1****

Equipo 3

**Tema:**

Delincuencia y víctimas de la CDMX entre 2019 y junio de 2021.

**Base de Datos:**

Delitos reportados en la CDMX de enero 2019 a junio 2021

12 de agosto del 2021

|  |  |
| --- | --- |
| Nombre | E-mail |
| Hugo Martínez Ibarra | [hugomtzib@gmail.com](mailto:hugomtzib@gmail.com) |
| Juan de Jesús Hernández Angulo | [jhernandezangulo@gmail.com](mailto:jhernandezangulo@gmail.com) |
| Jaime Israel Vazquez Bernal | [israfullshot@gmail.com](mailto:israfullshot@gmail.com) |
| Gilberto Silva Tijerina | [gilberto.silvat2812@gmail.com](mailto:gilberto.silvat2812@gmail.com) |
| Fernando Arriaga Palma | [arriaga141@gmail.com](mailto:arriaga141@gmail.com) |
| Bryan Daniel Moreno Abrego | [abre.go@outlook.com](mailto:abre.go@outlook.com) |

Data Science BEDU

Catedrático: Rafael Arias González.

**Índice**

I Resumen

II Objetivos

1. Identificación del problema (Postwork 1)

2. Planteamiento de preguntas (Postwork 2)

**Resumen**

En la Ciudad de México suceden hechos delictivos diariamente. Dichos suceden en diversas localidades en diferentes horarios. Y a través de los años se ha buscado recopilar la información de cada uno de ellos de la manera más detalla posible. Y una dependencia encargada de dicha tarea es el C5 de la Ciudad de México. *“El Centro de Comando, Control, Cómputo, Comunicaciones y Contacto Ciudadano de la CDMX (C5), es la dependencia del Gobierno de la Ciudad de México encargada de captar información integral para la toma de decisiones en materia de seguridad pública, urgencias médicas, medio ambiente, protección civil, movilidad y servicios a la comunidad en la capital del país a través del video monitoreo, de la captación de llamadas telefónicas y de aplicaciones informáticas de inteligencia, enfocadas a mejorar la calidad de vida de las y los capitalinos.” [1]*

Luego de recopilar los datos se necesita de personal capacitado para darles un adecuado tratamiento, con la finalidad de que la información resultante sea de ayuda para prevenir dichas situaciones. En el presente proyecto se desarrolló justamente esa parte. La que realiza un científico de datos.

El objetivo del proyecto es presentar el desarrollo de la limpieza de datos de un conjunto de datos emitido por el C5 de la ciudad de México. El cual empieza desde la detección de un problema, hasta la obtención del dataset limpio y adecuado para ser usado en el siguiente módulo de Data Science. Se pretende usar Python como herramienta principal.

**Palabras clave**

Delincuencia

Datos

Limpieza

**Objetivos**

1. **Identificación del problema**:  identificar y escoger un problema (justificar su elección). Presentar investigación preliminar completa del tema. Presentar información previa sobre las soluciones al problema
2. **Planteamiento de preguntas:**  Las son correctamente respondidas. La entrega tiene al menos 5 preguntas pertinentes sobre el problema.
3. **Colección de datos:**  Obtener una colección de datos completa y no previamente procesada que usará para los siguientes pasos.  
    Los datos recopilados pueden responder completamente las preguntas planteadas anteriormente.
4. **Análisis exploratorio de los datos:**  Convertir de forma adecuada su colección de Datos (en JSON, CSV, etc.…) en un DataFrame de Pandas. Realizar un análisis exploratorio de estos datos a través del uso correcto de Pandas.
5. **Limpieza de datos:** El nuevo Dataset no contiene ningún NaN.  
    El Dataset está correctamente indexado (en orden y sin índices incoherentes). Las columnas están correctamente nombradas.  
    Se aplicaron agregaciones al Dataset para poder comenzar a responder algunas preguntas planteadas en la sesión 2.
6. **Transformación de datos**: Todos los datos de cada columna tienen un valor correcto dependiendo del tipo de dato (fechas, objetos, ints, floats). Todas las columnas de texto están manipuladas para estar en el formato correcto.  
   Hay nuevas columnas con nuevos datos resultantes del procesamiento de una o más columnas originales.  
   Se crearon otros datasets/subconjuntos/series que presentan información relevante para una pregunta específica, o simplemente para exploraciones más profundas.
7. **Ordenamiento de código**: Entregar un Jupyter Notebook ordenado y limpio que contiene todos los pasos llevados a cabo para el preprocesamiento de datos. El Notebook hace uso de las diversas técnicas aprendidas a través de las clases para la exploración y procesamiento de datos. El Notebook contiene el código (nombrado correctamente) de todas las sesiones pasadas.
8. **Opcional uso de APIs**

**1. Identificación del Problema (Postwork 1)**

*“Hay que señalar que la violencia ha incrementado su presencia en nuestra vida cotidiana difícilmente parece novedoso, pero no deja de ser necesario. La violencia homicida ha arrebatado la vida de 200 mil personas durante los últimos 11 años, principalmente, de nuestros jóvenes. 2017 ha sido el año más violento del que se tiene registro, con más de 25 mil casos de homicidio y una tasa de 20.5 homicidios por cada 100,000 habitantes.* [Según datos del Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública (cifras al 20 de mayo de 2018, con corte al 31 de diciembre de 2017).] [2]

Es evidente que la inseguridad y la delincuencia se sigue elevando año tras año a nivel nacional. La sociedad se acostumbra a que hechos delictivos sean tan cotidianos, que se ha convertido en algo “común y normal” ver como suceden hasta el punto de permitir que sucedan. Cabe resaltar que esto es más reiterado en ciertos lugares del país como lo es la Ciudad de México (CDMX).

**1.1 Problemática**

El problema identificado es el siguiente:

**Alto grado de delincuencia en la CDMX entre 2019 y junio del 2021.**

**1.1.1 Justificación**

Se detecto un fuerte problema nacional: el aumento de la delincuencia. Como se menciono anteriormente, el proyecto se centra únicamente a las localidades de la CDMX. Esto porque en esta ciudad existe un alto grado de concentración de delincuencia. Es una cuestión que debe ser resuelta con las herramientas que estén a disposición.

El panorama de la Ciencia de Datos nos dice que este problema (como muchos otros) puede ser abordado y tratado. ¿Cómo? A través de la inmensa cantidad de datos que surgen día con día al suceder los hechos delictivos. Actualmente, en la CDMX, se cuenta con grandes y avanzados sistemas de monitoreo y reportes. Estos generan los datos que se van a necesitar para realizar modelos que puedan predecir o prevenir incidentes delictivos.

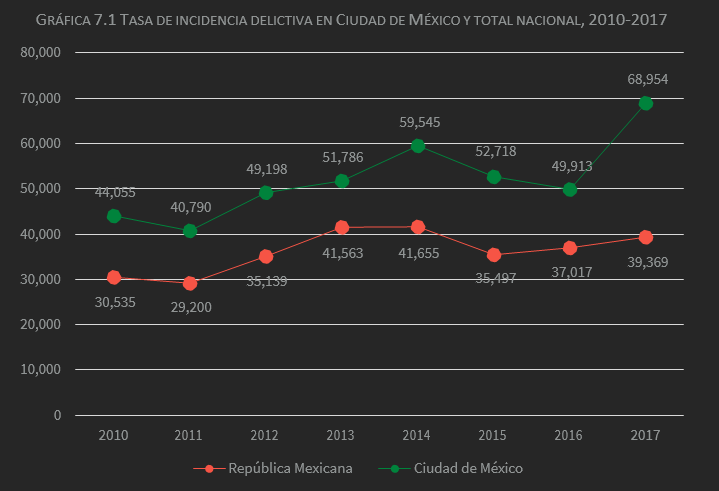
Es por estas razones que se decidió seleccionar este tema y esta problemática. Se cuenta con el tiempo y con capacidad necesaria para desarrollar el proyecto, es decir, que es un proyecto viable en todos sus ámbitos.

**1.2 Investigación del problema**

En varias investigaciones y documentos relacionado con el aumento de la delitos, violencia y denuncias coinciden en que dichos índices han ido creciendo con el paso del tiempo. Cada uno de ellos ha llegado a la misma conclusión. Fragmentos de ellos serán mostrados a continuación.

*“Las tasas a lo largo del tiempo presentan una tendencia ascendente tanto para la Ciudad de México como para el conjunto del país, pero notoriamente más elevada en el primer caso, como corresponde a sus tasas medias de crecimiento: entre 2010 y 2017 la tasa de incidencia delictiva en la capital del país ha aumentado a un ritmo interanual de 6.6%, mientras que para el total nacional el crecimiento medio de un año a otro ha sido de 3.7%. El segundo resultado que debe ser destacado es el aumento de la incidencia delictiva en la Ciudad de México entre 2016 y 2017, pues entre ambos años la tasa cre- ció 38.2%, la mayor variación relativa bianual registrada desde el año 2010.”* [3]

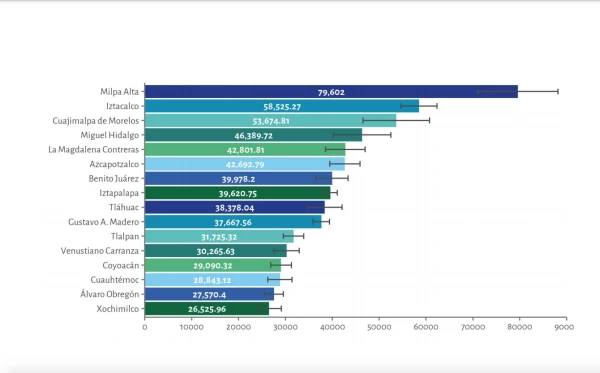
Tasa de incidencia delictiva en Ciudad de México y Total Nacional, 2010-2017



Fuente: Evalúa Ciudad de México a partir de ENVIPE (2011 a 2018).

Diversas fuentes de divulgación periodística han registrado de igual manera un notable crecimiento en el aumento de la delincuencia. También a su vez han realizado investigaciones para determinar en cuales alcaldías se han registrado un mayor número de incidentes delictivos. Los resultados fueron los siguientes:

“Las alcaldías con mayor cantidad de delitos por cada 100,000 habitantes fueron Milpa Alta, Iztacalco y Cuajimalpa, en ese orden; mientras, los que tuvieron una incidencia menor fueron Xochimilco, Cuauhtémoc y Álvaro Obregón.”



Del mismo sustrato se hace evidente que en los hogares es reportado en una gran parte de la población por lo menos algún incidente delictivo en contra de algún miembro de dicho hogar.

*“Durante 2020, 53.8% de los hogares reportó que al menos uno de sus miembros sufrió un delito, lo que resulta en alrededor de 2 millones 695,682 delitos en la capital de enero a septiembre pasado.*

*No obstante, 81.1% de las personas no denunció, mientras 11.5% de los denunciantes sí acudió a reportar ante las autoridades, pero se le negó la apertura de una carpeta de investigación.”*

El sentimiento inseguridad es una realidad que se comparte en la población nacional. Pero que, en ciertos lugares, como lo es la Ciudad de México se vive en escalas indignantes. Desde fraudes, que es el delito con mayor cifra negra en la ciudad, pues 97.2% de los casus no son denunciados, hasta robos de vehículos con un porcentaje de denuncia de 53.5% han sido las cifras que se han registrado en esta investigación.

Imágenes que representan el problema:

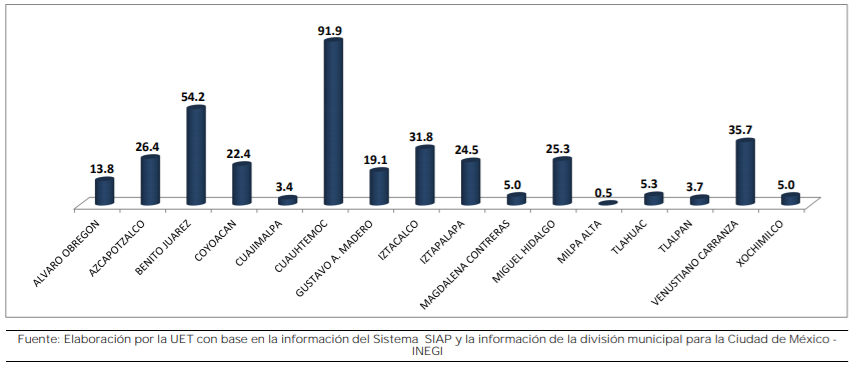


Imágenes de El Sol De México (www.elsoldemexico.com.mx), EXPANSIÓN politica (politica.expansion.mx) y Animal Politico (www.animalpolitico.com).

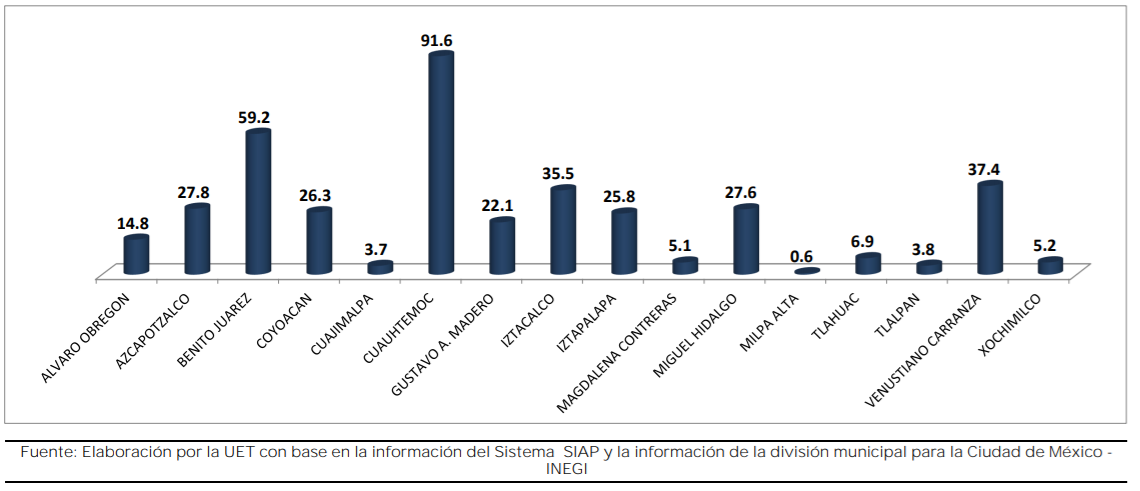
La Fiscalía General de Justicia de la Ciudad de México tiene en su sitio en internet estadísticas delictivas de cada mes de cada año. En ellas su pudo sustraer la siguiente información.

En cada gráfica se muestran los delitos cometidos por kilómetro cuadrado en cada una de las Alcaldías.

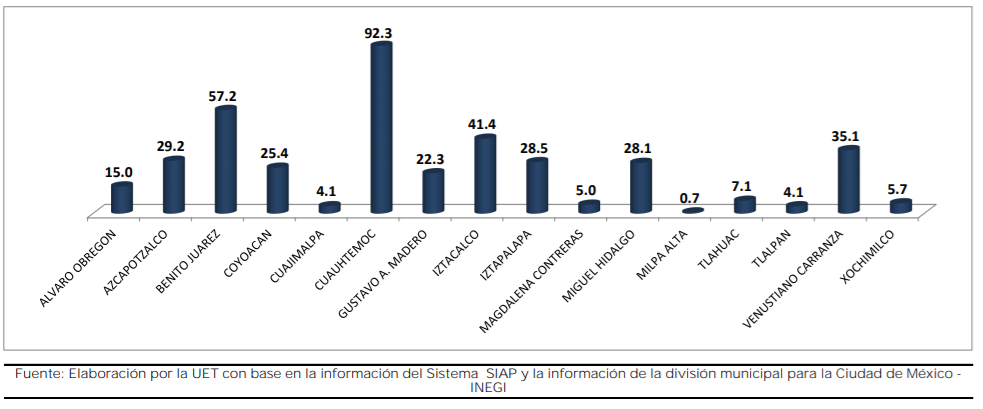
Enero 2020:



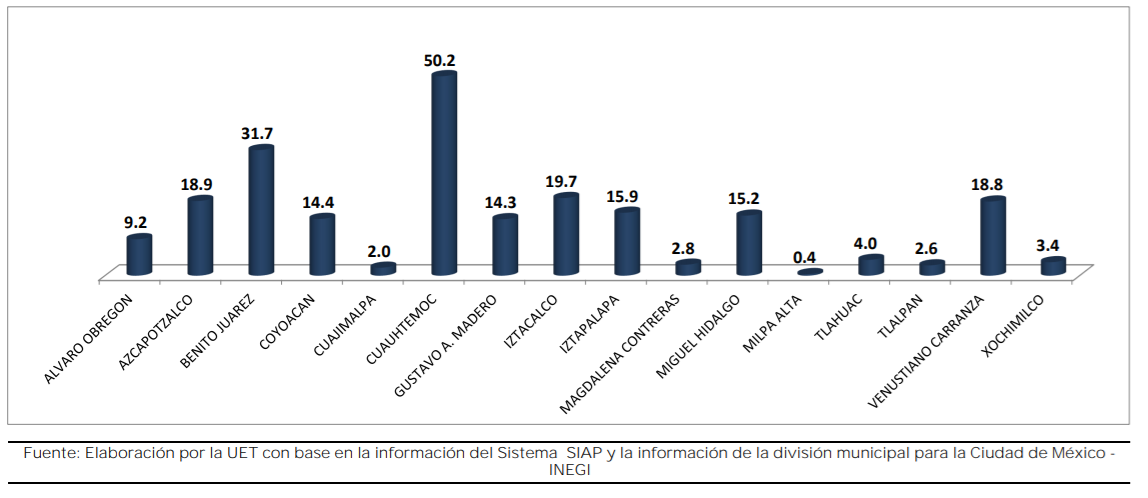
Febrero 2020:



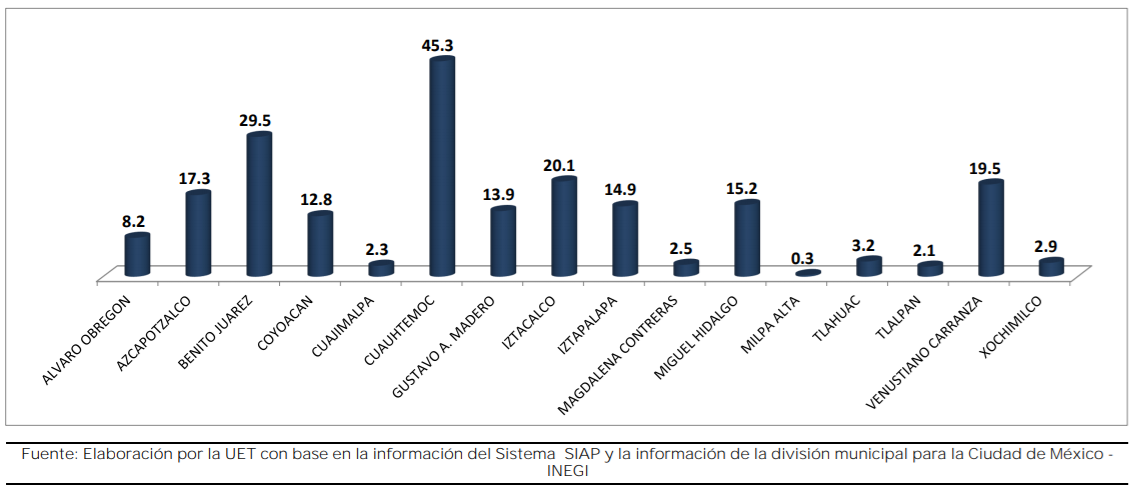
Marzo 2020:



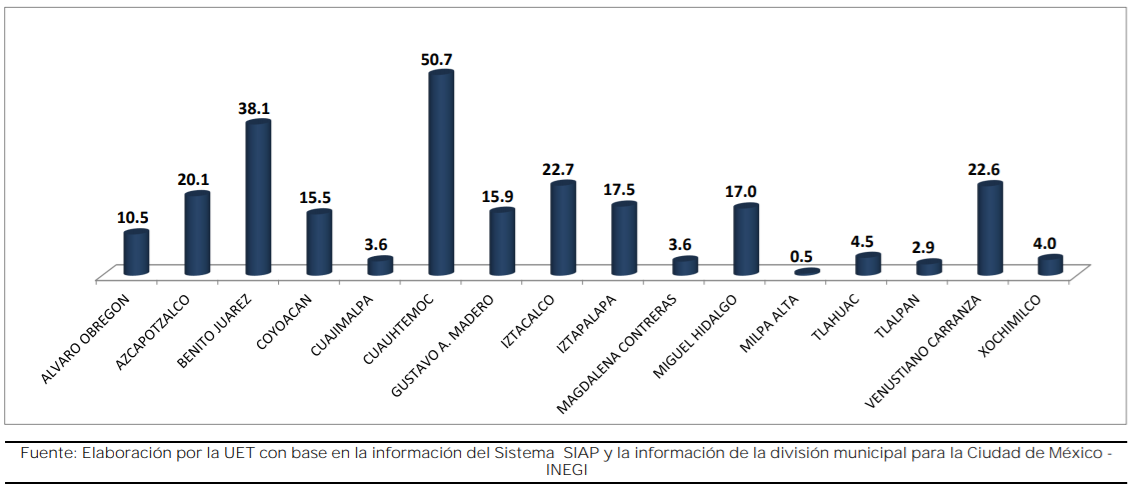
Abril 2020:



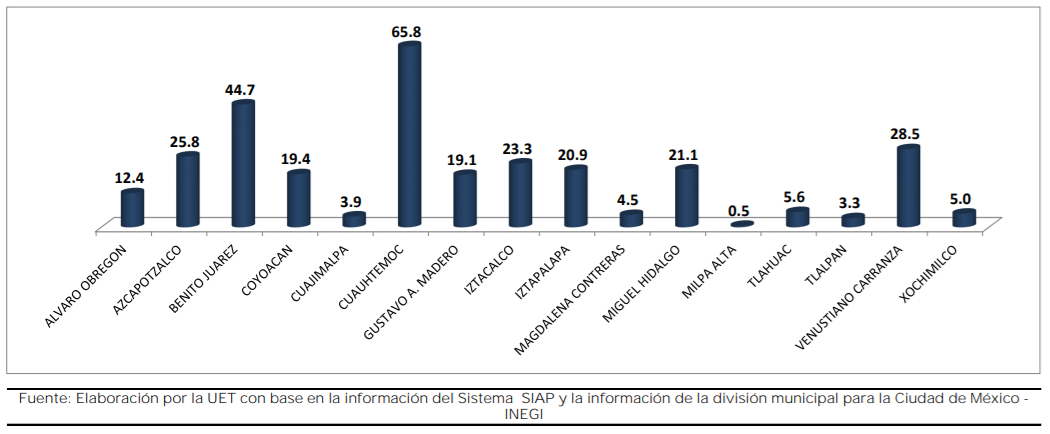
Mayo 2020:



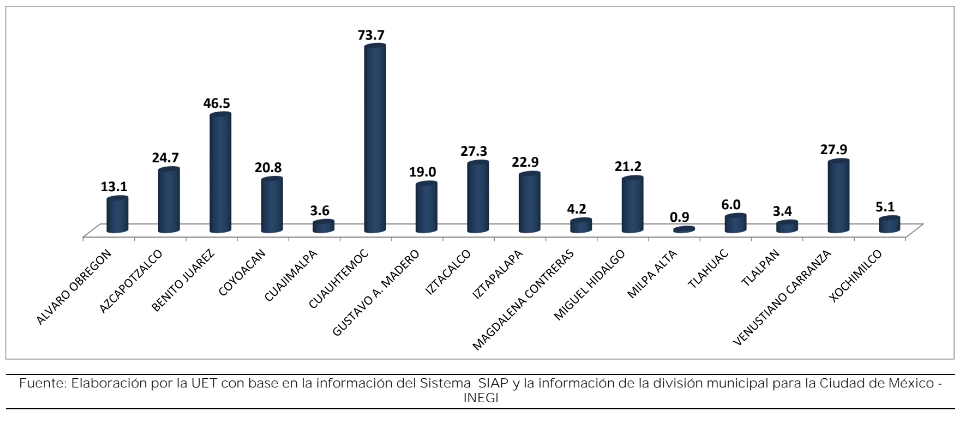
Junio 2020:



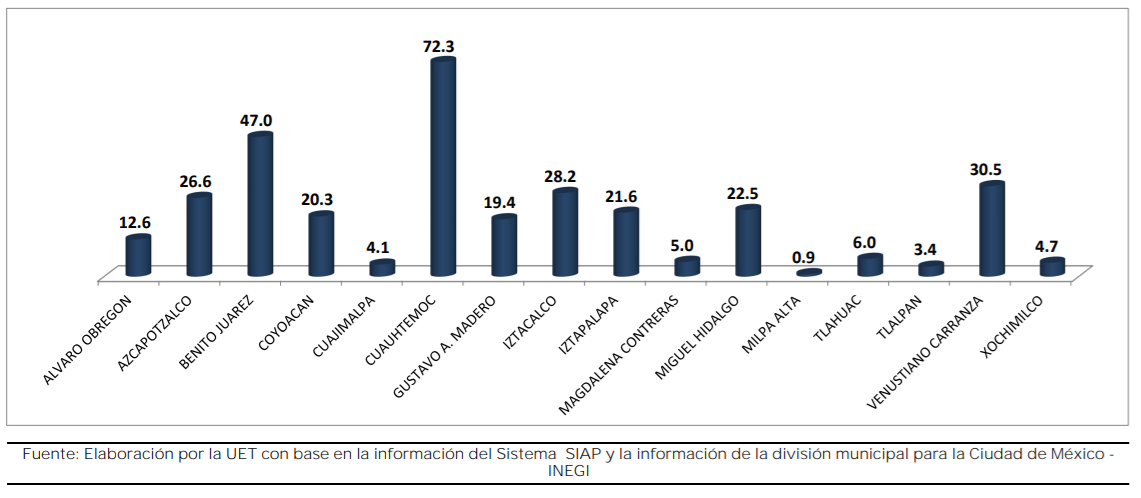
Julio 2020:



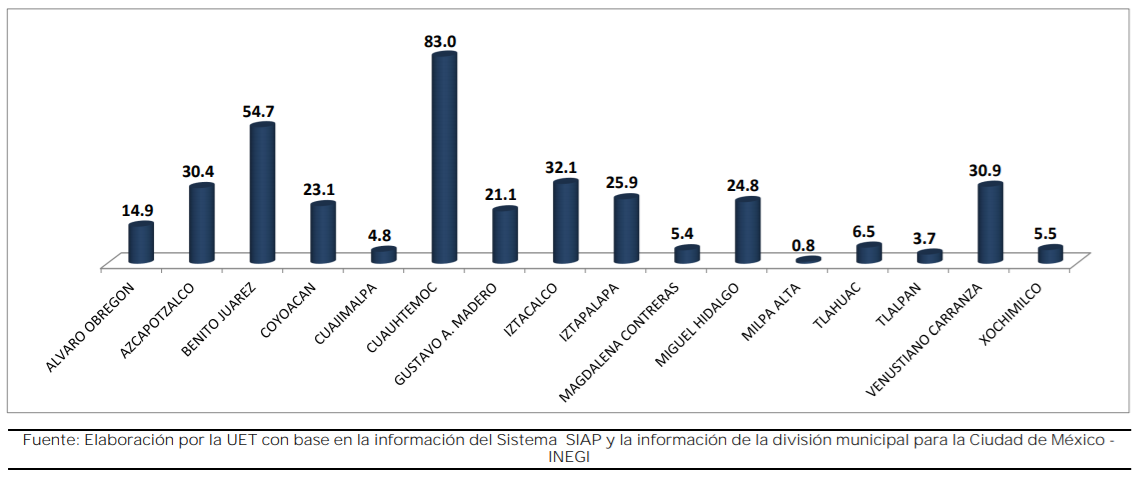
Agosto 2020:



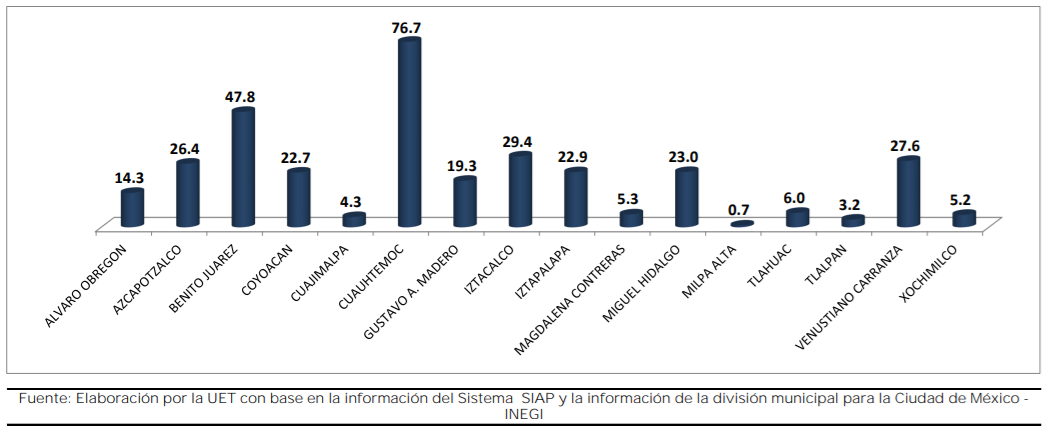
Septiembre 2020:



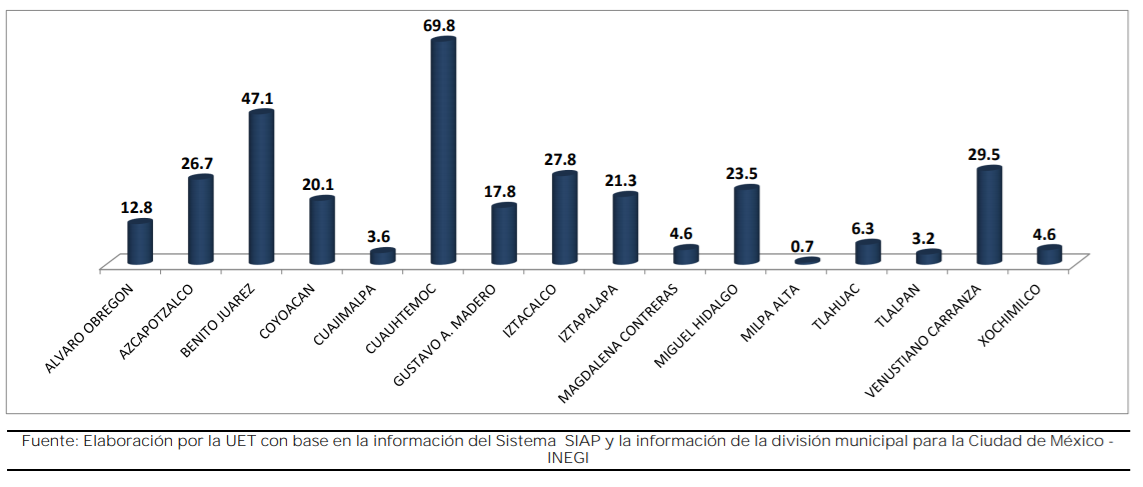
Octubre 2020:



Noviembre 2020:



Diciembre 2020:



**1.3 Búsqueda de soluciones anteriores**

La delincuencia no es un problema nuevo en la CDMX, y durante el paso de los años se han dado propuestas que intentan disminuir los índices de incidentes delictivos.

Según el sitio de Política Expansión se especifican estrategias que usa el gobierno durante la actual presidencia de Andrés Manuel López Obrador:

*“Nuevo Modelo Policial:*

*Si como objetivo se plantea la creación de la Guardia Nacional como un cuerpo permanente que pueda servir a futuros gobiernos, como una estrategia se prevé articular los esfuerzos y las aportaciones de los tres órdenes de gobierno y tomar en cuenta condiciones, contextos y necesidades locales para impulsar un Nuevo Modelo Policial. En él, se busca que los elementos deban tener una visión de derechos humanos, proximidad y participación ciudadana. El eslabón más importante es la policía municipal, pues es la que está más cerca de los ciudadanos.*

*Prevención del delito*

*Esto incluye a su vez varias acciones:*

* *Desarrollo alternativo. Plantea dar alternativas económicas para los hogares y las comunidades que dependen del ingreso provisto por actividades ilícitas.*
* *Prevención especial de la violencia y el delito. Son intervenciones restaurativas, orientadas a protección (en el caso de los menores que son forzados a cometer delitos), resocialización y reparación del daño.*
* *Disuasión focalizada. Plantea el diseño y la implementación de estrategias que permitan identificar dónde y quiénes se dedican a actividades ilícitas, comunicarles las consecuencias, movilizar a la comunidad para generar conciencia y ofrecer apoyo integral a quienes deseen abandonar la actividad ilícita.*
* *Cumplimiento regulatorio. Se promueve que dependencias o instancias que puedan estar involucradas con actos ilícitos se apeguen a estándares de operación y cumplimiento legal.*
* *Justicia procesal, proximidad y de atención a víctimas. Velar por que las autoridades de seguridad y justicia cuenten con herramientas y protocolos para consolidar su legitimidad frente a la sociedad.*

*Participación ciudadana*

*En esta estrategia se plantea recuperar espacios públicos, combatir los delitos que más afectan a la sociedad y conformar redes vecinales.*

*Cambio en el reparto recursos para seguridad*

*De acuerdo con esta estrategia, se hará un cambio en los criterios, las fórmulas y las variables para la asignación, la administración y el ejercicio del Fondo de Aportaciones para la Seguridad Pública (FASP), y se les dará prioridad a los siguientes lineamientos:*

* *Profesionalización, certificación y capacitación de policías estatales, municipales, custodios y ministeriales con énfasis en sistema de justicia penal acusatorio, derechos humanos y perspectiva de género.*
* *Equipamiento policial y de instituciones, con unidades especializadas en combate al secuestro, cibernética, búsqueda de personas y sistema penal.*
* *Fortalecimiento y/o creación de las unidades de inteligencia financiera en los estados.*
* *Construcción, mejora y equipamiento de los servicios médicos forenses.*
* *Fortalecimiento del sistema penitenciario nacional y de ejecución de medidas para adolescentes.*

*Combate al lavado de dinero*

*Las operaciones con recursos de procedencia ilícita, defraudación fiscal y finanzas de la delincuencia organizada son áreas a las que hasta el momento no se les había dado atención, según el gobierno, por lo que en la estrategia se prevé que la Unidad de Inteligencia Financiera (UIF) de la Secretaría de Hacienda (SHCP) tenga un papel más importante.*

*Combate al robo en carreteras*

*Se realizarán operaciones de revisión y supervisión del transporte en rutas con alta incidencia delictiva, así como filtros de revisión y arcos en centrales de autobuses, revisiones de afectaciones en tramos carreteros y blindaje de accesos a autopistas; también, se implementará el uso de arcos gamma para detectar armas y drogas y el acceso a cámaras de vigilancia.*”

[9]

**EL C5 de la Ciudad de México**

*“El Centro de Atención a Emergencias y Protección Ciudadana de la Ciudad de México (CAEPCCM), fue creado el 22 de Junio de 2009 y a partir del año siguiente comenzó a operar el programa “Ciudad Segura” a través de las cámaras de video vigilancia, que permite mejorar la reacción de las autoridades ante emergencias, situaciones de crisis y comisión de ilícitos. Ciudad Segura es el programa en materia de seguridad urbana más ambicioso a nivel mundial.*

*Para fortalecer y crecer los servicios que proporciona, así como su nivel de calidad en la atención ciudadana, el 23 de diciembre de 2015 el Gobierno de la Ciudad de México decretó la anexión de LOCATEL al CAEPCCM para conformar el Centro de Comando, Control, Cómputo, Comunicaciones y Contacto Ciudadano de la Ciudad de México “C5”, el cual actualmente ofrece los servicios de:*

* *Video Monitoreo*
* *Servicio de Atención de llamadas de emergencia 9-1-1 CDMX*
* *Denuncia Anónima 089*
* *LOCATEL 5658 1111*

*Todos los servicios del C5 operan las 24 horas, los 365 días del año.”* [10]

**Cámaras con IA de la CDMX ya podrán evitar delitos**

*“Jean Marie Letort, vicepresidente del consejo de ciberseguridad de Thales, indicó a Expansión que gracias a una licitación que ganó con el gobierno de la Ciudad de México ahora será posible “no sólo vigilar y seguir la huella de algún criminal, sino que podremos prevenir delitos a través de sistemas de videovigilancia que usan inteligencia artificial con el fin de prevenir y evitar que algún delincuente culmine un delito”.*

*Con aproximadamente 5,000 unidades, la empresa de origen francés será la indicada para soportar y apoyar al C5 en temas de prevención de delito, gracias a una licitación que ganó en la actual gestión de Claudia Sheinbaum. El proyecto en el que participa Thales forma parte de la modernización de la red de cámaras de videovigilancia del gobierno actual y que es denominado Ciudad Segura.*

*El C5 opera y monitorea las más de 15,000 cámaras de vigilancia que hay en el Sistema Tecnológico de Videovigilancia (STVs), así como las 6,000 cámaras instaladas en el Sistema de Transporte Colectivo Metro, “con la finalidad de prevenir y alertar inmediatamente a las autoridades de seguridad y de emergencias capitalinas sobre cualquier situación de riesgo” y por ello es que ha buscado proveedores que le ayuden a cumplir con este objetivo, entre ellos Thales, empresa de seguridad y logística de origen francés.*

*En el momento de su anuncio, Sheinbaum indicó que el plan de esta modernización es integrar al C5 como un organismo que opere de forma cercana en las tareas policiales y de protección civil que requiere la capital mexicana y destacó que existía obsolescencia tecnológica en 6,588 cámaras de videovigilancia, por lo que se harían mejoras en términos de monitoreo, procesamiento y almacenamiento de los datos que recopilan estas unidades.*

*De acuerdo con el gobierno de la Ciudad de México, el contrato de licitación está planteado en conjunto con Thales, Telmex y Cisco, quienes modernizarán esta red de cámaras y puertos de Wifi gracias a una inversión gubernamental de 500 millones de dólares.*

*La empresa de origen francés ha hecho inversiones interesantes en términos de infraestructura de seguridad y operación logística, entre ellas crecer el talento en México y en América Latina y en la compra de algunas empresas como Gemalto, la cual adquirió en abril del presente año por 4,800 millones de euros (5,372 millones de dólares) y con quien buscó aliarse para brindar mayor énfasis en temas de ciberseguridad y salvaguarda de datos.”*

[11]

**2. Planteamiento de preguntas**

Para darle solución al problema se propusieron las siguientes preguntar, para darles respuestas en un futuro con el tratamiento de los datos.

* ¿Cuáles son los delitos más comunes?
* ¿Cuál es la distribución de los delitos en el tiempo (día y mes)?
* ¿Cuáles son los lugares con mayor/menor incidencia de delitos?
* ¿Cuáles son las estadísticas de las víctimas (edad, sexo)?
* Diferencia de tiempo entre el momento de los hechos y el momento de la denuncia.
* Delitos denunciados en la CDMX, pero no ocurridos en la ciudad.

**3. Colección de datos**

Para darle solución al problema planteando a través de las herramientas que puede proporcionar un Científico de Datos, se ha elegido una colección de datos del gobierno de la CDMX, dentro del portal de datos abiertos en su sitio web.

El link para entrar a la siguiente colección de datos es el siguiente:

<https://datos.cdmx.gob.mx/dataset/victimas-en-carpetas-de-investigacion-fgj>

En él se encuentra un archivo de tipo de CSV, lleno de datos relacionados justamente con el tema seleccionado. También se encuentran archivos que explican los datos que encuentran en la colección. Dentro del sitio se menciona lo siguiente:

***Víctimas en carpetas de investigación FGJ***

*Esta base de datos contiene la información actualizada de las víctimas de los delitos en las carpetas de investigación de la Fiscalía General de Justicia (FGJ) de la Ciudad de México a partir de enero de 2019.*

*Para una correcta interpretación de la información, se realizan las siguientes aclaraciones.*

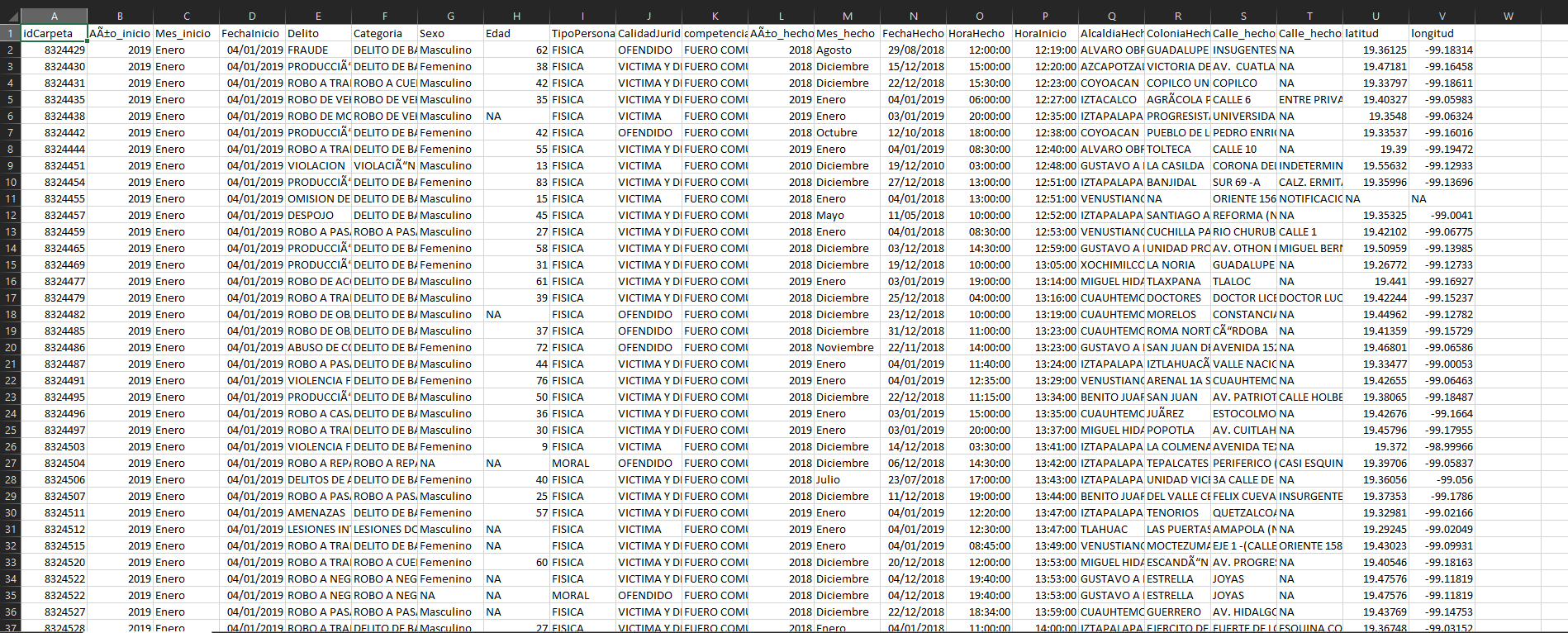
* *El campo "FechaHecho" representa la fecha en la que se cometió.*
* *El campo "FechaInicio" corresponde a la fecha de la apertura de la carpeta de investigación.*
* *En esta base se señala el sexo de la víctima, así como la fecha en que ocurrieron los hechos denunciados.*
* *Es importante destacar que, aunque en algunas ocasiones la víctima es la denunciante, en otras, denunciante y víctima son diferentes (por ejemplo, casos en los que menores de edad son víctimas).*
* *Es posible que en una misma denuncia se incluya a una o más víctimas.*
* *Estas y otras aclaraciones las puedes encontrar en el Diccionario que se adjunta a la presente base.*

*Los datos de esta base fueron actualizados por la FGJ el 29 de julio de 2020. En esta página se puede descargar una nota realizada por la FGJ en la que se explica el proceso de la actualización y la nueva información. Los datos anteriores a la podrán seguir consultándose.*

*Los datos antes de la actualización del 29 de julio de 2020 los puedes consultar en esta liga*

[***https://archivo.datos.cdmx.gob.mx/fiscalia-general-de-justicia/victimas-en-carpetas-de-investigacion-fgj/victimas\_ss\_junio2020.csv***](https://archivo.datos.cdmx.gob.mx/fiscalia-general-de-justicia/victimas-en-carpetas-de-investigacion-fgj/victimas_ss_junio2020.csv)

Un primer acercamiento de los datos, usando el software de Excel, es el siguiente:



**REFERENCIAS**

1. <https://www.c5.cdmx.gob.mx/dependencia/acerca-de>

2. <https://www.mexicoevalua.org/homicidioscdmx/documentos/5013HomicidiosCDMX.pdf>

3. <https://www.evalua.cdmx.gob.mx/storage/app/media/DIES20/8-cap-7-inseguridad-y-violencia-publicas.pdf>

4. <https://politica.expansion.mx/cdmx/2021/03/24/extorsion-y-fraude-los-delitos-que-crecieron-en-la-cdmx-durante-la-pandemia>

5. [www.elsoldemexico.com.mx](http://www.elsoldemexico.com.mx)

6. <https://politica.expansion.mx>

7. [www.animalpolitico.com](http://www.animalpolitico.com)

8. <https://www.fgjcdmx.gob.mx/procuraduria/estadisticas-delictivas>

9. [https://politica.expansion.mx/mexico/2019/02/04/una-a-una-estas-son-las-9-estrategias-especificas-de-amlo-contra-la-violencia](https://politica.expansion.mx/mexico/2019/02/04/una-a-una-estas-son-las-9-estrategias-especificas-de-amlo-contra-la-violencia%20)

10. <https://www.c5.cdmx.gob.mx/dependencia/acerca-de/el-c5-de-la-cdmx>

11. <https://expansion.mx/tecnologia/2019/06/05/camaras-con-ia-de-la-cdmx-ya-podran-evitar-delitos>