

# **PRODUCTOS 'ZTE'**

## **APRENDIZAJE DE MÁQUINA**

### **PRESENTAN:**

<b>Alejandra Santillan Palos</b>	<b>A01365626</b>
<b>Eda Eraldí Martínez Vázquez</b>	<b>A01365784</b>
<b>Justine Gutiérrez Mendoza</b>	<b>A01365878</b>
<b>José Luis Estrada Bastida</b>	<b>A01365811</b>
<b>Guillermo Martínez Ávila</b>	<b>A01365593</b>



ZTE es una marca china que busca posicionarse en el top 3 en el mercado en Estados Unidos y México, busca posicionarse de gran forma en México ya que sabe que así podrá tener una gran entrada en el mercado latino americano.

Sistema operativo de android

En 2016 la marca contaba con el 9% del mercado mexicano en ventas colocando en la cuarta posición

## ¿QUIENES SON ZTE?



# ¿CUÁL ES NUESTRO PROBLEMA?

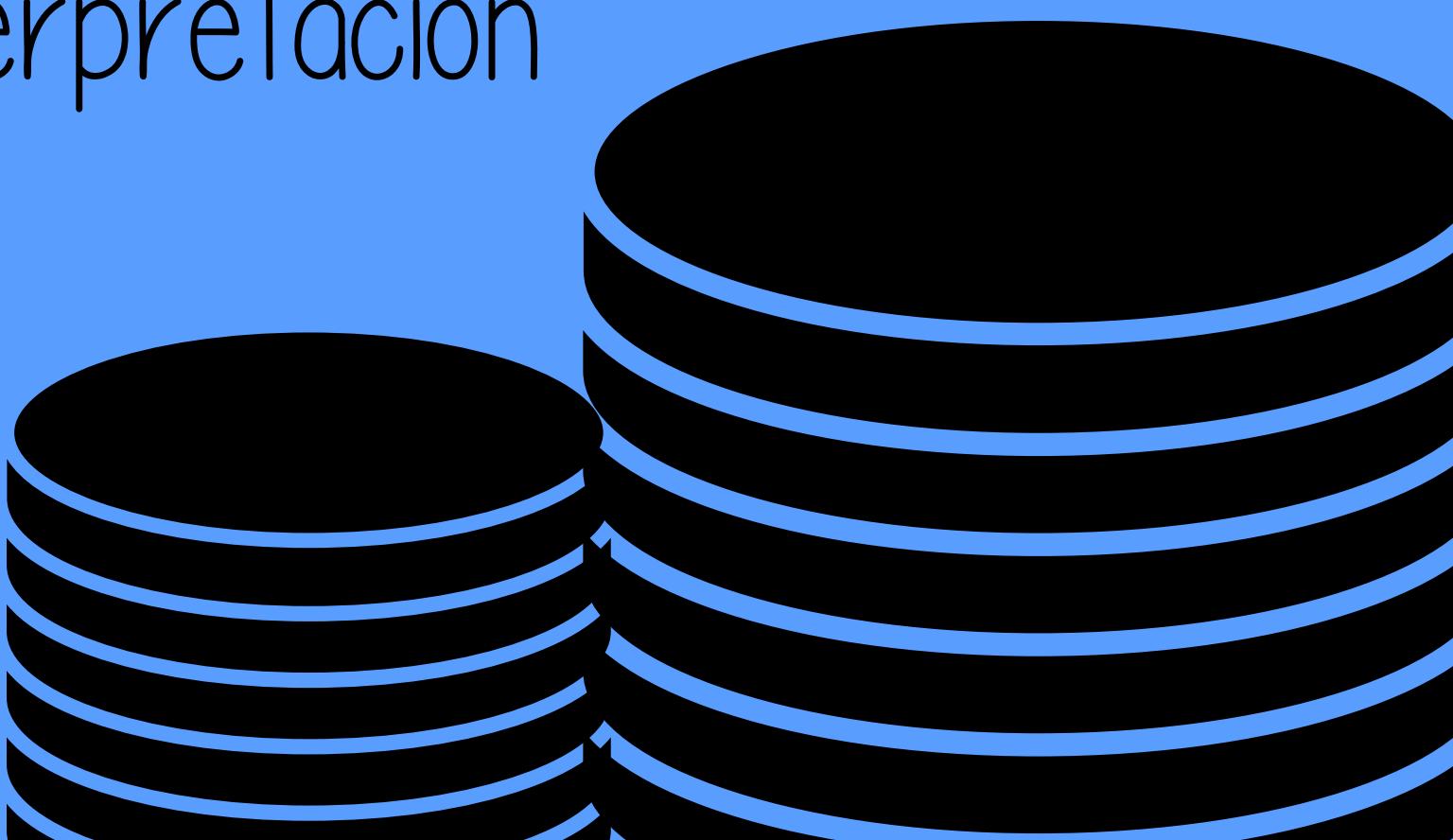


Principal problemática de esta empresa es el poder construir un eficiente portafolio de productos, mediante una predicción de la demanda.

Los rubros más importantes son el manejo de inventarios y el transporte, de ahí la necesidad de tener buenos pronósticos de demanda, ya que este pronóstico de la demanda afecta directamente la planificación estratégica del negocio ya que involucra la administración de la cadena de suministro, la compra de materias primas y la distribución de pedidos.

SINO TENEMOS UN PRONOSTICO DE DEMANDA EFICIENTE, PODRIAMOS ENFRENTAR...

- Disminución de la satisfacción del cliente: Datos
  - Mala planeación de materias primas y recursos: Limpieza de datos
  - Altos costos por inventarios: Cálculos
  - Deficiente planificación de la estrategia de ventas Interpretación
  - Falta de comunicación con proveedores y departamento de compras en cuanto a ticket de producción:



# NUESTROS OBJETIVOS

## Objetivo para la Empresa

El objetivo principal que buscamos para la empresa es el de construir un eficiente portafolio de productos mediante la predicción de la demanda de sus dispositivos móviles, esto provocará una disminución de costo de inventarios, transporte y logística, una mejor planeación del ticket de producción y compra de las materias primas, mejor comunicación con los demás departamentos y proveedores y sobre todo una excelente satisfacción al cliente.

## Objetivo para la materia

Ser capaces de aplicar los conocimientos que adquiramos durante el semestre en este proyecto de una forma dinámica, para poder poner en práctica la ciencia de datos y todo lo relacionado con ella en este proyecto, y también para ser capaces de utilizar dichos conocimientos en el resto de nuestra carrera y también para la aplicación en la vida profesional.

## Objetivo Personal

Utilizar con mayor profundidad de una manera práctica la ciencia de datos para poder responder las preguntas que se realicen al principio y así con los datos obtenidos, que ayuden a la toma de decisiones. También el poder aprovechar, identificar y utilizar las diversas herramientas, habilidades, tipos de tareas y conocimientos que se tengan que utilizar en el proceso.

# ETAPAS A DESARROLLAR

1

Plantear el problema  
(términos de ciencia de  
datos de negocio)

2

Familiarizarse con los  
datos, relacionarlos con  
nuestra principal  
problemática

3

Construcción del Dataset  
Final, seleccionar, limpiar,  
construir, integrar y dar  
formato a nuestra  
información.

4

Definir y realizar  
nuestras técnicas de  
modelado para abarcar  
nuestro principal  
problema.

5

Evaluar, revisar nuestros  
resultados y determinar  
cuales serían nuestros  
siguientes pasos

6

Representar el  
conocimiento adquirido,  
con un plan de monitoreo  
y mantenimiento,  
presentar nuestro  
reporte final

# ETAPA 2: DATOS

punto\_de\_venta —> Tipo de variable: Character  
fecha —> Tipo de variable: Character.  
mes —> Tipo de variable: Factor.  
anio —> Tipo de variable: Factor.  
num\_ventas —> Tipo de variable: Integer  
sku —> Tipo de variable: Factor.  
marca —> Tipo de variable: Character.  
gamma —> Tipo de variable: Character.  
costo\_promedio —> Tipo de variable: Numeric.  
zona —> Tipo de variable: Factor.  
estado —> Tipo de variable: Factor.  
ciudad —> Tipo de variable: Factor.  
latitud —> Tipo de variable: Numeric.  
longitud —> Tipo de variable: Numeric.



```
if _operation == "MIRROR_X":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = True  
    mirror_mod.use_z = False  
elif _operation == "MIRROR_Z":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = False  
    mirror_mod.use_z = True  
  
#selection at the end -add back the deselected mirror modifier object  
mirror_ob.select= 1  
modifier_ob.select=1  
bpy.context.scene.objects.active = modifier_ob  
print("Selected" + str(modifier_ob)) # modifier ob is the active ob  
    #mirror_ob.select = 0  
bone = bpy.context.selected_objects[0]  
bone.modifiers.new(name="mirror", type='MIRROR')  
    bone.modifiers["mirror"].use_x = False  
    bone.modifiers["mirror"].use_y = True  
    bone.modifiers["mirror"].use_z = False  
    bone.modifiers["mirror"].mirror_offset = 0  
    bone.modifiers["mirror"].mirror_min = -1  
    bone.modifiers["mirror"].mirror_max = 1  
    bone.modifiers["mirror"].use_clipping = False  
    bone.modifiers["mirror"].use_recrease = False  
    bone.modifiers["mirror"].use_recenter = False  
    bone.modifiers["mirror"].use_bounce = False  
    bone.modifiers["mirror"].use_cyclic = False  
    bone.modifiers["mirror"].use_axis = False  
    bone.modifiers["mirror"].axis = 'Z'
```

**Detectar problemas de calidad**

# ETAPA 3: PREPARACIÓN DE NUESTROS DATOS

24,089 registros y  
14 variables

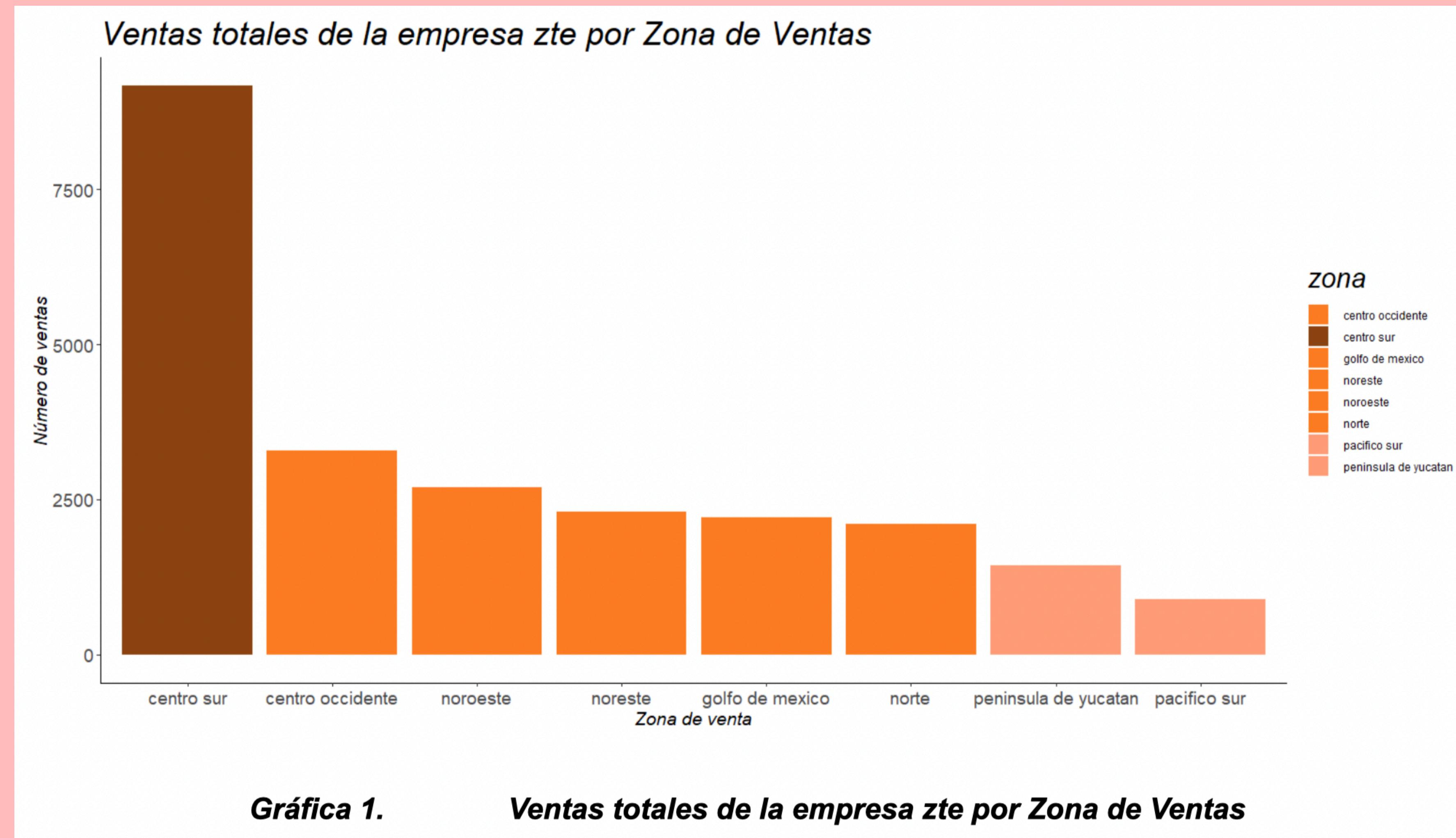
## Limpieza de Datos

Pudimos encontrar algunos problemas de calidad de los datos y variables con errores, durante esta actividad pudimos corregir dichas falta, entre las cuales se encontraron:

- ▶ En la primer variable de puntos de venta se encuentran problemas de que algunos puntos son escritos en mayúsculas y en otras en minúsculas
- ▶ En el caso de la variable de mes hay algunos meses que son escritos en letra y otros escritos en número.
- ▶ La variable de marca viene escrita de 6 diferentes formas, en esta caso mayúsculas, minúsculas, con doble t, escrita 2 veces en mayúscula y en minúscula, y escrita con doble Z.
- ▶ En la variable de zona la zona centro occidente está mal escrita, occidente no tiene i.
- ▶ Para la variable de estado hay varios que no son estados como por ejemplo Monterrey que es una ciudad en Nuevo León.
- ▶ Para las últimas variables de longitud y latitud hay errores de valores fuera de rango. Se identificaron los valores fuera de rango con la función.

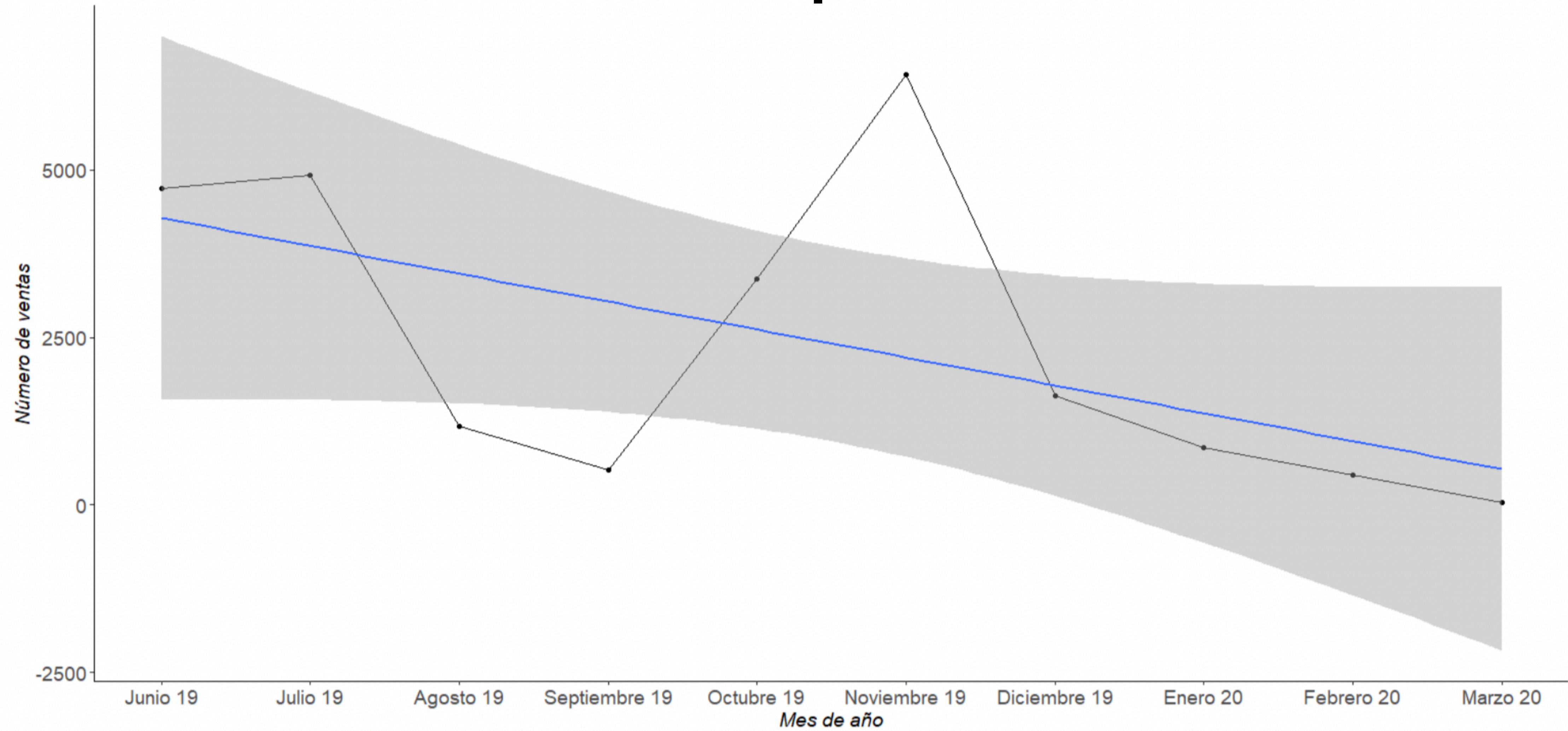
# Análisis Exploratorio de la situación actual

El objetivo consiste en comprender el comportamiento regular de los datos a través de su análisis y encontrar si hay relaciones entre ellos.



*Ventas totales por mes*

# Análisis Exploratorio de la situación actual

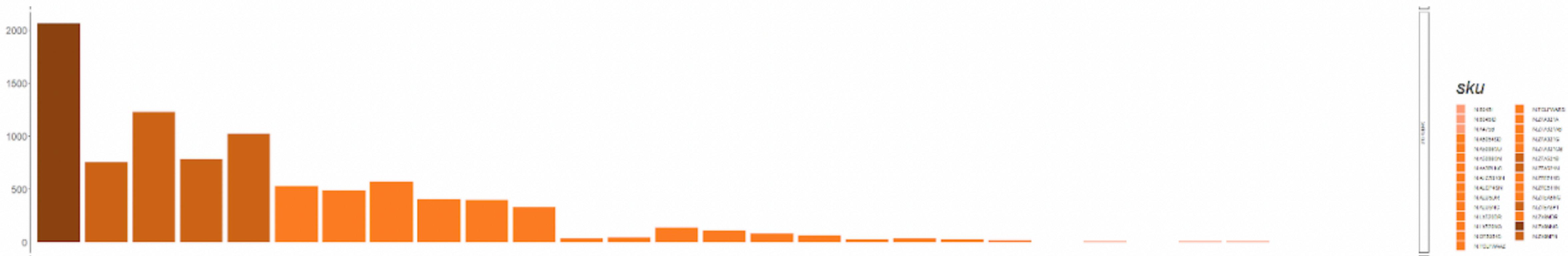


**Gráfica 2.**

*Ventas totales por mes de la empresa zte, de junio 2019 a marzo 2020*

# Análisis Exploratorio de la situación actual

## Ventas totales de la empresa zte por SKU y zona de ventas



*Extracto de la Gráfica 3. Ventas totales de la empresa zte por SKU y Zona de Ventas.  
Para consultar la gráfica completa, de favor recurrir a los anexos del presente documento*

# Selección y construcción de variables para la etapa de modelado

Con la ingeniería de características debemos ver cuales son las variables de mayor importancia para nuestro proyecto y crear las variables adicionales necesarias para poder proseguir con la siguiente etapa la cual consiste en el modelado de los datos.

¿Cuántas unidades de cada producto de mi marca, se van a vender en todos los puntos de venta al siguiente mes de registro?

ventas\_totales, ventas\_totales\_tienda\_y\_mes

ventas\_promedio\_tienda\_y\_mes

ventas\_totales\_tienda\_y\_sku

ventas\_promedio\_tienda\_y\_sku,

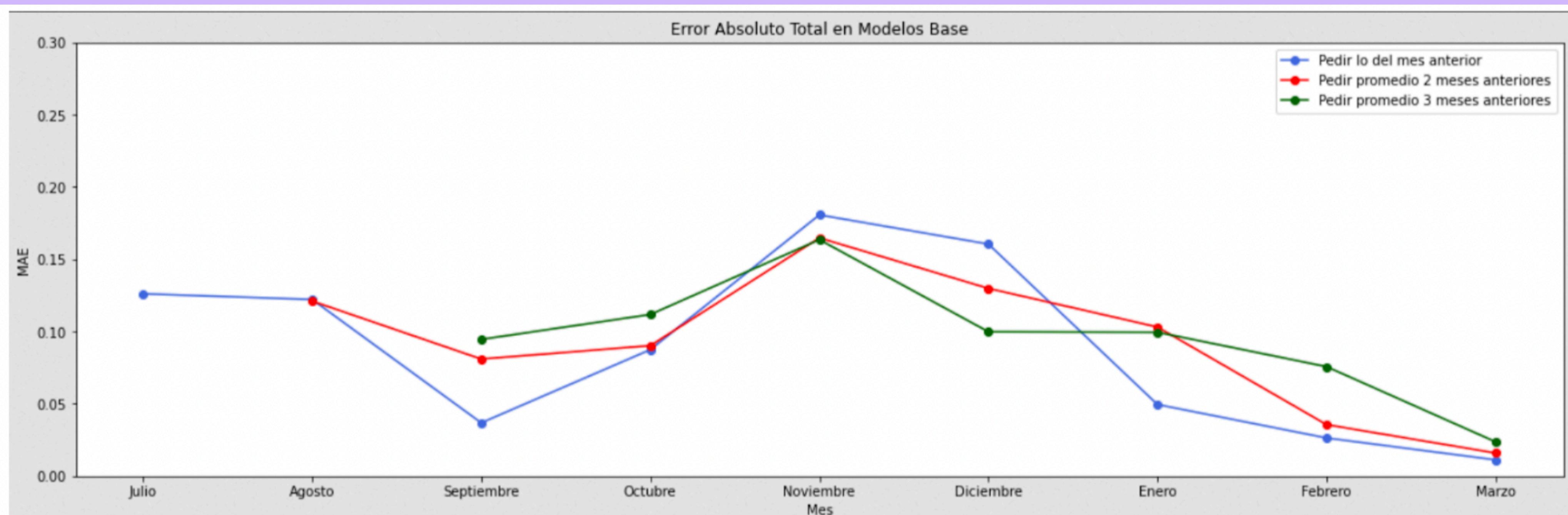
Para poder generar los pronósticos agregamos las columnas de 1, 2 y 3 meses pasados para las variables

Ingeniería de características, concluimos un Dataset, “Datos completos”, el cual cuenta con 380,277 registros y 24 variables

Construir el modelo de promedios móviles, cambiando 2 veces el periodo móvil

# ETAPA 4: MODELADO

¿Cuántas unidades de cada producto de mi marca, se van a vender en todos los puntos de venta al siguiente mes de registro?



Gráfica 4.

Error absoluto Total en Modelos Base

Concluimos que el mejor modelo podría ser el promedio móvil de un mes, es decir la opción de pedir lo del mes anterior

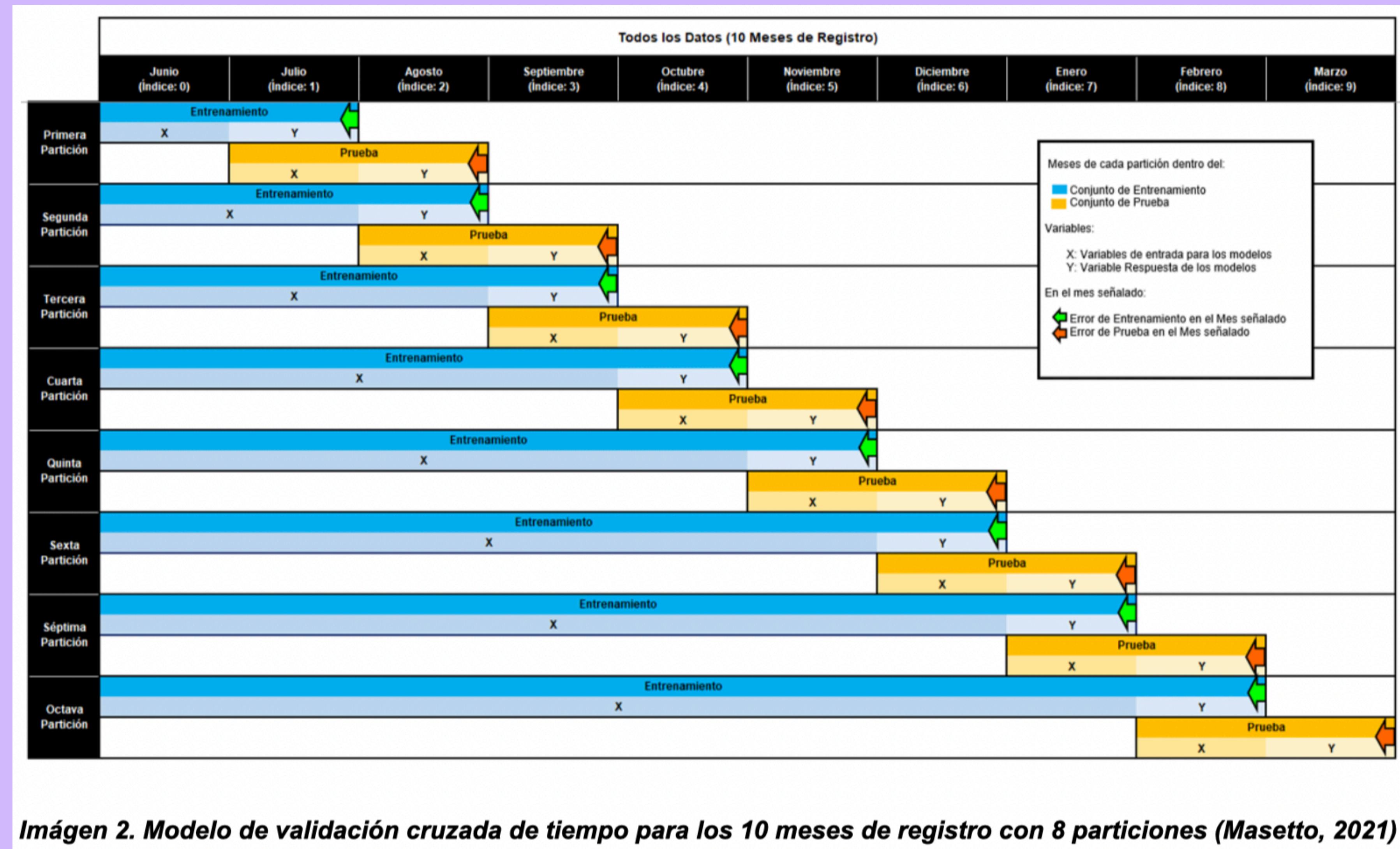
# Construir el modelo de aprendizaje de máquina.

Nos encontramos con un problema de aprendizaje supervisado ya que queremos conocer la variable de salida “y”(ventas siguiente mes por tienda) con base en varias variables de entrada “x” (ventas totales) y al ser la variable “y” del tipo numérico se traduce también en un problema de regresión. Dividiremos los datos en 8 conjuntos de entrenamiento que nos ayudarán a aprender de los datos, no memorizar, y 8 conjuntos de prueba para probar nuestro modelo a datos no antes vistos. También usaremos una validación cruzada de tiempo para evitar el sobreajuste, tomando en cuenta una secuencia periódica respecto al tiempo.

El modelo de máquina que seleccionamos fue el de árbol de decisión, para predecir el valor de una variable en un problema de aprendizaje supervisado de regresión.

Posteriormente procedimos a realizar la preparación de los datos con una validación cruzada de tiempo con 8 particiones para los conjuntos de entrenamiento y prueba de acuerdo a la cantidad de de datos que tenemos y por cada una de estas particiones se llevará a cabo un modelo

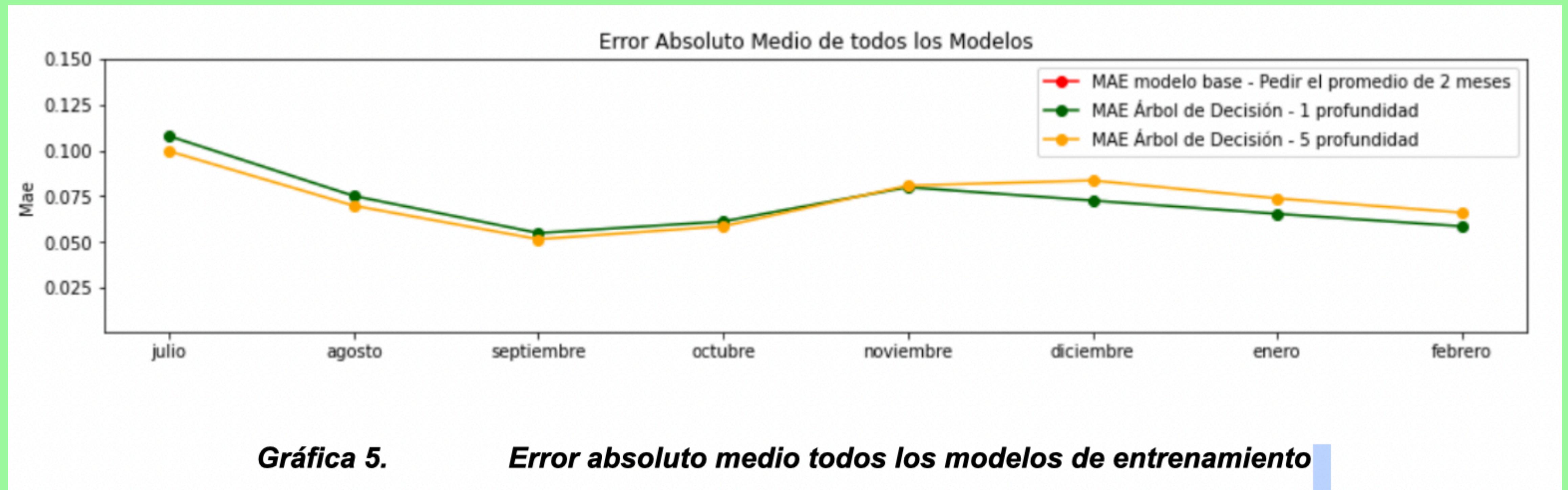
Para realizar la primera participación, designamos los datos del primer mes con información, junio, como “ $x_1$ ” de entrenamiento para estimar el siguiente mes, julio, como “ $y_1$ ” de entrenamiento. De igual manera designamos los datos de julio como “ $x_1$ ” de prueba, para estimar el siguiente mes agosto como “ $y_1$ ” de prueba

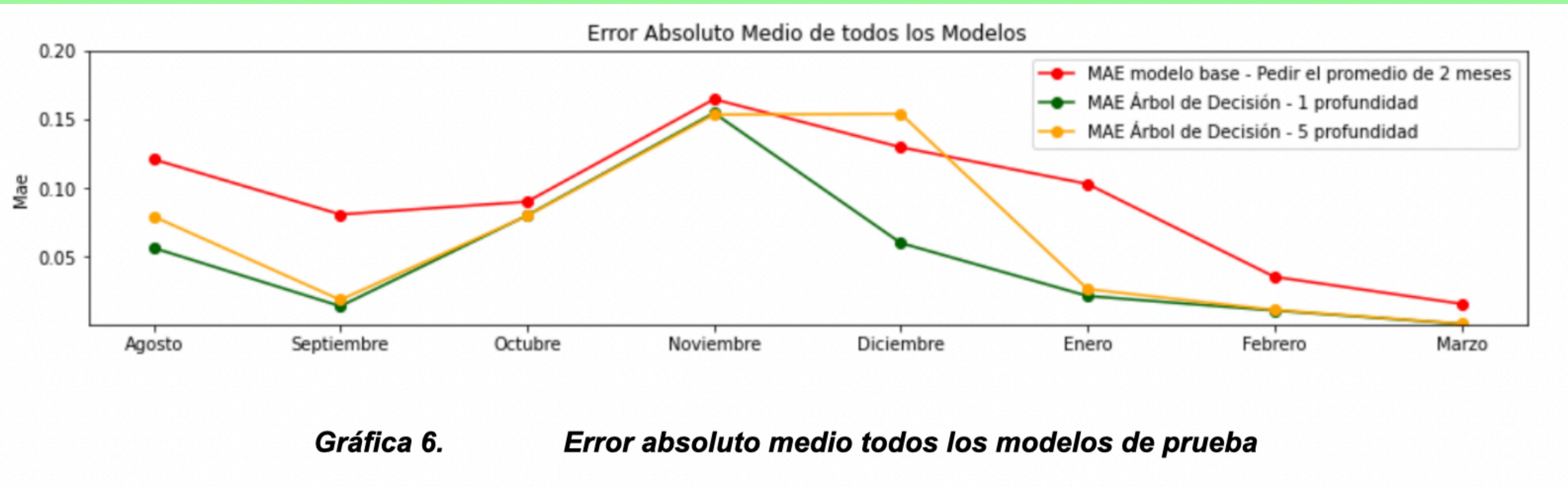


Optamos por probar el árbol de decisión con una diferente profundidad para poder comparar entre estos modelos y encontrar la mejor solución, por lo que repetimos todo el proceso desde de la etapa de modelado, cambiando solamente la profundidad del árbol de 1 a 5.

# ETAPA 5: EVALUACIÓN

Para la evaluación de los modelos de entrenamiento únicamente se pueden observar los errores de los árboles de decisión ya que para el modelo de promedios móviles simples no fue necesario un conjunto de entrenamiento para ser realizado.





Así que el mejor modelo para predecir la demanda del siguiente mes sería un modelo árbol de decisión con profundidad 1 ya que ofrece las mejores predicciones, esto evaluando con el indicador de desempeño MAE, por lo que es el que le recomendamos a la empresa.

# ¿Cuántas unidades de cada producto de mi marca, se van a vender en todos los puntos de venta, al siguiente mes de registro?

Así que en el mes de Abril podemos predecir que se van a vender las siguientes unidades

Punto de Venta	mes	sku	ventas_por_mes_
bca aristeum ermita	Abril	N.LX520DR	1
fgt san esteban naucalpan	Abril	N.ZTEA6PT	1
tda cdmx ecatepec las americas	Abril	N.ZTEA6PT	1
tda cdmx guadalupe fortaleza	Abril	N.ZTEA6PT	1
tda cdmx tintoreto	Abril	N.ZV8MNG	1
		N.ZV8MPN	1
tda cordoba arco	Abril	N.OT5054S	1
tda guamuchil rosales	Abril	N.OT5054S	1
tda los mochis plaza sendero	Abril	N.ZTA521N	1
tda monterrey plaza cumbres	Abril	N.ZV8MPN	1
tda nuevo laredo	Abril	N.A5098ON	1
tda san cristobal de las casas ii	Abril	N.ALC5010N	1
tda tepic los fresnos	Abril	N.ZTA521N	1
		N.ZTA521B	1
tda valle de bravo	Abril	N.ZTA321G	1
	Abril	N.ZTA521B	1
tda veracruz	Abril	N.A5098ON	1
walmart miramontes	Abril	N.LX520DR	1
		N.LX520NG	1
	Total		19

Tabla 1. Predicción de ventas en el mes de Abril

# NUESTRAS RECOMENDACIONES...

