



**UNIVERSIDAD AUTONOMA DE NUEVO LEON**  
**FACULTAD DE CIENCIAS FISICO MATEMATICAS**



**MINERIA DE DATOS**

**GRUPO: 003**

**Maestra: Mayra Cristina Berrones Reyes**

**Minutas Base de Datos**

**Alumna: Marlene Calderón Rangel**

**Matricula: 1811330**

## Google Play Store Apps

**Objetivo:** Segmentar el tipo de aplicaciones y acomodarlas dependiendo de los reviews de los usuarios en categorías para así poder acceder a las apps de una manera más fácil y el usuario pueda encontrarlas y llegar a ellas de una manera más sencilla.

**Problema planteado:** Que los usuarios descargan la aplicación innecesariamente y no hacen uso de ella, o que los usuarios no pueden encontrarla o no la descargan ya que no esta visible para ellos de una manera muy notoria. Así mismo muchos de los usuarios no dan el review necesario.

**Solución:** Segmentar el tipo de aplicaciones dependiendo de la edad del usuario, por niños, adolescentes, adultos o por intereses de trabajo, escuela o entretenimiento y estas puedan aparecer dependiendo del rango en el que se encuentre para así poder encontrarlas más fácilmente.

## Novel Corona Virus 2019 Dataset

**Objetivo:** Probabilidades, numero y causas de muerte. Como es que se correlacionan los fallecimientos con la edad, causas, enfermedades y como afecta en la probabilidad de muerte.

**Problema planteado:** Problemática es que existen casos jamás conocidos ya que son asintomáticos, debemos de saber si estos también son probables a contagiar e incluso a fallecer. Que tan probable es que una persona fallezca.

**Solución:** Realizar una regresión lineal con alguna ecuación como es que se relacionan las variables de edad, peso, enfermedades con las muertes, en donde se puedan registrar los datos necesarios de los enfermos que nos ayudarían a saber si aumenta la probabilidad de muerte, como edad, enfermedades, e ir resumiendo y acercándonos a las probabilidades exactas y quienes son los mas propensos a morir o poder encontrar una solución en como cortar la cadena de transmisión.

## **Wine Reviews**

**Objetivo:** Encontrar un modelo de producción efecto para identificar la región exacta para extraer el vino, el precio de la botella y a donde distribuirlo de manera que los clientes consuman el producto que sea mas de su preferencia y cumpla con las características que el cliente prefiere.

**Problema planteado:** Que el no se identifica a que región se debe de mandar y es así que el precio de la botella de vino puede aumentar o puede aumentar la merma, dependiendo de si no se consume el producto debido a la producción que no es asertiva o no es bien distribuida.

**Solución:** Realizar una recolección de datos poder identificar por medio de algún diagrama de árbol como es que se designan los viñedos así buscando los que mas se consuman en las regiones específicamente.

## **Iris Especies**

**Objetivo:** Identificar el tipo de planta o flores que presentan algún comportamiento diferente que se pueda tomar como dato atípico.

**Problema planteado:** Identificar anomalías en las características de las plantas, ya sea grosor o altura de la planta y así comparar y encontrar las características específicas y normalizar los datos de ellas.

**Solución:** Crear agrupaciones con las características específicas de las dimensiones obtenidas de las plantas y plantear rangos para cada tipo de planta y si es que no cumplen con los rangos y características específicas clasificarlas como datos atípicos e identificar la razón del problema en ciertas plantas.

## **Netflix Movies and TV shows**

**Objetivo:** Identificar que el tipo de películas y series que mas prefieren los usuarios para así poder concentrar su contenido en lo que mas le interesa al usuario dependiendo del rating que tenga la clasificación de programa o película.

**Problema planteado:** Netflix ha fallado con el contenido en el que se concentra ya que el usuario no lo termina, esto hace que el rating baje y el dinero invertido en concentrarse en realizar su propio contenido y la creación de nuevas plataformas o aplicaciones lo cual ha quitado mucho contenido de agrado para el usuario.

**Solución:** En base al rating y las clasificaciones dependiendo de la temporada del año, Netflix debería de crear más películas o series con respecto a los temas de interés actuales o así mismo ir introduciendo material dependiendo de la época y de la misma manera dejar el contenido que mas tenga rating o que mas sea de agrado para el usuario.