

WAREHOUSE AND RETAIL SALES Exploratory Data Analysis by Angelica Adjei-Kwarteng

Date: 20th October 2024

INSPECTION

- Column heads had spaces which could pose a challenge when querying so were renamed.
- Worked with total entries of 307585.
- Previewed unique values:
 1. Year: 4 years; 2017, 2018, 2019 and 2020.
 2. Month: 12 months.
 3. Suppliers: 397 unique suppliers identified.
 - 1 empty string unique supplier.
 - 1 supplier without item_type. Name: Republic National Distribution Co.
 4. Item codes: 1 code without item_type.
 - Code: 347939.
 5. Item description: 1 item_desc without item_type.
 - Item_desc: FONTANAFREDDA BAROLO SILVER LABEL 750 ML.
 6. Item type: 9 unique item types.
 - Liquor, Beer, Wine, Non-alcohol, Str_supplies, REF, Dunnage, Kegs, space/blank type.
 - space/blank Type match item code 347939.
 7. Retail sales: many retail sales have no supplier match records.
 8. Retail Transfers: many transfers were made explaining many negatives.
 9. Warehouse sales: records show one item sold without type.

Inspection Summary

Inspecting the data revealed missing values, empty strings and wrong format. Column heads were altered to help with query. A new table called '**cleaned dataset**' was created to keep the original dataset untouched.

Data Cleaning

Some values did not show as missing when we previewed dataset by columns because they contained empty strings.

- Updated columns and set values with Trim (): Preview shows Trim did not work indicating there are true empty strings.
- Set empty strings to NULL.
- Missing data/ nulls count:
 1. Suppliers: 117
 2. Item_type: 1
 3. Retail_sales: 121,645
 4. Retail_transfer: 188,408
 5. Warehouse_sales: 96,943

Dealing with missing data

Setting threshold at 5% of total count to drop nulls. If total nulls < 5%, drop nulls else, impute values.

- Threshold: 15,379.25.
- Dropped missing data for suppliers, item type

- Imputed retail sales, retail transfers and warehouse sales with median values.
- Total rows remaining: 307,467.
- Total unique: 396 suppliers and 8 unique item types.

Questions to answer

Let's identify patterns, gain some insights and answer questions like

1. Which year recorded the highest retail sales and warehouse sales?
2. Which months consistently recorded very high sales?
3. Which item types sell more relative to others?
4. Which supplier is consistent with the supplies and services that boost sales?
5. How many retail transfers and warehouse sales were made in a month?

DATA ANALYSIS

- Total retail sales by year:
 1. 2020 recorded 540,342.87
 2. 2019 recorded 1,493,252.16
 3. 2018 recorded 240,413.28
 4. 2017 recorded 1,002,003.68
- Total number of suppliers by year:
 1. 2020 recorded 294 suppliers
 2. 2019 recorded 324 suppliers
 3. 2018 recorded 301 suppliers
 4. 2017 recorded 328 suppliers.
- Average retail sales per month: Top 3 months are,
 1. December ranked first with an average sale of 12.1
 2. March followed with an average sale of 11.88
 3. May ranked third with 11.33
- Average sales per supplier: Top 3 suppliers whose supplies have consistently boosted retail sales throughout the years are
 1. FIFTH GENERATION INC: retail average sales of 302.33
 2. CROWN IMPORTS: retail average sales of 70.69
 3. YUENGLING BREWERY: retail average sales of 53.43
- Average retail sales by Item types: Ranking the 8 item types sold in order of highest to lowest,
 1. Beer: 18.47
 2. NON-ALCOHOL: 15.5
 3. LIQUOR: 12.66
 4. STR_SUPPLIES: 11.5
 5. KEGS and Dunnage: 9.23
 6. REF: 8.52
 7. WINE: 8.22

- Total retail transfers by year:
 1. 2020 recorded 624,913.81 transfers
 2. 2019 recorded 1,792,664.98 transfers
 3. 2018 recorded 1,243,709.75 transfers
 4. 2017 recorded 308,337.42 transfers.
- Warehouse sales by sales year: So far, 2020 recorded the highest warehouse sales.
 Insights: The high warehouse sales are a direct reflection of the drop in total sales in the year 2020 as warehouse sales are without profits.
 Recommendation: Warehouse sales must be reduced to make more profits.

INSIGHTS AND RECOMMENDATIONS

- Total retail sales by year:
Insights: This means that retail sales have not been consistent in growth over the years. It rises in one year and falls in the next.
Recommendation: Drop in retail sales is significantly very low hence further investigations must be made.
- Average retail sales per month:
Insights: This means that despite the same number of item types supplied throughout the month, retail sales on average have been consistent in December, March and May.
Recommendation: Staff and inventory levels must be consistent as well to match customer demands in these months.
- Average sales per supplier:
Insights: Despite the fact that some suppliers supplied few quantities, their items gave us more sales on average.
Recommendation: These suppliers whose items gave us more revenue should be kept and those with revenue below an average of \$50 should be reconsidered.
- Average retail sales by Item type:
Insights: Some item types sold a few quantities but on average, gave us more sales.
Recommendation: These items should inform our restocking plans and especially the top three items should always be in stock. Those items which gave an average sale below \$10 yet recorded higher quantities should be reconsidered.
- Retail transfers by sales year:
Insights: Retail transfers have a direct effect on retail sales recorded. As they increase, retail sales decrease.
Recommendation: Higher retail transfers indicate that these transfer locations need more of these items so in the subsequent years, it is advisable for items to be stocked at those locations to increase retail sales at both locations and accountability.
- Warehouse sales:
Insights: There is a relationship between warehouse sales and retail sales. They explained why some item types recorded higher counts sold and yet they did not increase retail sales. However, much

cannot be said of 2020 since warehouse sales could have been as a result of the global pandemic. Perhaps, to avoid expiry and spoilage, the team decided to make more warehouse sales.

Recommendation: Since warehouse sales mean no profits at all, it is advisable to reduce warehouse sales. Should the reason be expiry and spoilage, the team need to check for these qualities when purchasing from suppliers.