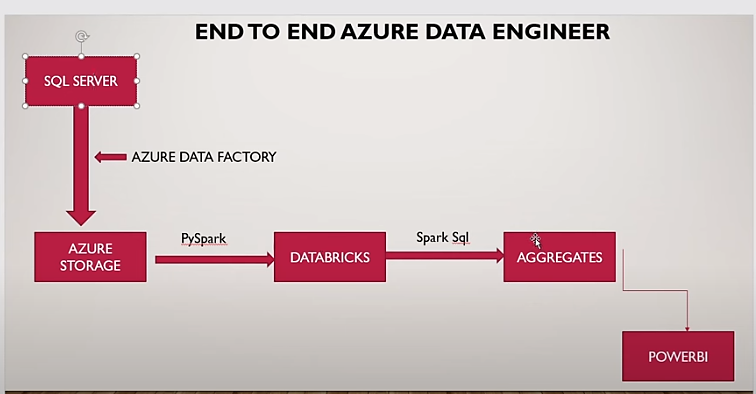
**BUSINESS REQUIRMENT:**

The company requirements are as follows:

* We have data available in SQL server and we need to move into cloud storage.
* We have to build one pipeline which will transfer data and run daily
* Connect the Databricks to storage account.
* BUILD one aggregate table for the reporting with all the Business related KPI
* Build a dashboard using Power BI

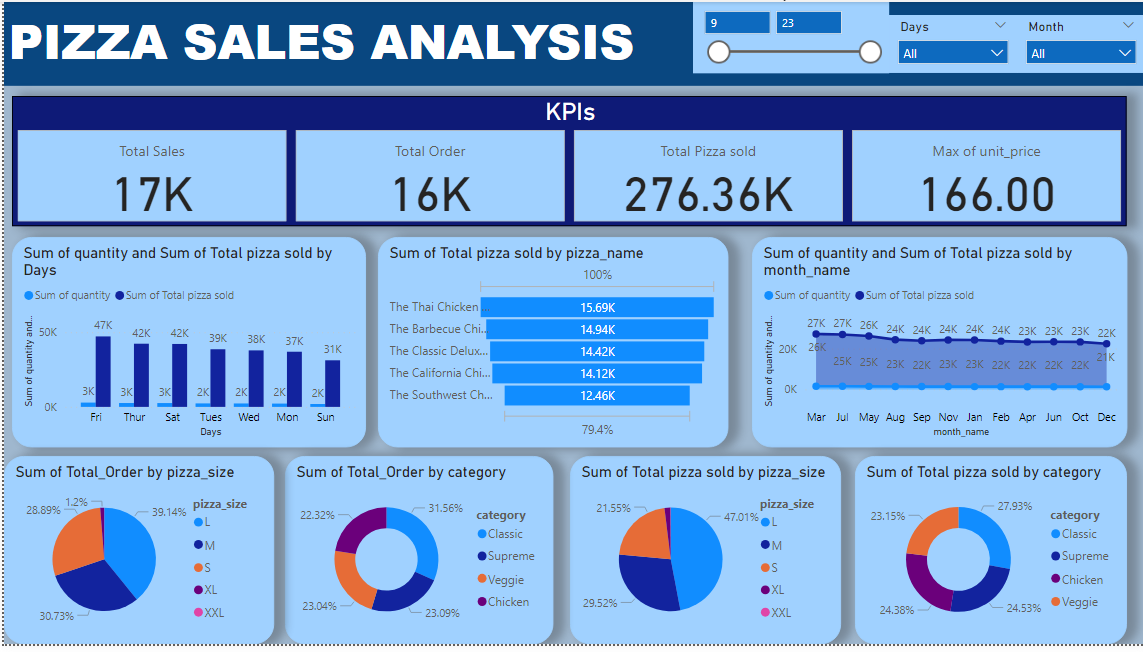
**STEPS TAKEN**

* The dataset was imported into the Microsoft Sql server
* To build a pipeline which will transfer data and run daily, I connected the Microsoft SQL Server to my Azure SQL Database via Azure Data Factory, from where I copied the file to my Azure storage blob. This is to enable my connect it to my Databricks account.
* The dataset was connected to the DataBricks successfully.
* On the Databricks, i carried out series of analysis to uncover the different trends on the pizza sales and to build one aggregate table for the reporting as shown here: https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/1333937627440061/1740020138823503/8336241901072640/latest.html
* Finally, the cleaned dataset was sent to Powerbi for visualization.



***fig 1***

**INSIGHTS GOTTEN FROM THE PIZZA ANALYSIS**



***fig 2: PowerBi dashboard***

From the analysis in fig2

1) **Total pizza sold**: The total pizza sold was 276.36k

2) **Total Orders:**  The total order was 16k

3) **Sales Over Month:** The highest pizza sales were made in March.

4) **Sales over Days:** Sales were made more on Friday.

5) **Sales by pizza category:** Classic made more sales, followed by Supreme, then Chicken and finally, Veggie

**6) Sales by Pizza size:** Size L made the highest sales with a percentage of 47.01%

**7) Top 5 Ordered pizza:** The top 5 ordered Pizzas are Thai Chicken, Barbecue Chicken, Classic Delux, California Chicken and Southwest Chicken.