

ir a <https://finance.yahoo.com/quote/COP%3DX/history?p=COP%3DX> y descargar el csv

ir a s3 y guardar el csv en el bucket "dolar3"

crear una base de datos en aws glue "dolar3"

crear un crawler para tener la estructura del csv como tabla "dolar3", debe buscar el csv del s3 dolar3

se corre el crawler

se hace una consulta en athena, para obtener la desviacion estandar movill, como una columna, el preceding de 19 se debe a que las franjas de bollinger se hallan en periodos de 20



```
SELECT date,close,STDDEV_POP(close) OVER (ORDER BY date ROWS BETWEEN 19
PRECEDING AND CURRENT ROW) AS std
FROM "AwsDataCatalog"."dolar3"."dolar3";
```

Se crea la tabla con la nueva columna llamada dolar final3

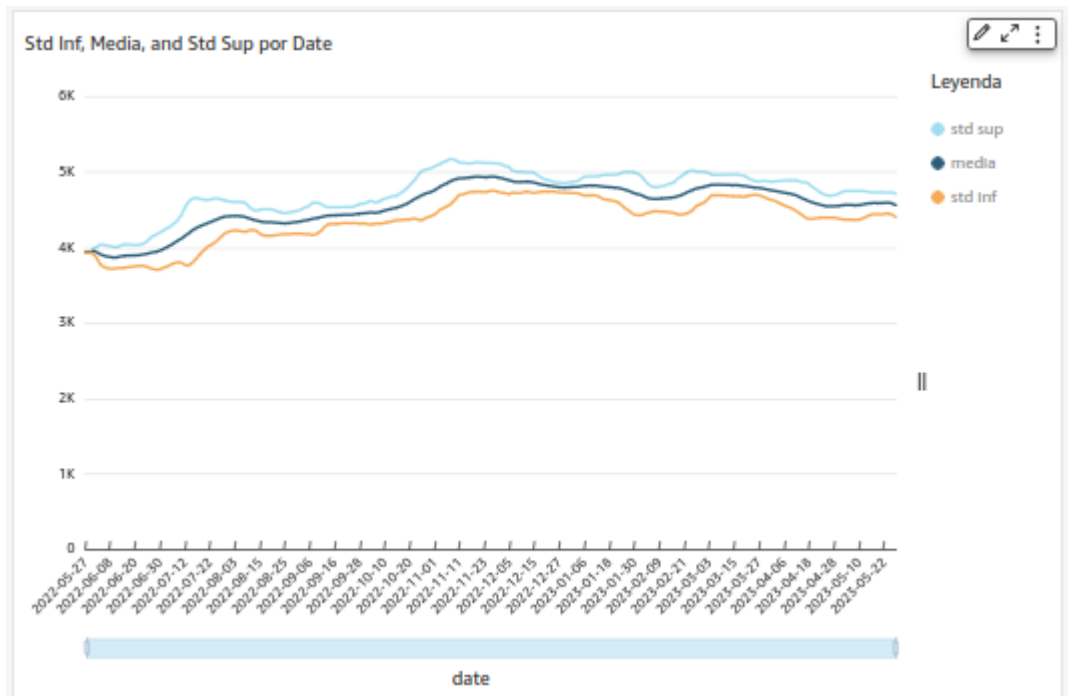
ir a quicksight, nuevo analisis, nuevo conjunto de datos, athena, se elige la tabla dolarfinal3, se visualiza y se hallan las franjas de bollinger y la media movil, en los campos calculados

Media = `windowAvg(sum(close), [date ASC], 19, 0)`

std inferior: `{media} - (sum({std}) * 2)`

std superior:  $\{media\} + (\sum(\{std\}) * 2)$

Se hace el grafico de lineas en el eje x las fechas y en los valores, es stdinf, sup y la media



Segundo punto en EMR

se clona un cluster y se crea un jupyter

## Crear un bloc de notas

### Asigne un nombre al bloc de notas y configúrelo

Asigne un nombre al bloc de notas de Jupyter administrado por EMR, elija o cree un clúster y personalice las opciones de configuración, si lo desea. [Más información](#)

Nombre del bloc de notas\*

Los nombres solo pueden contener letras (a-z), números (0-9), guiones (-) o caracteres de subrayado (\_).

Descripción

256 caracteres como máximo.

Cluster\* ☒ Elegir un clúster existente

Mi clúster [j-UAMP24QG75E0](#)

☐ Crear un clúster

Grupos de seguridad ☒ Usar los grupos de seguridad predeterminados

☐ Choose security groups (vpc-0eb3357806c25ae7f)

Rol de servicio de AWS\*

Ubicación del bloc de notas\* Choose an S3 location where files for this notebook are saved.

☒ Use the default S3 location  
s3://aws-emr-resources-729788758200-us-east-1/notebooks/

☐ Choose an existing S3 location in us-east-1

Se toma el csv de las noticias del tiempo, y se crea el notebook que está en la carpeta. Para ver los resultados, se guardan en el mismo bucket en formato de parquet, al abrirlo, se puede hacer una consulta s3 en json, y se ven los resultados.

se guarda en ("s3://noticias-parcial99/headlines/noticiasToken/")

```
"_c1": "economia",
"_c2": "/economia/sectores/los-entresijos-de-la-eleccion-de-bahamon-en-la-fnc-763519",
"label": 3,
"words": [
  {
    "element": "economia"
  }
],
"rawFeatures": {
  "indices": [
    {
      "element": 253015
    }
  ],
  "size": 262144,
  "values": [
    {
      "element": 1
    }
  ],
  "type": 0
},
"features": {
  "indices": [
    {
      "element": 253015
    }
  ],
  "size": 262144,
  "values": [
    {
      "element": 1.8562979903656263
    }
  ],
}
```

3. Se crea un kinesis y se vincula en cloud9, solo con el nombre en kinesis, se crea el código del productor y del consumidor, ambos productores son los mismos, solo cambia el consumidor inferior y superior, ya que uno alerta cuando está por debajo de la franja(-) y el otro cuando está encima de la franja (+)

Se debe ejecutar primero el productor y luego el consumidor ya que el consumidor recibe los datos del productor

En total hay 4 entornos 1 para cada código

en kinesis de secuencia de datos se llama parcial3

<https://docs.aws.amazon.com/kinesisanalytics/latest/java/gs-python-createrecordapp.html> Write Sample Records to the Input Stream .(Hasta ejecutar el producer.py) como hacer lo de kinesis, con el código, eso lo subió el profesor al classroom