Análisis de Datos

*Proyecto final*

Modelos de predicción para la tasa de Cambio Representativa del Mercado (TRM)

Jose Alexander Ríos Trespalacios

Estefany Alejandra Jaimes León

María Celeste Zuluaga Giraldo

# Universidad de Medellín Facultad de Ingeniería

Ingeniería de Telecomunicaciones

# 2022



**Introducción**

La Tasa de Cambio Representativa del Mercado (TCRM), corresponde al promedio ponderado de las operaciones de compra y venta de contado de dólares de los Estados Unidos de América a cambio de moneda legal colombiana.

Este documento modela los tipos de cambio anuales entre USD/CO, y compara los datos reales con pronósticos desarrollados utilizando análisis de series de tiempo durante el período de 1992 a 2022. Los datos semanales oficiales del Banco Nacional de la República de Colombia se utilizan para el presente estudio. El objetivo principal de este documento es aplicar el modelo ARIMA para la previsión de tipos de cambio semanales de USD/CO. La precisión del pronóstico se compara con el error absoluto medio (MAE) y el error cuadrático medio (MSE).

1. **Entendimiento de negocio**

Esta etapa se refiere a la predicción del precio de dólar. En esta etapa se necesita una comprensión de los antecedentes y objetivos de los procesos relacionados con la bolsa de valores, incluyendo:

1. **Establecer objetivos comerciales**

* El objetivo comercial de hacer investigación es reconocer el precio del dólar de meses anteriores para determinar la predicción de la moneda en la posteridad para que se puedan tomar políticas y acciones si las predicciones no son apropiadas y conocer la relación de otros factores que afectan su valor.

1. **Evaluar la situación**

* El precio del dólar con respecto a la moneda colombiana está influenciado por factores internos y externos. Factores internos, exportaciones, el precio del petróleo y factores externos, inversionistas extranjeros y los capitales golondrina.

1. **Determinar el propósito de la minería de datos**

* El propósito de la minería de datos o el propósito de este estudio es explorar el conocimiento sobre los patrones de influencia de los factores externos en el precio del dólar con respecto a al peso colombiano para la predicción del valor futuro.

1. **Preguntas que el modelo responderá**

* ¿Puedo predecir el precio del dólar con respecto al peso colombiano con los datos de hace una semana?
* ¿Cuántos datos debería entrenar para poder obtener el precio del dólar del día siguiente?
* ¿Cuál será el precio del peso colombiano frente al dólar mañana y en la próxima semana?
* ¿Es buen momento para hacer negocios en esta divisa?

1. **Riesgos**

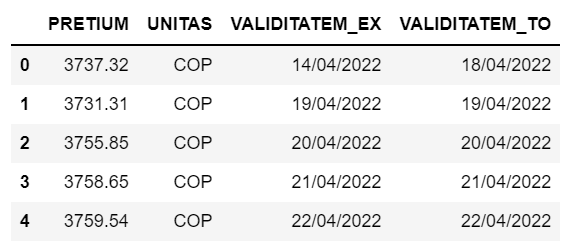
* Los datos no presentan riesgos, pues son extraídos de una fuente confiable, este sistema lo establece el mercado de divisas sobre la oferta y la demanda del peso colombiano en particular en relación con el dólar estadounidense. Además, el tipo de cambio se guía por el impacto significativo de las actividades de los bancos centrales y otras instituciones financieras.

1. **Entendimiento de datos**

* En los datos obtenidos existen pocas variables, pues solo se requiere la fecha y el valor del dólar actual para poder predecir el valor futuro, es difícil que tengamos datos faltantes, ya que los datos se actualizan constantemente (lunes - viernes) y es imposible que haya algún error en la toma de datos.
* Los datos de series temporales que estamos utilizando se pueden encontrar en https://www.datos.gov.co/Econom-a-y-Finanzas/Tasa-de-Cambio-Representativa-del-Mercado-TRM/32sa-8pi3. Los datos históricos se pueden recuperar utilizando la hoja de Excel ingresando la abreviatura de moneda estándar, fecha de inicio y fecha de finalización. Usando la API anterior, hemos recopilado datos de los últimos 30 años a partir de 1992 para el dólar con respecto al peso colombiano.
* La fecha se entregó en un formato distinto al "datetime", por lo tanto, debemos realizar las conversiones necesarias para poder poner este formato estándar en todas las fechas, luego, esta fecha será el índice de nuestra tabla para poder graficar bien nuestros datos.
* Para poder modelar se necesita también ajustar el periodo de las fechas, ya que como los datos son entregados de lunes a viernes, y el modelo no responde a este orden, porque se necesita una secuencia uniforme. Por lo tanto, se reorganiza de manera semanal para que puedan ser modelados.

1. **Preparación de los datos**

* La etapa de preparación de datos es la preparación del conjunto de datos, los datos utilizados para el modelado. En esta etapa se selección y construyen los datos, se elige la tabla relacionada para simplificar el proceso de selección de datos. Esta tabla contiene el valor del peso colombiano frente al dólar estadounidense con su respectiva fecha de registro. Contiene 7431 datos con variaciones, se re-muestrean para que queden con una frecuencia semanal.



1. **Modelación**

**Construcción del modelo ARIMA**

Recordemos que los hiperparámetros de un modelo ARIMA son (p,d,q), donde p es para la parte "AR" (cuántos rezagos incluir), d es para la parte "I" (cuántas veces integrar o diferenciar la serie) y la q es para la parte "MA" (cuántos medias móviles del error se incluirán).

Se tuvo que diferencia la serie para poder volverla estacionaria, en el modelo ARIMA para aplicar este cambio la "d" será igual a 1. Nos falta determinar entonces valores para la p y la q.



Este gráfico es más "conservador": intenta modelar muy bien la mayoría de los puntos, incluso en los picos altos.

Podemos por ejemplo ver que:

- Se tienen datos de forma mensual, desde 1992 a 2022, para el valor diario del precio del peso colombiano frente al dólar estadounidense.

- Se nota que hay una tendencia creciente clara en los datos, por lo tanto, no es estacionaria (pero se puede convertir).

- Se nota claramente que las observaciones están relacionadas entre sí, es decir, hay dependencia (común en una serie de tiempo).

- Hay outliers evidentes en la serie de datos, sin embargo, son congruentes, ya que le dólar en cualquier momento puede alcanzar un punto máximo y volver a caer, no afecta el modelado y predicción de datos.

**Construcción del modelo SARIMA**

Claramente habíamos notado que la serie tiene una estacionalidad. Por dicha razón, incluyamos entonces el modelo SARIMA para los efectos estacionales.

Agreguemos dos rezagos a los efectos estacionales.

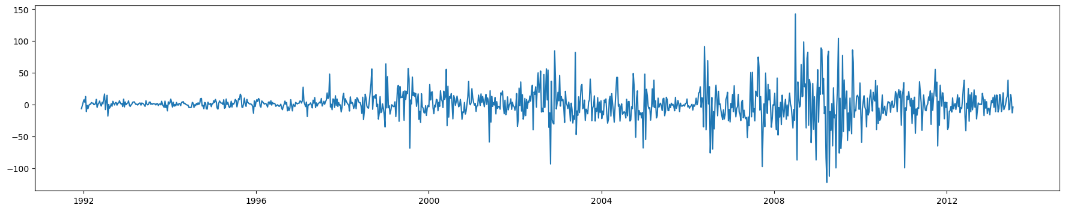
****

1. **Evaluación**

Podemos observar que el modelo que más exacto dio, fue el ARIMA, por lo tanto, seguiremos con este, sin embargo, sus resultados fueron muy parecidos, procedemos a realizar la predicción de los datos futuros.

Ya seleccionamos nuestro primer ARIMA como el mejor modelo. Probemos que sus residuos si sean ruido blanco y validemos sus métricas finales en el conjunto de prueba.

Empecemos primero verificando: los residuos de un modelo ARIMA deben ser ruido blanco. Si no lo son, no cumple lo que dice la estadística, por lo que lo ideal es reajustar el modelo con otros **hiperparametros.**

****

**p = 2**

**d = 1**

**q = 2**

1. **Despliegue**

Generamos a futuro las predicciones del modelo ARIMA, con un intervalo de confianza del 95%, para los siguientes meses a futuro:



**Conclusión**

* En este proyecto la técnica ARIMA fue la que se usó para la modelación, fue mejor esta que la técnica SARIMA, en todos los casos no será así, por eso se deben probar diferentes técnicas para el modelado de los datos.
* La técnica ARIMA para pronosticar los tipos de cambio de divisas del peso colombiano frente al dólar estadounidense (USD), se aplicó durante el período de 1992 a 2022. Se utilizó el software Python para la predicción. de los tipos de cambio. Se presentó la técnica ARIMA y se identificaron tres pasos principales para construir el modelo, a saber, Identificación, Estimación y Verificación del modelo. Además, se estimó el modelo de pronóstico y se comparó con los datos reales de la moneda. La eficacia de los resultados del modelo de pronóstico se comparó con el error absoluto medio (MAE) y el error cuadrático medio (MSE).