

MINERIA DE DATOS

Ejercicio de bases de datos

José Andrés García Ortega
Diego Adrián Cruz Martínez
Diana Gabriela Ramírez Moreno
Ofelia Velazquez Montemayor
Jose Miguel Martínez Reyes



Wine Reviews Data set

Hipotesis

Se puede estimar la calidad relativa de un vino a partir de sus características de origen , tipo y precio

Problema planteado

Para el consumidor común resulta complicado saber la calidad de un vino debido a la gran variedad y sutiles diferencias según el origen, por lo que es probable que no elijan la mejor calidad por el precio que estan pagando

Objetivo

Brindar a los consumidores un estimador de calidad del vino que están interesados, a partir de datos como el pais de origen, precio y tipo de vino



Shows de netflix data set

Hipotesis

Se puede evaluar el tiempo y la categoría a partir de sus características de las cuales la gente frecuenta para realizar unos shows ideales para la audiencia.

Problema planteado

La plataforma ha creado un nuevo mercado y es el crear nuevos show propios, por lo cual el tener un estudio de lo que sus usuarios más les satisface ver con sus tipo de audiencia y la duración, lo cual podria reducir gastos en inversión y mayores beneficios monetarios.

Objetivo

Crear series y películas con las temporadas y duración que sus usuarios esten dispuestos a brindar para dar una mejor servicio y calidad.



Google Play Store Apps Data set

Hipotesis

Se puede crear un algoritmo que permita dar sugerencias de aplicaciones que sean de mayor interés para el usuario en base a sus búsquedas o aplicaciones instaladas.

Problema planteado

Google Play a pesar de ser una de las tiendas de aplicaciones más importantes del mundo no ha logrado dar recomendaciones en base a los perfiles de sus usuarios. Estas aplicaciones se podrían asignar en base a su edad, género, búsquedas recientes, incluso por recomendar las mejores aplicaciones en base a su calificación obtenida por los usuarios (rating).

Objetivo

Facilitar al usuario a encontrar aplicaciones idóneas con sugerencias basadas en sus búsquedas recientes o su interacción en general con la plataforma.

Iris Species dataset

Classify iris plants into three species in this classic dataset.

Hipotesis

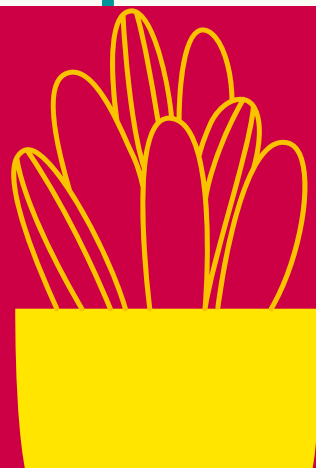
Se pueden conocer las propiedades que poseen las plantas de cada especie a partir de sus características.

Problema planteado

Las plantas pueden llegar a tener propiedades de gran ayuda para la producción de productos, desgraciadamente las personas no conocen y por lo tanto no aprovechan los remedios que nos da la naturaleza.

Objetivo

Bridar a la población los datos necesarios para que puedan hacer uso de estas plantas, aprovechando sus propiedades.



COVID-19



Hipotesis

Se puede evitar sobrepasar la capacidad hospitalaria debido al Covid-19

Problema planteado

Las personas al enfermarse, un porcentaje de ellas requiere hospitalización, lamentablemente personas han fallecido por no tener esta atención medica adecuada dentro de un hospital por la gran cantidad de gente dentro de los hospitales

Objetivo

Realizar un modelo de predicción de casos para tener previsto cuanta gente necesitará hospitalización y a que ritmo se estarán rotando los pacientes