



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FCFM



FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

Universidad Autónoma de Nuevo León
Facultad de Ciencias físico Matemáticas

Minería de Datos
M.C. Mayra Cristina Berrones Reyes

Análisis de Bases de datos.

Grupo 003

Alumno
Isidro Garza Villarreal 1818012

Base de datos:

Google Play Store (<https://www.kaggle.com/lava18/google-playstore-apps>)

Objetivo: Ayudar a los desarrolladores al elegir si quieren hacer una aplicación de pago o una aplicación gratis, esto con el fin de visualizar ingresos a futuro.

Problema Planteado: Encontrar una relación entre las 100 aplicaciones mas populares y que tanto el influye el precio en estas.

Solución: Realizar análisis estadísticos y poder ver los patrones importantes que nos pueden mostrar los datos, y que esta herramienta ayude a los desarrolladores a que sea más atractiva esta plataforma para poder tener una visualización a largo plazo

Base de datos:

Coronavirus (<https://www.kaggle.com/sudalairajkumar/novelcorona-virus-2019-dataset>)

Objetivo: Predecir el número de contagios en un lugar determinado y que esta información le pueda ser útil al sector económico para reactivar la economía.

Problema Planteado: El tomar la decisión durante este tiempo de pandemia para reactivar la economía depende mucho sobre que tan afectado se vio el dicho lugar por el virus.

Solución: Realizar una herramienta en base a métodos estadísticos que nos ayude a predecir los valores de los infectados en un lugar determinado, ya que puede que ya vaya disminuyendo la curva o que estemos en un punto alto, esto ayudara a las empresas del sector económico a tomar una decisión.

Base de Datos:

Criticas de vinos (<https://www.kaggle.com/zynicide/wine-reviews>)

Objetivo: Encontrar una relación puntuación y precio, brindar al usuario una experiencia lo más justa posible.

Problema Planteado: Muchas veces un vino puede ser muy caro y tal vez la puntuación no puede ser de acuerdo con su precio o viceversa y esto puede causar alguna controversia entre los clientes.

Solución: Diseñar una herramienta que nos ayude a calcular en base a la puntuación- precio, esto les brindara a los clientes un dato que les permitirá tomar una decisión lo mas justa posible, y para que las vinas puedan mejorar sus servicios y atraer mas clientes.

Base de datos

Clasificación de plantas (<https://www.kaggle.com/uciml/iris>)

Objetivo: Encontrar la relación que tiene el sépalo con el largo del pétalo.

Problema Planteado: Para los biólogos que se encargan de estudiar este tipo de plantas puede resultarles útil la relación que tiene el sépalo con el largo del pétalo y poder hallar algunos factores que pudieran hacer que la flor fuera mas perfecta.

Solución: Diseñar una herramienta que le ayude a los biólogos a poder saber la relación que tienen el sépalo y el largo del ancho, realizar análisis y encontrar patrones que les permitan predecir el comportamiento de dicha planta esto a su vez podía ayudarnos a encontrar un subtipo de plantas.

Base de datos:

Shows de Netflix (<https://www.kaggle.com/shivamb/netflix-shows>)

Objetivo: Analizar el impacto que tienen los actores dentro de una serie o película.

Problema Planteado: Muchas veces las películas o series no son las mejores por diferentes razones y muchas veces los actores que aparecen en estas mismas son los que levantan el éxito de las películas.

Solución: observar el éxito que tienen estas películas a pesar de que no puede llegar a ser considerado un buen material avalado por los expertos en cine o filmación, esto nos permitirá ver que los actores influyen mucho en este tipo de filmaciones

