

Materia: Gestión de Análisis y Diseño de Comercialización (COM145)

Profesor: Sarahí Aguilar González

Fecha de entrega: 24 de mayo

Ciclo: 1222

Nombre del proyecto: Índice de criminalidad en los estados de la República Mexicana

Miembros del Equipo		
ID	Nombre	Carrera
0215397	Mauricio Torres Minguez	Ing TI
0213360	Diego Gilberto Guzmán González	Ing TI
0220791	Carlos Manuel Villalobos Santiago	Ing TI

Rúbricas				
ID	2-social		7-knowledge	
	D	C	A	JI

Abstract

A lo largo de las últimas décadas, el crimen y la delincuencia en el territorio mexicano han ido en aumento, llegando a tocar valores preocupantes y a ser noticia no solo en el país, sino en todo el mundo. El proyecto tiene como objetivo el poder analizar los datos históricos de los crímenes que han sido registrados en el país, así como ofrecer con detalle las relaciones que guardan las variables con el índice delictivo, por último, se busca obtener una predicción de dicho índice para el primer mes del año del 2019, utilizando los datos históricos del año 2018.

Introducción

Entre los hechos que en las colectividades dañan, disocian o violan leyes y/o normas, hay una clase especial llamada criminalidad, que habitualmente es concebida como la que tiene consecuencias más graves. (Quiroga, 89)

A causa de los reportes y a la mayor difusión que tienen los medios informativos debido a las redes sociales, es notorio que la delincuencia en el territorio mexicano tristemente es algo cotidiano y se ha normalizado en prácticamente la totalidad del país. Año con año se han dado a conocer cifras y estadísticas peores que el año anterior, por lo que es prácticamente un hecho que la violencia y la delincuencia han ido en aumento, por lo menos en lo que va de este siglo.

Gracias a esta información y a las cifras que son de carácter público, se decidió buscar patrones sociales en los estados, para encontrar si hay una correlación entre los delitos de la región con respecto a deficiencias en la composición humana, se va a predecir los datos del 2019 referentes a los valores del 2018, con el objetivo de poder entrenar el modelo para predicciones futuras, esto dará facilidad de encontrar puntos clave para mermar en principio el índice de criminalidad para posteriormente realizar programas sociales para hacer la delincuencia mínima en los estados.

Desarrollo

Para el desarrollo del proyecto se utilizaron los siguientes x conjuntos de datos abiertos:

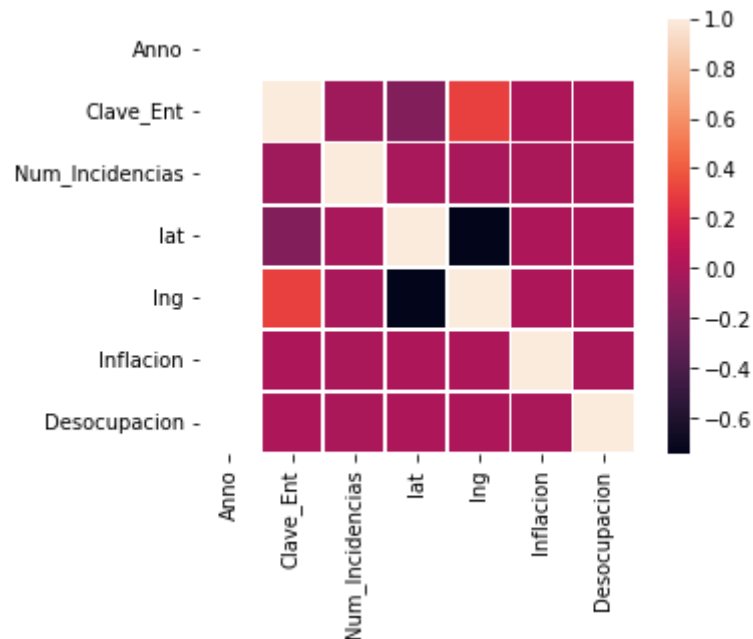
1. Set de datos con número de incidencias históricas en la República Mexicana, proporcionados por el INEGI.
2. Set de datos geográficos de la República Mexicana, proporcionados por INEGI
3. Set de datos con los índices de inflación por estado por mes
4. Set de datos con los índices de desocupación por estado por trimestre

Después de obtener todos los datos necesarios para buscar la solución, se cargó toda esta información al ambiente de desarrollo que inicialmente fue un notebook de google colab.

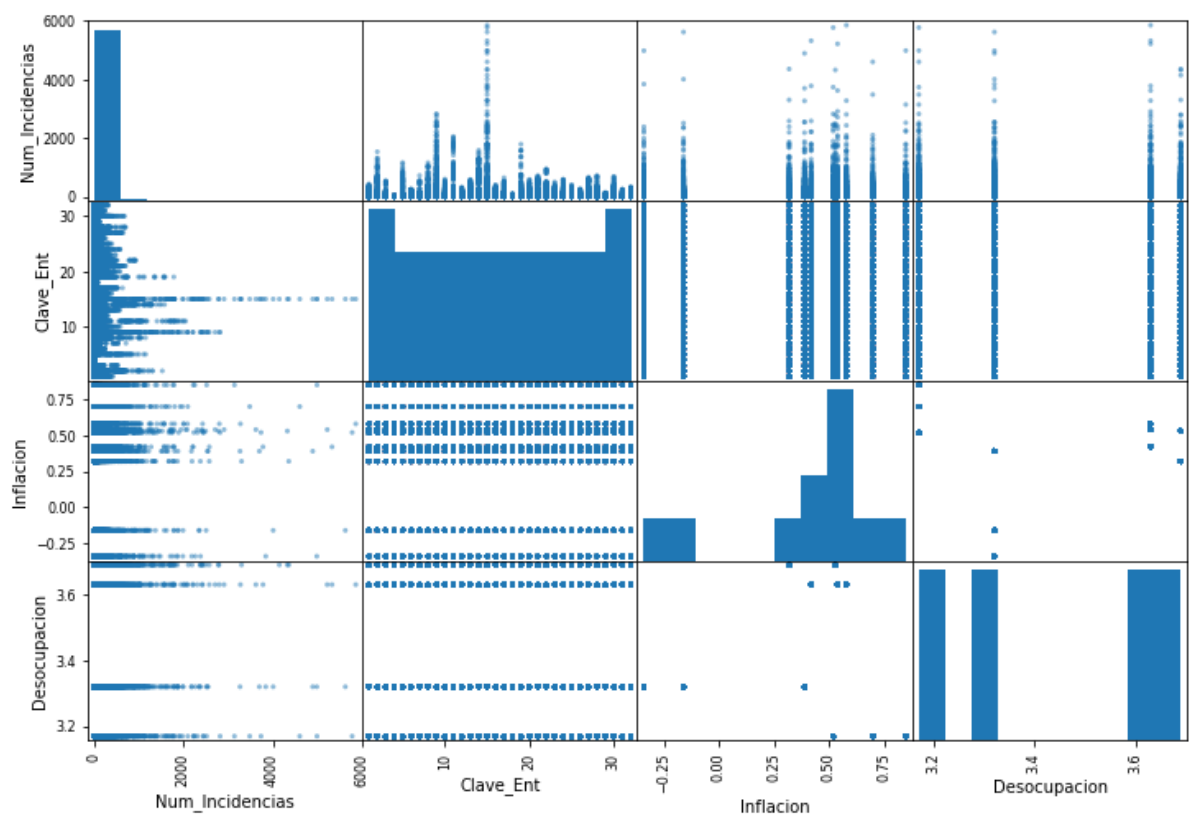
A todos los conjuntos de datos se les hizo algún tipo de procesamiento o limpieza, esto para mejorar la calidad de los resultados que se fueran a obtener y para poder continuar con el desarrollo del modelo, las operaciones que se realizaron fueron melts, merge, drops de columnas y joins.

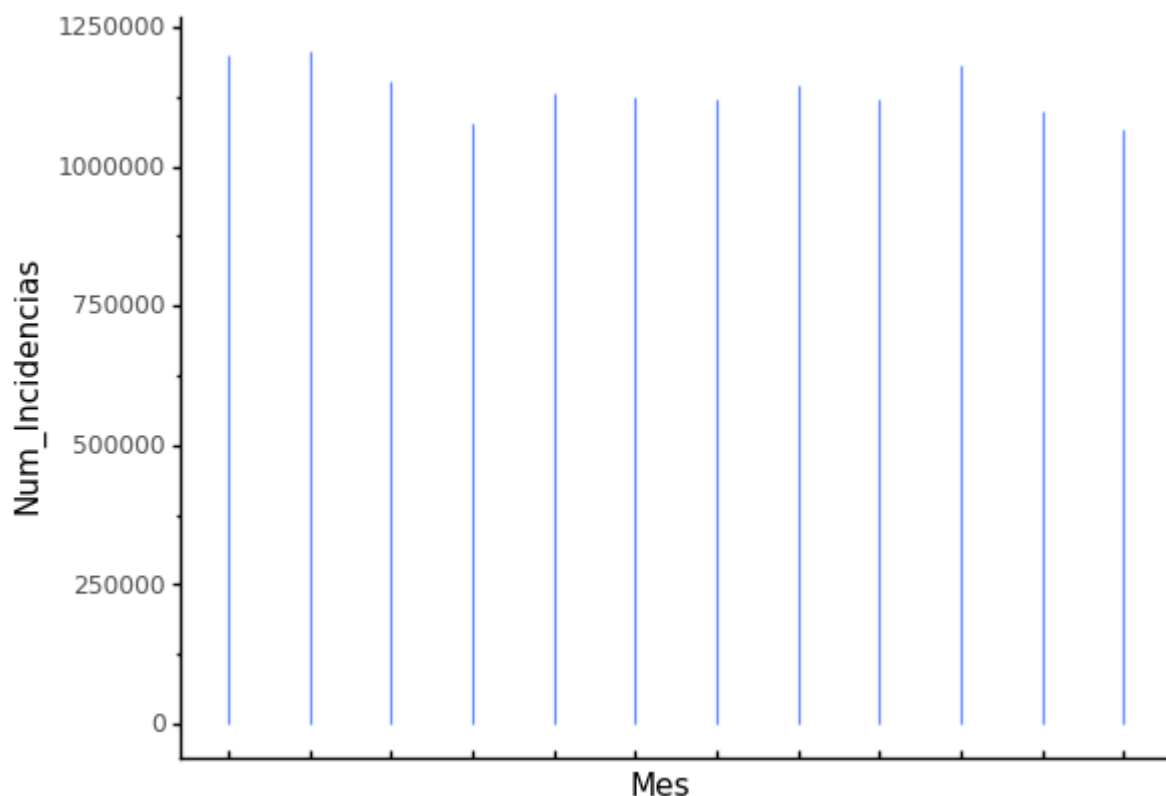
Una vez teniendo toda esta información lista para visualizar y analizar, se generaron las siguientes gráficas y tablas:

Para el análisis exploratorio inicial, se generó un histograma de nuestra tabla principal para poder visualizar los valores iniciales con los que se cuentan.



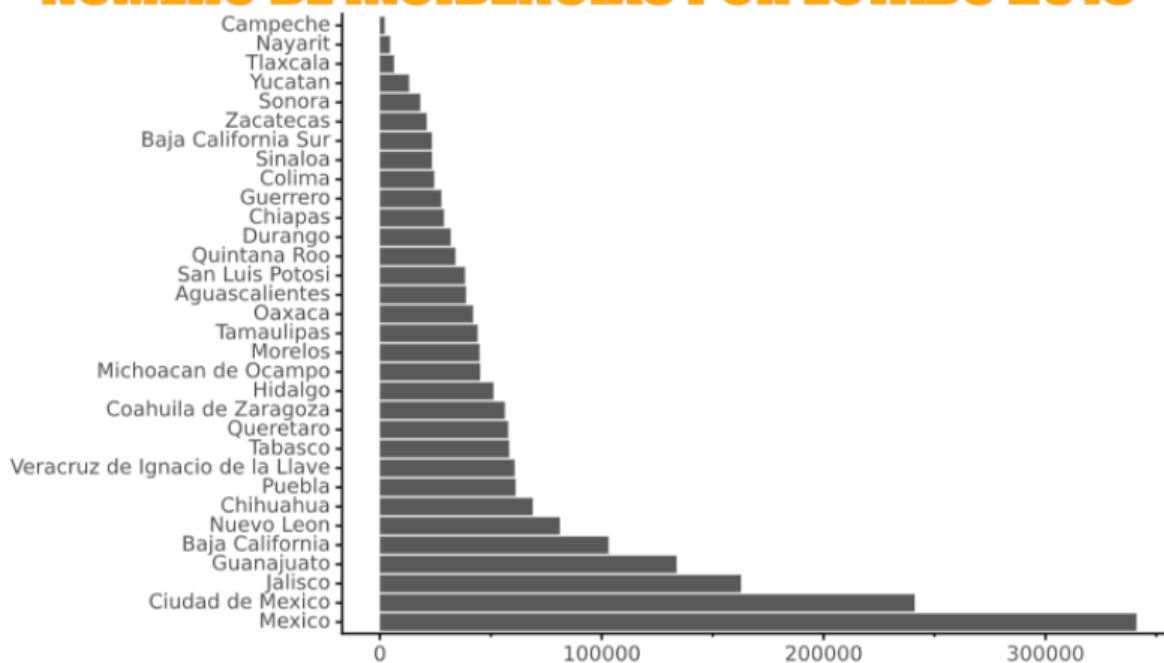
Posteriormente se generó un mapa de calor en el que podemos observar las relaciones que tienen unas variables de nuestra tabla con otras variables. Recordando que lo que queremos es tener colores oscuros en la gráfica, no podemos obtener gran insight acerca del efecto de unas variables en otras.





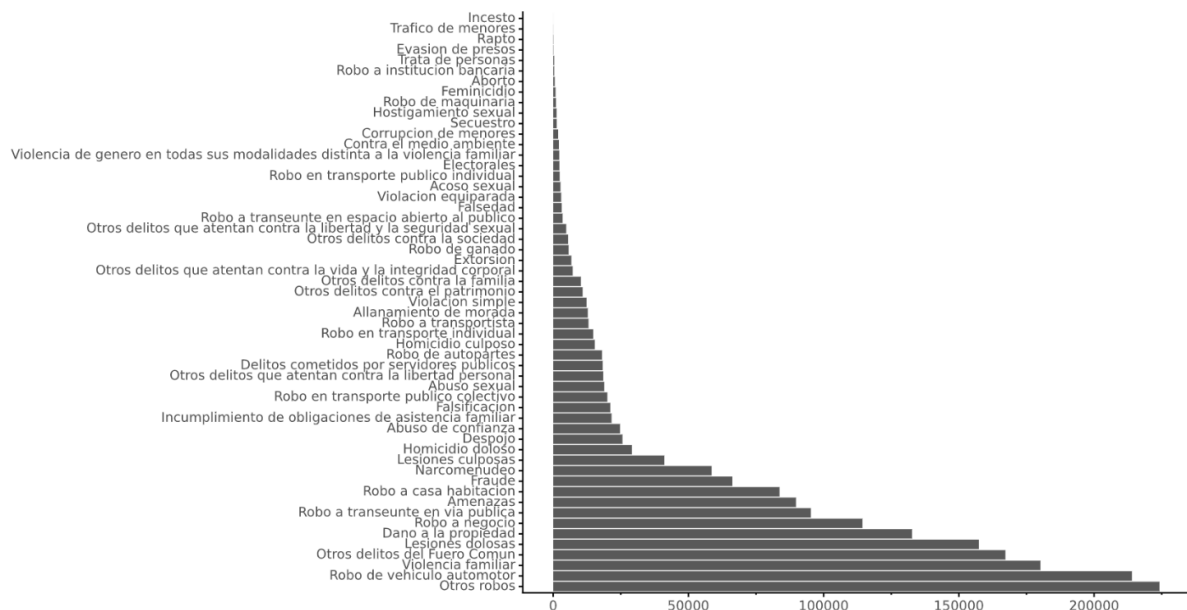
En esta gráfica podemos ver el acumulado de incidencias en cada uno de los meses del año 2018, podemos tomar como nota importante el hecho de que las incidencias nunca han estado por debajo del millón de unidades.

NÚMERO DE INCIDENCIAS POR ESTADO 2018



El primer análisis que podemos realizar sobre esta gráfica es el número de incidencias por estado ya que los que se encuentran más habitados son donde suceden más delitos, encontrando algunas excepciones donde no ocurre esto en concreto, debido a que no se encuentran normalizados los valores, no se encuentran un resultado objetivo ya que los estados con poca población muestra un bajo índice pero no necesariamente significa que tengan nivel delictivo bajo.

SUBDELITO MÁS COMÚN A NIVEL NACIONAL



Como segundo análisis se puede observar los sub delitos más comunes, siendo el más prominente la categoría de otros robos, esto es muy general y no nos dice gran cosa, los siguientes dos puestos son el robo de vehículos y la violencia familiar.

Para el entrenamiento se dividió el set en dos partes 70% fueron el train set mientras que el 30% fueron de prueba, las variables categóricas para el modelo tuvieron que ser transformadas a numéricas con onehotencoder, después se añaden las variables numéricas originales para formar una misma tabla, esta tabla se convierte a ndarray para meterlo al modelo de entrenamiento el cual es un modelo de regresión polinomial. Esto dio como resultado un vector con las thetas las cuales son los coeficientes por los cuales llegaremos a la variable dependiente (número de incidencias). Después se calcula el error cuadrático medio, el cual fue 4.4.

Implicaciones sociales

Se encontró una correlación entre las afecciones sociales y los aumentos delictivos, el desempleo y la inflación aumentan con respecto a la criminalidad, en nuestro punto de vista faltaron incluir más variables que se implican en este problema, para poder tener un análisis completamente detallado en su reciprocidad, a sí mismo saber si hay aspectos más influyentes en el que la inseguridad en México vaya en aumento

Para el entrenamiento se dividió el set en dos partes 70% fueron el train set mientras que el 30% fueron de prueba, las variables categóricas para el modelo tuvieron que

Conclusiones

Gracias al modelo,

se pudo obtener el valor del error cuadrático, el cual tiene un valor de 4.4, esto quiere decir que, el modelo puede predecir el número de incidencias normalizadas ± 4.4 . Esto puede considerarse como un resultado positivo, pues creemos que este valor es bajo y no representa gran margen de error, por lo que se puede concluir que el modelo quedó bien entrenado y da resultados posibles y que pueden tomarse en cuenta para poder generar políticas públicas para mejorar la seguridad en dichos estados.

Para delitos con muchas incidencias el error cuadrático medio puede representar un error muy pequeño, pero para delitos menos comunes o con menos incidencias, puede considerarse un error muy grande al solo tener un par de incidencias a lo mucho cada mes.

Bibliografía

Quiroga, Héctor. *La delincuencia como hecho social*. Revista mexicana de sociología, 1960.