|  |
| --- |
|  |
|  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| REVISION HISTORY | | | | | |
| Ver. | Description of Change | Author | Date | Approved | |
| Name | Effective Date |
| 1.0 | Initial status | [Kiryl Bucha](mailto:Kiryl_Bucha@epam.com) | 12-JAN-2012 |  |  |
| 2.0 | Updated in accordance with renewed content | [Elias Nema](mailto:Elias_Nema@epam.com) | 20-JAN-2014 |  |  |

Contents

[1. Data Modelling Task 3](#_Toc444190110)

[1.1. Detail your diagrams for 3NF and Star layers. 3](#_Toc444190111)

[1.2. Create visual and textual description of layers of your data warehouse. 3](#_Toc444190112)

[2. Data Vault 3](#_Toc444190113)

# Data Modelling Task

## Detail your diagrams for 3NF and Star/snowflake layers.

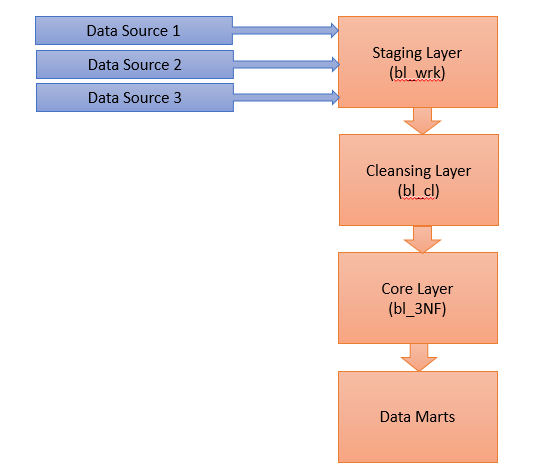
Reminder:

* You need to have at least 5 dimensions.
* **Dim\_Products** contains information about products, their manufacturers, and their stock availability.
* **Dim\_Customers** contains information about customers, their contacts and their locations.
* **Dim\_Sellers** contains information about company’s employees, their contacts and their locations.
* **Dim\_Date** contains date information.
* **Dim\_PaymentMethod** contains information about methods of payment of orders.
* **Dim\_DeliveryMethod** contains information about methods of delivering of orders.
* **Dim\_Stocks** contains information about stocks and their locations.
* You need to use at least 2 type SCD2 attributes in different dimensions.

1. Dim\_Customers
2. Dim\_Employees

## Create visual and textual description of layers of your data warehouse.

* Image (PowerPoint or whatever tool do you prefer) of your data warehouse layers.



**Staging Layer (bl\_wrk).**The main purpose of the Staging Layer is to load source data into the Data Warehouse environment for further processing (the processing from source to staging). In other words, the Staging Layer is responsible for the physical movement of data from the source into the Data Warehouse platform.

**Cleansing Layer (bl\_cl).** The Cleansing Layer is used for data cleansing, filtering wrong data, replace missing values with singletons and performing transformations like code lookups or currency conversions. As the Staging Area, the Cleansing Area contains only data of the last delivery, and data from different sources is not integrated.

**Core Layer(bl\_3NF).** The Core has two purposes: First, it is the integration layer for the data from all source systems of the Data Warehouse. The data in the Core is not stored separately for each source like in the previous layers, but stored in a "subject-oriented" form. The second purpose is to store the history of all data changes in a form that allows to retrieve previous states of the data at any time in the past. This is an important requirement for traceability of data changes, and there are different approaches to store historic data.

**Data Marts.** The Data Marts provide the data in an adequate form that allows easy access for the front-end applications of the BI platform.

# Data Vault

* Get back to the task 5 from and refactor sample business schema into the Data Vault model.