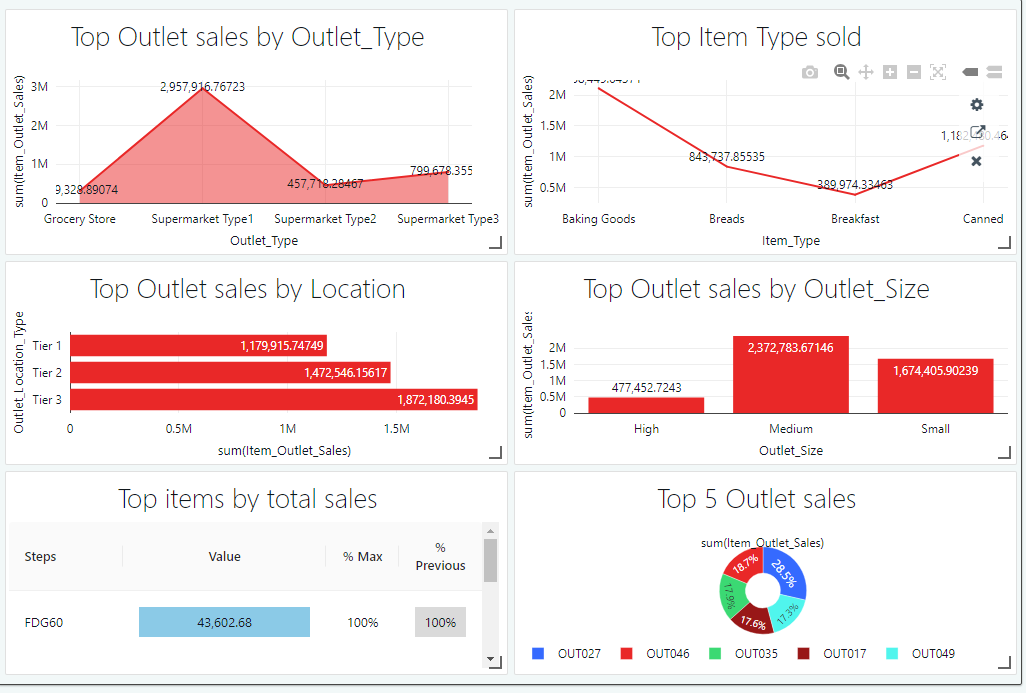
**BUSINESS REQUIRMENT:**

The company requirements are as follows:

* We have sales outlet data available in azure sql database and we need to transform, clean this data and move to ADLS storage
* Business and data scientists want to use this data and they want it in csv format

**STEPS TAKEN**

* The dataset was imported into the Microsoft Sql server which was linked to Azure
* I built a pipeline transfer the data 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 like filling in the missing values and computing the median scores for some missing values in some columns to uncover the different trends on the pizza sales and to build one aggregate table for the reporting as shown here: <https://community.cloud.databricks.com/?o=1333937627440061#notebook/2313895306547024/command/2084274002527348>
* Visuals were made on Databricks: https://community.cloud.databricks.com/?o=1333937627440061#notebook/2313895306547024/dashboard/2084274002527350
* Finally, the cleaned dataset was sent to Powerbi for visualization.

***fig 1***

**INSIGHTS GOTTEN FROM THE OUTLETS ANALYSIS**

From the analysis in fig1

1) **Top 5 Outlet sales**: Among the Top 5 Outlets, Outlet with this identification number, OUT027 mad the most sales

2) **Top Outlet sales by Location:**  By location, Tier 3 made the most sales

3) **Top Item Type sold:** Baking goods sold enough.

4) **Top Outlet sales by Outlet\_Type:** Supermarket Tier 3 made the most sales

5) **Top Outlet sales by Outlet\_Size:** From all three sizes of outlet, medium size made the most sales

**6) Top Items by total sales:** Among the top items sold, FDG60 made the most sales.