ETL - Telecom - SSIS

A telecommunications company has contacted you as an ETL Developer/Consultant to carry out the data storage process as follows:

1. The company has a system that saves the CSV file periodically every 5 minutes, this file includes basic data of the various transactions made by customers during a specific period of time.

The company provided you with the following table which shows the data stored in 2.

the CSVfile and

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Column  Name | Data  Type | Length | Is  Nullable | Sample |
| ID | Int |  | False | 156 |
| IMSI | String | 9 | False | 310120265 |
| IMEI | String | 14 | True | 490154203237518 |
| CELL | Int |  | False | 123 |
| LAKE | Int |  | False | 12 |
| EVENT\_TYPE | String | 1 | True | 1 |
| EVENT\_TS | Timestamp |  | False | 16/1/2020 7:45:43 |

1. The processing required to be completed on this data before it is stored in the database

|  |  |  |
| --- | --- | --- |
| Column  Name | Mapping Rules | Target  Model |
| ID | As-is | Transaction\_id |
| IMSI | As-is, **reject the record if null** | IMSI |
| IMSI | Join with **IMSI reference** and get subscriber id, **add -99999 if null** | subscriber\_id |
| IMEI | Substring first 8 chars, **if null or size is less than 15 add -99999** | TAC |
| IMEI | Substring last 6 chars, **if null or size is less than 15 add -99999** | SNR |
| IMEI | As-is | IMEI |
| CELL | As-is, **reject the record if null** | CELL |
| LAKE | As-is, **reject the record if null** | LAKE |
| EVENT\_TYPE | As-is | EVENT\_TYPE |
| EVENT\_TS | Validate the timestamp format to be YYYY-MM-DD HH:MM:SS, **reject the record if null** | EVENT\_TS |

1. Rejected records will be stored in a separate table, with the original CSV file name recorded.
2. After completing the process of storing the data in the database according to the previous conditions, transfer the csv file to another folder