



# **La violencia en México durante 2020: un análisis de los principales delitos en los estados del país**

Curso propedéutico de Probabilidad y Estadística

Dra. Gudelia Figueroa P.

Equipo:

Mario Estrada Ferreira

Luis Federico Olivarría Ávila

# 1. Selección del problema

Identificar los principales delitos con incidencia en las capitales de los estados de la Republica Mexicana en el año 2020.

## 2. Descripción del problema

La violencia en México involucra una problemática multifacética y compleja que abarca diversas áreas como la social, económica y política. Esta situación ha sido resultado de una larga cadena de acontecimientos y factores que la han hecho persistir a lo largo del tiempo. En esta guía, se exploran los aspectos más relevantes del problema para poder abordarlo con métodos estadísticos

# Contexto histórico y social

## Revolución Mexicana

El largo conflicto armado que aconteció en el periodo 1910-1920 contribuyó al clima de violencia que perdura hasta ahora.

## Narcotráfico

La existencia del tráfico de drogas hacia Estados Unidos ha sido un importante motor de la violencia en México.

## Pobreza y Desigualdad

La pobreza y la desigualdad son factores estructurales que han contribuido a la violencia en diversas formas.

# Tipos de violencia y manifestaciones actuales en México

Violencia Física	Violencia Sexual	Violencia Psicológica
Asesinatos, Tortura, Lesiones	Acoso Sexual, Trata de Personas	Discriminación, Humillación, Maltrato

Actualmente, el homicidio doloso es la manifestación de violencia más destacada con una tasa de 27 por cada 100,000 habitantes.

# Estado actual de la violencia en el país



En 2020 se registraron 36,579 homicidios en México, lo que representa un aumento respecto al año anterior.



El estado de Guanajuato es el que registra la mayor tasa de homicidios dolosos del país.



La participación de las fuerzas armadas en acciones de seguridad ha sido controversial y ha generado críticas a nivel nacional e internacional.

# Variables relevantes para el análisis estadístico

Las variables que definimos de interés para alcanzar los objetivos fueron:

1. Entidad
2. Municipio
3. Bien jurídico afectado
4. Tipo de delito
5. Mes del año en que se sumaron denuncias
6. Total de eventos



### 3. Objetivos

- hay diferencias significativas en el total de eventos de violencia entre diferentes entidades federativas
- Identificar cuáles son los principales delitos en los estados de México.
- Determinar si alguna o algunas de las variables están correlacionadas.
- Poder ofrecer recomendaciones a partir de los resultados del análisis estadístico.

## 4. Recopilación y limpieza de datos

De acuerdo con la página del Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública (SESNSP), La información publicada se refiere a los presuntos delitos registrados en las averiguaciones previas o carpetas de investigación iniciadas, y que son reportados por las Procuradurías Generales de Justicia y Fiscalías Generales de las entidades federativas, en el caso de los delitos del fuero común, y por la Fiscalía General de la República, en el caso de los delitos del fuero federal. En ambos casos, dichas instancias son las responsables de la veracidad y actualización de la información.

Los presuntos delitos reportados pueden estar involucrados con una o más víctimas u ofendidos, o afectar más de un bien jurídico.

Los delitos no denunciados o aquellas denuncias en las que no hubo suficientes elementos para presumir la comisión de un delito, fueron excluidos en el reporte.

De acuerdo con los postulados anteriores, se puede afirmar que la información proviene de una metodología estandarizada entre los distintos niveles de competencia en la materia de seguridad publica.

En nuestro análisis, decidimos convertir las variables **AÑO**, **CLAVE\_ENT** y **Cve.Municipio** de int.64 a tipo object, ya que son variables cualitativas y no tienen significancia al momento de incluirlas en los análisis estadísticos.

# Metodología para abordar el problema

En nuestro proyecto, seguimos el método de análisis de datos, que consiste en el análisis de la información estadística publicada en el portal de datos abiertos del SESNSP y con esto identificaremos patrones y correlaciones entre las distintas variables.

# 5. Estadística descriptiva

De acuerdo con el análisis en la libreta jupyter de la base de datos, nos arroja el siguiente resumen por mes del año 2020.

```
df_violencia.describe()
```

	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
count	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000	242158.000000
mean	0.684722	0.692135	0.707749	0.494747	0.503233	0.585448	0.636378	0.653330	0.665371	0.705589	0.640549	0.633999
std	7.913517	7.874925	8.217510	6.023346	6.055441	6.879104	7.500991	7.690377	7.580925	7.951791	7.417011	7.707886
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
50%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
max	965.000000	906.000000	1060.000000	840.000000	774.000000	958.000000	1019.000000	1140.000000	1099.000000	1023.000000	1156.000000	1394.000000

Figura 1. Tabla de estadística descriptiva de la base de datos.

Podemos apreciar que el mes con la media más alta a nivel nacional fue durante Octubre.

También se aprecian como mínimos y cuartiles 1, 2 y 3 en cero porque recordemos que la base de datos viene el registro por municipio y varios de los tipos de delitos no cuentan con valor porque no se abrieron nuevas carpetas de investigación.

Por otro lado, en valor máximo se despliegan valores que corresponden al mayor numero de carpetas de investigación abiertas por tipo de delito y municipio, teniendo como el más alto durante diciembre.

## 6. Análisis estadístico inferencial

En las siguientes diapositivas se muestran algunos análisis que desarrollamos con las variables.

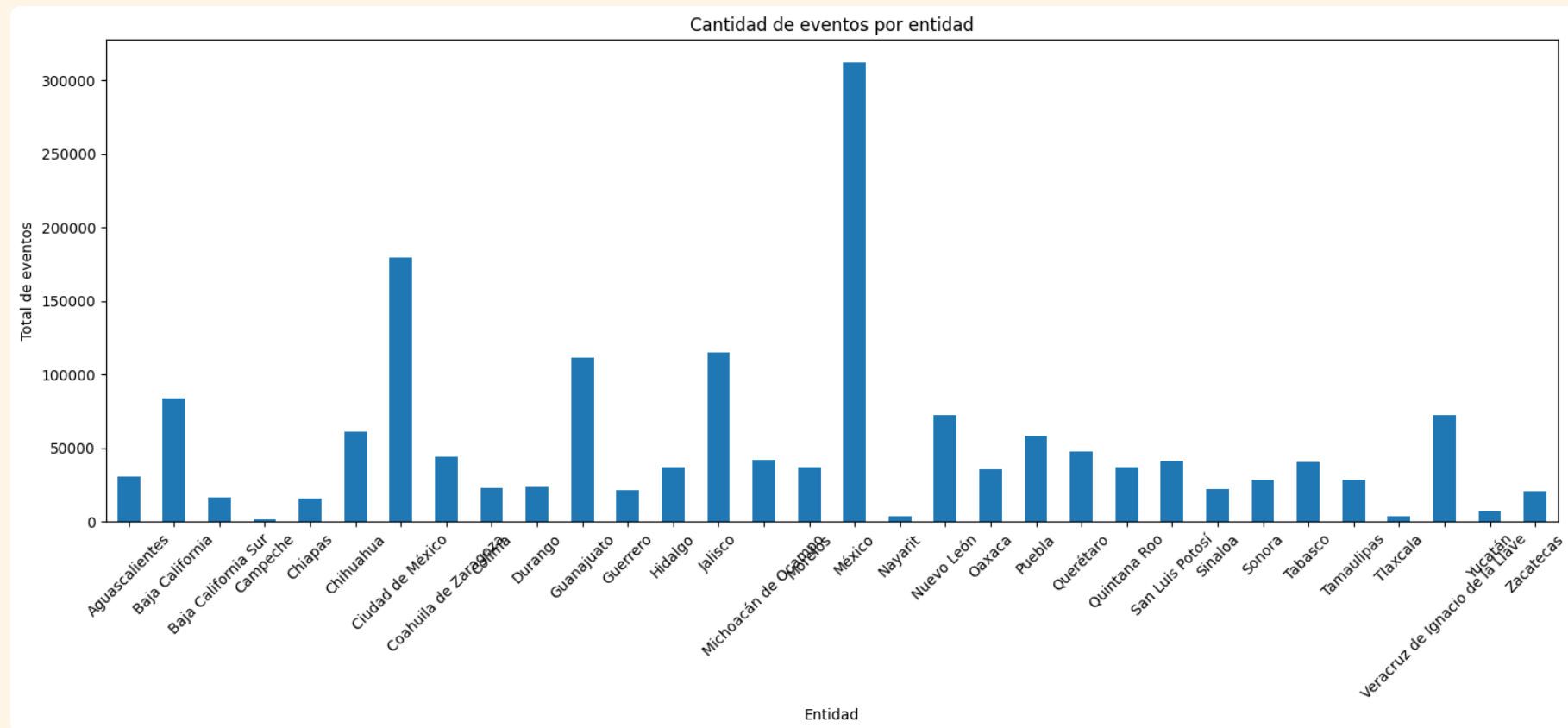


Figura 2. Eventos delictivos por entidad federativa durante el año 2020.

En la figura anterior podemos ver un resumen de la cantidad de nuevas carpetas de investigación por entidad federativa, apreciando su dependencia con la población que habita en cada una de ellas.

## 6. Análisis estadístico inferencial

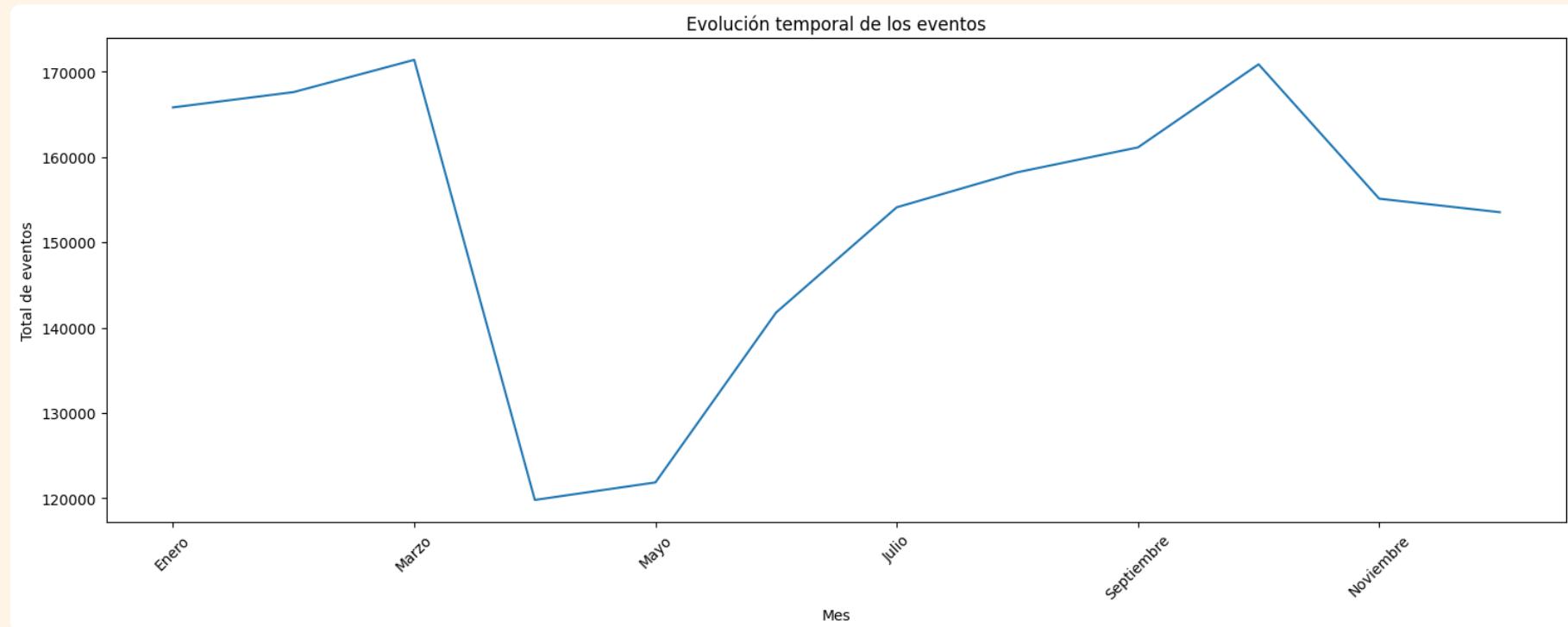


Figura 3. Evolución mensual de eventos delictivos.

En la figura anterior podemos apreciar la curva que describe el número de delitos a nivel nacional durante el transcurso del año 2020. Podemos notar una sensible disminución de los mismos durante las primeras semanas en que se decretó el inicio de la pandemia por la enfermedad COVID-19 a finales de marzo y cómo después de principios de mayo comenzó a aumentar hasta llegar a octubre casi alcanzando el nivel máximo de marzo.

## 6. Análisis estadístico inferencial

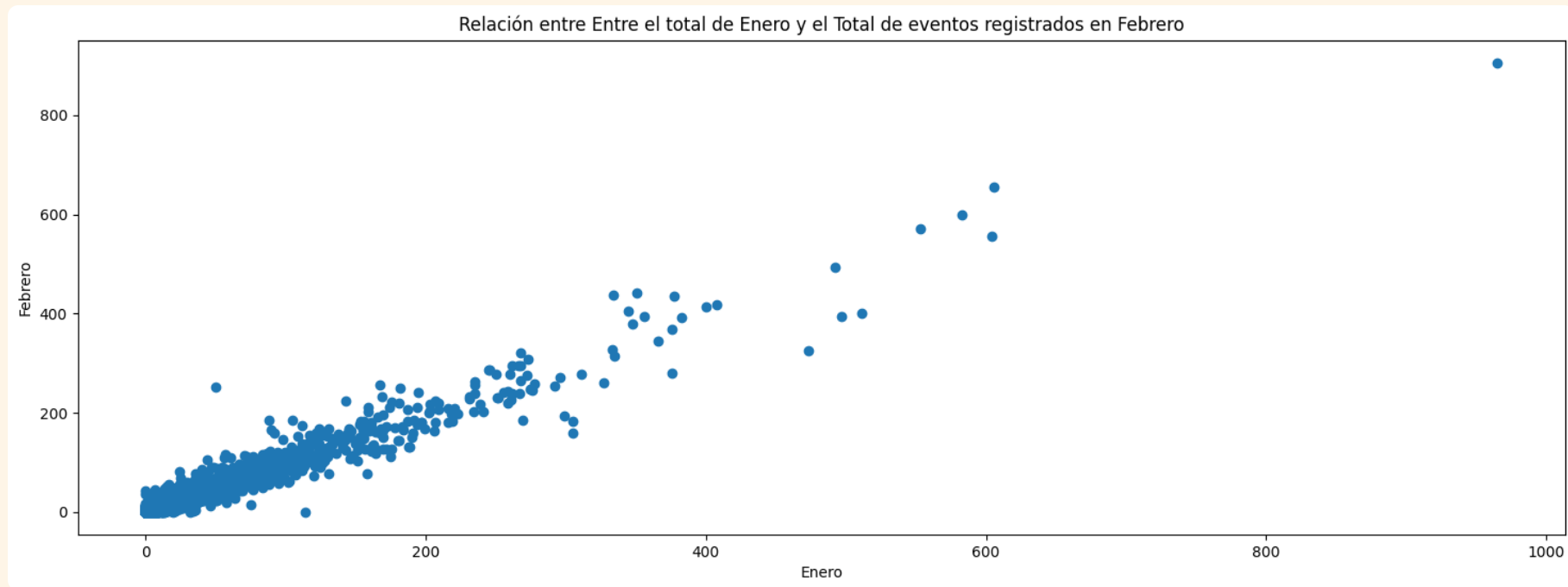


Figura 4. Relación de delitos registrados en enero y febrero 2020.

En la figura vemos un diagrama de dispersión relacionando los delitos registrados en enero y febrero. Los valores son muy parecidos y se agrupan casi en mismas categorías.



## 6. Análisis estadístico inferencial

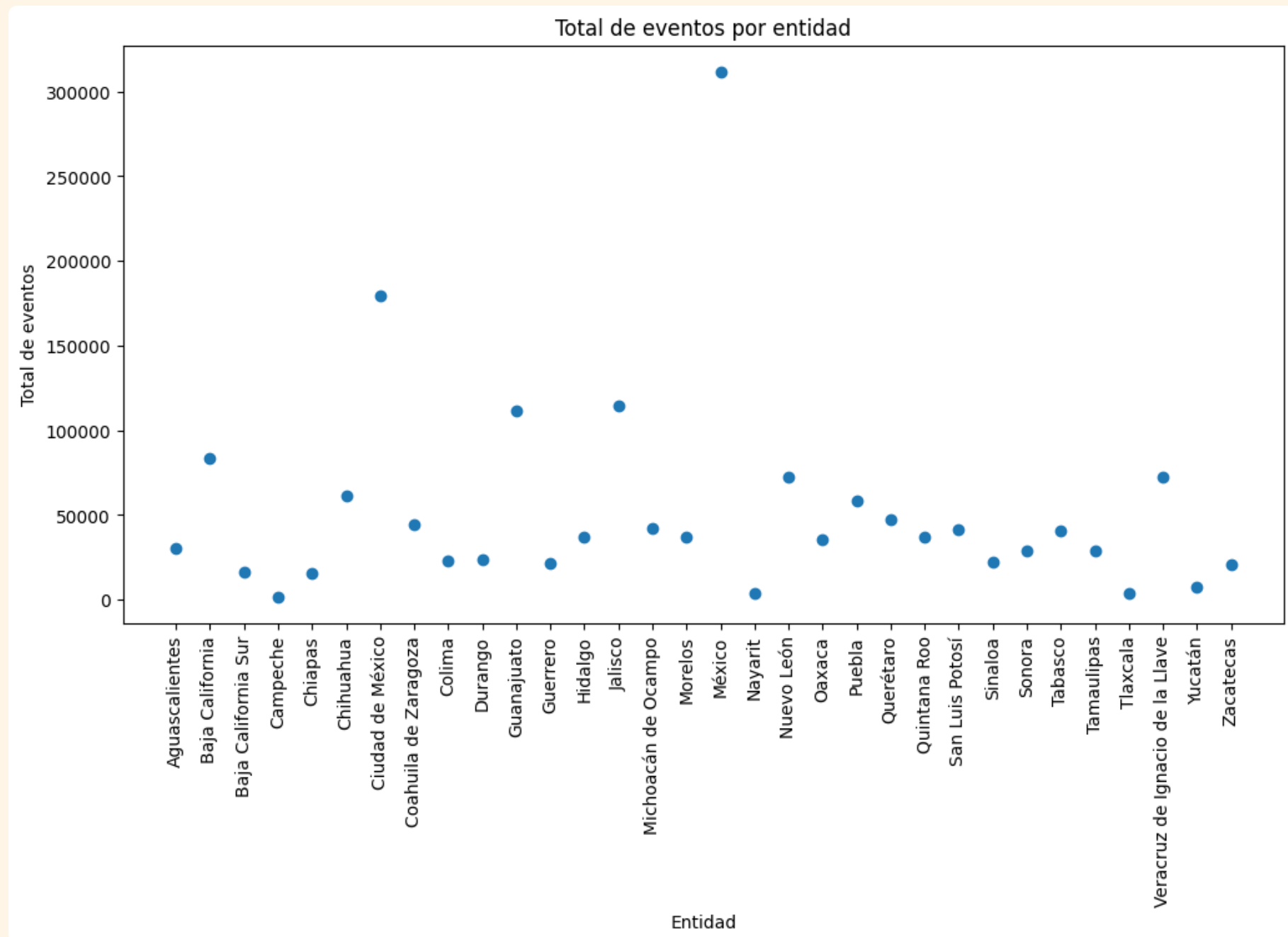


Figura 5. Diagrama de dispersión total de eventos por entidad en el año 2020.

En la figura anterior mostramos la suma de delitos generados durante el año 2020 por entidad federativa. Como se mencionó anteriormente, hay una relación entre el numero de reportes y la población que habita en cada estado de la República, contabilizando el mayor numero de delitos el Estado de México y la Ciudad de México en segundo lugar.

## 6. Análisis estadístico inferencial

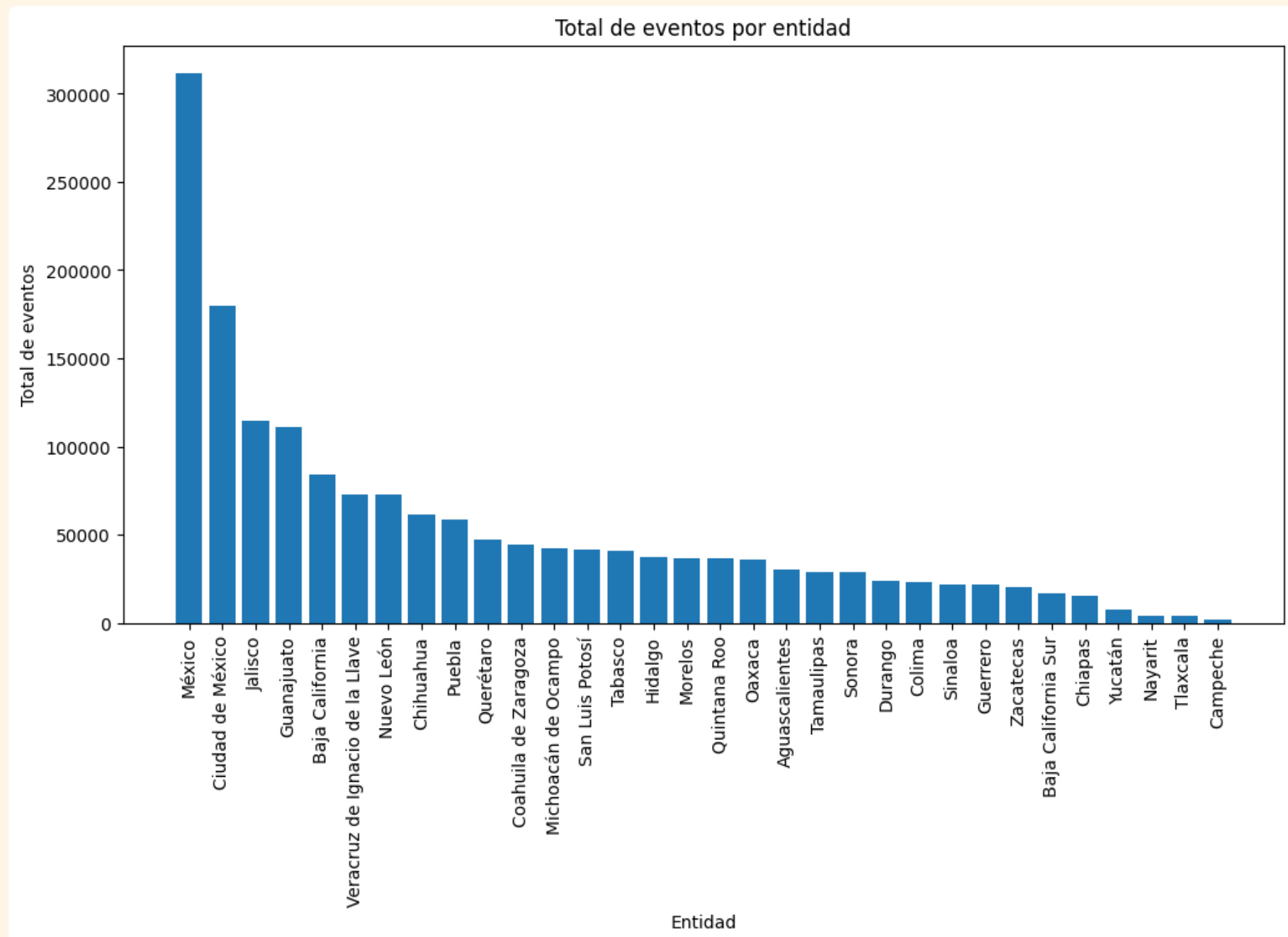


Figura 6. Total de eventos por entidad en orden descendente.

En este diagrama de barras ordenamos los datos según el numero de nuevas carpetas de investigación generadas durante el año 2020 destacando el Estado de México con más de 300,000 registros, seguido de la Ciudad de México, Jalisco, Guanajuato y Baja California.

# 6. Análisis estadístico inferencial

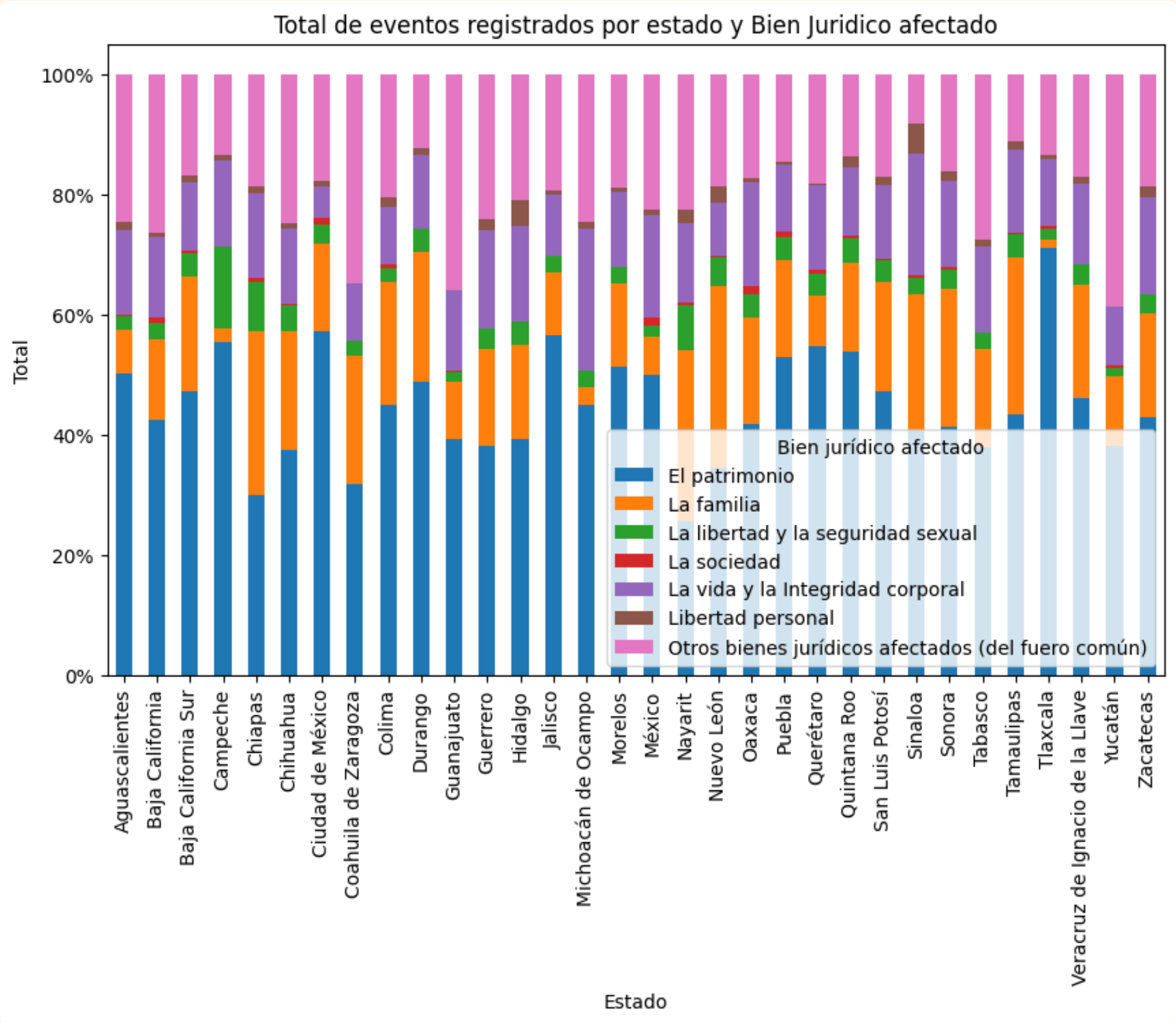


Figura 7. Eventos registrados por estado y Bien jurídico afectado

En esta figura hay mejor representatividad de los datos ya que son barras agrupadas por porcentaje por estado. De esta manera podemos apreciar la variación por estado y tipo de bien jurídico afectado, conservando la predominancia de Patrimonio en primer lugar en todos los estados.

## 6. Análisis estadístico inferencial

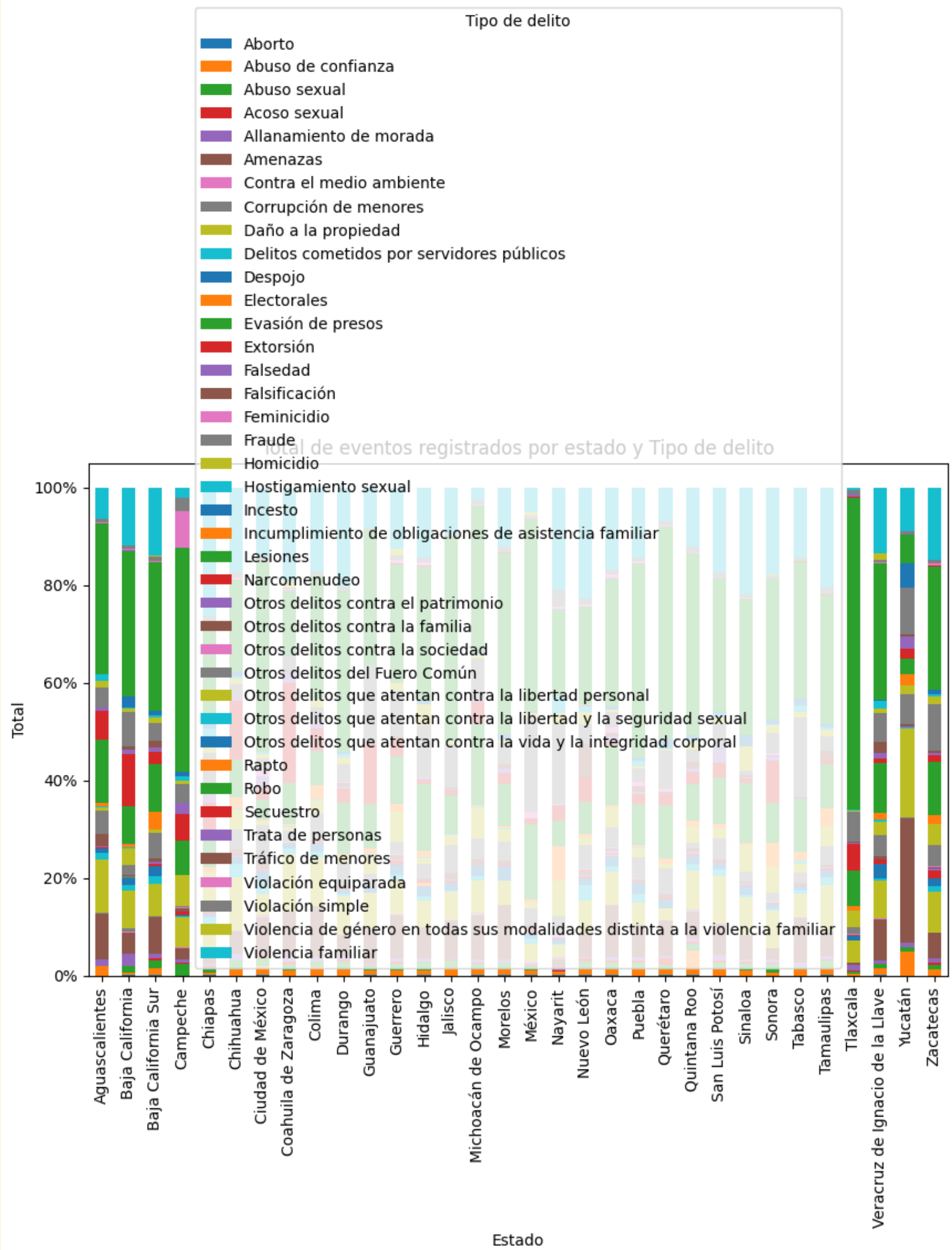


Figura 8. Eventos registrados por estado y Tipo de delito.

En esta figura otro tipo de representatividad de los datos, siendo barras agrupadas por porcentaje por estado. De esta manera distinguimos la variación por estado y tipo de delito, predominando el tipo robo a nivel nacional.

## 6. Análisis estadístico inferencial

```
# Se crearon dos grupos de prueba entre dos estados de distintas regiones
grupo1 = df_violencia[df_violencia['Entidad'] == 'Sonora']['Total']
grupo2 = df_violencia[df_violencia['Entidad'] == 'Michoacán de Ocampo']['Total']

# Se realiza la prueba t de Student
t_stat, p_value = stats.ttest_ind(grupo1, grupo2)

print("Estadística de prueba (t) test 1:", t_stat)
print("Valor p:", p_value)

# Se crearon dos grupos de prueba entre dos estados de distintas regiones
grupo3 = df_violencia[df_violencia['Entidad'] == 'Nuevo León']['Total']
grupo4 = df_violencia[df_violencia['Entidad'] == 'Chiapas']['Total']

# Se realiza la prueba t de Student
t_stat, p_value = stats.ttest_ind(grupo3, grupo4)

# Imprimir el resultado
print("Estadística de prueba (t) test 2:", t_stat)
print("Valor p:", p_value)
```

Estadística de prueba (t) test 1: 0.29623551262261305  
Valor p: 0.7670535872527449  
Estadística de prueba (t) test 2: 15.022916954020568  
Valor p: 1.096771083820585e-50

Figura 9. Pruebas t de Student y resultados de valor p para dos parejas de estados.

Haciendo una exploración de los datos, ejecutamos 2 pruebas t de Student entre 2 parejas de estados: Sonora y Michoacán; y Nuevo León y Chiapas. La primera arrojó un valor  $p = 0.767$  y la segunda un valor  $p = 1.096$ . Debido a que son valores  $p$  altos, podemos asegurar que la hipótesis nula es verdadera en ambas pruebas, por lo tanto si existe relación entre los datos de ambas parejas de estados.

## 6. Análisis estadístico inferencial

```
# Ejemplo de cálculo del intervalo de confianza para la media
datos = df_violencia['Total']

# Calcular la media y la desviación estándar
media = np.mean(datos)
desviacion = np.std(datos)

# Definir el nivel de confianza (por ejemplo, 95%)
nivel_confianza = 0.95

# Calcular los límites del intervalo de confianza
intervalo = stats.norm.interval(nivel_confianza, loc=media, scale=desviacion)

# Imprimir el intervalo de confianza
print("Intervalo de confianza:", intervalo)

Intervalo de confianza: (-148.00233900889054, 161.83939580651855)
```

Figura 10. Cálculo de límites de intervalo de confianza con un nivel de 95%.

Se realizó el cálculo de los límites del intervalo de confianza del Total de delitos durante el año, con un nivel de confianza del 95%, teniendo como resultado los valores (-148.0023, 161.839).

# 7. Resultados y conclusiones

Analizando la información de datos abiertos del Gobierno de México para los índices de delincuencia durante el año 2020, logramos cumplir con los objetivos planteados al inicio del proyecto.

Verificamos la preponderancia de ciertos delitos a nivel nacional como son el robo y la violencia familiar.

Otro dato curioso fue observar la tendencia en el crecimiento de las nuevas carpetas de investigación y como éstas presentaron una disminución drástica al inicio de la pandemia por COVID-19 y cómo lamentablemente en cuestión de semanas volvió a crecer a cifras semejantes al inicio del año.

Es importante seguir trabajando con los reportes que publica el gobierno federal para dar un seguimiento puntual a los índices de delincuencia y de esta manera exigir como ciudadanía cambios a las políticas públicas para vivir en una sociedad con mayor seguridad y bienestar integral.


## 8. Referencias

Incidencia delictiva (Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública). 2023


**Incidencia delictiva**



## 9. Anexos

 Enlace donde se hicieron las pruebas estadísticas



 colab.research.google.com



**Google Colaboratory**

**"<https://colab.research.google.com/drive/1g1M4aFBiZhrCmX0mj-Z0p9sX8JJYiDnX?usp=sharing>"**