

## Taller 3: Gestión del Dato

Jesús Ernesto Suárez Triana.

Para el desarrollo del taller, realizamos los siguientes pasos:

### 1. Validación de las BBDD:

Para poder construir las consultas solicitadas, realizamos una conexión previa a la base de datos y luego de explorar las tablas de cada ejercicio armamos las siguientes consultas. Las cuales nos dan los resultados solicitados.

```
/*
 * 1. Cuáles son los 10 productos más vendidos en AdventureWorks utilizando Production.Product y Sales.SalesOrderDetail.
 */

● SELECT TOP 10
  p.ProductID, p.Name, count(sod.SalesOrderID) as veces_vendido
FROM Production.Product p
LEFT JOIN Sales.SalesOrderDetail sod ON (sod.ProductID = p.ProductID)
GROUP BY p.ProductID, p.Name
ORDER BY Count(sod.SalesOrderID) DESC

/*
 * 2. Determinar los vendedores estrella de los que más venden a los que menos venden por año en Adventure Works utilizando Sales.SalesPersonQuotaHistory
 */

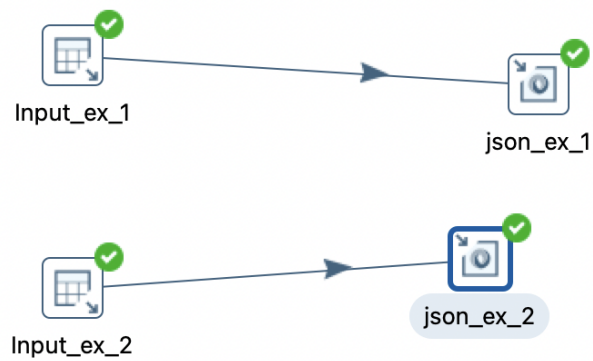
● SELECT p.FirstName, p.LastName,
  YEAR(spqh.QuotaDate) as año,
  SUM(spqh.SalesQuota) as total_vendido
FROM Person.Person p
LEFT JOIN Sales.SalesPersonQuotaHistory spqh on (spqh.BusinessEntityID = p.BusinessEntityID)
where spqh.QuotaDate IS NOT NULL
GROUP BY p.FirstName, p.LastName, YEAR(spqh.QuotaDate)
ORDER BY YEAR(spqh.QuotaDate),
  SUM(spqh.SalesQuota) DESC
```

**Nota:** Para el punto 1 hacemos el filtro de 10 productos debido a que son los solicitados en el ejercicio, mientras el segundo enviamos el dataset completo. Adicional en el segundo agregamos el filtro de que la unión de año sea no nula para tener todos los registros tal como la imagen del ejercicio reportado.

### 2. Construcción ETL co Pentaho.

Para la construcción de la ETL, realizamos todo el proceso en el mismo archivo, el cual se divide en dos partes.

- Lectura de la información, para ello, usamos los objetos **input\_table**, el cual considera la información construidas en el literal anterior.
- Exportamos la información a un archivo Json. El cual genera los archivos deseados.



```

2022/11/05 19:10:12 - Spoon - Save file as...
2022/11/05 19:10:12 - Spoon - Running transformation using the Kettle execution engine
2022/11/05 19:10:12 - Spoon - Transformation opened.
2022/11/05 19:10:12 - Spoon - Launching transformation [etl_taller_3]...
2022/11/05 19:10:12 - Spoon - Started the transformation execution.
2022/11/05 19:10:13 - etl_taller_3 - Dispatching started for transformation [etl_taller_3]
2022/11/05 19:10:13 - Input_ex_2.0 - Finished reading query, closing connection
2022/11/05 19:10:13 - Input_ex_1.0 - Finished reading query, closing connection
2022/11/05 19:10:13 - Input_ex_2.0 - Finished processing (I=58, O=0, R=0, W=58, U=0, E=0)
2022/11/05 19:10:13 - Input_ex_1.0 - Finished processing (I=10, O=0, R=0, W=10, U=0, E=0)
2022/11/05 19:10:13 - json_ex_1.0 - Finished processing (I=0, O=10, R=10, W=10, U=0, E=0)
2022/11/05 19:10:13 - json_ex_2.0 - Finished processing (I=0, O=58, R=58, W=58, U=0, E=0)
2022/11/05 19:10:13 - Spoon - The transformation has finished!!
  
```

### 3. Revisión de los archivos

Los archivos quedan guardados en dos carpetas por separado, la cual tiene la información solicitada en las consultas

#### archivo 1:

```

ex_1_0.js  x  ex_2_0.js  x
1  {"data":[{"ProductID":870,"veces_vendido":4688,"Name":"Water Bottle - 30 oz."}, {"ProductID":712,"veces_vendido":3382,"Name":"AWC Logo Cap"}, {"ProductID":873,"veces_ve
  
```

## archivo2:

```
1 {"data":[{"FirstName":"Tsvi","total_vendido":2366000.0000,"LastName":"Reiter","año":2011}, {"FirstName":"Jillian","total_vendido":2283000.0000,"LastName":"Carson","año":2011},
```