# BlinkIt Data-Analysis

By – Saifee Raza

## # KPI's Requirements

SELECT * FROM [BlinkIt Dataset]

# first step is Data Cleaning

```
UPDATE blinkit_data
SET Item_Fat_Content =
    CASE
        WHEN Item_Fat_Content IN ('LF', 'low fat') THEN 'Low Fat'
        WHEN Item_Fat_Content = 'reg' THEN 'Regular'
        ELSE Item_Fat_Content
    END;

      After executing this query check the data has been cleaned or not using
below query
```

SELECT DISTINCT Item_Fat_Content FROM [BlinkIt Dataset]

Results Messages

| | Item_Fat_Content |
|---|---|
| 1 | Low Fat |
| 2 | Regular |

1. Sum of total sales in millions

```
SELECT CAST(SUM(Total_Sales)/1000000 AS DECIMAL(10,2)) AS TOTAL_SALES
FROM [BlinkIt_Dataset]
```

Results Messages

| | TOTAL_SALES |
|---|---|
| 1 | 1.20 |

2. Average Sales

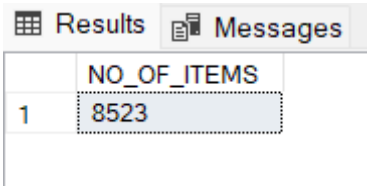
SELECT CAST(AVG(TOTAL_SALES) AS DECIMAL(10,1)) AS AVERAGE_SALES FROM [BlinkIt Dataset]

Results Messages

| | AVERAGE_SALES |
|---|---|
| 1 | 141 |

3. Number Of Items

```
SELECT COUNT(*) AS NO_OF_ITEMS FROM [BlinkIt Dataset]
```

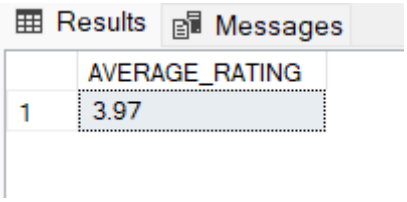


Results Messages

| | NO_OF_ITEMS |
|---|---|
| 1 | 8523 |

4. Average Ratings

```
SELECT CAST(AVG(RATING) AS DECIMAL(10,2)) AS AVG_RATING FROM[BlinkIt Dataset]
```



Results Messages

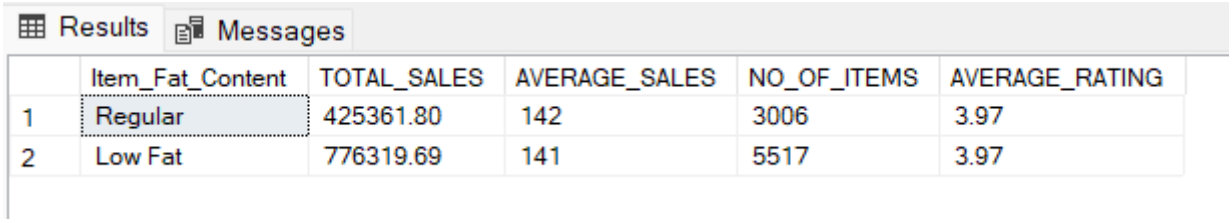
| | AVERAGE_RATING |
|---|---|
| 1 | 3.97 |

/* Granular Requirements */

5. Total Sales by Fat_Content

```
SELECT Item_Fat_Content ,
      CAST(SUM(Total_Sales)AS DECIMAL(10,2)) AS TOTAL_SALES,
       CAST(AVG(Total_Sales) AS DECIMAL(10,0)) AS AVERAGE_SALES,
       COUNT(*)AS NO_OF_ITEMS,
       CAST(AVG(Rating) AS DECIMAL (10,2))AS AVERAGE_RATING

FROM [BlinkIt Dataset]
GROUP BY ITEM_FAT_CONTENT
ORDER BY ITEM_FAT_CONTENT DESC
```



Results Messages

| | Item_Fat_Content | TOTAL_SALES | AVERAGE_SALES | NO_OF_ITEMS | AVERAGE_RATING |
|---|---|---|---|---|---|
| 1 | Regular | 425361.80 | 142 | 3006 | 3.97 |
| 2 | Low Fat | 776319.69 | 141 | 5517 | 3.97 |

/* for establishment year 2020 */

```
SELECT Item_Fat_Content ,
      CAST(SUM(Total_Sales)AS DECIMAL(10,2)) AS TOTAL_SALES,
       CAST(AVG(Total_Sales) AS DECIMAL(10,0)) AS AVERAGE_SALES,
       COUNT(*)AS NO_OF_ITEMS,
       CAST(AVG(Rating) AS DECIMAL (10,2))AS AVERAGE_RATING

FROM [BlinkIt Dataset]
WHERE Outlet_Establishment_Year = 2020
GROUP BY ITEM_FAT_CONTENT
ORDER BY ITEM_FAT_CONTENT DESC
```

Results Messages

| | Item_Fat_Content | TOTAL_SALES | AVERAGE_SALES | NO_OF_ITEMS | AVERAGE_RATING |
|---|---|---|---|---|---|
| 1 | Regular | 46156.91 | 139 | 332 | 4.00 |
| 2 | Low Fat | 82947.05 | 140 | 594 | 3.97 |

/*Total Sales By Item Type */

```
SELECT Item_Type ,
       CAST(SUM(Total_Sales)AS DECIMAL(10,2)) AS TOTAL_SALES,
       CAST(AVG(Total_Sales) AS DECIMAL(10,0)) AS AVERAGE_SALES,
       COUNT(*)AS NO_OF_ITEMS,
       CAST(AVG(Rating) AS DECIMAL (10,2))AS AVERAGE_RATING
FROM [BlinkIt_Dataset]
GROUP BY Item_Type
```

Results Messages

| | Item_Type | TOTAL_SALES | AVERAGE_SALES | NO_OF_ITEMS | AVERAGE_RATING |
|---|---|---|---|---|---|
| 1 | Snack Foods | 175433.92 | 146 | 1200 | 3.95 |
| 2 | Seafood | 9077.87 | 142 | 64 | 3.96 |
| 3 | Breads | 35379.12 | 141 | 251 | 3.88 |
| 4 | Canned | 90706.73 | 140 | 649 | 3.99 |
| 5 | Dairy | 101276.46 | 148 | 682 | 3.97 |
| 6 | Baking Goods | 81894.74 | 126 | 648 | 3.98 |
| 7 | Others | 22451.89 | 133 | 169 | 3.95 |
| 8 | Breakfast | 15596.70 | 142 | 110 | 3.93 |
| 9 | Fruits and Vegetables | 178124.08 | 145 | 1232 | 3.96 |
| 10 | Frozen Foods | 118558.88 | 139 | 856 | 3.97 |
| 11 | Health and Hygiene | 68025.84 | 131 | 520 | 3.99 |
| 12 | Meat | 59449.86 | 140 | 425 | 4.02 |
| 13 | Starchy Foods | 21880.03 | 148 | 148 | 3.92 |

/* Top 5 Item Type */

```
SELECT Top 5 Item_Type ,
       CAST(SUM(Total_Sales)AS DECIMAL(10,2)) AS TOTAL_SALES,
       CAST(AVG(Total_Sales) AS DECIMAL(10,0)) AS AVERAGE_SALES,
       COUNT(*)AS NO_OF_ITEMS,
       CAST(AVG(Rating) AS DECIMAL (10,2))AS AVERAGE_RATING
FROM [BlinkIt Dataset]
GROUP BY Item_Type
```

Results Messages

| | Item_Type | TOTAL_SALES | AVERAGE_SALES | NO_OF_ITEMS | AVERAGE_RATING |
|---|---|---|---|---|---|
| 1 | Snack Foods | 175433.92 | 146 | 1200 | 3.95 |
| 2 | Seafood | 9077.87 | 142 | 64 | 3.96 |
| 3 | Breads | 35379.12 | 141 | 251 | 3.88 |
| 4 | Canned | 90706.73 | 140 | 649 | 3.99 |
| 5 | Dairy | 101276.46 | 148 | 682 | 3.97 |

/* Fat Content By Outlet For Total Sales */

```
SELECT Outlet_Location_Type, Item_Fat_Content,
       CAST(SUM(Total_Sales)AS DECIMAL(10,2)) AS TOTAL_SALES,
       CAST(AVG(Total_Sales) AS DECIMAL(10,0)) AS AVERAGE_SALES,
       COUNT(*)AS NO_OF_ITEMS,
       CAST(AVG(Rating) AS DECIMAL (10,2))AS AVERAGE_RATING
FROM [BlinkIt Dataset]
GROUP BY Outlet_Location_Type, Item_Fat_Content
ORDER BY Outlet_Location_Type, Item_Fat_Content DESC
```

Results Messages

| | Outlet_Location_Type | Item_Fat_Content | TOTAL_SALES | AVERAGE_SALES | NO_OF_ITEMS | AVERAGE_RATING |
|---|---|---|---|---|---|---|
| 1 | Tier 1 | Regular | 121349.90 | 143 | 848 | 3.97 |
| 2 | Tier 1 | Low Fat | 215047.91 | 140 | 1540 | 3.98 |
| 3 | Tier 2 | Regular | 138685.87 | 142 | 976 | 3.95 |
| 4 | Tier 2 | Low Fat | 254464.78 | 141 | 1809 | 3.97 |
| 5 | Tier 3 | Regular | 165326.04 | 140 | 1182 | 3.97 |
| 6 | Tier 3 | Low Fat | 306807.00 | 142 | 2168 | 3.96 |

/* Total Sales by outlet Establishment */

```
SELECT OUTLET_ESTABLISHMENT_YEAR,
       CAST(SUM(TOTAL_SALES) AS DECIMAL(10,2)) AS TOTAL_SALES
FROM [BlinkIt Dataset]
GROUP BY OUTLET_ESTABLISHMENT_YEAR
ORDER BY OUTLET_ESTABLISHMENT_YEAR ASC;
```

Results Messages

| | OUTLET_ESTABLISHMENT_YEAR | TOTAL_SALES |
|---|---|---|
| 1 | 2011 | 78131.57 |
| 2 | 2012 | 130476.86 |
| 3 | 2014 | 131809.02 |
| 4 | 2015 | 130942.78 |
| 5 | 2016 | 132113.37 |
| 6 | 2017 | 133103.91 |
| 7 | 2018 | 204522.26 |
| 8 | 2020 | 129103.96 |
| 9 | 2022 | 131477.78 |

/* Percentage of Sales by Outlet Size */

```
SELECT OUTLET_SIZE,
	CAST(SUM(TOTAL_SALES) AS DECIMAL(10,2)) AS TOTAL_SALES,
	CAST(SUM(TOTAL_SALES) * 100.0 / SUM(SUM(TOTAL_SALES)) OVER() AS DECIMAL(10,2)) AS
SALES_PERCENTAGE
FROM [BlinkIt Dataset]
GROUP BY OUTLET_SIZE
ORDER BY TOTAL_SALES DESC;
```

Results Messages

| | OUTLET_SIZE | TOTAL_SALES | SALES_PERCENTAGE |
|---|---|---|---|
| 1 | Medium | 507895.74 | 42.27 |
| 2 | Small | 444794.17 | 37.01 |
| 3 | High | 248991.59 | 20.72 |

/* Sales by outlet Location */

```
SELECT OUTLET_LOCATION_TYPE,
	CAST(SUM(TOTAL_SALES) AS DECIMAL(10,2)) AS TOTAL_SALES,
	 CAST(SUM(TOTAL_SALES) * 100.0 / SUM(SUM(TOTAL_SALES)) OVER() AS DECIMAL(10,2)) AS
SALES_PERCENTAGE,
CAST(AVG(TOTAL_SALES) AS DECIMAL(10,1)) AS AVG_SALES,
COUNT(*) AS NO_OF_ITEMS,
CAST(AVG(RATING) AS DECIMAL(10,2)) AS AVG_RATING
FROM [BlinkIt Dataset]
GROUP BY OUTLET_LOCATION_TYPE
ORDER BY TOTAL_SALES DESC;
```

Results Messages

| | OUTLET_LOCATION_TYPE | TOTAL_SALES | SALES_PERCENTAGE | AVG_SALES | NO_OF_ITEMS | AVG_RATING |
|---|---|---|---|---|---|---|
| 1 | Tier 3 | 472133.03 | 39.29 | 140.9 | 3350 | 3.96 |
| 2 | Tier 2 | 393150.65 | 32.72 | 141.2 | 2785 | 3.96 |
| 3 | Tier 1 | 336397.81 | 27.99 | 140.9 | 2388 | 3.98 |

/* All metrics by Outlet Type */

```
SELECT OUTLET_TYPE,
	CAST(SUM(TOTAL_SALES) AS DECIMAL(10,2)) AS TOTAL_SALES,
	CAST(SUM(TOTAL_SALES) * 100.0 / SUM(SUM(TOTAL_SALES)) OVER() AS DECIMAL(10,2)) AS
SALES_PERCENTAGE,
	CAST(AVG(TOTAL_SALES) AS DECIMAL(10,1)) AS AVG_SALES,
	COUNT(*) AS NO_OF_ITEMS,
	CAST(AVG(RATING) AS DECIMAL(10,2)) AS AVG_RATING
FROM [BlinkIt Dataset]
GROUP BY OUTLET_TYPE
ORDER BY TOTAL_SALES DESC;
```

Results | Messages

| | OUTLET_TYPE | TOTAL_SALES | SALES_PERCENTAGE | AVG_SALES | NO_OF_ITEMS | AVG_RATING |
|---|---|---|---|---|---|---|
| 1 | Supermarket Type1 | 787549.89 | 65.54 | 141.2 | 5577 | 3.96 |
| 2 | Grocery Store | 151939.15 | 12.64 | 140.3 | 1083 | 3.99 |
| 3 | Supermarket Type2 | 131477.78 | 10.94 | 141.7 | 928 | 3.97 |
| 4 | Supermarket Type3 | 130714.67 | 10.88 | 139.8 | 935 | 3.95 |

If you want more information you can get from this BlinkIt dataset using SQL Query.