Tasks

I. Introduction

In this Task you will have the opportunity to get familiar with and manipulate the main components of a Data Warehouse (**Data Mart**). The DW is implemented in Access, an easy-to-use tool for creating **business applications**. The case study used here is a **grocery store application** (sales data). By consulting the related tables created for this example, you will learn about the typical modelling of a DW (**star schema**) and the different way to explore its dimensions.

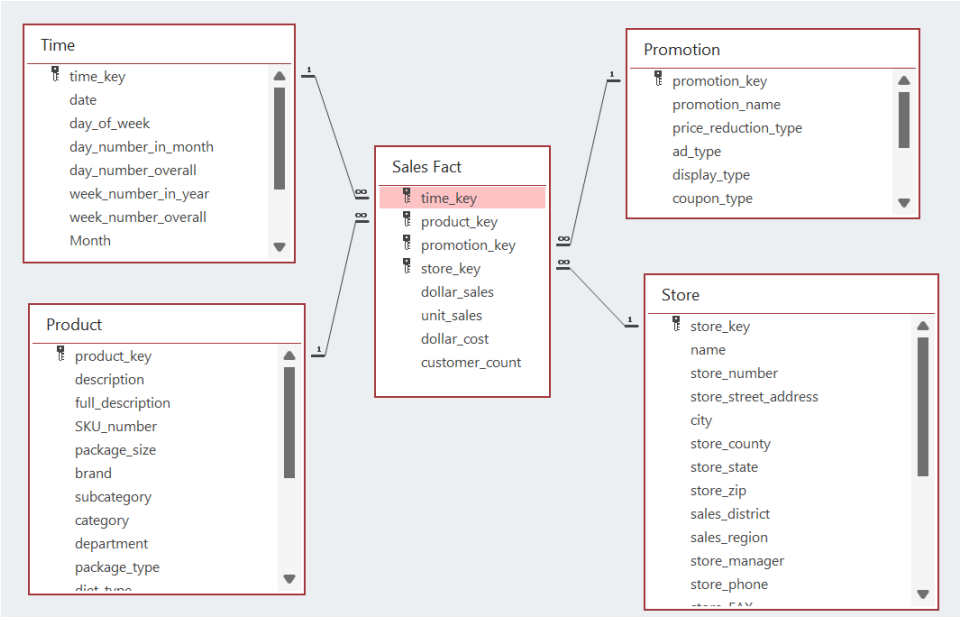
The objective of this tutorial is to understand the flexibility of the Cube modelling to summarize and consult a large quantity of data according to different business perspectives. To do so, you will have to select and display the relevant content of the DW in an Excel pivot table. You will realise that the pivot table is a very useful data processing tool to extract information from a large and detailed dataset in a quick and easy way.

II. Task

II.1. Tutorial

Complete the video tutorial available here: https://www.youtube.com/watch?v=eGhjklYyv6Y

The video is presenting a grocery DW (actually, it is a Data Mart), which is modelled through a **Star schema.**



Description of the figure above:
- **1 Fact Table:** 'Sales Fact' which stores measures (dollar_sales, unit_sales, dollar_cost, customer_count).
- **4 Dimensions:** When (Time table), What (Product table), Where (Store table), and Why (Promotion).

- Link between the tables: The **primary key** for all the four **dimensional tables** has been placed as **foreign key** in the **fact table.**

1) Explore the several tables and try to understand their content and how they are related to each other.

2) Connect an Excel file to the DW and create **Pivot table** as demonstrated on the video.

3) Understand how the several dimensions can be used and manipulated to analyse the business performance of the company according to different views (**roll-up, drill-down, slice and pivot**).

II.2. Exercises

Use the DW and the Pivot table to answer the questions as follows:

- How many products did the store No. 19 sale in 1994?

  Ans:- **Store No. 19 sold 14,016 products (units) in the year 1994.**

	A	B
1	Question 1	
2	year	1994
3	store_number	19
4		
5	**Sum of SumOfunit_sales**	
6	14016	

- Which type of promotion provides the highest amount of unit sales?

  Ans:- Among the promotional campaigns, **POS Grabbers** provides the highest amount of unit sales, with a total of **41,949 units sold**,

3	Question 2	
9	**Row Labels**	**Sum of SumOfunit_sales**
1	POS Grabbers	41949
2	**Grand Total**	**440959**

- Which are the top 10% best stores in terms of dollar sales in 1995?

  Ans:- The **top 10% best-performing stores** in terms of dollar sales are **Store No. 8** and **Store No. 12**, with total dollar sales of **21,563.91** and **22,647.59** respectively for the year 1995

3		
4	Question 3	
5	year	1995
6		
7	**Row Labels**	**Sum of SumOfdollar_sales**
8	8	21563.91
9	12	22647.59
0	**Grand Total**	**44211.5**
1		

- Which one of the stores has the highest amount of customers?

  ANS:- The analysis shows that **Store No. 8** and **Store No. 10** have the **highest and equal number of customers**, with a total customer count of **21,843** each.

Question 4	
Row Labels	**Sum of SumOfcustomer_count**
8	21843
10	21843
Grand Total	**43686**

- Which product is the most lucrative?

  ANS:-Based on the profit analysis using the Pivot Table, **Buffalo Jerky** is the **most lucrative product**, generating a total profit of **11,022.52**.

Question 5	
Row Labels	**dollor profit**
Buffalo Jerky	11022.52
Grand Total	**11022.52**

- What was the most lucrative day, month, and year?

  ANS:- Using the profit-based analysis derived from the Pivot Table, it was observed that the most lucrative year was **1994,** generating a **total profit of 54,282.79.** The most lucrative **month was November**, with a profit **of 37,367.45**, and the most lucrative day **was the 21st**, yielding a **profit of 3,838.97.**

Question 6					
Row Labels	**dollor profit**	**Row Labels**	**dollor profit**	**Row Labels**	**dollor profit**
1994	54282.79	**Nov**	**37367.45**	21	3838.97
1995	42318.26	**Grand Total**	**37367.45**	**Grand Total**	**3838.97**
Grand Total	**96601.05**				

Draw conclusions about the pros and cons regarding the use of a DW to support decision making.

Pros of Using a Data Warehouse for Decision Making

- Integrates data from multiple sources into a **single, consistent repository**
- Provides **historical data**, enabling trend and pattern analysis
- Supports **fast and complex analytical queries**
- Enables **multidimensional analysis** (roll-up, drill-down, slice, pivot)
- Improves **data accuracy and consistency** for reporting
- Does not affect **operational (OLTP) system performance**
- Helps management make **data-driven and strategic decisions**

Cons of Using a Data Warehouse for Decision Making

- **High initial setup cost** (hardware, software, and development)
- Requires **time-consuming ETL processes**
- Data may not be **real-time** due to batch updates
- Needs **skilled technical resources** for maintenance
- High **storage requirements** for large historical datasets
- Changes in business requirements may require **redesign**