**Evaluación Sumativa 1**

**Modulo 1: Modelado de Datos Avanzados**

**Objetivo de la evaluación**

El objetivo de esta evaluación es aplicar los conceptos avanzados de modelado de datos en Power BI Desktop. Los participantes deberán buscar un dataset relevante que contenga varias pestañas o varios datasets que compartan variables comunes. Utilizarán Power BI Desktop para cargar, limpiar y transformar los datos, configurar relaciones entre tablas, y preparar el modelo para un análisis más detallado.

**Instrucciones**

1. **Búsqueda y selección de dataset (20 puntos)**
   * **Busqueda de Dataset:** Encuentra un dataset con varias pestañas o varios datasets relacionados que contengan información del mismo negocio. Ejemplos incluyen datos de ventas, productos, clientes y fechas.
   * **Relevancia:** Asegúrate de que los datasets compartan al menos una variable común que pueda utilizarse para establecer relaciones entre ellos (por ejemplo, ProductID, CustomerID).
2. **Carga de datos en Power BI Desktop (20 puntos)**
   * **Importación de Datos:** Abre Power BI Desktop e importa el dataset o los datasets seleccionados.
   * **Verificación:** Confirma que todas las pestañas o archivos se han importado correctamente y que los datos son visibles en el Editor Power Query.
3. **Transformación y limpieza de datos (30 puntos)**
   * **Visualización en Power Query:** Explora los datos en el Editor Power Query y realiza una revisión inicial.
   * **Eliminación de filas duplicadas:** Identifica y elimina filas duplicadas en las tablas según sea necesario.
   * **Cambio de tipos de datos:** Asegúrate de que los tipos de datos en cada columna sean correctos (fechas, números, texto).
   * **Relleno de valores faltantes:** Gestiona los valores faltantes, ya sea reemplazándolos o eliminándolos, según el contexto de los datos.
   * **Creación de columnas calculadas:** Si es necesario, crea nuevas columnas calculadas para enriquecer el dataset (por ejemplo, margen de beneficio, categorías de ventas).
4. **Configuración de relaciones (20 puntos)**
   * **Establecimiento de relaciones:** Configura relaciones entre las tablas utilizando las variables comunes encontradas en el paso 1 (por ejemplo, relaciona la tabla de ventas con la tabla de productos mediante ProductID).
   * **Cardinalidad y direccionalidad:** Revisa y ajusta la cardinalidad de las relaciones (uno a uno, uno a muchos) y verifica la direccionalidad para asegurar la correcta integración de los datos.
5. **Preparación y validación del modelo (10 puntos)**
   * **Revisión de relaciones:** Asegúrate de que todas las relaciones entre las tablas estén configuradas correctamente y que el modelo de datos esté listo para el análisis.
   * **Visualización básica:** Crea una visualización simple para comprobar que los datos se integran correctamente y que el modelo funciona como se espera.

**Criterios de evaluación**

1. **Búsqueda y selección de dataset (20 puntos)**
   * Dataset relevante y bien seleccionado con varias pestañas o datasets que compartan variables comunes (10 puntos).
   * Claridad y justificación en la elección del dataset (10 puntos).
2. **Carga de datos en Power BI Desktop (20 puntos)**
   * Correcta importación de datos y verificación de la integridad de las pestañas (10 puntos).
   * Confirmación de la visibilidad de los datos en Power Query (10 puntos).
3. **Transformación y limpieza de datos (30 puntos)**
   * Eliminación efectiva de filas duplicadas (6 puntos).
   * Correcto cambio de tipos de datos (6 puntos).
   * Gestión adecuada de valores faltantes (6 puntos).
   * Creación de columnas calculadas relevantes y correctas (6 puntos).
   * Documentación de los pasos de transformación y limpieza realizados (6 puntos).
4. **Configuración de relaciones (20 puntos)**
   * Establecimiento correcto de relaciones entre tablas (10 puntos).
   * Configuración adecuada de cardinalidad y direccionalidad (10 puntos).
5. **Preparación y validación del modelo (10 puntos)**
   * Revisión y validación del modelo de datos (5 puntos).
   * Creación y funcionalidad de una visualización básica para comprobar la integración de datos (5 puntos).

**Conclusión**

Esta evaluación proporciona una oportunidad para aplicar habilidades avanzadas en el modelado de datos en Power BI Desktop. La capacidad de buscar, cargar, transformar y relacionar datos es esencial para construir modelos de datos robustos y realizar análisis profundos. La evaluación está diseñada para fomentar la competencia en la gestión y el análisis de datos, preparando a los participantes para enfrentar desafíos más complejos en el uso de Power BI Desktop.

**Plazo**: [Entrega previo a la siguiente clase]

**Entrega**: Enviar el documento Word documentando sus procedimientos al mail [profesoravivianfierro@gmail.com].

¡Éxito en el desarrollo!

**Desarrollo**

**3- Transformación y limpieza de Datos**

Tabla

Descripción generada automáticamente

Una vez seleccionada la bavse de dato, seleccionamos las pestañas que vamos a trabajar y nos vamos al paso de “Transformar Datos”, que es donde depuraremos la data, categorizaremos nuestras columnas, etc.

**3.1 Limpieza de Data**

**Interfaz de usuario gráfica, Aplicación, Tabla

Descripción generada automáticamente**

En la pestaña “Sales\_Data”, notamos que la columna “Unit Price discount PCT”, no existen datos relevantes, solo sale el valor de “0”, por lo que eliminamos la columna al no aportarnos valor en la lectura de los datos.

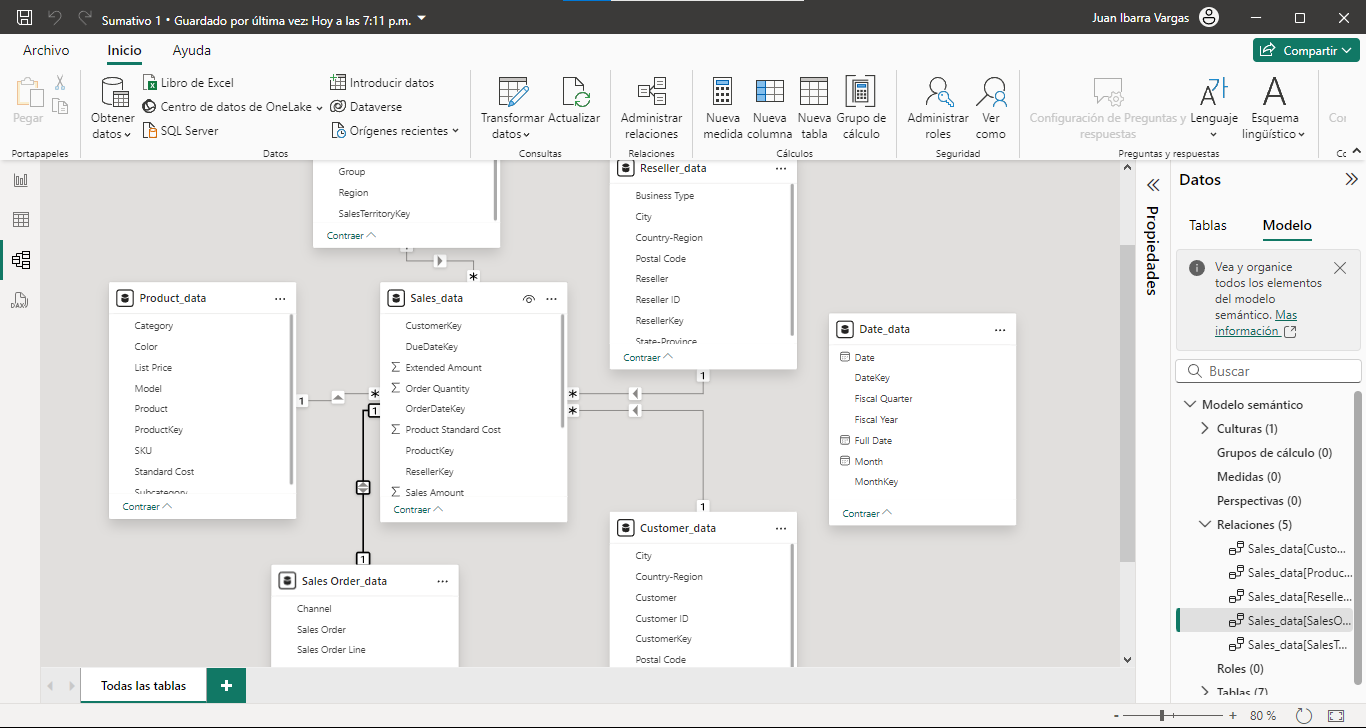
**3.2 Cerrar y aplicar**

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Posteriormente, le damos al botón de “Cerrar y aplicar” para cargar nuestros datos en Power BI.

4 Configuración de relaciones



En este paso, nos dirigimos a la Vista de Modelo, donde analizaremos si las relaciones entre nuestras tablas está correcta. En primer lugar administramos la relación desde la tabla Sales Order\_Data a nuestra tabla de hecho “Sales\_Data”, dejándolo como se muestra en la siguiente imagen:

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

Podemos apreciar entonces, que la relación entre estas dos datas a cambiado teniendo solo dirección a nuestra tabla de hecho en una cardinalidad de “uno es a muchos”.

Interfaz de usuario gráfica, Aplicación

Descripción generada automáticamente

A continuación, creamos una nueva relación entre nuestra data “Date\_data” hacia nuestra tabla de hecho “Sales\_data”. En la imagen vemos como creamos esta nueva relación, vinculando la columna “Datekey” con la columna “OrderDateKey”, luego seleccionamos la cardinalidad (uno es a muchos) y finalmente la dirección la dejamos como única.

Interfaz de usuario gráfica, Diagrama

Descripción generada automáticamente

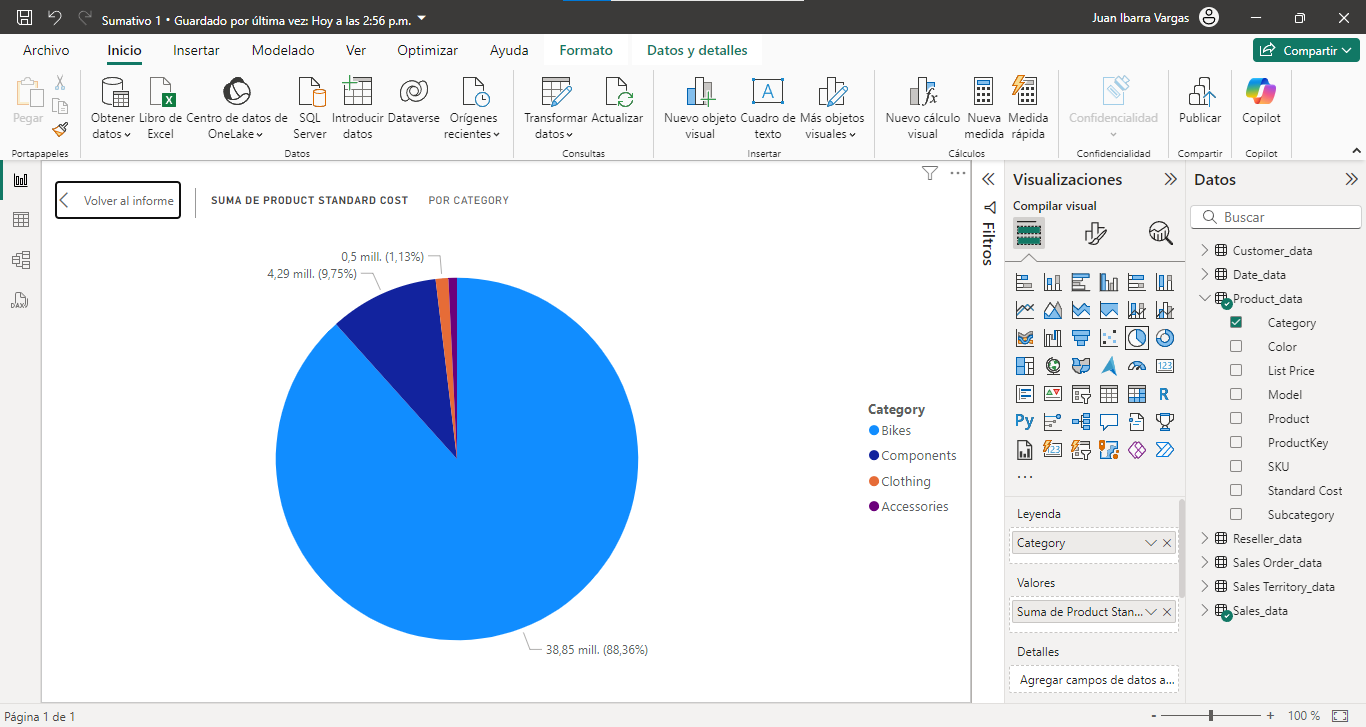
Apreciamos entonces como queda creada esta nueva relación. Las tablas nos quedan con una configuración de estrella.

**5 Preparación y validación del modelo**

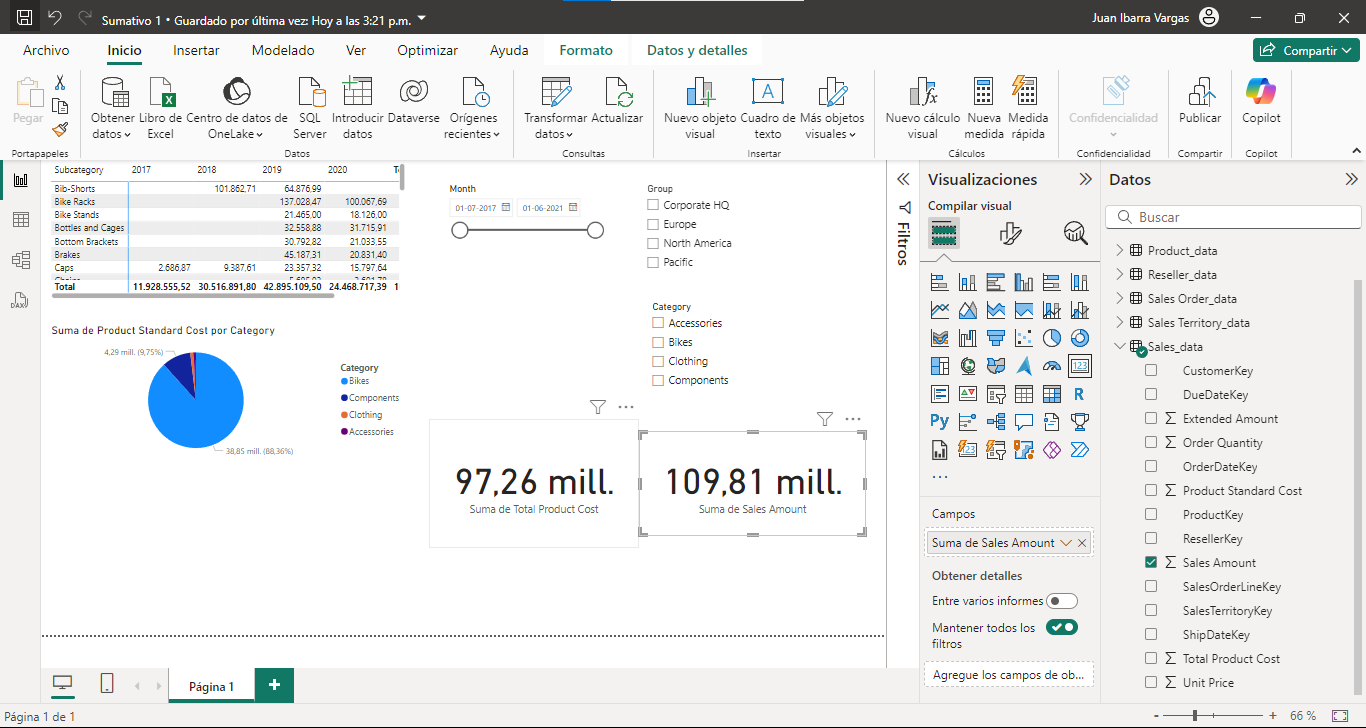
**Interfaz de usuario gráfica, Tabla, Excel

Descripción generada automáticamente**

Para comprobar y validar el modelosinsertamos una matriz, en la cual podemos analizar el comportamiento en las ventas de cada subcategoría a lo largo de los años.



En el siguiente gráfico circular, podemos ver los costos implicados en cada categoría.



Agregamos además, tres segmentadores; uno para segmentar por fecha, otro para segmentar por categoría y otro para segmentar por Grupo geográfico. A su vez, agregamos dos tarjetas, una con los cotos totales y otra donde se pueden ver las ventas.