

Laboratorio 2 - creación de una visualización

Objetivos

En esta actividad, utilizará Tableau Desktop Public, una versión gratuita del producto Tableau Desktop para crear visualizaciones significativas para responder preguntas.

Parte 1: Usar un conjunto de datos públicos en Tableau

Parte 2: Usar datos para crear visualizaciones significativas

Información básica / situación

Las visualizaciones se utilizan a menudo para Descubrir tendencias o descubrir correlaciones previamente desconocidas dentro de un conjunto de datos. La visualización correcta puede mostrar áreas de una empresa que necesitan atención o hacia dónde deben dirigirse los programas de marketing.

En este lab, usará un conjunto de datos de la empresa ficticio que está disponible públicamente en Tableau. Con diferentes tipos de visualizaciones, creará tablas y gráficos significativos que pueden ilustrar los problemas existentes en el negocio.

Recursos necesarios

- Computadora con acceso a Internet
- Aplicación Tableau Desktop

Instrucciones

Parte 1: Uso de un conjunto de datos públicos en Tableau

Paso 1: Conectarse al conjunto de datos.

1. Abra Tableau Desktop. En la página de inicio, seleccione **Conjuntos de datos de muestra** en el panel **Descubrir** a la derecha.
2. Desplácese hacia abajo hasta la categoría **Entretenimiento**. Descargue el conjunto de datos de las **historias más rentables de Hollywood** (csv) haciendo clic en **conjunto de datos (csv)** a la derecha de la ventana.
3. Cuando se complete la descarga, regrese a Tableau Public Desktop. En el panel **Conectar**, seleccione **Archivo de texto**.
4. Busque y abra el archivo **HollywoodsMostProfitableStories.csv** que acaba de descargar.

Paso 2: Explorar el conjunto de datos públicos.

El análisis de datos comienza con una pregunta o un problema. Con este conjunto de datos de películas, un analista puede responder preguntas como "¿Qué género de películas obtiene las mejores puntuaciones de audiencia?" o "¿Qué estudio de películas produjo más ingresos en un año específico?"

1. Cuando carga el conjunto de datos, aparece la página de resumen **Origen de datos**. Aquí puede verificar los tipos de datos contenidos en cada campo y realizar cambios rápidamente en cualquier tipo que no sea correcto.
2. Desplácese hacia la derecha en la tabla de la derecha hasta que vea la columna **Año**. Haga clic en el icono **Tipo (#)** encima del **Año**.

¿Cuál es el tipo de datos que realmente contiene la columna Año?

Fecha

¿Por qué el tipo que se muestra para el campo Año no es correcto?

Porque Tableau reconoce los años como números.

3. Cambie el tipo en el campo **Año al tipo correcto**.

¿Qué sucede con los datos de la columna?

Los datos cambian al 1 de enero de cada año.

4. Vuelva a cambiar el tipo del campo Año a **Número (completo)**.
5. Revise los campos incluidos en este conjunto de datos.

¿Cuál podría ser otra pregunta que podría responder mediante el análisis de estos datos?

¿Qué años presentaron el mayor margen entre popularidad e ingresos, y qué factores podrían explicarlo?

Parte 2: Usar datos para crear visualizaciones significativas

Paso 1: Pregunta: ¿Qué género de película obtiene las mejores puntuaciones?

Comencemos con la pregunta: "¿Qué género de película obtiene las mayores puntuaciones de audiencia?"

1. Haga clic en la ficha **Hoja 1** en la parte inferior de la pantalla para abrir el espacio de trabajo. Los campos de la tabla de datos están en el panel de la izquierda.

Nota: Si Tableau aún muestra la base de datos de películas de la práctica de laboratorio anterior, haga clic con el botón derecho en el título de la base de datos en la parte superior izquierda y seleccione **Cerrar**. Si aparece una ventana emergente, seleccione **Aceptar**.

2. Arrastre el campo Género al estante Columnas en la parte superior del espacio de trabajo.
3. Arrastre el campo Porcentaje de puntuación de audiencia al estante Filas en la parte superior del espacio de trabajo. La puntuación de la audiencia se muestra automáticamente como una suma cuando la mueve al estante Filas. Sin embargo, para este análisis se necesita la puntuación media de cada género. Use la flecha hacia abajo junto a SUMA (% de puntuación de audiencia) en el estante Filas para seleccionar Medida (Suma)> Promedio.
4. Tableau ahora debería presentar un gráfico de columnas que muestre las puntuaciones promedio de la audiencia para cada género.

Según el gráfico de columnas resultante, ¿Qué género tiene la mayor puntuación de audiencia?

Fantasía

5. Para profundizar un poco más en este análisis, es necesario asegurarse de que la puntuación de audiencia del género más alta se base en una gran cantidad de puntuaciones altas y no una cantidad muy pequeña de puntuaciones muy altas, lo que haría que los resultados no fueran confiables. Agregue un recuento de puntuaciones en el gráfico arrastrando el campo **Audience score%** al estante **Row**. Esto hará que se abra un segundo gráfico que muestra la cantidad de calificaciones por género.
6. Utilice la flecha de la lista desplegable en la nueva entrada **Audience score%** para cambiar la medida de SUM a COUNT seleccionando **Medida (Suma)> Recuento**.

Preguntas:

¿Qué género tiene la mayor cantidad de puntuaciones?

Comedia

¿Cómo afecta el análisis el conteo de puntuaciones para el género con la mayor puntuación de audiencia?

El género de fantasía tiene muy pocas puntuaciones de audiencia y todas son altas, por lo que es probable que el análisis no sea preciso. Si se tratara de un análisis de cliente real, sería necesario obtener un conjunto de datos con un número similar y alto de calificaciones para cada género.

Pregunta: ¿Qué estudio de películas produjo más ingresos en un año específico?

Para responder la segunda pregunta, deberá iniciar una nueva hoja de trabajo para este análisis.

7. Abra una segunda hoja con el icono + más a la izquierda en la parte inferior de la pantalla.

8. Comience arrastrando **Lead Studio** al estante Columns. Luego, agregue **Año** al estante Filas. Una cuadrícula que muestra qué estudios tenían datos en qué años se abren.
9. Busque las barras asociadas con el año 2010 en la cuadrícula; hay 11. Selecciónelas dibujando un cuadro sobre todas ellas con el mouse o seleccionándolas individualmente. Haga clic con el botón derecho o pase el cursor sobre las barras seleccionadas y elija **Mantener solo**. Las otras partes de la cuadrícula desaparecen.
10. Ahora arrastre **Bruto mundial** al estante Filas y elimine Año (haga clic con el botón derecho y seleccione Eliminar o haga clic en la burbuja y presione la tecla Eliminar). Debería ver un gráfico de columnas. En este contexto, “Bruto” significa la cantidad total de dinero que el estudio ganó ese año con las películas del conjunto de datos.

Nota: Asegúrese de que la ventana sea lo suficientemente amplia para ver todos los estudios del conjunto de datos.

¿Qué estudio obtuvo más dinero en las películas del conjunto de datos durante 2010?

| Warner Bros.

11. Ahora ajustemos las medidas para mostrar el promedio bruto por imagen: en el estante Filas, cambie la función **SUMA (bruto mundial)** por una función de AVG mediante la flecha desplegable para seleccionar **Medida (suma)> Promedio**.

Ahora, ¿qué estudio obtuvo el mayor ingreso bruto promedio por imagen?

| Disney

12. Puede cambiar el nombre de las hojas haciendo doble clic en el nombre de la hoja ("Hoja 3", por ejemplo) e ingresando nuevos nombres.

Pregunta de reflexión

Piense en cómo las diferentes vistas del mismo conjunto de datos brindan diferentes conocimientos y cuentan diferentes historias sobre los datos. ¿Por qué es importante tener datos representativos y relevantes al comenzar el análisis?

| Tener datos representativos y relevantes desde el inicio del análisis es crucial porque garantiza que las conclusiones reflejen la realidad del fenómeno estudiado. Diferentes vistas pueden destacar patrones, anomalías o relaciones ocultas, pero si los datos están sesgados o incompletos, esas historias serán engañosas. Datos bien seleccionados permiten construir visualizaciones precisas, tomar decisiones informadas y evitar interpretaciones erróneas. Además, aseguran que el análisis sea reproducible y útil en contextos reales.