

### **Examen Parcial # 1**

#### **Serie 1**

Desarrolle con según su criterio el caso definido, debe justificar cada una de sus respuestas.

1. La empresa “Los Mejores” desea contratar sus servicios para definir el alcance de su proyecto de inteligencia de negocios (Business Intelligence). En la actualidad la empresa cuenta con una base de datos única llamada [“NorthWind”](#) la cual debe analizar e iniciar el planteamiento de su proyecto de datos haciendo uso del ciclo de vida CRISP-DM. Plantee el escenario global del proyecto definiendo cada una de las fases.

1. Comprender el negocio
2. Comprender los datos
3. Preparar los datos
4. Modelado
5. Evaluar
6. Implementar

#### **Comprender el negocio**

##### Objetivos de negocio

Vender la mayor cantidad de productos de calidad.

Generar ganancias.

##### Evaluación de la situación

- Riesgos:  
Seguir con gastos innecesarios que impidan el desarrollo de la compañía.
- Recursos:  
Analista, Contacto con conocimiento del negocio, Asesor
- Contingencias:  
Se está analizando data desactualizada (1996-1998)
- Proyecto costo-efectivo:
  - Costos  
Software de análisis  
Analista
  - Beneficios  
Conocimiento del negocio  
Aumentar ganancias al conocer productos más vendidos

### Objetivo del análisis

Fundamentar la mejor decisión sobre las estrategias que utilizará la compañía en relación a sus productos y tipo de clientes.

### Plan de proyecto

#### *Comprender el negocio*

Determinar los requisitos del proyecto y los objetivos del negocio. Entender la necesidad del usuario para resolver el problema. Generar el plan de proyecto. Charter del proyecto, involucrados, objetivos, etc.

#### *Comprender los datos*

Recopilar datos iniciales. Tomar datos necesarios y trasladarlos a un repositorio que nos apoye. Describir los datos. Examinar los datos y definir sus propiedades, formato, fuentes, etc.

#### *Preparar los datos*

Se estructura la fuente de datos de la cual se forma el modelo.

#### *Modelado.*

Diseño de prueba. Modelo de compilación. En este caso, la generación de reporte en Power BI.

#### *Evaluación*

Proceso de revisión. Repasar trabajo realizado. Determinar pasos siguientes.

#### *Implementación de solución*

Darle continuidad al proyecto.

### Herramientas y tecnologías

Microsoft SQL Server

Power Bi

## Comprender los datos

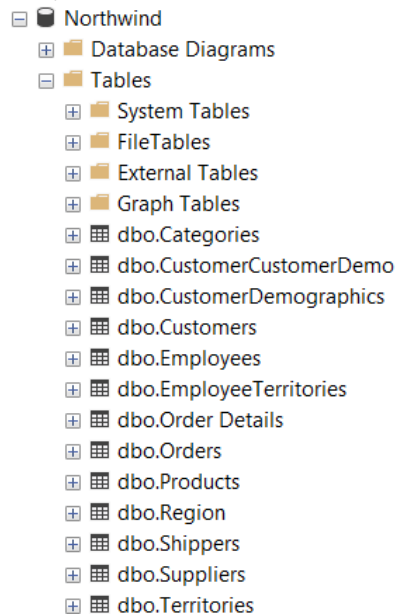
### Recopilar datos iniciales

Descarga de base de datos Northwind.

<https://github.com/Microsoft/sql-server-samples/tree/master/samples/databases/northwind-pubs>

### Describir datos

La base de datos cuenta con 13 tablas. Entre las que se incluye la de ventas, territorios, clientes, empleados, categorías, productos, entre otros.



### Explorar datos

Se debe comprobar que los datos aportan a la hipótesis y permiten llegar a una conclusión basado en el giro de negocio.

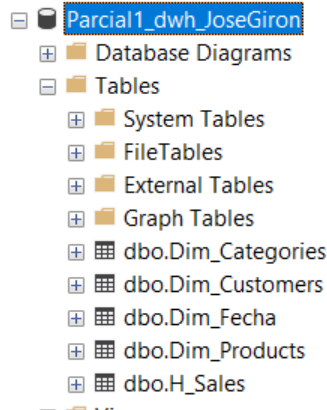
### Verificar calidad de datos

En este apartado se verificará que los datos a analizar no tengan valores nulos.

## Preparar los datos

### Selección de datos

Para el estudio, se seleccionaron 5 tablas, las cuales son las más relevantes para resolver el problema.



### Limpieza de datos

El proceso de limpieza de datos consistirá en ubicar las incoherencias y los datos no útiles para su análisis aplicando las siguientes soluciones para cada caso:

- Se excluirán los datos que tengan un valor nulo en alguno de sus campos.
- Se aplicarán filtros para detectar errores y se intentarán corregir.
- Se examinarán manualmente algunos de los campos no documentados y se comprobará el verdadero significado de estos.

### Construir datos

En este se pueden agregar nuevas columnas que se considere pertinente para el análisis y la toma de decisiones.

Por ejemplo: El total de la venta.

---

```
MeasureTotalSale = H_Sales[UnitPrice] * H_Sales[Quantity] - H_Sales[Discount]
```

---

### Integrar datos

No aplica. Los datos vienen de una sola fuente (Base de datos Northwind).

## Formatear datos

División de fechas en trimestre, mes, año y día.

CalendarDate	Month	Year	Day	Quarter	Weekday
Monday, 1 de January de 1996	1	1996	1	1	2
Tuesday, 2 de January de 1996	1	1996	2	1	3
Wednesday, 3 de January de 1996	1	1996	3	1	4
Thursday, 4 de January de 1996	1	1996	4	1	5
Friday, 5 de January de 1996	1	1996	5	1	6
Saturday, 6 de January de 1996	1	1996	6	1	7
Sunday, 7 de January de 1996	1	1996	7	1	1

## **Modelado**

### Técnica de modelado

Se realizó una gráfica de tendencia, una gráfica de pie, y diversos cálculos, tales como total de productos vendidos o total de ganancias en los 2 años.

### Diseño de prueba

No aplica. No existe forma objetiva de probar el modelo, pero se pueden establecer una guía de la efectividad de la gráfica generada, tales como si muestra una tendencia clara o si aporta información nueva.

### Modelo de compilación

Se necesitan realizar varias iteraciones para poder llegar a conclusiones definitivas, se espera que en esta fase compare resultados de los modelos para poder llegar a conclusión.

### Evaluar modelo

En esta fase se debe comprobar con el cliente si las gráficas son correctas.

## **Evaluar**

### Evaluar resultados

Se observa el resultado final del análisis.

### Proceso de revisión

Responder a preguntas:

¿Se pasó algo por alto?

¿Se ejecutaron correctamente todos los pasos?

### Determinar los pasos siguientes

El paso para seguir será refinar los cálculos.

## **Implementar**

### Planificación

Esta fase se debe tener resumidos los resultados y tenerlos en mente para implementar la solución.

### Generar informe

Se deberá detallar un informe final con la siguiente información:

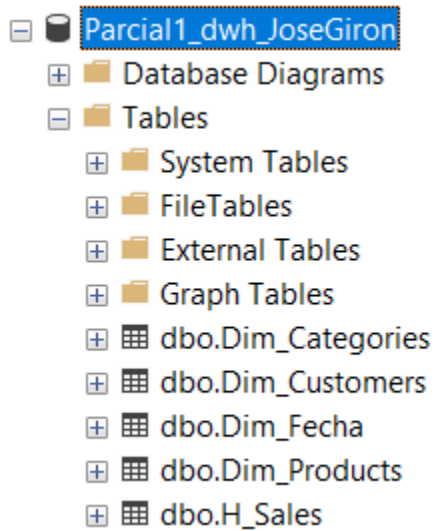
- Descripción detallada del problema
- El procedimiento utilizado, en este caso las gráficas y cálculos realizados
- El costo del proyecto
- Un resumen de los resultados
- Un resumen del plan propuesto
- Recomendaciones

### Revisión del proyecto

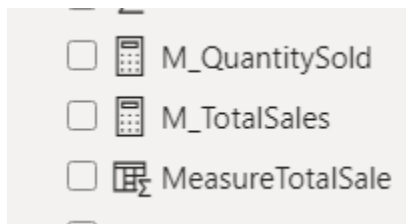
Se debe realizar una revisión general del proyecto con los involucrados.

## Serie 2

1. Su proyecto fue aceptado por lo que teniendo analizado el panorama general de la compañía “Los Mejores”, se le ha solicitado la construcción de un data warehouse empresarial. Para ello, debe establecer cómo mínimo 2 de las principales dimensiones que den sentido a los datos de su tabla de hechos. Dentro de estas 2 dimensiones a elegir no se toma en cuenta la dimensión de tiempo.



2. Elaborado el data warehouse, desarrolle dentro de PowerBI un esquema tabular que le permita definir un esquema semántico para que los usuarios finales puedan realizar consultas de manera sencilla hacia su información. Dicho esquema semántico debe contener al menos 2 campos calculados y 1 columna calculada.



3. Cree un tablero utilizando buenas prácticas de visualización de datos. El mismo deberá contener cómo mínimo 2 cards, 1 gráfico lineal, 1 gráfico de pie y tabla con formato condicional aplicado.

