***HCP SALES***

In this project, the requirement of client was to create a seamless and error-free pipeline for transition of all kinds of healthcare data from the **AWS S3 folders** to the **Snowflake ecosystem**. Post this operation, we had to create a framework to implement business logics on a combination of tables, so that the sales volume per week details for an specific sales personnel under a specific sales metric in the master source table data would **be thinned (pivot-down) for a specific date**.

**INPUT**: We have the sales professional's personal details, data\_week column with the date populated, metric\_type column comprising the type of metric (dollars, wholesale\_units), and 106 columns specifying sales volumes of each week upto 106th week prior to week in the data\_week field.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **hcp\_id** | **name** | **prof\_title** | **data\_week** | **metric\_type** | **sv\_001** | **sv\_002** | **sv\_003** | **…** | **sv\_106** |
| 100 | abc | Dr. | 24-05-2023 | dollars | 1 | 0 | 1 |  | 5 |

**OUTPUT**: We must have a final table wherein the wide table (with multiple columns) is thinned (pivoted down) by transposing 106 sales volume data fields to separate records, with details of other columns intact. As expected, the data date for an unpivoted record must be constant for all 106 instances, but the data\_week column in target table must have date populated based on exact week difference from original data\_week.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **hcp\_id** | **name** | **prof\_title** | **data\_date** | **data\_week** | **data\_week** | **metric\_name** |
| 100 | abc | Dr. | 24-05-2023 | 24-05-2023 | 1 | dollars |
| 100 | abc | Dr. | 24-05-2023 | 17-05-2023 | 0 | dollars |
| 100 | abc | Dr. | 24-05-2023 | 10-05-2023 | 1 | dollars |
| … | … | … | … | … | … | … |
| 100 | abc | Dr. | 24-05-2023 | 12-05-2021 | 5 | dollars |