***Project DW***

**Describing dataset and its columns:**

The Australian agriculture Composition Database is a reference database that contains data on the nutrient content of Australian agriculture. It is referred to as a reference database because it contains mostly analyzed data. Only a small proportion of data in the database come from other sources such as recipe calculations, food labels, imputing from similar foods or by borrowing from other countries.

Release 2 of the Australian agriculture Composition Database contains nutrient data for 1,616 plants available in Australia and up to 256 nutrients per plant. It is our most recent reference database with data preparation completed in 2021

**Description:** Detailed information about the food, including its appearance, texture, production and preparation.

**Food Name**: Scientific name of the planet or animal food.

**Derivation:** Indicates whether the majority of data for the particular product was analyzed, calculated, borrowed or from labels etc.

**Nitrogen Factor:** Used to calculate protein content. Note that all foods with a derivation of recipe have been allocated a nitrogen factor of 0 by default, as individual nitrogen factors have been applied at the ingredient level.

**Fat Factor:** Used to calculate mass of fatty acids, together with other data. Note that all planets with a derivation of recipe have been allocated a fat factor of 0 by default, as individual fat factors have been applied at the ingredient level.

**Specific Gravity:** Density of a substance relative to that of water (grams per milliliter); provided for liquid foods. Note that all solid foods have been allocated a specific gravity of 0 by default.

**Sampling Details:** Includes information on where the nutrient data were obtained, such as the number of samples purchased for analysis, the date and place of purchase, whether the data were imputed or borrowed etc.

**Analyzed Portion:** Part and/or proportion of the food that is typically consumed, e.g. 80% (flesh and skin) and has been analyzed.

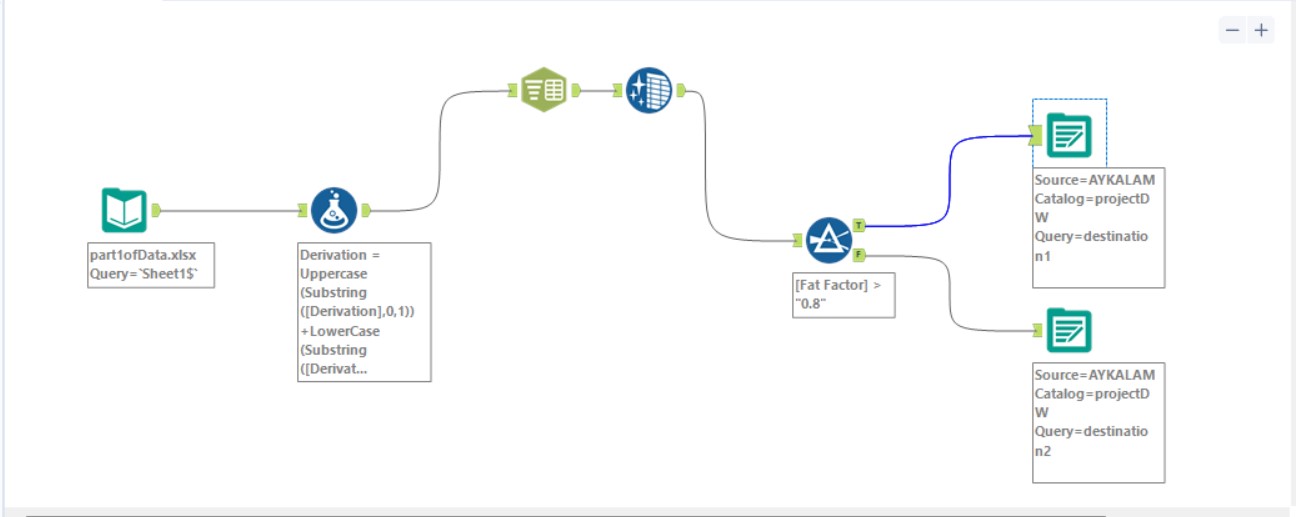
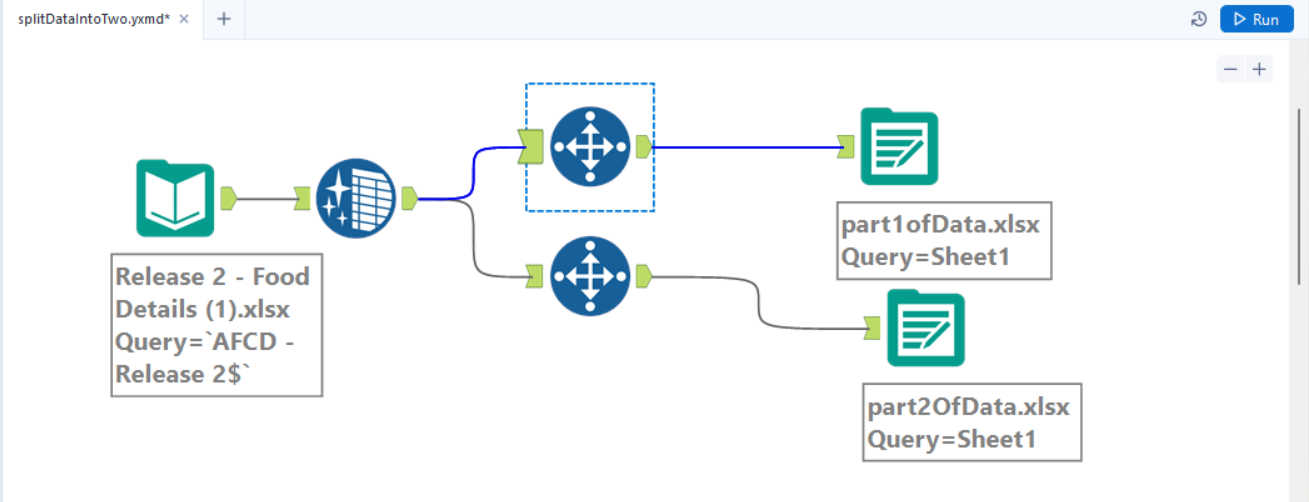
**Unanalyzed Portion:** Part and/or proportion of the food that is typically not consumed, e.g. 20% (core and seeds) and has not been analyzed.

**Components used in the (Alteryx):**

1-Input Data 2-Data Cleansing 3-Select record

4-Formula 5-Text To Columns 6-Filter

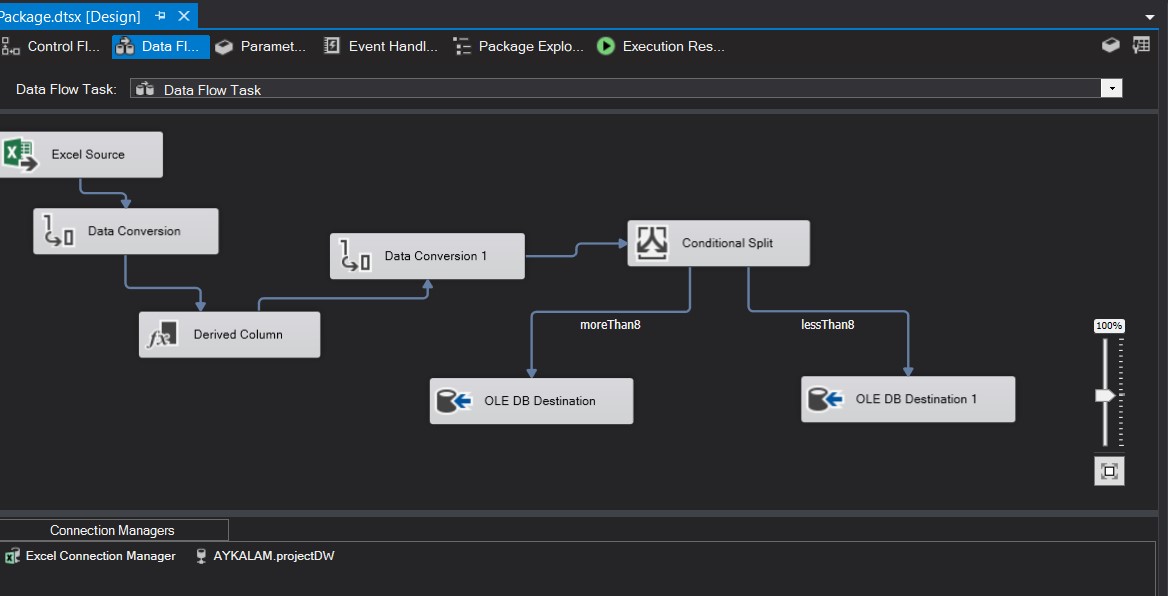
7-Output Data 8-select columns 9-Join (Used for load data to star schema)

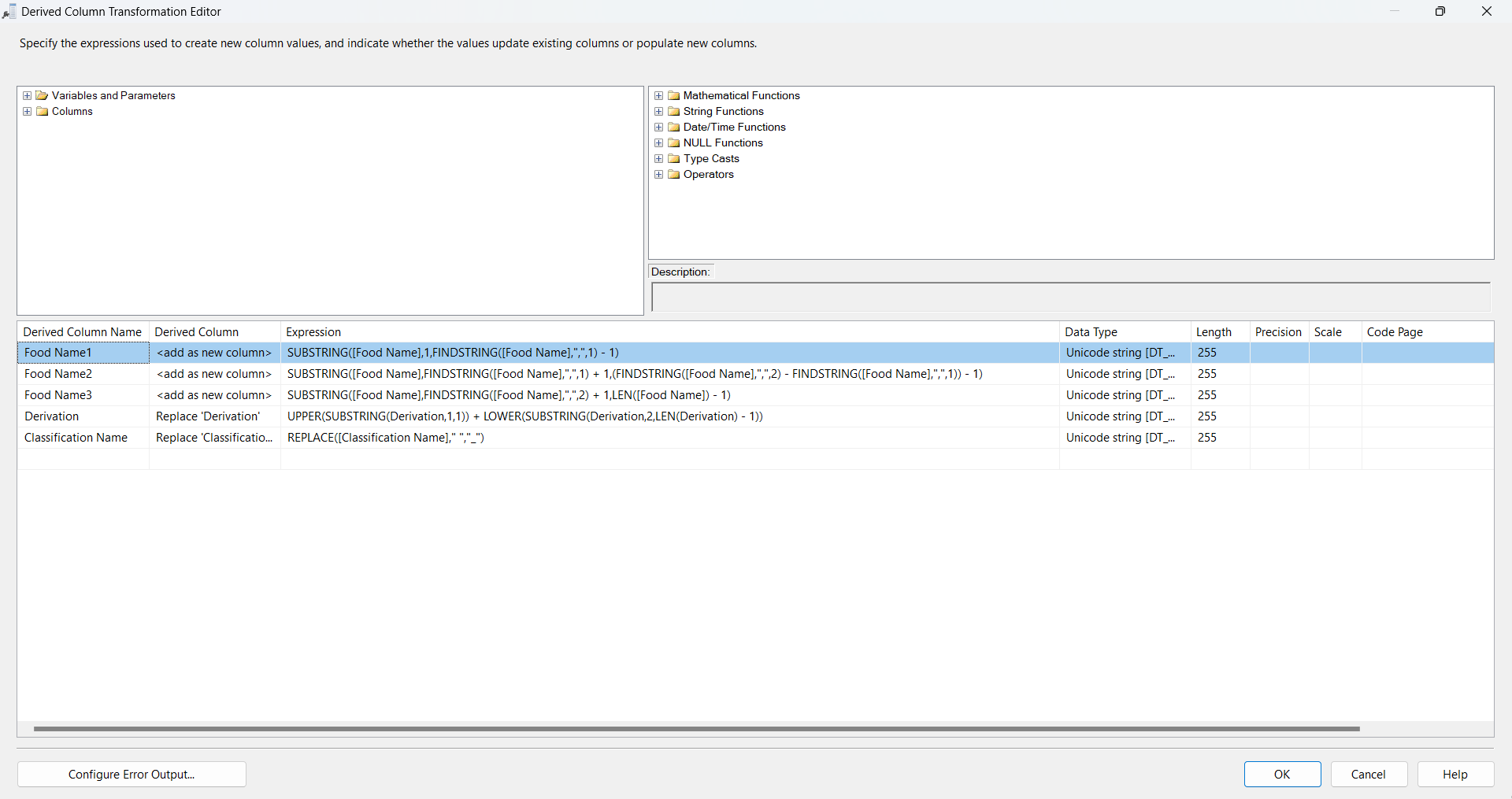
****

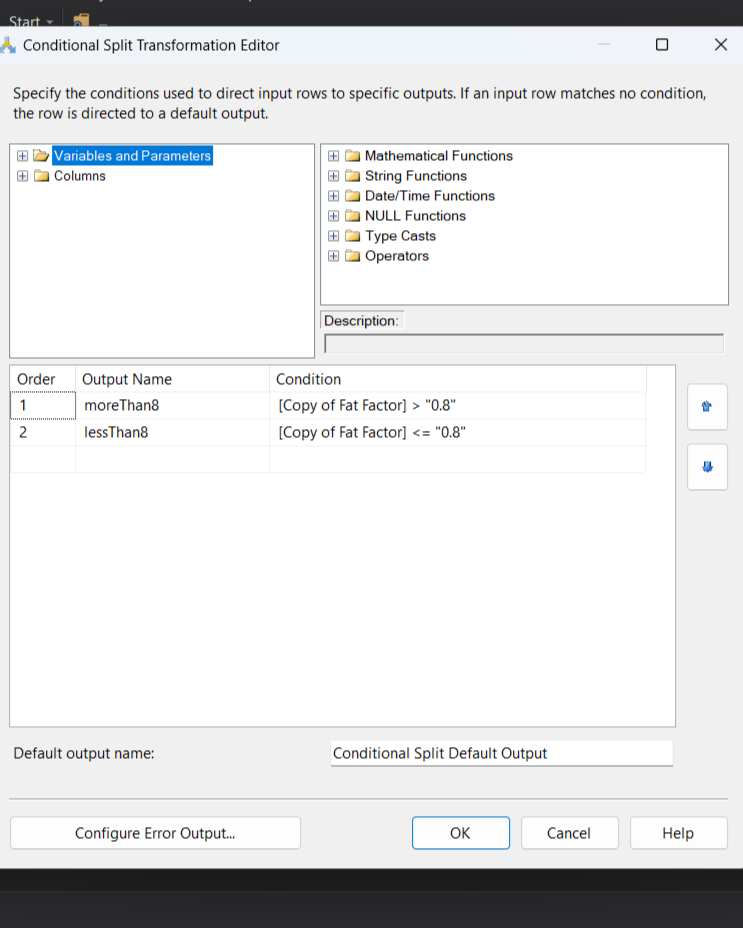
**Components used in the (SSIS):**

1-Excel Source 2-Data Conversion 3-Derived Column

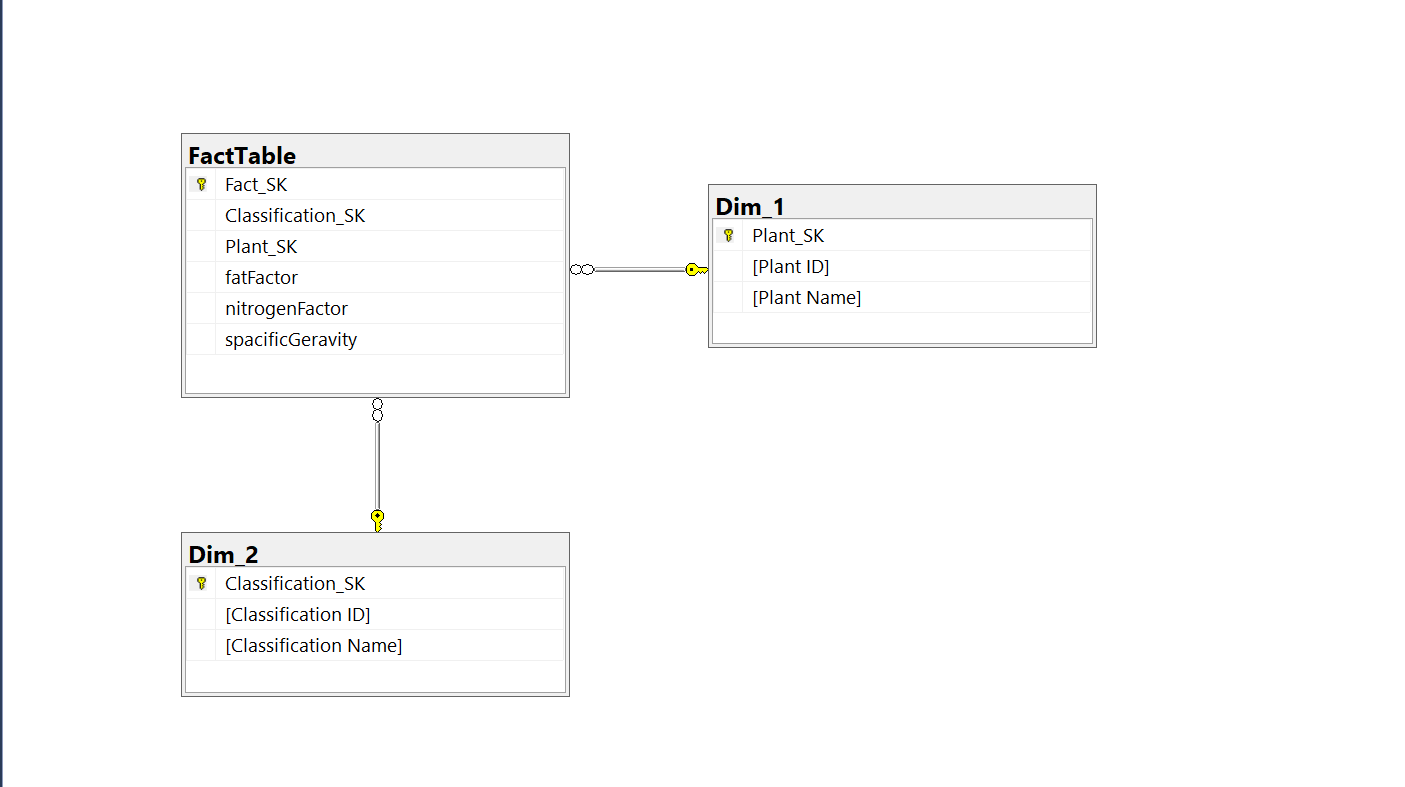
4-Conditional Split 5-OlE DB Destination

****

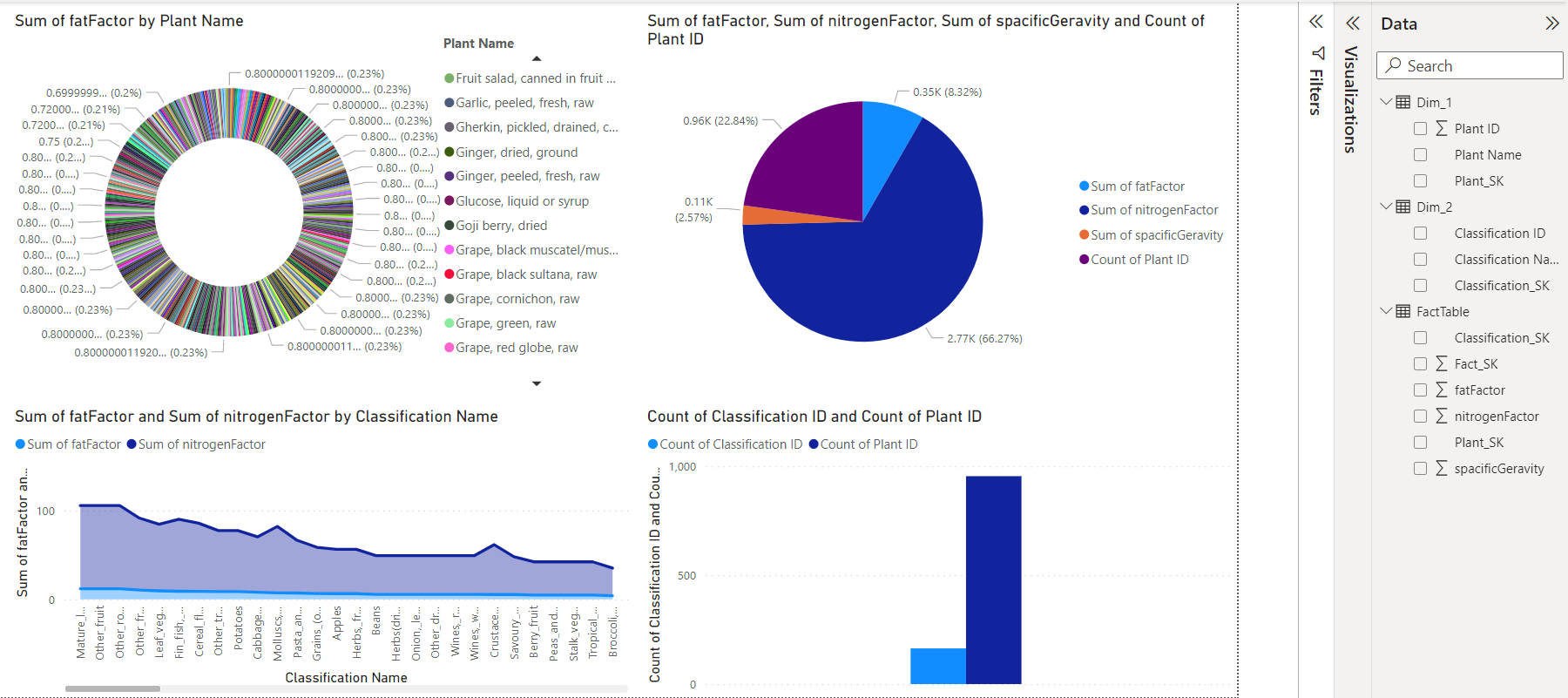
****

****

**Star Schema Design:**

****

**Using Power BI to generate charts from the star schema:**

****

* We found 164 classification for 955 Planet & each of them has numbers for fat & nitrogen factor and specific gravity.
* So try more to classify plants
* Plants are rich in many and varied components, whether it is fat or nitrogen from which the well-known vegetable protein comes, so the addition of plants that contain fat is important and beneficial to health.
* More plants saturated with fats, such as olives and others, must be planted
* The sum of protein\nitrogen is much more compared to the sum of fat factor.