**Logotipo, nombre de la empresa

Descripción generada automáticamenteImagen que contiene Logotipo

Descripción generada automáticamente**

**UNIVERSIDAD AUTONOMA DE NUEVO LEON**

**FACULTAD DE CIENCIAS FISICO MATEMATICAS**

**MATERIA:** Minería de Datos

**A CARGO DE:** Mayra Cristina Berrones Reyes

**TEMA:** Ejercicios de bases de datos

**POR:** Mariann Adaliz Avila Rios

*Siguiendo el ejemplo dado en clase, el alumno utilizará las ligas proporcionadas para analizar en equipo las bases de datos:*

**a)Google Play Store** [**https://www.kaggle.com/lava18/google-playstore-apps**](https://www.kaggle.com/lava18/google-playstore-apps)

*Objetivo*: Crecer la descarga de apps de pago

*Problema planteado:* Actualmente mas del 95% de las apps descargadas dentro de la playstore son apps gratuitas. Existen funciones que permiten omitir las apps de pago y cuando se hacen recomendaciones en la store muchas son apps gratuitas

*Solución:* Para las recomendaciones que hace playstore sobre las apps podemos limitar a que en su mayoría sean apps de paga utilizando árboles de decisión según las aplicaciones que el usuario ha instalado en su dispositivo entonces recomendamos apps de la misma gama ya sea por categoría o uso pero apps de paga. Podemos incluir además que de las recomendadas, las que sean gratis incluyan compras dentro del app

**b) Coronavirus corona-virus-2019-dataset**

*Objetivo:* Determinar si las medidas por país y por provinvia son efectivas para combatir el covid

*Problema planteado:* Revisando los datos se observan diferencias entre países y en muchos casos el nivel económico del país no interfiere, véase EUA como claro ejemplo de un país con posición alta económicamente. Por lo que el problema recae en si las medidas de salud son efectivas o no y cuales se debe replicar.

*Solución* Se propone una regla de asocicion por país y por provincia asociando medias como uso de cubrebocas y que tanta difusión y alcance se le da a la medida con el numero de infectados. Después otra asociación de poder económico vs el numero de muertos para identificar si la diferencia de muertos y recuperaciones cae en el poder económico del paois.

**c)Criticas de vinos** [**https://www.kaggle.com/zynicide/wine-reviews**](https://www.kaggle.com/zynicide/wine-reviews)

*Objetivo: Crecer la compra de ciertos vinos*

*Problema planteado:* La base de datos obtiene clasificaciones por twiter y entrega el handle de usuario además clasifica los vinos con “variety, location, winery, price, and description”. Como se puede entregar al usuario mejor recomendaciones de vino de acuerdo a sus compras y a su descripcion

*Solución:* Creo que deberíamos asociar el precio promedio de vinos que consumen usuarios de la variedad de vino y su descripción para poder ofrecer otras entradas de vinos del mismo promedio y descripción, sin embargo de otra variedad. Se puede hacer con un árbol de decisiones para poder además ofrecer publicidad de complementos para el vino, es decir quesos o tablas y carnes frías que vayan con la variedad determinada de vino

**d)Clasificación de plantas** [**https://www.kaggle.com/uciml/iris**](https://www.kaggle.com/uciml/iris)

*Objetivo:* Encontrar las plantas mas grandes o con mejores características

*Problema planteado* Las plantas de la familia Iris crecen de diferente forma dependiendo de su especifico, además varían en tamaño y alto. La base de datos falta información para este problema pero si se diera ubicación de la planta pudiéramos resolverlo

*Solución* Hacer técnica de clustering para agrupar el tipo de planta con la ubicación y con el mejor resultado. Teniendo los grupos de los mejores resultados podemos encontrar las diferencias y así buscar obtener esos mejores resultados desde el inicio

**e)Shows de Netflix** [**https://www.kaggle.com/shivamb/netflix-shows**](https://www.kaggle.com/shivamb/netflix-shows)

*Objetivo* Identificar si Netflix debe invertir en producciones propias y compra

*Problema* planteado De los insigths ya otorgados observamos un crecimiento exponencial en el consumo de series de Tv Vs películas. Se debe entender al usuario y la demanda para saber si Netflix deberá seguir creciendo esta industria o invertir en películas.

*Solución* Para entender al usuario y su demanda propongo crear perfiles d eusuario y por ubicación ver cuantos de estos perfiles existen. Usando la técnica de clustering formamos estos grupos en base al tiempo que gastan viendo series o películas y si visitan series viejas o buscan nuevas series