* Please reiterate the purpose of staging area and explain why it is not possible to load data directly into the 3NF layer?

The staging area in a data warehouse serves as an intermediate storage area where data is initially loaded before being processed and transformed for integration into the target data model. Since we get data from various sources the data has to be at first de-duplicated to be able to follow 3NF rules. There are also other steps before loading that the data form the source should be passed, it’s about data cleansing, data integration and data validation.

For de-dublication we need to create MAP\_table in BL\_CL layer which we will populate using SQL scripts like union. But it depends how the data comes from the source system, it the attributes are the same then we need to decide which source data is more reliable. For example, the customer address from billing source is more reliable than the date from other sources.