

---

**UANL**



---

**FCFM**

# **Ejercicio Práctico de Base de Datos**

Minería de Datos

**Fase II**

---

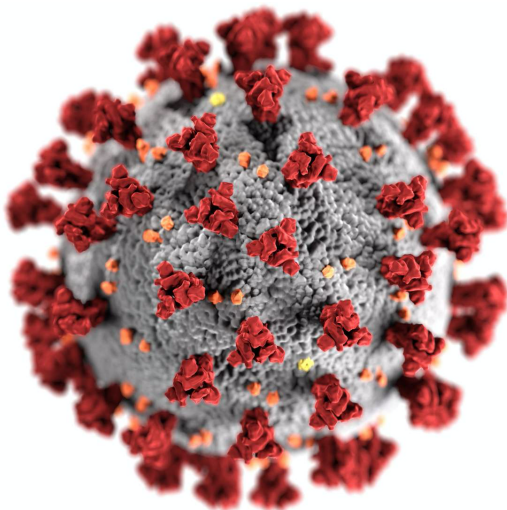
Jorge Noé Zúñiga Gómez 1565813



Objetivo: Determinar la influencia que tiene el numero de reseñas (Reviews) y la Calificación General del Usuario (Rating) a la hora de elegir descargar una app u otra de determinada categoría o funcionalidad.

Problema Planteado: Se requiere analizar cada subgrupo de los diferentes tipos de aplicaciones y solo analizarlos en función de Reviews y Rating y de este modo considerar más atención en estas secciones en la compra o descarga, fomentando la valoración de la app, por medio de un descuento o membresía gratis en el caso de ser una app de pago.

Solución: Determinar la proporcionalidad estadística del tipo de muestra por subgrupo y encontrar si hay alguna correlación que nos permita determinar en qué tipos de apps si es más influyente esta información.



Objetivo: Determinar que zonas geográficas en el mundo son mas adversas al cambio, de modo que podamos llegar a la conclusión de que si hubiera un futuro suceso podríamos brindar mucho más información y medidas a ciertas áreas que son más adversas a un cambio.

Problema Planteado: Hay una resistencia al cambio en diferentes zonas geográficas, determinadas por su cultura, clase social y económica, a su mismo modo este resistencia pudo haber sido contribuyente directo en la propagación del virus, y podríamos determinar por zonas esa aversión para considerarse en futuras anomalías.

Solución: Teniendo un conocimiento basado en estadística de resultados, brindado por el tiempo que le tomo a una zona cumplir con las medidas de precaución requeridas podríamos evitar que a esas zonas en un futuro conociendo su aversión tengan que resentir las mismas consecuencias.



**Objetivo:** En general esta persona busca crear un modelo que pueda identificar la variedad, la bodega y la ubicación de un buen vino en función de una descripción de modo que el modelo se transforme en una especie de sommeliers digital.

**Problema Planteado:** La determinación de un buen vino primero está determinada por el gusto de cada persona puesto que al menos esta base de datos maneja casi la misma puntuación para todos los vinos

esto sin si quiera saber de que manera fueron determinados, lo que se podría hacer es en base a los datos recabados, es decir, crear una app que te haga una encuesta creando rango de presupuestos calculados en base a la zona de donde se ubican y al precio que aparece aquí al mismo tiempo el tema de inventarios podría ser un gran problema.

**Solución:** Podríamos fomentar este modelo en las diferentes viñetas y casas de vinos de las diferentes zonas para que al mismo tiempo parte del incentivo sea otorgarnos sus bases de datos como pago simbólico del marketing y alcance que se puede brindar a nivel mundial, de modo que podamos enlazar algunas casas de vinos no muy conocidas. Al mismo tiempo proporcionar una encuesta de los datos y características que necesitamos para enriquecer este modelo.



**Objetivo:** Determinar el crecimiento de las flores Iris, en las diferentes categorías:

- Longitud del sépalo
- Longitud del pétalo
- Ancho del pétalo

Para determinar el crecimiento de las flores en base a un inventario desde que son cultivos y fijar una fecha de entrega al portador para cuando alcancen ciertas características.

**Problema Planteado:** Por ser un tema de plantas el promedio de crecimiento no está muy relacionado con la varianza del crecimiento de estas plantas por las diferentes zonas donde se plantan.

**Solución:** Realizando un muestreo como el que tenemos en esta base de datos si podríamos determinarlo por incluso sus categorías para eliminar esa parte sesgada que pueda causar ruido en nuestros Datos.



Objetivo: Determinar en base a los datos que países ven más Netflix, y en que proporción el contenido que ven son Películas y TV Shows, de modo que pudiéramos incluir edades de los suscriptores para centrarnos directamente en el cliente potencial, y en el caso de que la minoría sean películas ver cual es la duración máxima que están dispuestas a ver seguridad.

Problema Planteado: Además de que Netflix cuenta con diferentes perfiles de entrada que determinan las categorías y gustos en función de un perfil que se genere en base a lo que más ven, esto podría representar un sesgo puesto que muchas de las personas mayores que ven Netflix sólo ingresan a la Smart TV y ponen lo que el perfil que sea y ven lo que sea que les aparezca y son datos que estarían dañando nuestra muestra.

Solución: No solo el tema de los Usuarios es suficiente, podríamos hacer una pregunta al ingresar al Netflix inclusive a cierto tiempo también de estar en la cuenta, Sigues siendo "Juanita" ? De modo que si la esta viendo la mamá de Juanita se le presente la opción de decir, No, soy su mamá y cambiar de usuario para recabar mas información.