



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS



Avance I Proyecto Integrador

Materia: Minería de Datos

Grupo: 002

Frecuencia: Miércoles 19:00-22:00

Maestra: Mayra Cristina Berrones Reyes

Equipo: 8

Integrantes:

Francisco Javier Hernandez Ortiz

1599327

Mirthala Nallely Cantú Cortina

1614768

Raúl Ismael Martínez Tavasci

1331284

1) Nombre de la base de datos:

Homicide Reports, 1980-2014

Liga de descarga: <https://www.kaggle.com/murderaccountability/homicide-reports>

2) Descripción de los datos:

✓ Tipo de datos:

Los datos son del tipo encuesta, recopilados del Informe Suplementario de Homicidios del FBI, incluyen las características de los homicidios registrados en cada estado.

✓ Descripción de las columnas:

En las columnas se describen distintos puntos, se muestra el nombre de la agencia que estudió el caso, el tipo de agencia que es, la ciudad y el estado en que pasó el homicidio, la fecha y el tipo de crimen que fue cometido, nos da la información sobre si se resolvió o no el crimen, el sexo y edad de la víctima, si tenía o no parentesco el autor del crimen con la víctima, el arma que se utilizó, etc.

En general las columnas nos muestran datos de medición y de características, con texto corto.

3) Justificación del uso de los datos:

- ✓ Lo que más nos llamó la atención de los datos es la forma en que vienen estructurados, creemos que resultaría muy útil hacer clasificaciones para poder comprender y organizar la información, ya que es demasiada, otro punto importante es que en la actualidad es muy común ver en las noticias que se habla de homicidios, es algo alarmante y nos gustaría estudiar un poco los datos para ver si podemos de alguna manera predecir cuantos homicidios se espera que pasarán en los próximos meses o años, qué características tienen comúnmente y así se puedan quizá planear mejores estrategias dependiendo del lugar.

✓ Beneficios:

Al trabajar con estos datos encontramos distintos beneficios, primero que nada, se habla de una base de datos muy extensa, por lo que nos resulta más fácil el hacer agrupaciones o trabajar los datos, creemos que podemos sacar más provecho de una base de datos con estas características que mencionamos, queremos utilizar distintas herramientas como lo es la regresión, hacer gráficas que nos permitan hacer más visuales y fácil de entender los datos. Podemos sacar conclusiones a partir de esta base para predecir y prevenir, saber en qué puntos se debe reforzar la seguridad y qué grupos de personas son más vulnerables.

4) Planteamiento del problema:

- ✓ Cada vez son más los homicidios que ocurren, por lo que, a partir del estudio de los datos presentados, queremos crear estrategias que nos ayuden a disminuir el índice de homicidios por entidad, creemos que entre más conozcamos sobre los casos, más podremos saber cómo enfrentarlos.

5) Objetivo final:

- ✓ El objetivo de nuestro proyecto es la temprana detección del número estimado de delitos cometidos en los estados con el fin de proporcionar atención preventiva mediante diferentes estrategias propuestas por el gobierno; como la destinación de recursos, campañas deportivas, antidrogas, etc.
- ✓ Basándonos en el pronóstico del modelo, podremos tomar decisiones más certeras a la hora de destinar distintos programas sociales orientados a la prevención de actos criminales en el país. De igual manera, se podrá ver los resultados de dichos experimentos para la mejora continua del plan de acción para dicha problemática.

6) Planeación de la herramienta a utilizar

- ✓ Se contempla inicialmente utilizar el método de regresión lineal con el fin de encontrar la tendencia que seguirá en el futuro nuestra muestra, de esta manera estimar los índices de homicidios esperados.
- ✓ El método de visualización de datos será un de amplia ayuda en nuestro análisis, ya que nos ayudará a comparar por los segmentos necesarios como entidades u otras variantes.