



# {Avance de TOG}

Luis Alfonso Guerrero Montaño  
Maestría en Ciencia de Datos



# Sinodales

Diana Paola Montoya

Cesar Nuñez

Asesor

Juan Franciso Muñoz



# Antecedentes e Interés en el proyecto

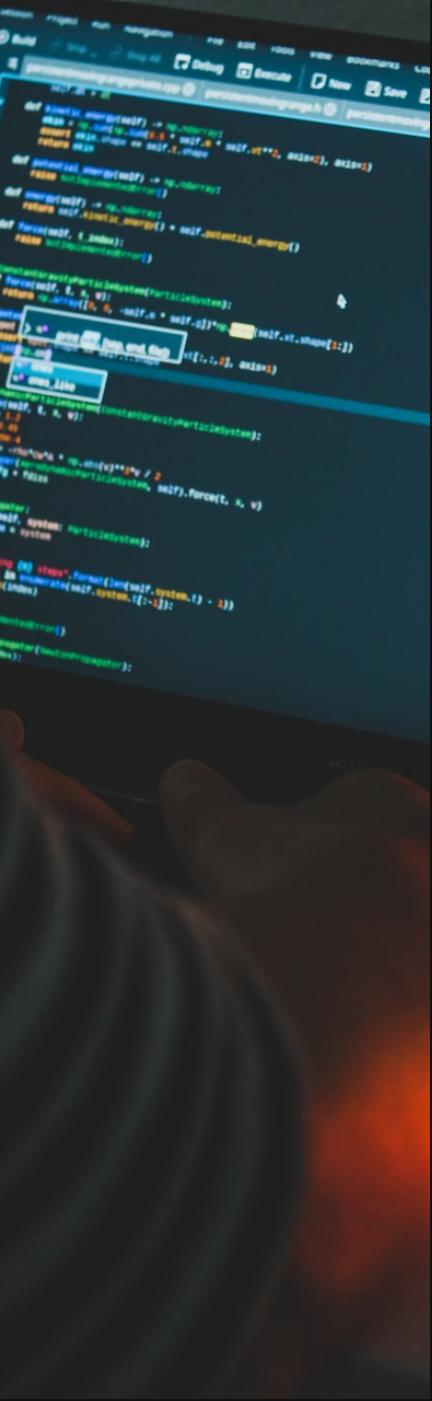
Soy Luis Guerrero, estudiante egresado de Ingeniería en sistemas computacionales (TESO 1994-1998) y MBA (ITESO y Regis University dual degree program 2011-2013)

Mi interés en el tema de finanzas descentralizadas y blockchain, surge primero por ver que tan seguro es invertir en el tema, específicamente en los “nuevos mercados” de cripto-monedas.

Mi objetivo es invertir en un futuro y poder aconsejar y ayudar a tomar decisiones a este respecto a más personas.



# Problema a resolver y base de trabajo



El problema que intento resolver es que tan volátil o incierto es este mercado comparado con instrumentos tradicionales, por lo que me enfocare en medir esa volatilidad con la diferencia de precios de apertura y cierre de una cripto-moneda en intervalos definidos de tiempo.

El trabajo de obtención de grado se basa en los mercados de finanzas descentralizadas (DEFI por sus siglas en inglés), que se refiere a mercados que no hay una persona o entidad los controla, el trabajo está enfocando en mercados de alta frecuencia (HFT) que es una plataforma automatizada de inversión, que permite un gran número de transacciones a muy alta velocidad o frecuencia.

# Taxonomía de los datos y finalidad

El Enfoque principal es en dos estructuras de datos:

- Libros de órdenes (“Order books”): Es una lista electrónica de órdenes de compra y venta para un instrumento financiero.
- Libro de operaciones públicas (“Public trades”): Una vez que una orden es ejecutada aquí se almacenan los datos de la transacción.

**El trabajo es proceso de modelado predictivo para series de tiempo financieras. Para el cual usaré una regresión, teniendo como variable objetivo la diferencia entre el valor del precio de cierre y de apertura de la moneda seleccionada.**



# Extracción de datos



El principal reto de este periodo fue aprender del tema y la correcta extracción de los datos. En este último en particular tuve que trabajar con procesos asíncronos para bajar información de la plataforma y manejar altos volúmenes de datos.

La fuente de los datos es pública, y se extraen usando la plataforma “binance”. para esta etapa en específico de la cadena de bloques (“blockchain”) Ethereum “ETH”, usando la moneda USDT (Tether).



# Ejemplo de obtención de datos

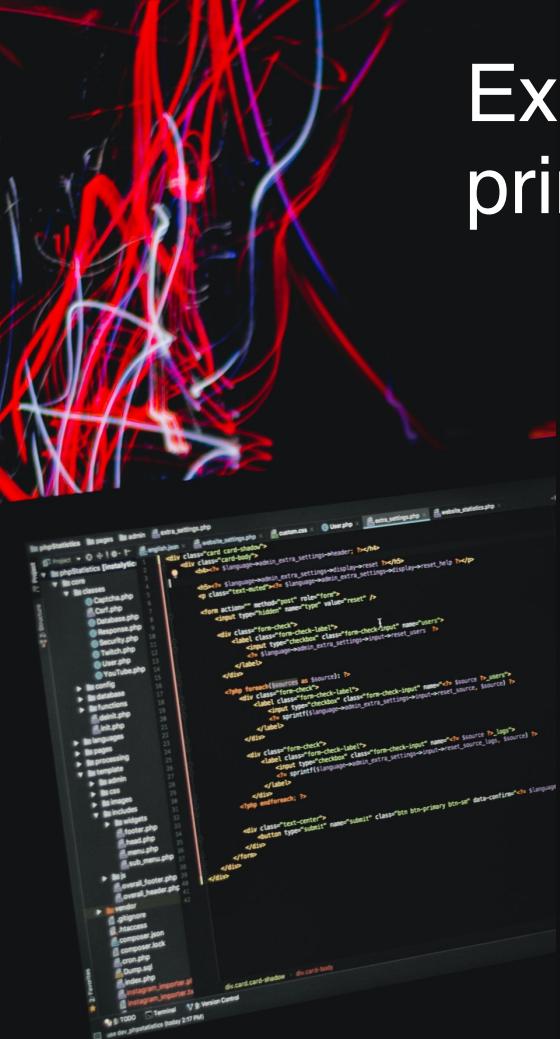


```
1 r_data = {i_exchange: {} for i_exchange in exchanges}
2
3     async def async_client(exchange, symbol, data_type):
4
5         client = getattr(ccxt_async, exchange)({'enableRateLimit': True})
6         await client.load_markets()
7
8         # Check for symbol support on exchange
9         if symbol not in client.symbols:
10             raise Exception(exchange + ' does not support symbol ' + symbol)
11
12
13 ● ● ●
```



```
1 exchanges = ['binance']
2 symbol = 'ETH/USDT'
3
4 # Fetch realtime orderbook data until timer is out (60 secs is default)
5 orderbooks = dt.async_data(symbol=symbol, exchanges=exchanges, output_format='inplace',
6                             data_type='orderbooks', file_route='Files/OrderBooks', stop=True,
7                             elapsed_secs=3600, verbose=2)
8
9 orderbks = pd.DataFrame()
10 y = pd.DataFrame()
11
12 for key in orderbooks['binance']:
13     print(key)
```

# Exploración de Datos y primeras pruebas 1



- Distribución de precios y volúmenes
- "Resampling" (Intervalos de tiempo)
- Creación de variable objetivo ("close - open" price)
- Features: Variables de función de costo
  - Mid\_price
  - Spread
  - Vol Imbalance
  - Weighted mid Price
  - VWAP
- Graficas de series y primera validación

# Exploración de Datos y primeras pruebas 2

- Modelo OLS para ver comportamiento de los datos en un modelo sencillo
  - Mid Price vs ask high
  - Vwap vs ask close
- Grafica de correlación
- Analisis de serie de tiempo para la variable target
  - Acf, pacf, df
  - En primero y segundo orden
  - Modelo ARIMA de prueba.



```
<?php
    // ...
    $this->_data['extra_settings'] = [
        'help' => [
            'label' => 'Help',
            'type' => 'button',
            'value' => 'reset'
        ],
        'reset' => [
            'label' => 'Reset',
            'type' => 'button',
            'value' => 'reset'
        ]
    ];
    // ...
}>
```

# Code Transitions



```
1 # Target
2 data_ask_bid['Target'] = data_ask_bid[(
3     'Ask', 'close')] - data_ask_bid[('Ask', 'open')]
4
5 # Features
6
7 # MidPrice
8 data_ask_bid['mid_price'] = (
9     data_ask_bid[('Bid', 'high')] + data_ask_bid[('Ask', 'low')])
10
11 # Spread
12 data_ask_bid['spread'] = abs(
13     (data_ask_bid[('Bid', 'high')] - data_ask_bid[('Ask', 'low')]))
14
15 # Volume Imbalance
16 data_ask_bid['Vol_Imbalance'] = abs(
17     data_ask_bid[('Volume', 'bid_vol')] - data_ask_bid[('Volume', 'ask_vol')])
```

# Retos

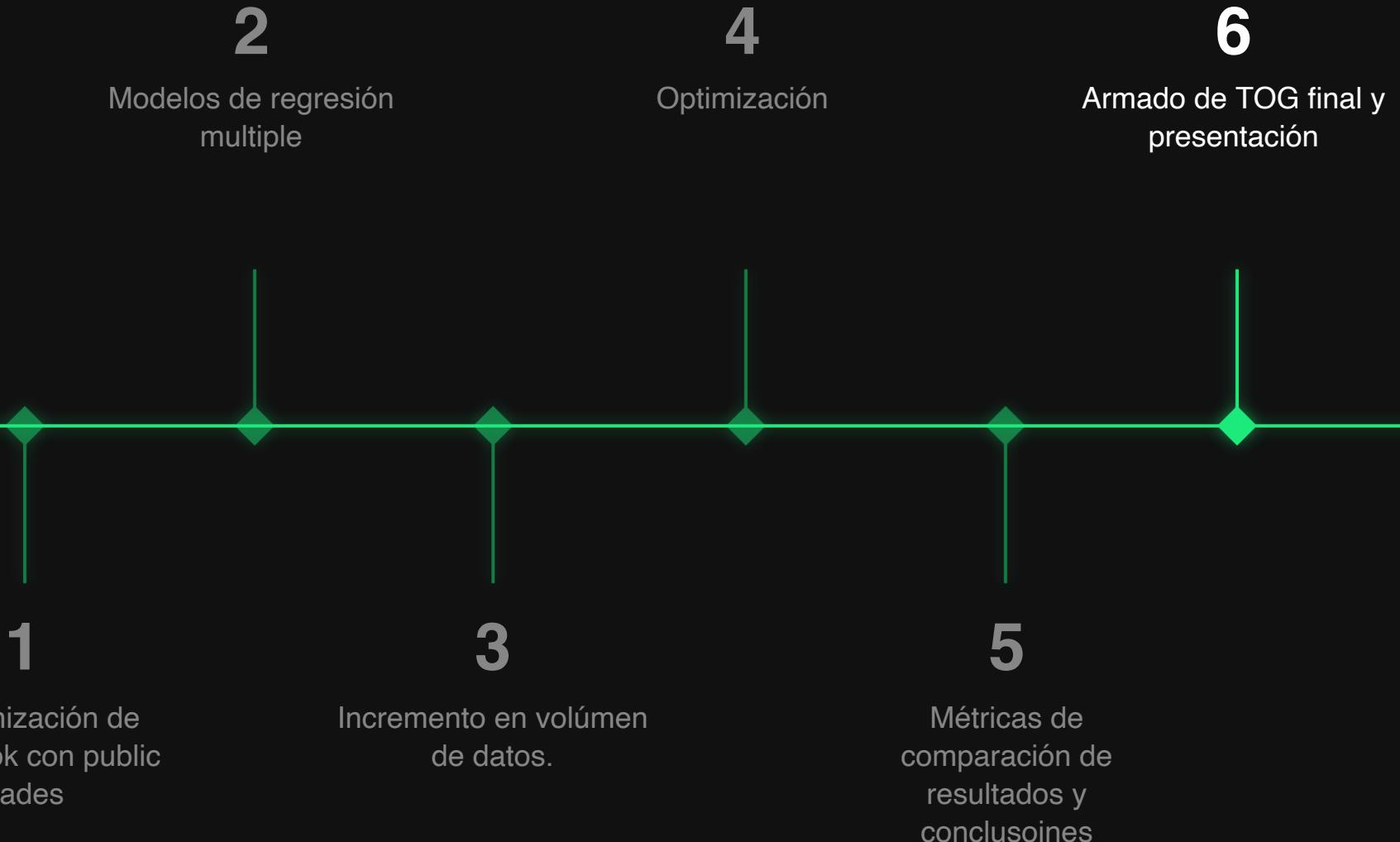
- Novedad del tema y desconocimiento
- Código de obtención de datos y manejo de altos volúmenes de información
- Administración de tiempo





# Cambio de ruta

- Una materia en verano para dejar una sola en otoño e IDI4
- Mejor comunicación con mi asesor
- Administración de tiempo





Les agradezco de antemano su tiempo y disponibilidad para esta revisión, espero poder darles un panorama general de mi avance y próximos pasos en mi trabajo, estaré atento también a sus recomendaciones y sugerencias para incrementar la calidad del mismo.





# Get in touch

<https://github.com/lagrom/TOG>

luis.guerrero@iteso.mx

luisalfonsoguerero@gmail.com

33 1329 2556



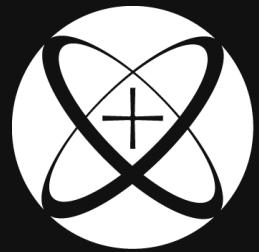
Screenshot of a GitHub repository page for 'lagrom / TOG' (Public). The repository was generated from IFFranciscoME/python-project.

Code Issues Pull requests Actions Projects W

master 1 branch 0 tags

lagrom Narrativa

File	Description
files	Initial commit
tog_venv	Last Version
.gitignore	testing git ignore
LICENSE	Initial commit
Narrativa_guia.docx	Narrativa
Preparation on data handling.ipynb	uodate 2022-04-10
Pruebas.py	Avances al 2022_04_30
README.md	Initial commit
TOG_1.py	Avances al 2022_04_30
TOG_2_EDA.py	Avances al 2022_04_30
colab.ipynb	Initial commit
data.py	Extraccion Funcional
dictionaries.py	uodate 2022-04-10



ITESO, Universidad  
Jesuita de Guadalajara