

Universidad Autónoma de Nuevo León

Facultad de Ciencias Físico Matemáticas

Séptimo Semestre  
Agosto 2020 ~ Enero 2021

Licenciatura en Actuaría  
MINERÍA DE DATOS  
Docente: Mayra Berrones

“Análisis de Bases de Datos”

Alumno: Daniel de la Rosa Coss  
Matricula: 1723784  
Grupo: 003

16/10/2020

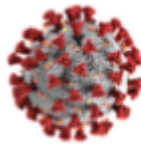


**Base de datos:** Google Playstore Apps.

**Objetivo:** Encontrar las mejores aplicaciones de juegos en su relación rating y peso de la aplicación.

**Problema:** Hay gran variedad en juegos que dirigidos a distintos tipos de usuarios, como niños, adolescentes por lo que los datos podrían verse sesgados.

**Solución:** Desarrollar un programa que separe los juegos por clasificación de contenido y evaluar la relación entre rating y peso.



**Base de datos:** Novel Coronavirus 2019 Dataset

**Objetivo:** Realizar una muestra para estimar la tasa de mortalidad.

**Problema:** La relación de personas muertas entre personas contagiadas no es igual en cada país por lo que el número podría verse sesgado.

**Solución:** Desarrollar un programa que divida los datos por estratos para poder aplicar un muestreo estratificado, donde cada país será un estrato.



**Base de Datos:** Wine Reviews.

**Objetivo:** Hacer un catálogo con recomendaciones de los mejores vinos.

**Problema:** Hay una gran variedad de vinos con distintos precios, si los mejores vinos son caros no cualquiera podría comprarlos.

**Solución:** Crear un programa que haga una clasificación de los mejores vinos en base a distintos rangos de precios.



**Base de Datos:** Iris Species.

**Objetivo:** Determinar si existe una relación entre el tamaño y la anchura del pétalo.

**Problema:** Tenemos datos de 3 distintas especies de plantas por lo que no se puede buscar la relación de forma directa.

**Solución:** Realizar un programa que agrupe los datos por especie y después verifique si existe una relación.



**Base de Datos:** Netflix Movies and TV Shows.

**Objetivo:** Buscar si hay una relación entre la clasificación de contenido y duración de la película o TV show.

**Problema:** Dentro de la base de datos no solo hay películas, también hay series y son de distintos países.

**Solución:** Realizar un programa que agrupe entre países y clasifique entre series y tv shows y después vea si hay relación entre la duración y la clasificación de contenido.