Project Guide 02

Decision Support Systems (DSS), 2022-23

Degrees on Computer Systems Engineering and
Medical Informatics Engineering

Introduction

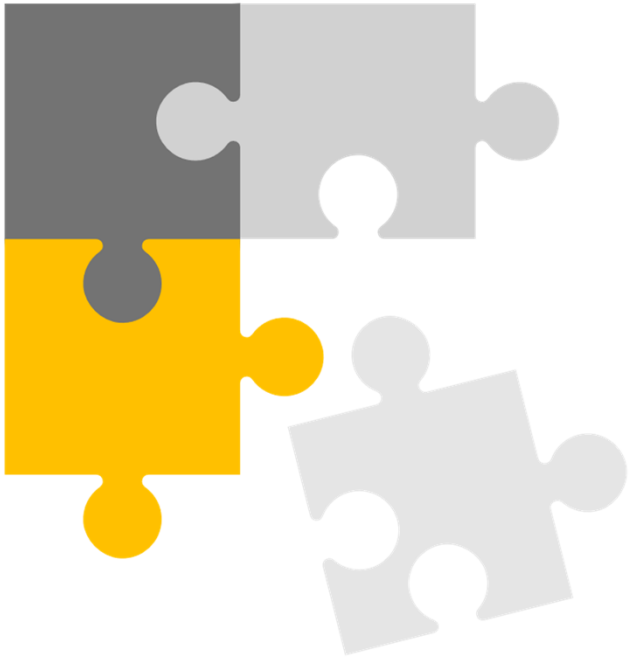
- The goal of project 02 is the development of a data processing and visualization software solution

- It must implement operations of
  - Data preparation (Power Query)
  - Columns and measures calculation
  - Data presentation and visualization

- The solution should integrate several data sources
  - Main source:
    - The data mart created on project 01 or
    - The WWI data mart (PG database server on IPCA)
  - Auxiliary data from other sources

Project team and submission

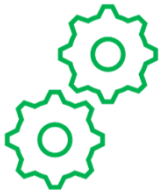
- The project must be carried out in teams of three members
  - The project grade is individual, based on the discussion of the project
- The project must be submitted on the Moodle platform as a **zipped file** that must include
  - (1) the project report and (2) the developed solution file (*.pbix or other)
  - It should be named **DSS_###_P02**, where ## corresponds to the team code. It also applies to (1) and (2).
  - The report must be submitted in PDF format and according to the template available on Moodle
- Projects submitted after the deadline may not be evaluated



Tasks to be developed

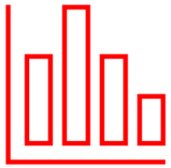
Data Acquisition and Preparation

- Data integration from more than one source
- Data preparation (Power Query)



Modeling and processing

- Definition of relationships and hierarchies
- Creation of columns and calculated measures
- Include the DAX expressions in report
- Etc.



Data Visualization

- Creation of two or more dashboards with one or more visual elements
- Use of relevant elements, e.g. slicers, KPIs or influencers

Patterns finding (optional, for fearless explorers)

- Definition of objectives (association, classification or segmentation)
- Preparation of the data and application of the mining function
- Analysis of the obtained results
- https://www.forbes.com/sites/cognitiveworld/2020/05/10/finding-patterns-and-anomalies-in-your-data/

Data sources

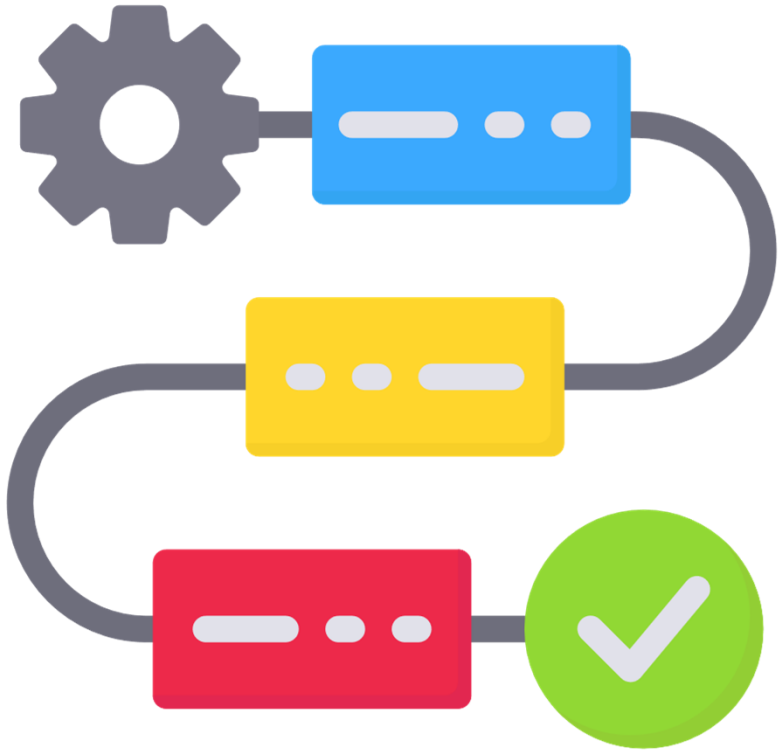
- The solution should combine data from at least two data sources:
  - Mandatory: WWI data mart
  - CSV, JSON, Excel, text files, etc.
  - Html tables or other data gathered directly from Web/cloud

- The Wide World Importers (WWI) sample database
  - WWI fictitious company is a wholesale novelty goods importer and distributor operating from the San Francisco, US
  - WWI's customers are mostly companies who resell to individuals.
  - WWI sells to retail customers across the United States including specialty stores, supermarkets, computing stores, tourist attraction shops, and some individuals.
  - .
- More information
  https://docs.microsoft.com/en-us/sql/samples/wide-world-importers-what-is?view=sql-server-ver15

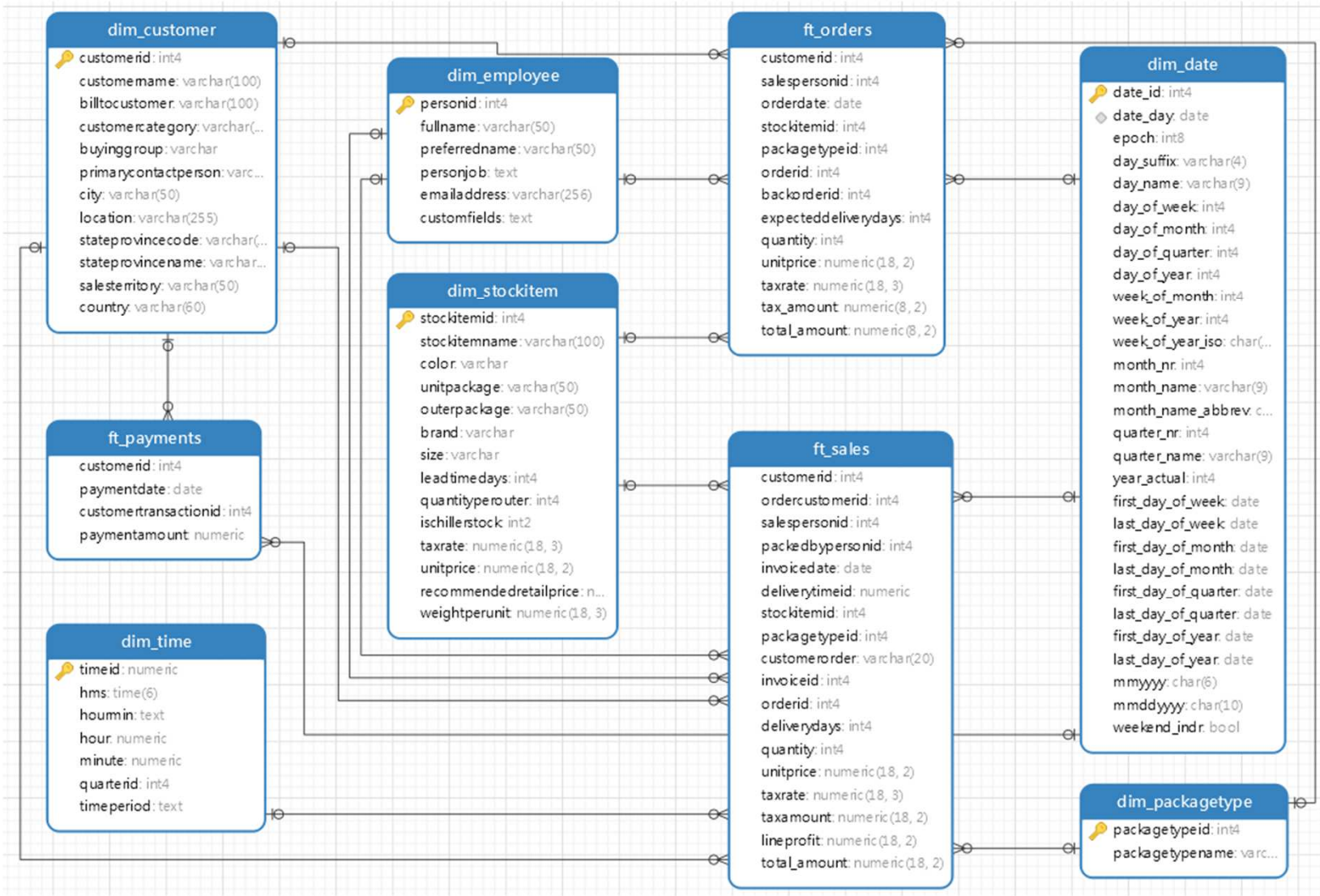
2021.05.12 from
https://www.sisense.com/blog/not-big-data-data-sources/

WWI business workflow

- Customers order items from WWI
- When WWI don't have sufficient stock, they order the additional stock from the suppliers.
  - If customers don't want to wait for items that aren't in stock, the item would then be sent later in a separate shipment.
- WWI invoices customers for the stock items, typically by converting the order to an invoice.
- Customers might order items that aren't in stock. These items are backordered.
- WWI delivers stock items to customers either via their own delivery vans, or via other couriers or freight methods.
- Invoices and the customer payments are recorded as customers transactions



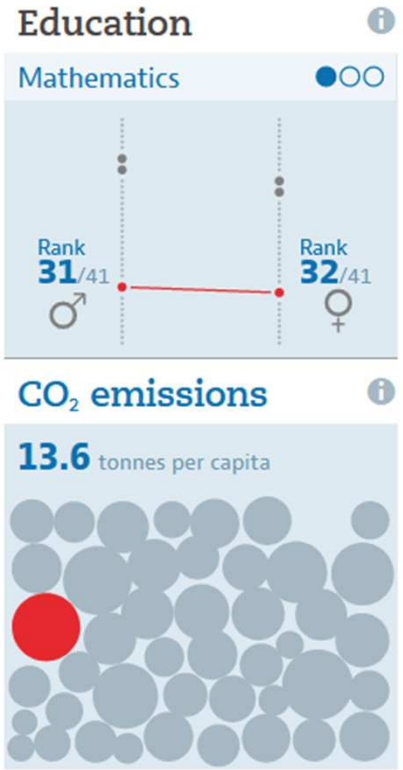
WWI Sales Data mart [adapted]



Examples of external data sources to combine with AW data mart

- World Bank Open Data
  - https://data.worldbank.org/indicator
- OECD Data
  - https://data.oecd.org/united-states.htm

- These are just a few examples.
- The project should try to combine external data with WWI sales data to get some insights.
- Compare your data on sales with some other country indicators, such as, population, GDP per capita, urban population ratio, etc.



Final remarks

The main concepts and techniques covered during the classes on analytical processing and data visualization should be put into practice.

The evaluation of projects will ponder the application of skills, team commitment and the matching of the solution to the project objectives.

Thank you!