



Universidad Autónoma de Nuevo León
Facultad de Ciencias Físico Matemáticas



Unidad de Aprendizaje: Minería de datos

Evidencia de aprendizaje: Ejercicio de bases de datos

Docente: Lic. Mayra Cristina Berrones Reyes

Grupo: 001

Equipo 4:

Sandra María Cavazos Huerta	1877283
Alberto Elizondo Villarreal	1869764
Arturo Isaac Sánchez Tovar	1867031
Marlene Michel Sepúlveda Bermudez	1862646
Diana Villarreal Garza	1877498

Fecha de entrega: 16 de agosto del 2021

A) Google Play Store

Hipótesis o pregunta de interés

Se pueden identificar áreas de oportunidad en relación a la creación de aplicaciones con ciertas características (categoría, género, precio, público quien puede acceder al contenido, tamaño de descarga) debido a que se presentan bajas calificaciones.

Problema planteado

Actualmente, Google Play Store posee una gran variedad de aplicaciones con distintas características las cuales se pueden resumir en la categoría, género, tipo, tamaño de descarga, precio y público a quien va dirigido.

Las aplicaciones que se pueden descargar dentro de Google Play Store obtienen diferentes calificaciones y reseñas por parte del público día con día. Estas calificaciones y reseñas resultan ser de gran importancia ya que mediante ellas se pueden identificar las aplicaciones con las calificaciones más bajas o críticas negativas y a su vez esto permite conocer las áreas de oportunidad donde nuevas aplicaciones se pueden integrar y tomar gran fuerza dentro del mercado considerando sus características para que se sigan ofertando aplicaciones de calidad que puedan mejorar las calificaciones.

Objetivo

Identificar las aplicaciones dentro de Google Play Store con las calificaciones más bajas según sus características.

B) Coronavirus

Hipótesis o pregunta de interés

Se puede reducir el porcentaje de fallecimientos tras contraer COVID-19.

Problema planteado

La pandemia causada por COVID-19 ha cobrado millones de vidas a nivel mundial, causando graves problemas en el ámbito económico y social. En el año y medio que lleva la pandemia, el virus ha causado más de 4.3 millones de muertes.

A pesar de que es un fenómeno a nivel mundial y se ha estado luchando diariamente para combatirlo, no todos los países o regiones se ven afectados de la misma manera o han logrado avanzar con el mismo ritmo. Con el fin de reducir las muertes ocasionadas por la pandemia, se han realizado campañas de vacunación y programas de salud, pero no todos han tenido acceso a este tipo de medidas contra el COVID-19.

Objetivo

Identificar los países y regiones con menor mejoría en relación de casos confirmados y personas recuperadas.

C) Críticas de vinos

Hipótesis o pregunta de interés

Es posible identificar la calidad de los vinos a través de un modelo predictivo sin necesidad de la cata de un maestro sumiller.

Problema planteado

Actualmente, la cata de vinos es degustada por un maestro sumiller que es un experto en el estudio y en la práctica de una gran variedad de habilidades relacionadas con el vino. Sin embargo, durante la cata encontramos ciertas dificultades: puede ser que tengamos una percepción dirigida (haciendo un esfuerzo orientado hacia determinado carácter, no permitiendo apreciar los otros), puede ser que no podamos transmitir en palabras lo que sentimos, o podemos tener falta de elementos suficientes memorizados como para emitir un juicio.

Las principales causas de error que tiene una cata son la propia personalidad del experto (tanto a experiencia como el vocabulario que disponen, la importancia que le dan a determinados defectos o si se dejan sugerir por el juicio de otros), las condiciones externas (como hora del día, fatiga de las papilas gustativas, o gustos que se suman), o el momento en el que se cante (por los efectos de contraste, ya que por ejemplo un vino tánico se parece más recio si antes probando uno ligero, la temperatura de servicio, etc.).

Objetivo

Crear un modelo que pueda brindar una breve descripción de la calidad del vino tomando en cuenta la variedad, la bodega, el precio y la ubicación de éste.

D) Clasificación de Plantas

Hipótesis o pregunta de interés

Se puede reducir el tiempo en la identificación del tipo de especie de la planta "Iris" identificando sus medidas de anchura y longitud de su sépalo y del pétalo.

Problema planteado

Es de suma importancia saber el tipo de especie de iris, ya que se tiene un cuidado completamente diferente dependiendo del tipo de especie (los iris de jardín tienen una menor resistencia a las heladas, las especies de estanques necesitan un lugar húmedo y la iris germánica se desarrolla en un ambiente seco y árido).

Una vez identificado el tipo de especie, se pueden tomar las medidas de cuidados necesarias para que esta especie tenga el mejor desarrollo posible.

Objetivo

Identificar las proporciones de los tipos de especies para poder distinguir las diferencias y relaciones entre ellas mismas.

E) Shows de Netflix

Hipótesis o pregunta de interés

Se puede reducir el número de series canceladas tras una o dos temporadas de transmisión, identificando las características de series exitosas previas.

Problema planteado

Actualmente, la plataforma de streaming Netflix ha optado por cancelar series de televisión cuando han alcanzado su primera o segunda temporada de transmisión. Se ha destacado que las razones de esto se basan en captar más espectadores con nuevo material televisivo, reducir costos que se elevan considerablemente por cada temporada, así como el bajo crecimiento de espectadores a lo largo de ellas.

En consecuencia de lo anterior, los usuarios de la plataforma de streaming han mostrado un descontento, justificando que las historias terminan de manera inconclusa o inesperadamente. Aunado a esto, se trae efectos en la disminución de usuarios de la plataforma y bajas en ventas de Netflix, así como una imagen desfavorable del manejo de transmisión de las series de televisión en manos de esta compañía.

Objetivo

Identificar los géneros de series exitosas dependiendo del número de temporadas de duración.