

Descripción del proyecto

Como analista de datos, tu trabajo consiste en analizar datos para extraer información valiosa y tomar decisiones basadas en ellos. Esto implica diferentes etapas, como la descripción general de los datos, el preprocesamiento y la prueba de hipótesis.

Las hipótesis son suposiciones hechas sobre una población con base en una muestra de datos. Estas suposiciones se prueban utilizando métodos estadísticos para determinar si son verdaderas o falsas. Una prueba de hipótesis consiste en hacer una declaración sobre la población, recopilar datos y usar métodos estadísticos para probar esa declaración.

Siempre que realizamos un análisis, tenemos que formular hipótesis que luego podamos poner a prueba. A veces aceptamos estas hipótesis; otras veces, las rechazamos. Para tomar las decisiones correctas, una empresa debe ser capaz de entender si está haciendo las suposiciones correctas.

Descripción del proyecto

En el contexto de este proyecto, probarás una hipótesis relacionada con las preferencias musicales de dos ciudades. Esto se hará mediante el análisis de datos reales de transmisión de música online para probar la hipótesis a continuación y comparar el comportamiento de los usuarios de estas dos ciudades.

Esto implicará analizar los datos reales de transmisión de música online para comparar el comportamiento de los usuarios en Springfield y Shelbyville. El proyecto se divide en tres etapas, cada una con objetivos específicos.

En la etapa 1, proporcionarás una descripción general de los datos y anotarás tus observaciones. En la etapa 2, preprocesarás los datos al limpiarlos. Finalmente, en la etapa 3, pondrás a prueba la hipótesis siguiendo los pasos de programación necesarios para probar cada declaración y comentando tus resultados en los bloques apropiados.

Al completar estas etapas, podrás extraer información valiosa de los datos y tomar decisiones basadas en datos.

Hipótesis

Para este proyecto, recopilamos los requisitos y desarrollamos una hipótesis que debe ser confirmada o negada.

Al probar hipótesis, es importante tener en cuenta que pueden aceptarse por completo, aceptarse parcialmente, rechazarse parcialmente o rechazarse por completo.

Cuando una hipótesis se acepta por completo, significa que los resultados de la prueba respaldan sin ambigüedades la declaración hecha sobre la población.

Si se acepta solo parcialmente, significa que los resultados respaldan la declaración en cierta medida, pero no lo suficiente como para aceptarla por completo.

Por el contrario, si una hipótesis se rechaza por completo, significa que los resultados de la prueba no respaldan la declaración hecha sobre la población.

Por último, una hipótesis también puede ser rechazada parcialmente si los datos indican que es falsa, pero no puedes rechazarla por completo. Cuando se interpretan los resultados de una prueba de hipótesis, es importante tener en cuenta estas diferentes posibilidades.

Aquí está nuestra hipótesis para aceptar o rechazar:

1. La actividad de los usuarios y las usuarias difiere según el día de la semana y dependiendo de la ciudad.

Diccionario de datos

Los datos se almacenan en el archivo /datasets/music_project_en.csv. También puedes descargarlo [aquí](#).

Descripción de las columnas:

- 'userID': identifica de forma exclusiva a cada usuario o usuaria;
- 'Track': título de la canción;
- 'artist': nombre del artista;
- 'genre': género musical;
- 'City': ciudad del usuario o la usuaria;
- 'time': hora del día en la que se reprodujo la pista (HH:MM:SS);
- 'Day': día de la semana.