Walk Through steps for Data Tutorials, Real Time SQL Project (Data Analysis) by Yeo Chee En Luke

NOTE:

1. Where to get the data: <https://drive.google.com/drive/folders/1jcM1HxhAunaWUPUkD_6v2rcni0wUJQOm>
2. For this project , I will reference the video from Data Tutorials in terms of the scope of the project , however, the way my thought process or even some SQL queries will be different : <https://www.youtube.com/watch?v=9A9jSvf6KZc&list=PLNr6y7fJuf_f9wCIPQTun4pMosf5e4fFk&index=6>
3. **THE QUERIES (the one I use and the one Data Tutorial use) WILL BE IN A NOTEPAD/SQL FILE !!!!, THIS IS JUST TO SEE THE RESULT AND QUERY BY PICTURE**

KPI Requirements: (info can be found in the powerpoint slides)

1. Total sales: overall revenue from all items
2. Average sales: the total count of different items sold
3. Number of items: the total count of different items sold
4. The average customer rating for items sold

Granular Requirements: (info and more details can be found in the powerpoint slides)

1. Total sales by fat content
2. Total sales by item type
3. Fat content by outlet for total Sales
4. Total Sales by outlet Establish

Chart Requirements:

1. Percentage of Sales by outlet size
2. Sales by outlet location
3. All metrics by outlet type

Scanning of the CSV data file:

* Has a LF instead of Low Fat
* Item weight has some missing numbers
* There seem to have an outlet and item ID
* Outlet type column has a ‘grocery store’, maybe an entry error

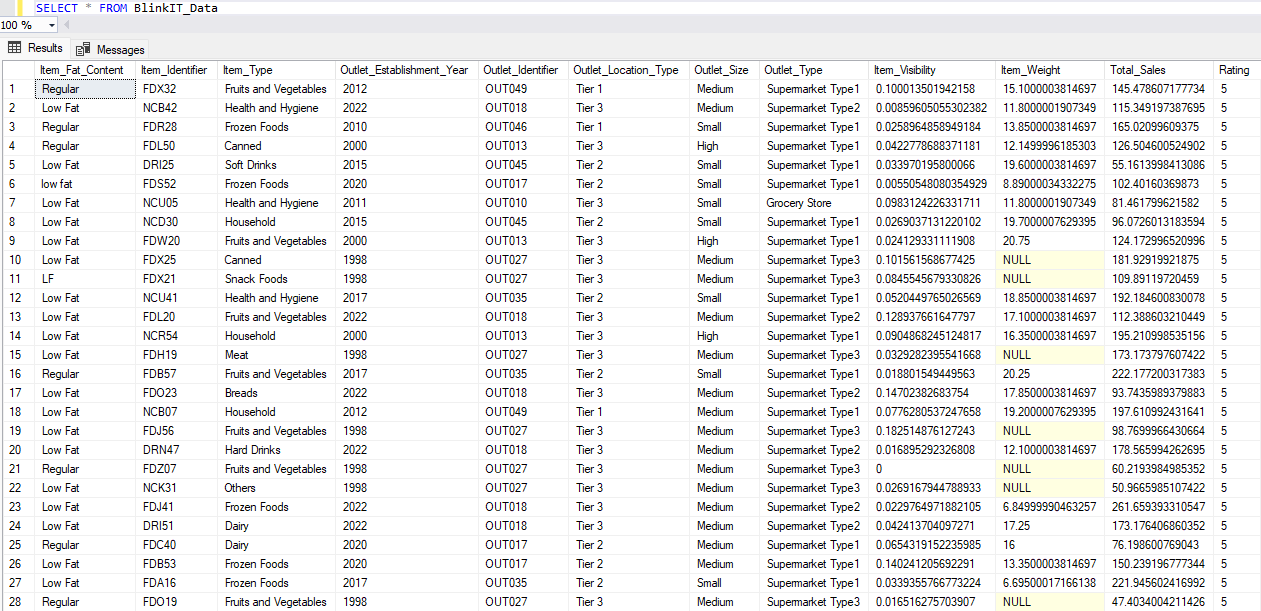
Importing data:

* Open SSMS (Sql management server)
* Connect as per normal
* Create a new database (right click) , name it as ‘BlinkitDB’ , click OK .
* Close the csv file if its still open
* Right click the BlinkitDB -> go to task -> import flat file -> rename the table name to ‘Blinkit\_dataset’ -> go to preview data and modify data to see if any changes to be made if want
* In Modify column , change the data types to what you feel is suitable/correct ( like smallint to int , nvarchar to varchar )

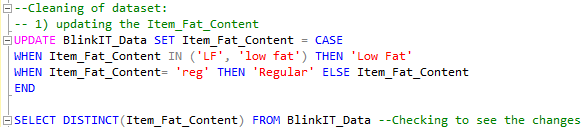
Queries:

* Right click BlinkitDB , then ‘New Query’ -> should give a blank screen
* If want a Specific Query, highlight it then Execute . if you don’t highlight , it will run all

Checking importing correctly:

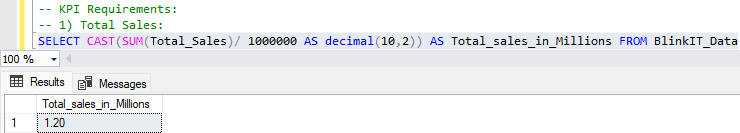


Cleaning dataset using SQL:

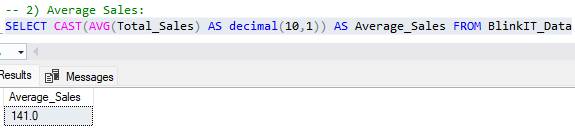


KPI Requirements:

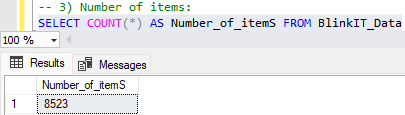
* 1) Total Sales:



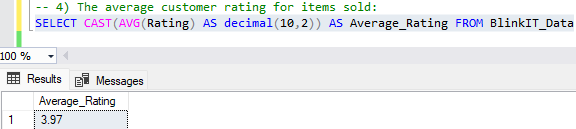
* 2) Average Sales:



* 3) Number of Items: (there is a typo)

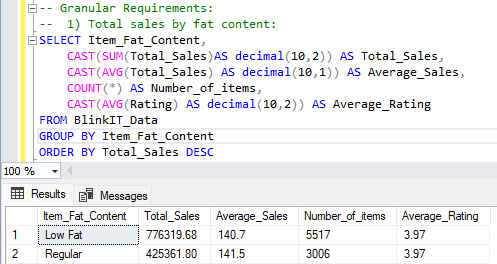


* 4)The average customer rating for items sold:

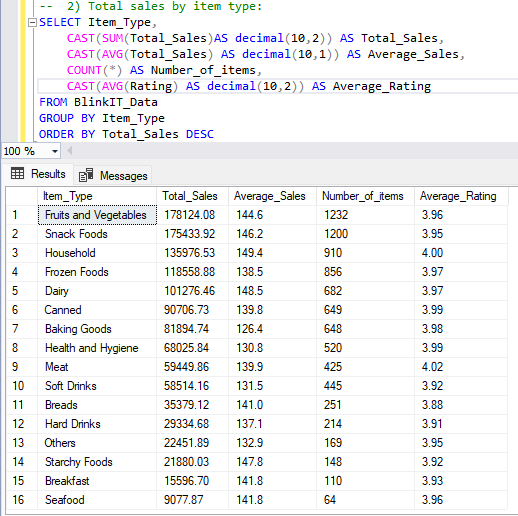


Granular Requirement:

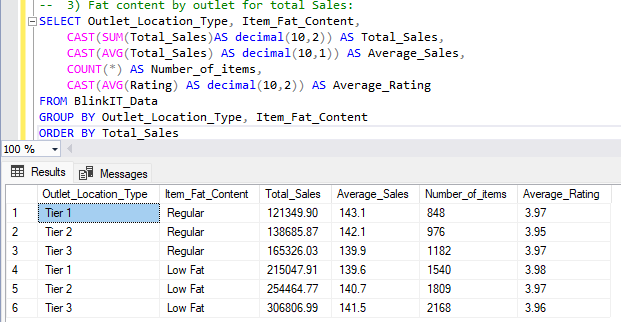
* 1) Total sales by fat content:

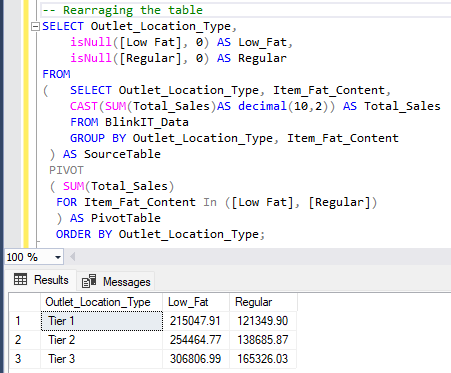


* 2) Total sales by item type:

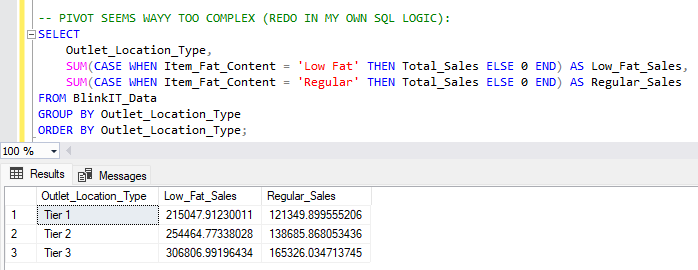


* 3) Fat content by outlet for Total Sales:





**(PIVOT WAS ALITTLE COMPLEX SO I REDID IT):**



* 4) Total Sales by outlet Establish:

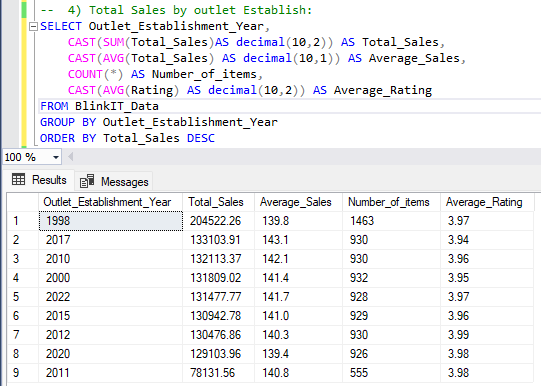
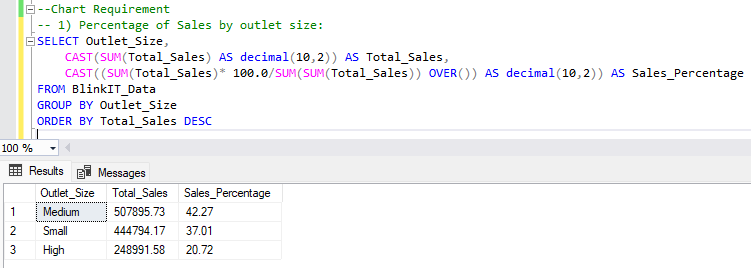
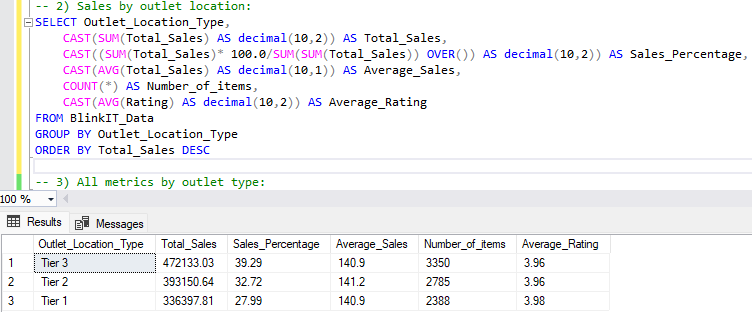


Chart Requirements:

* 1) Percentage of Sales by outlet size:



* 2) Sales by outlet location:



* 3) All metrics by outlet type:

