



Centro de Investigación en Matemáticas, A.C.

MODELACIÓN ESTADÍSTICA II

TAREA I

ESTADÍSTICA DESCRIPTIVA - ANÁLISIS EXPLORATORIO DE DATOS

Unidad Aguascalientes

PROFESOR:

HUMBERTO MARTÍNEZ BAUTISTA

ALUMNO:

IVÁN VEGA GUTIÉRREZ

27 de enero de 2022

Sección I – Las películas de Hollywood más taquilleras

INTRODUCCIÓN



Actualmente la industria cinematográfica ha sido afectada por diversos factores, uno de ellos son las plataformas de streaming que han surgido tales como Netflix, Prime Video, Disney+, HBO, etc. Sin embargo, a lo largo de la historia, uno de los gigantes invencibles ha sido Hollywood, películas de James Cameron como Avatar o Titanic tuvieron una recaudación mundial superior a los 2mil millones de dólares.

Otro dato relevante es que México se encuentra en la cuarta posición de países que más cine consumen, las salas de cinépolis son de las mejores a nivel mundial, es por esto que la industria del cine en México es tan importante, incluso aunque sea solo como consumidor, pues las producciones nacionales de cine independiente no han tenido buen recibimiento y plataformas como Netflix han sabido tomar ventaja de esto y ponerlo en su catálogo.

Indudablemente en la industria cinematográfica se maneja un gran flujo de dinero, pues producir películas a nivel Hollywood resulta caro, a cualquier cinéfilo o interesado en el negocio cinematográfico le resultaría interesante las siguientes preguntas: ¿Qué tan grandes son las recaudaciones de Hollywood? ¿Cuál es la recaudación promedio de las películas más populares de Hollywood? ¿Cuáles son las variables que influyen para que una película se vuelva taquillera? ¿Cuáles son los estudios de cine mas dominantes en Hollywood?

A continuación se presenta un breve análisis sobre las películas más taquilleras de hollywood. Trataremos de responder las preguntas planteadas mediante un análisis de estadística descriptiva con respecto a una base de datos que contiene alrededor de las mil películas más taquilleras de Hollywood. La base de datos fue tomada de los sitios IMDb y Rotten Tomatoes.

RESULTADOS

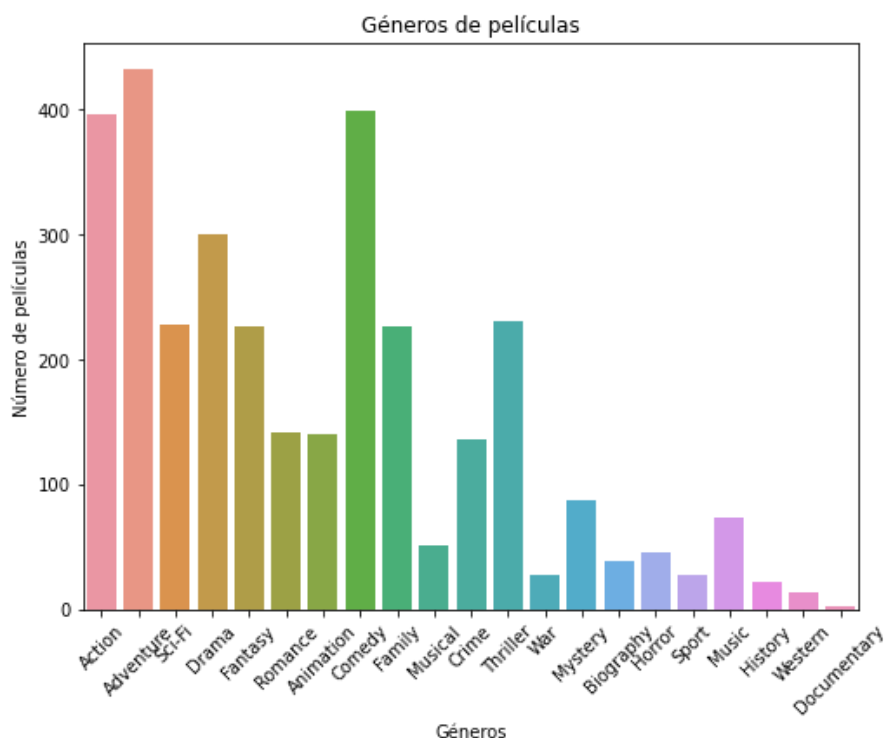
La base de datos fue proporcionada en un archivo con extensión csv y el software con el cuál se hizo el análisis de los datos fue python3. El análisis de los datos se encuentra en el archivo con extensión ipynb. A continuación se muestra únicamente los resultados del análisis.

Dentro de las variables que se analizaron se encuentran:

- Fecha de lanzamiento de las películas
- Distribuidor (estudio de cine)

- Recaudaciones nacionales, internacionales y totales
- Género
- Duración de las películas
- Clasificación

En la siguiente imagen, se muestra la frecuencia con la que aparecen los géneros en las películas rankeadas. Cada película puede tener más de un género.



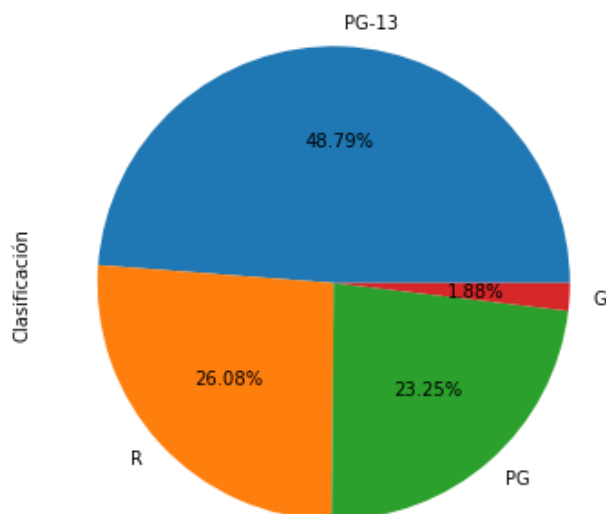
Podemos observar que los géneros más populares en Hollywood son Aventura, Acción y Comedia. Por otro lado, los géneros menos populares son el género Documental, Western e Historia. El resultado parece ser bastante congruente, pues la mayoría de las películas que están en cartelera son de Acción o Comedia, por eso universos cinematográficos como Marvel o Rápidos y furiosos han dominado carteleras hasta por meses.

Otro aspecto importante en una película es la clasificación. Las clasificaciones son:

PG-13: Se advierte enfáticamente a los padres. Puede ser inapropiada para niños de 12 años y menores.

G : Adecuada para público en general.

PG: Se sugiere orientación de los padres



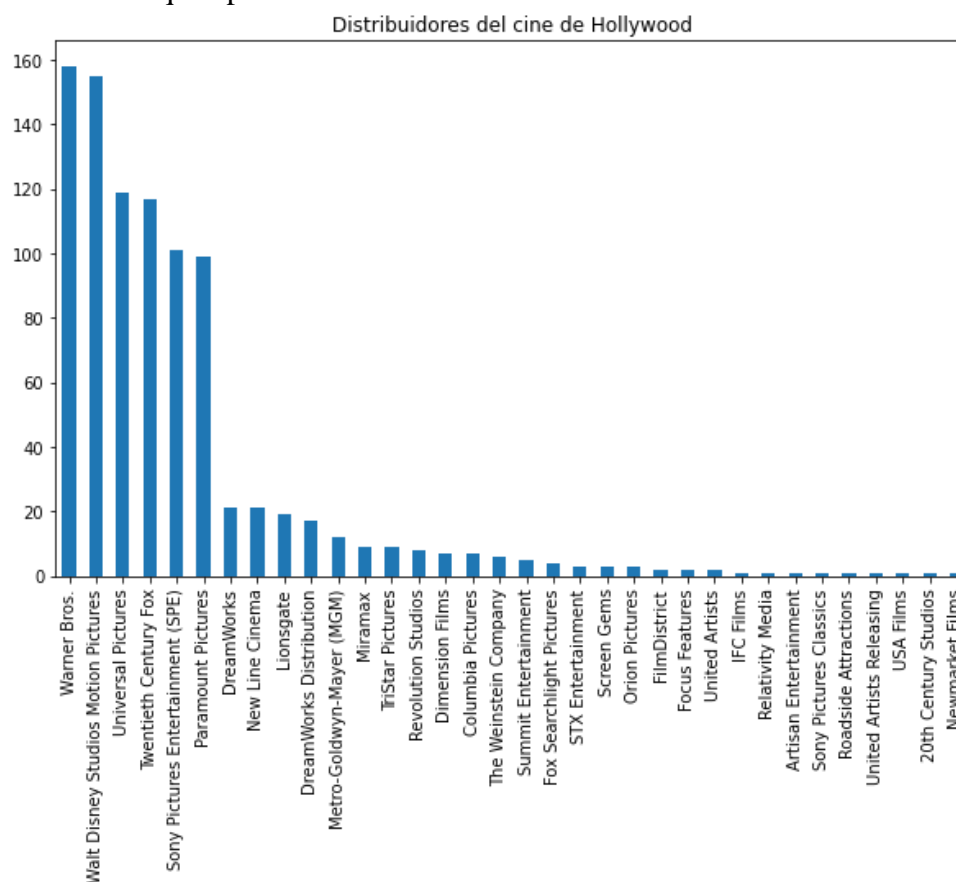
R: Restringida. Puede ser inapropiada para niños de 17 años y menores.

De la gráfica de pastel se observa que casi el 50% de las películas son clasificación PG-13 y solo casi el 2% son clasificación G. Por lo tanto, se puede inferir que el público a quien va orientado Hollywood es a un público adulto-adolescente.

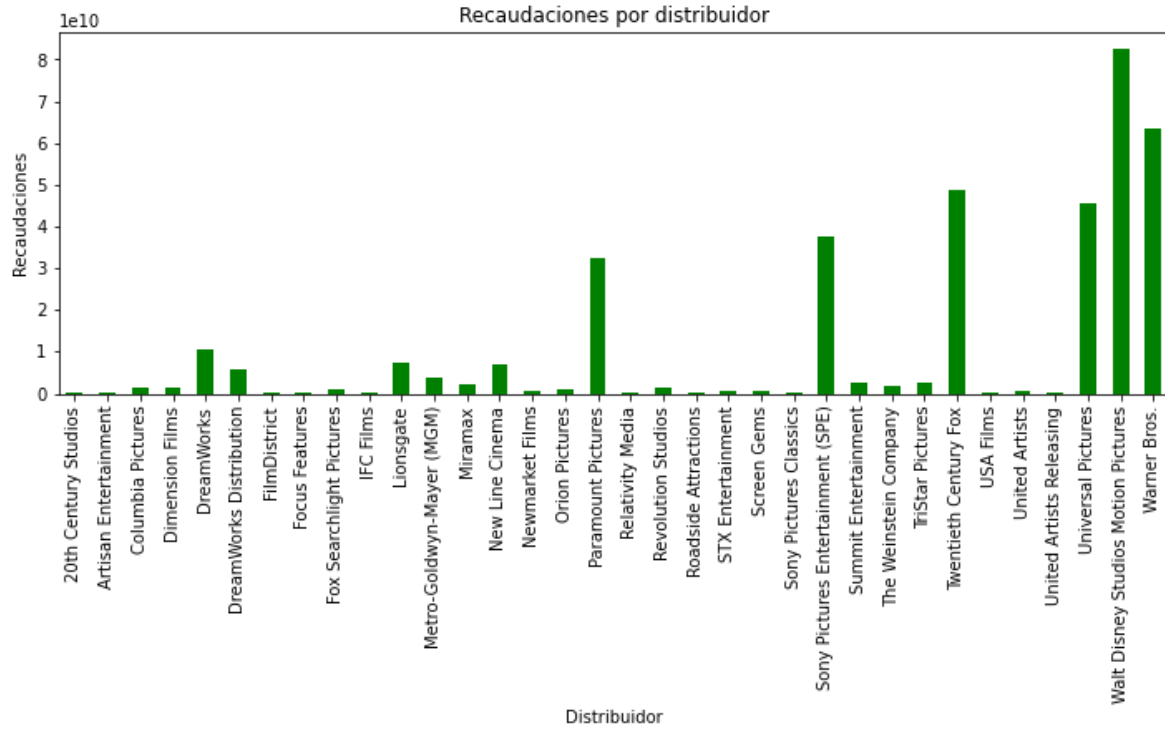
Con respecto a los estudios de cine tenemos que los más importantes son:

- Warner Bros
- Walt Disney Studios
- Universal Pictures
- Twentieth Century Fox
- Sony Pictures Entertainmet

En la siguiente figura, se muestra la lista completa de los estudios de cine en Hollywood y la frecuencia con la que aparecen.

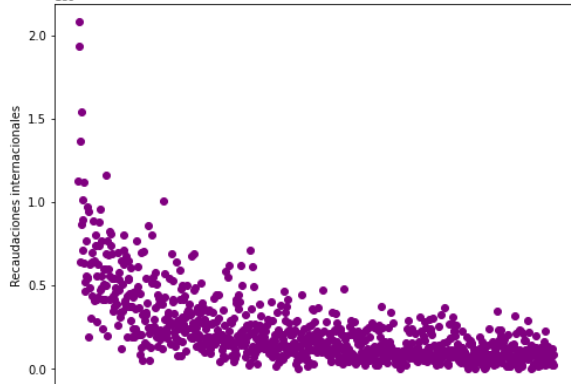


Como se puede ver en la gráfica anterior hay distribuidores que aparecen con más frecuencia en las películas de Hollywood. La siguiente imagen muestra las recaudaciones (en cientos de miles de millones de dólares) que tiene cada casa de estudio. Como es de esperarse los estudios que más películas producen son los que más recaudaciones obtienen. Sin embargo pese a que Warner Bros produce un mayor un número de películas, es Walt Disney Studios, el que más recaudaciones tiene, con un aproximado de 80 mil millones de dólares.

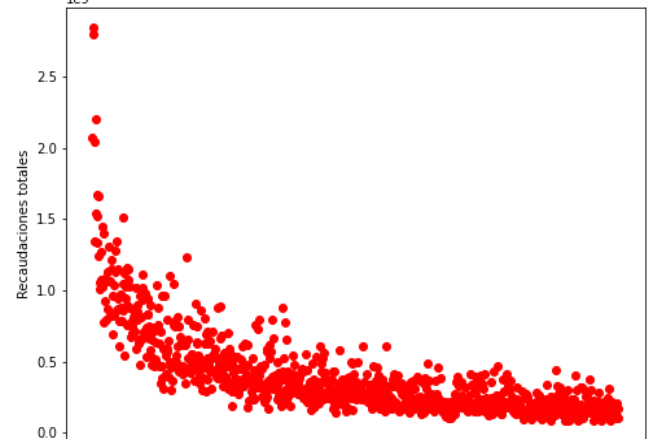


Por otro lado, un aspecto muy importante son las recaudaciones nacionales, internacionales y totales.

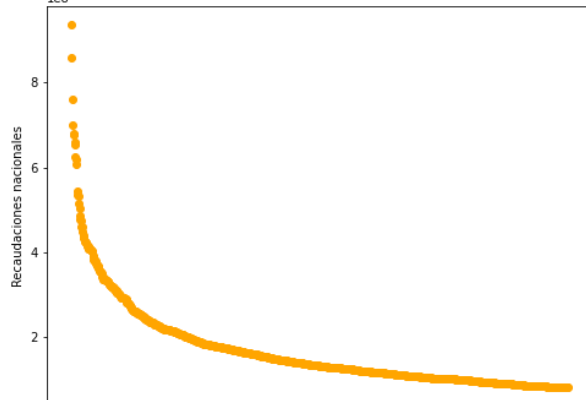
Recaudaciones internacionales en miles de millones de dolares



Recaudaciones totales en miles de millones de dolares

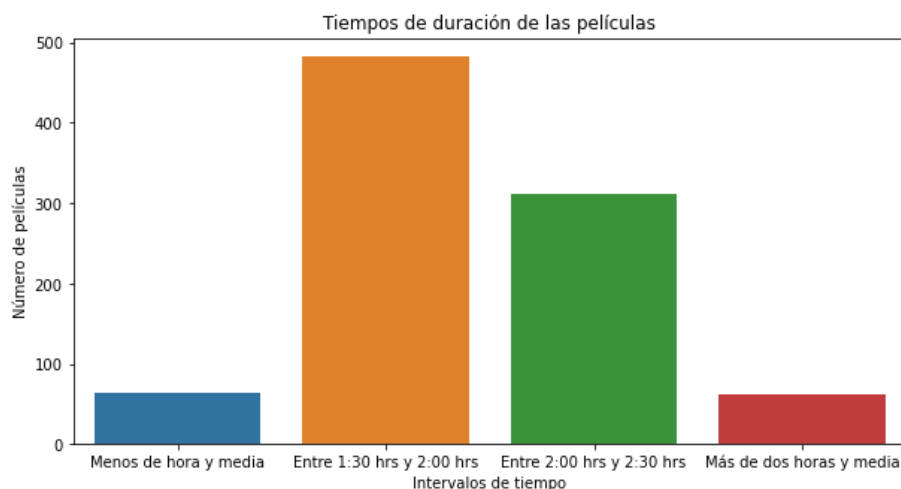


Recaudaciones nacionales en cientos de millones de dolares



Se observa que el comportamiento de las recaudaciones nacionales, internacionales y totales es muy similar. Para entender mejor el comportamiento, los datos de las recaudaciones nacionales se ordenaron y de esta forma se observa que el comportamiento de las recaudaciones es no lineal y se asemeja a un comportamiento logarítmico. Así mismo, son visibles los datos outliers que se presentan en películas que superaron por mucho a las demás, aproximadamente las 5 películas más taquilleras de Hollywood se encuentra muy por encima de las demás del ranking.

En otro sentido, como espectador de cine, el tiempo de duración es determinante para ver una película o no y la siguiente gráfica muestra justamente eso. Se observa la distribución de las películas con respecto a su duración. Como es esperarse, las películas que duran mucho o duran muy poco son las menos frecuentes, mientras que las películas que duran entre una hora y media y dos horas son las más vistas.



Por último, se muestra la relación que existe entre la fecha de lanzamiento y las recaudaciones totales.



Se observa que en las fechas vacacionales como diciembre y junio es cuando más recaudaciones existen, mientras que al inicio de año existe muy poca recaudación, esto debe tener una relación directa con respecto a las festividades de navidad y año nuevo y como se comporta la economía en esa época del año.

Por último, se muestra una tabla con datos que contineen información sobre las recaudaciones generadas en Hollywood.

	Recaudaciones		
	Nacionales	Internacionales	Totales
Media	1.68E+08	2.30E+08	3.97E+08
Desv. Estándar	1.04E+08	2.18E+08	3.08E+08
Mínimo	8.04E+07	2.60E+02	8.16E+07
25%	1.02E+08	8.85E+07	2.01E+08
50%	1.32E+08	1.64E+08	3.02E+08
75%	1.91E+08	2.89E+08	4.75E+08
Máximo	9.37E+08	2.09E+09	2.85E+09

CONCLUSIONES

Algunos datos interesantes que se obtuvieron son los siguientes:

- La película más taquillera a nivel nacional es Star-Wars episodio VII con una recaudación de casi 937 mil millones de dólares.
- La películas más taquillera a nivel internacional es Avatar con una unra recaudación que se acerca a los 3 mil millones de dólares.
- El tiempo promedio que dura una película taquillera de Hollywood es 116 minutos, aproximadamente dos horas.
- Los estudios de cine más importantes de Hollywood son Warner Bros y Walt Disney Studios.
- La recaudación promedio a nivel internacional de una película de Hollywood taquillera es de 200 millones de dólares.
- Los géneros más populares son los de Acción, Comedia y Aventura.
- La claasificación que más vende Hollywood es PG-13, teniendo en su mayoría un publico adolescente/adulto.

SECCIÓN II – Un breve análisis del municipio de Ocotlán de Morelos Oaxaca

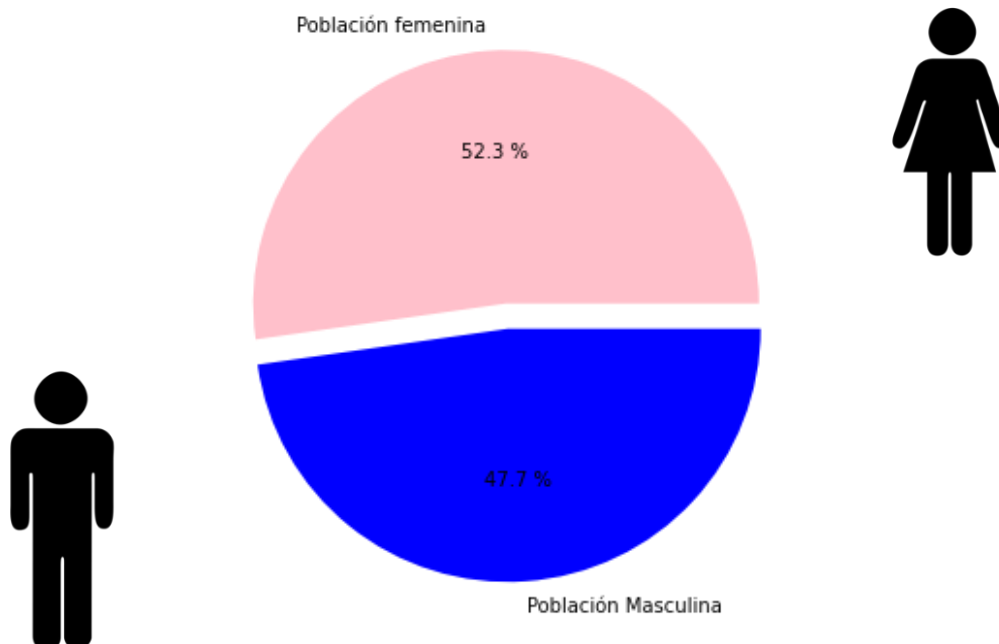
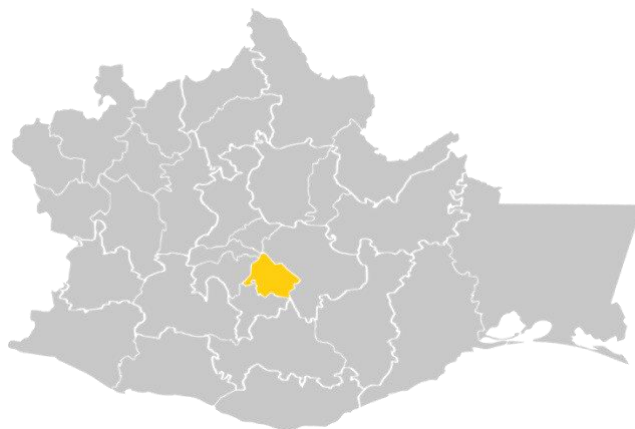
En México cada 10 años se hace el Censo de Población y Vivienda. El año pasado se dieron a conocer los resultados del Censo de Población y Vivienda 2020. En México existen 32 entidades federativas, de las cuales el estado de Oaxaca es la que tiene el mayor número de municipios, conformada por 570 municipios.

Debido a la gran cantidad de municipios, encontrar información particular sobre algunos municipios resultar ser una tarea difícil. Por ejemplo, en Wikipedia que es una de las páginas a las que la mayoría de las personas acude, no está actualizada, si buscamos información del municipio de Ocotlán de Morelos, los datos que se muestran son del año 2010. Por lo tanto, el objetivo de este breve análisis, es mostrar información actualizada del municipio de Ocotlán de Morelos.

POBLACIÓN TOTAL: 23751

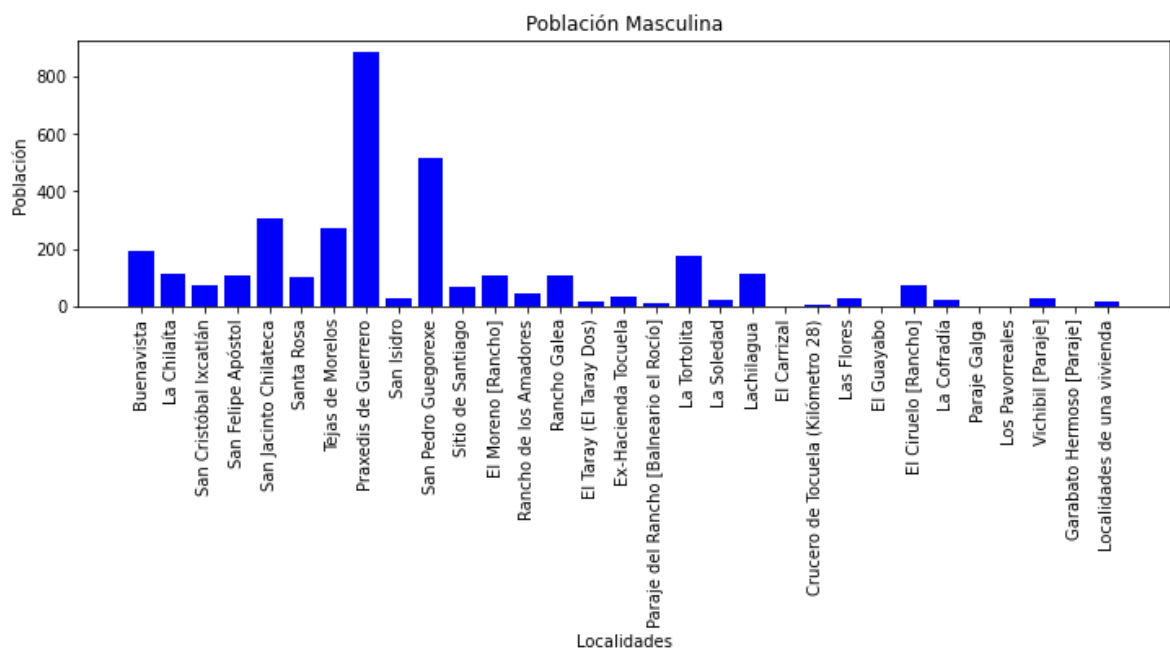
POBLACIÓN MASCULINA: 11337

POBLACIÓN FEMENINA: 12414

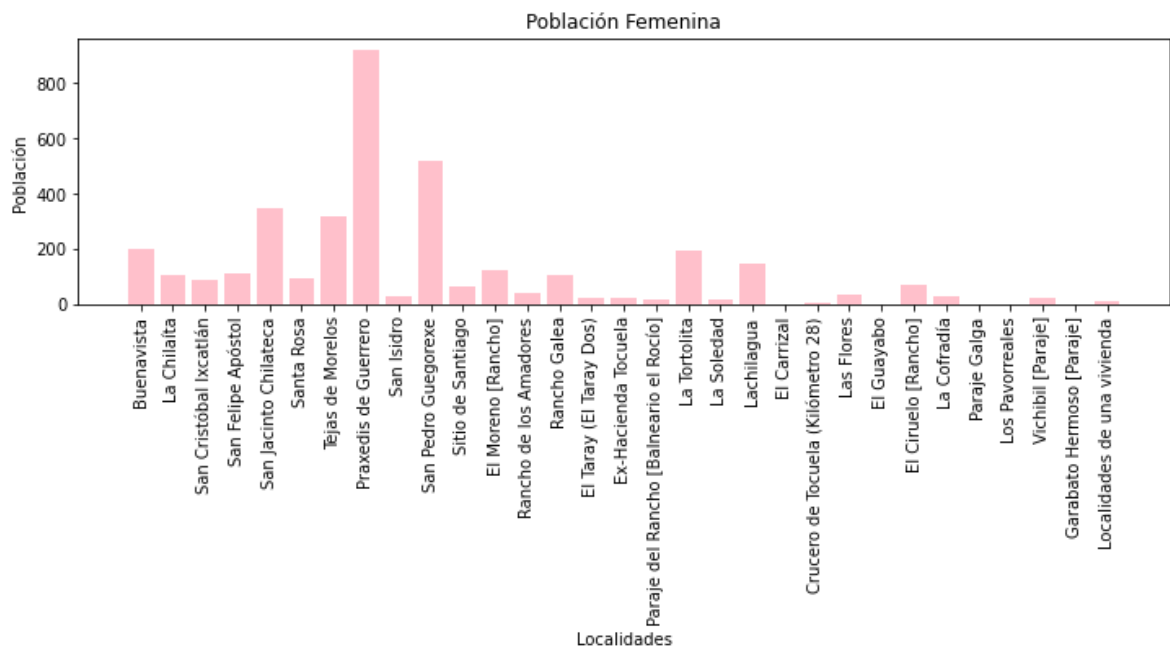


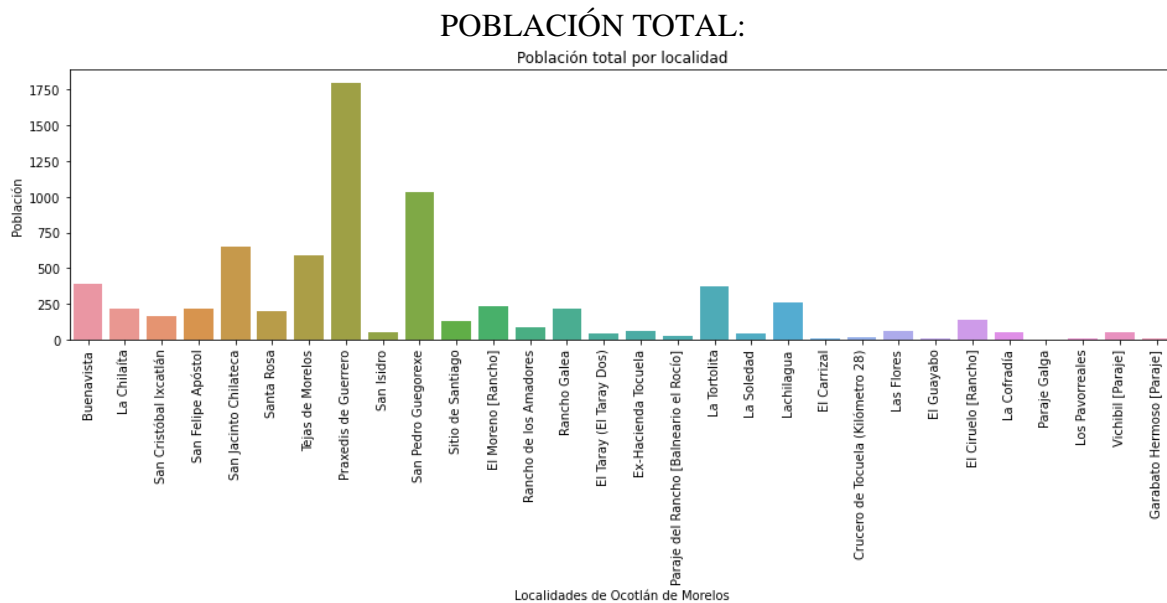
A continuación se muestran algunas imágenes que ilustran la distribución de la población en las localidades del municipio de Ocotlán de Morelos. Cabe aclarar, que dentro de las gráficas se eliminó la Localidad Ocotlán de Morelos, ya que dicha localidad abarca el 70% de la población total.

POBLACIÓN MASCULINA:



POBLACIÓN FEMENINA:





Por último se muestran datos con respecto a la media y la desviación estándar de la población.

	Población total	Población Masculina	Población Femenina
Media	791.6	354.28	387.94
Desv. Estándar	3017.34	1385.91	1538.97

Como se puede notar, la distribución de la población en las localidades de Ocotán de Morelos es muy dispersa, por lo que la media no es un buen indicador para saber el número promedio de habitantes por localidad, esto se ve reflejado directamente en la desviación estándar de los datos. El principal motivo es porque la mayoría de la población se encuentra concentrada en las localidades más céntricas del municipio, específicamente en la localidad Ocotlán de Morelos.