite extraer  
datos de manera  
accesible



Creación de un  
API que utiliza  
un query para  
encontrar data:

- Fecha
- Ticker
- Intervalo

Implementación  
de Dashboard  
para visualizar  
la data.



El resultado es un  
pronóstico de  
serie de tiempo  
univariada.



# RESULTADOS

*ejemplo obtenido por AAPL como ticker*

## Datos crudos

	Date	Open	High	Low	Close	Adj Close	Volume
1135	2021-12-27	177.0900	181.3300	177.0700	177.5700	177.5700	340248100
1136	2022-01-03	177.8300	182.9400	171.0300	172.1700	172.1700	481820000
1137	2022-01-10	169.0800	177.1800	168.1700	173.0700	173.0700	422569900
1138	2022-01-17	171.5100	172.5400	164.1800	164.5100	164.5100	276920900
1139	2022-01-20	166.9800	169.6400	164.2300	164.5100	164.5100	91420515

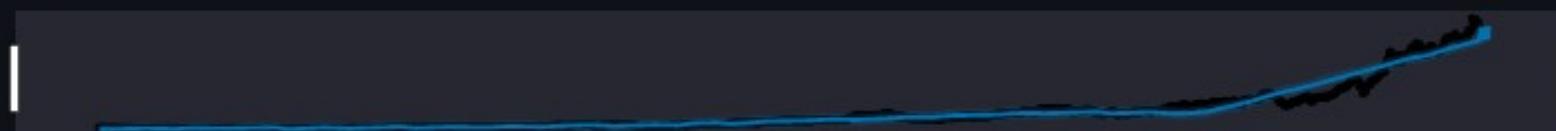
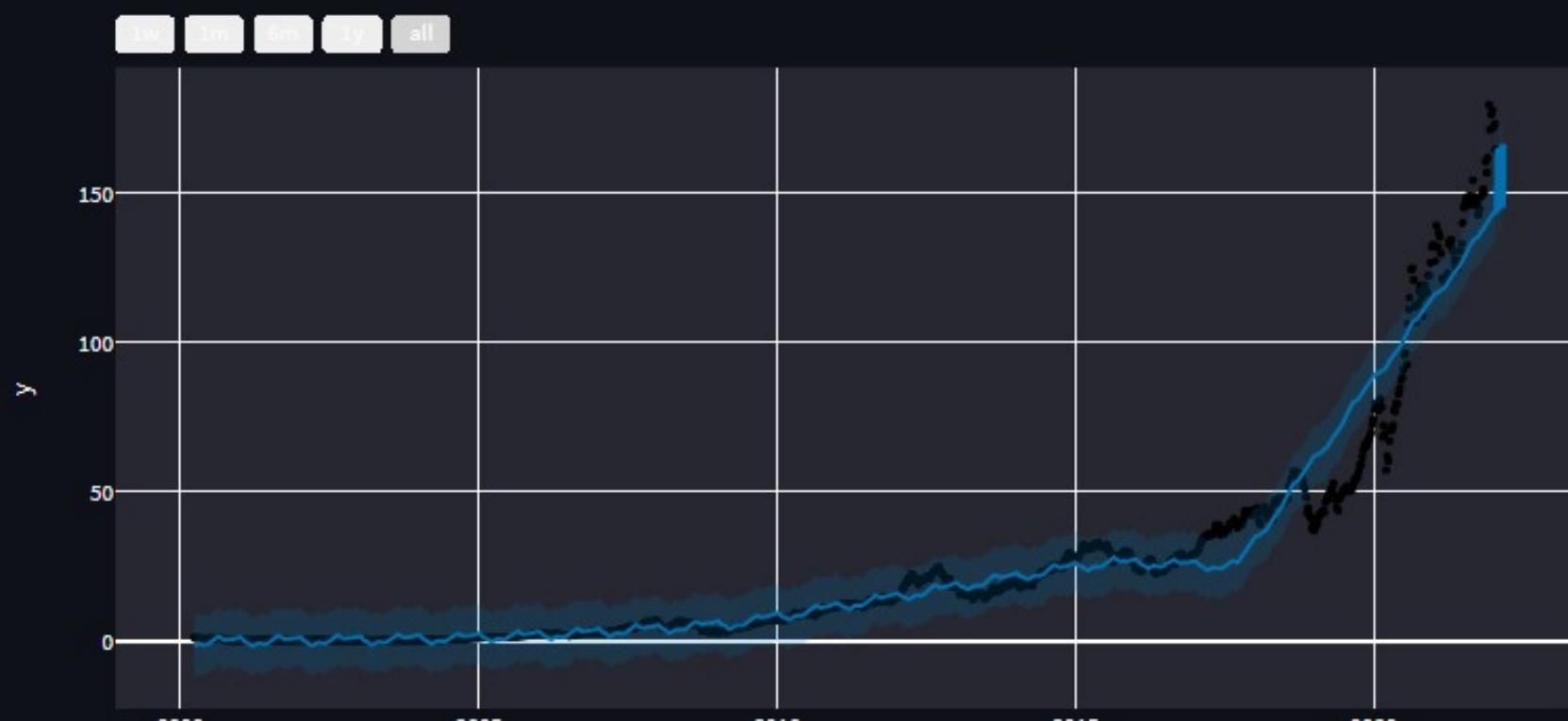
[Descarga los datos en formato Excel](#)

## Datos pronosticados

	ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper
1187	2022-03-09T00:00:00	152.3267	138.9684	158.1139	152.3267	152.3267
1188	2022-03-10T00:00:00	152.4008	156.4812	175.9143	152.4008	152.4008
1189	2022-03-11T00:00:00	152.4750	138.8155	158.5278	152.4750	152.4750
1190	2022-03-12T00:00:00	152.5491	138.4361	158.8078	152.5491	152.5491
1191	2022-03-13T00:00:00	152.6232	139.2258	158.5220	152.6232	152.6232

Pronóstico para 1 años

Gráfico de candelas interactivo para AAPL



# CONCLUSIONES

- Se cumplió el objetivo para poder extraer los datos de Yahoo Finance.
- Se realizó el pronóstico de tiempo univariada.
- Se logró crear un Dashboard para unificar la data recopilada
- Se obtuvo una interfaz dinámica sin utilizar la herramienta que se planteó originalmente.