Prueba OPI Analytics

Luisa Rojas

**Caso I**

**Observaciones: Se realiza el análisis a partir desde 2016, dado que es la fecha en donde se incrementan los delitos**

1. ¿Qué pruebas identificarías para asegurar la calidad de estos datos? No es necesario hacerlas, sólo describe la prueba y lo que te dice cada una.

* Identificar cantidad de nulos en cada variable (Tabla de frecuencias)
* Identificar datos atípicos en las variables numéricas (Boxplot)
* Revisar distribución de frecuencias
* Identificar valores nulos de las variables con el fin de obtener variables consistentes

1. Identifica los delitos que van al alza y a la baja en la CDMX (ten cuidado con los delitos con pocas ocurrencias).

* Delitos al alza:
  + VIOLENCIA FAMILIAR
  + AMENAZAS
  + CONTRA LA INTIMIDAD SEXUAL
  + ROBO S/V DENTRO DE NEGOCIOS, AUTOSERVICIOS, CONVENIENCIA
  + ROBO A TRANSEUNTE EN PARQUES Y MERCADOS CON VIOLENCIA
  + SUSTRACCIÓN DE MENORES
  + PRODUCCIÓN, IMPRESIÓN, ENAJENACIÓN, DISTRIBUCIÓN, ALTERACIÓN O FALSIFICACIÓN DE TÍTULOS AL PORTADOR, DOCUMENTOS DE CRÉDITO PÚBLICOS O VALES DE CANJE
* Delitos a la baja
  + DENUNCIA DE HECHOS
  + ROBO A NEGOCIO SIN VIOLENCIA
  + ROBO DE OBJETOS
  + ROBO A TRANSEUNTE DE CELULAR SIN VIOLENCIA
  + CONTRA FUNCIONARIOS PUBLICOS
  + DAÑO EN PROPIEDAD AJENA CULPOSA
  + DAÑO EN PROPIEDAD AJENA CULPOSA POR TRÁNSITO VEHICULAR A AUTOMOVIL
  + LESIONES CULPOSAS POR TRANSITO VEHICULAR EN COLISION
  + ROBO A CASA HABITACION SIN VIOLENCIA

1. ¿Cuál es la alcaldía que más delitos tiene y cuál es la que menos? ¿Por qué crees que sea esto?

La alcaldía que más delitos tiene es CUAUHTEMOC con 182659. Con respecto las alcaldías de menor delitos existen 229 que registran únicamente 1 delito asociado, como por ejemplo COYUCA DE BENITEZ. Si solamente se toman las alcaldías que registran más de 1000 delitos, NEZAHUALCOYOTL es quien obtiene la menor cantidad de delitos con 1227 casos.

Adicionalmente, CUAUHTEMOC es el centro de ciudad de México lo que implica un mayor índice de delitos al tener mayor población. El delito que más se presenta en esta alcaldía es el fraude con 17949 casos

1. ¿Existe alguna tendencia estacional en la ocurrencia de delitos (mes, semana, día de la semana, quincenas) en la CDMX? ¿A qué crees que se deba?

Los días de mayor cantidad de delitos son el 1, 15, 30 y 31(El día 31 se incluye en el día 30 dado que no todos los meses tienen este día), dado que son los pagos de las nóminas, lo que implica que las personas obtengan más dinero y propiedades en esas fechas.

Adicional, el día que mayor se presentan delitos son los viernes dado que este día es más propenso a que las personas salgan a eventos con amigos o familiares dado que es el último día de la semana, seguido del sábado. Por otra parte, el domingo presenta la menor cantidad de delitos, dado que es un día más familiar y muchas veces permanecen dentro del hogar. El resto de los días (lunes a jueves) se presenta una distribución similar en la cantidad de delitos.

Con respecto al mes, enero lidera la cantidad de delitos, seguido de febrero. La menor cantidad de delitos se presenta en abril y junio y diciembre. Es llamativo que en diciembre se presente una baja cantidad de delitos dado que es un mes que implica muchas reuniones familiares y salidas a eventos sociales, sin embargo, puede estar asociado a que las personas salgan de la ciudad por vacaciones, lo que influye una menor cantidad de densidad poblacional en las grandes ciudades.

5. ¿Cuáles son los delitos que más caracterizan a cada alcaldía? Es decir, delitos que suceden con mayor frecuencia en una alcaldía y con menor frecuencia en las demás.

Se identifican 4 grupos de alcaldías:



* Grupo 1 (CUAUHTEMOC, BENITO JUAREZ y MIGUEL HIDALGO): Se evidencia mayor participación de fraude y robo de objetos y menor participación de robo de vehículo de servicio particular sin violencia.
* Grupo 2 (IZTAPALAPA, GUSTAVO A MADERO y ALVARO OBREGON): Se evidencia mayor participación de violencia familiar y menor participación de robo de accesorios de auto y robo de objetos del interior de un vehículo.
* Grupo 3 (COYOACAN y TLALPAN): Se identifica mayor participación de violencia familiar y menor participación de robo a transeúnte en vía pública con violencia, denuncia de derechos.
* Grupo 4 (VENUSTIANO CARRANZA y AZCAPOTZALCO): Se presenta mayor participación de violencia familiar y menor participación de robo de objetos del interior de un vehículo y robo de vehículo de servicio particular sin violencia.

6. Diseña un indicador que mida el nivel de “inseguridad”. Genéralo al nivel de desagregación que te parezca más adecuado (ej. manzana, calle, AGEB, etc.). Analiza los resultados ¿Encontraste algún patrón interesante? ¿Qué decisiones se podrían tomar con el indicador?

Se realiza indicador en base a la ocurrencia de las categorías de delito, la colonia y la alcaldía donde se presentaron los hechos.

La alcaldía se incluye dado que existe una correlación entre concentración de población, violencia y delito. Las ciudades con mayor densidad de población muestran altos índices de criminalidad violenta[[1]](#endnote-1) . Con respecto a la colonia, se incluye en el indicador dado que no es lo mismo una colonia más comercial, como por ejemplo el centro de las ciudades, que una colonia apartada o rural, lo que influiría en la frecuencia de los delitos.

El tipo de delito no se incluye dentro del indicador dado que, sin importar el origen del delito, todos son considerados eventos de inseguridad.

El indicador se calcula en función de los últimos años (2019-2020), dado que es la información más actualizada y no se incluye 2021 dado que no ha finalizado todo el año.

A continuación, la formula del indicador (Ind\_F):

Ind = +

Ind\_F = Promedio de cada periodo j(Ind)

El calculo del indicador se encuentra en el archivo de Excel en la hoja Punto6

Se pueden tomar decisiones financieras dado que la inversión económica en la seguridad del país debe estar en función de la cantidad de delitos por zonas geográficas. Adicional, el indicador debe realizarse anualmente para que se revise si las estrategias de disminución de inseguridad funionan o no, y en base a esto calcular nuevamente la inversión económica.

También es importante generar políticas publicas que permitan minimizar este indicador. Este debe estar de la mano con inversión social.

**Caso II**

La Michoacana

**Objetivo:** Definir qué máquinas deben ser surtidas y la cantidad optima de paletas a surtir en cada una, manteniendo un costo de transporte y energía bajos, sin superar su cuota de indisponibilidad.

**Proceso:**

**Restricciones:**

* Cantidad de máquinas 4000 máquinas expendedoras
* No se puede surtir más de una vez la maquina
* No deben existir casos de máquinas sin paletas disponibles
* Cantidad de días/máquina donde hubo indisponibilidad de paletas menor a 2% al mes

**Indicadores actuales:**

* La cantidad de paletas retiradas para cada máquina, cada día desde que empezó la operación (hace 5 años) hasta el día anterior.
* La cantidad de paletas disponibles en cada máquina, cada día (a final del día, medianoche) desde que empezó la operación (hace 5 años) hasta el día anterior.
* La capacidad, en cantidad de paletas, de cada máquina expendedora.
* El costo de surtir cada máquina ($100)
* El costo de mantener una paleta por un día en cualquier máquina expendedora ($1)
* No hay límite de carga en los camiones repartidores

**Antecedentes:**

* La estrategia de llenar a tope la máquina expendedora y surtirla el día que observe que las paletas están a punto de acabarse para llenar de nuevo la máquina a tope tiene un costo muy alto de energía y un costo bajo de transporte.
* Para cada máquina hacer viajes diarios con el contenido de paletas necesarias para el siguiente día tiene un costo de energía bastante bajo pero un costo de transporte muy alto.
* El consumo de paletas cambia mucho de máquina a máquina y día con día

**Solución**

**Una vez se proyecten las cifras, cada cierto periodo de tiempo se deben alimentar las cifras reales al histórico de información, para realizar la proyección y optimización nuevamente.**

**Planteamiento matemático optimización**

**, con**

Con,

* = Costo óptimo de cada maquina i
* = Costo de transporte de cada maquina i
* = Costo energético de las paletas de cada maquina i

Sujeto a:

* = Cantidad de días/máquina donde hubo indisponibilidad de paletas menor a 2% al mes

**Métricas**

* Bondad de ajuste
* Criterios de AIC, BIC
* Variables significativas
* Método del codito para clústeres
* Afinidad de los clústeres

**Evaluación del modelo**

* Realizar paralelamente la distribución usual sugerida por Luis y la solución planteada por el modelo, con el fin de evaluar el impacto en los costos (variación de costos) durante al menos 3 meses.

**Despliegue de modelo**

* Una vez se proyecten las cifras, cada cierto periodo de tiempo se deben alimentar las cifras reales al histórico de información, para realizar la proyección y optimización nuevamente.
* Cada 6 meses realizar calibración de clústeres

**Riesgos**

* Variables exógenas al modelo (Por ejemplo, nueva heladería cerca de las neveras)
* Riesgos operativos asociados a la plataforma de análisis (Servidores)

1. https://onuhabitat.org.mx/index.php/violencia-en-inseguridad-en-las-ciudades [↑](#endnote-ref-1)