Warehouse and Retail Sales Mini-Project

By Connor Stanley

Introduction

The primary objective of this project is to analyze Warehouse and Retail Sales data to identify significant patterns and trends. While this project specifically analyzes sales data for the alcohol industry (including alcoholic beverages, mixers, and related accessories), the analytical approach can be adapted to data from various sectors.

The data was initially obtained from Data.gov, but after looking at the data, I further investigated and identified the direct source of the data as the Montgomery County MD website. In their own words, “this dataset contains a list of sales and movement data by item and department appended monthly” [1].

This analysis is crucial for understanding product popularity, enabling the optimization of inventory levels, and informing future purchasing decisions. This analysis covers warehouse and retail sales in Maryland from June 2017 to September 2020, with notable data gaps in 2018, 2019, and 2020.

It is important to note that ‘Sale’ in this data refers to a case, not a dollar amount. For example, 1 sale of a 1.75 L bottle of Tito’s Vodka does not equate to $1, but rather 1 case.

Data Features and Cleaning

All of the data cleaning/analysis was done in Microsoft Excel.

- Year – refers to the year for the product’s sales
- Month – refers to the month for the product’s sales
- Supplier – refers to the supplier that provided the given product
- Item Code – specific numerical code used to identify each product
- Item Description – description of each item that corresponds to each item code
- Item Type – refers to what type of product is being sold
- Retail Sales – refers to how many “cases of product were sold from Department of Liquor Control (DLC) dispensaries” [1]. I’m assuming this refers to shopping at a liquor store.
- Retail Transfers – refers to how many “cases of product transferred to DLC dispensaries” [1].

- Warehouse Sales – refers to how many “cases of product sold to MC licensees” [1]. I’m assuming this refers to when restaurants and bars order in bulk.

We then created a new sheet to contain the cleaned data. This was different from the original data as we cleaned the column names, supplier names, item descriptions, and item types. This was done as each was initially in all caps, and the title case format would allow users to read the data more easily.

We then used the Remove Duplicates function in the Data Tools section in the Data tab to check for any duplicate rows in the data (there were none). We then identified 167 missing values in the Supplier column, one missing value in the Item Type column, and three missing entries in Retail Sales.

- Filtering through the Supplier column to select all rows with missing values, it seems like many of these items are wine or beer credit, coupons, ice, plastic shot glasses, whiskey tasting journals, tumblers, empty kegs, etc.
- Filtering the Item Type column revealed that the missing value corresponds to a Fontanafredda Barolo. After researching, I determined this was a red wine, and could therefore impute “Wine” into the missing value.
- Filtering through the Retail Sales column to select the rows with missing values, all for non-alcohol item types (rms item and coupon). As there are only 3 rows, I’ll simply remove them entirely.

For the missing supplier values, I determined to leave them as is, as there may not be a traditional supplier for items such as beer credit or whiskey tasting journals.

Analysis

Summary statistics for retail sales, retail transfers, and warehouse sales are below:

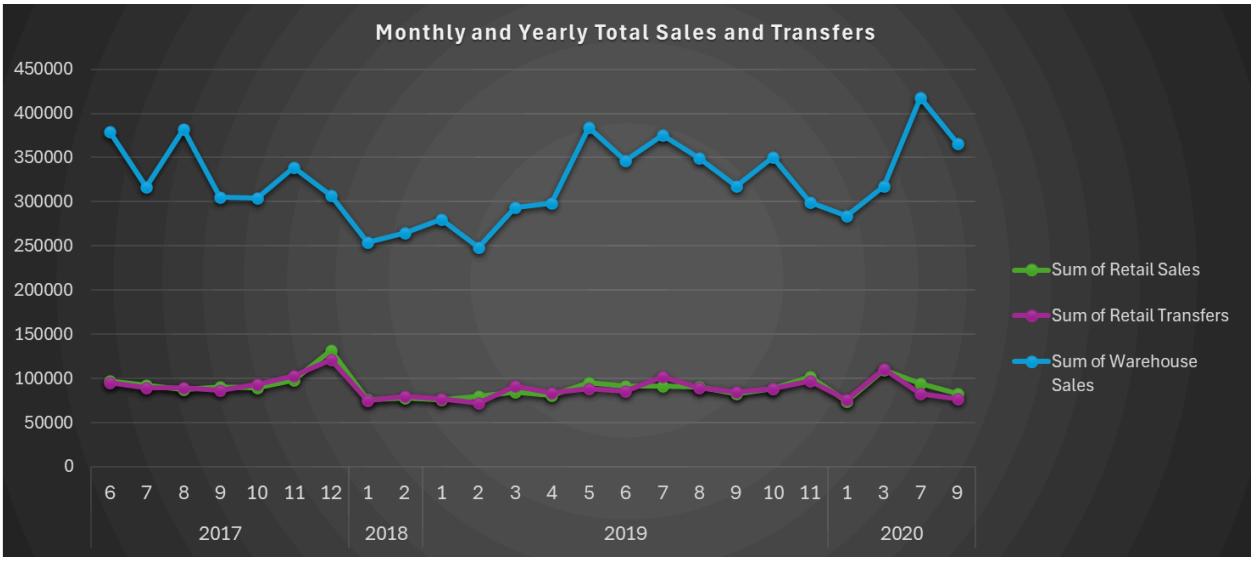
	Retail Sales	Retail Transfers	Warehouse Sales
Minimum	-6.49	-38.49	-7800
Average	7.02	6.94	25.29
Median	0.32	0	1
Max	2739	1990.83	18317
Sum	2,160,899.37	2,133,968.63	7,781,756.28
Standard Deviation	30.99	30.24	249.92

Retail sales had 0.04% negative values, retail transfers had 0.33% negative values, and warehouse sales had 0.23% negative values. While negative values are undesirable, it’s reassuring that they represent only a small portion of each column

The dataset contains 8 different item types, 34,045 different items, and 397 different suppliers.

The five most frequent suppliers are Republic National Distributing Co, Legends Ltd, Southern Glazers Wine and Spirits, E & J Gallo Winery, and The County Vinter, Llc Dba Winebow. The four most frequent Item Types are Wine, Liquor, Beer, and Kegs.

Monthly and Yearly Sales and Transfers



The y-axis represents how many cases were transferred or sold, while the x-axis represents the year and month. We have data from 2017 from June through December, data from 2018 in January and February only, data from 2019 from January through November, and data from 2020 in January, March, July, and September. Total warehouse sales are consistently higher overall through all the years, while retail sales and retail transfer are fairly similar to each other.

Due to missing monthly data, seasonality remains unclear. However, in between the last few months of 2017 and first month of 2018, it does seem that all total sales and transfers decreased, potentially indicating new year's resolutions where the demand dropped. A similar trend appears between November 2019 and January 2020 (but without the data from December, we can't know for sure).

Item Type Sales and Transfers

Moving on to item type performance:

- Beer: Warehouse sales (around 6,500,000 cases) significantly outpace both retail sales and retail transfers (both around 560,000 cases).
- Dunnage: There are no retail transfers or retail sales recorded for dunnage. However, a significant negative value (-120,000) in warehouse sales suggests a financial outflow in this category. Since most dunnage items are empty beer and wine kegs, the negative value likely reflects reimbursements for breakage or returns, resulting in a financial loss.
- Kegs: Retail sales for kegs are nonexistent, with only a minor negative amount recorded for retail transfers. Conversely, warehouse sales of kegs are substantial (around 120,000), which aligns with their primary use in restaurant and bar settings.
- Liquor: Both retail sales and retail transfers are significantly higher than warehouse sales (both at around 800,000 while warehouse is around 95,000). This disparity likely stems from how liquor is typically sold; individual bottles (not multi-packs like beer) are common in retail, while warehouse orders for licensees might also prioritize individual bottles or smaller case sizes compared to beer cases.
- Non-Alcohol: This category sees the highest cases for retail sales (around 34,000). Retail transfers and warehouse sales are roughly equivalent (26,000 cases). This makes sense given the product types — simple syrups, bitters, Bloody Mary mix, margarita mix, and ginger beer — which are frequently purchased by individual consumers and also ordered in bulk by restaurants and bars. Overall, this category shows the smallest gap between sales and transfers.
- Ref : Retail sales and transfers for ref products are barely positive. However, warehouse sales for this category show a significant negative value (around -20,000 cases). While the category includes positive-selling items like 'store special wine' and 'corkscrews', the negative values specifically correspond to 'beer and wine credit' items. This strongly suggests that these negative entries represent credits or reimbursements, likely due to damaged or returned beer and wine products.
- Str_supplies: Retail transfers for str_supplies are the highest (around 10,500), with no warehouse sales recorded. This category includes items like bartending books, wine tote bags, shot glasses, and receipt paper, which are primarily moved to retail locations for sale or internal use. Retail sales are around 2,500.
- Wine: Warehouse sales account for the highest total volume for wine (around 1,900,000). While retail sales and transfers occur at a substantial rate, they are both lower compared to the volume moved through the warehouse channel (both around 750,000).

The analysis above was done in a similar way as in this YouTube Video.

Top Suppliers by Total Retail Sales and Total Warehouse Sales

Moving onto supplier performance:

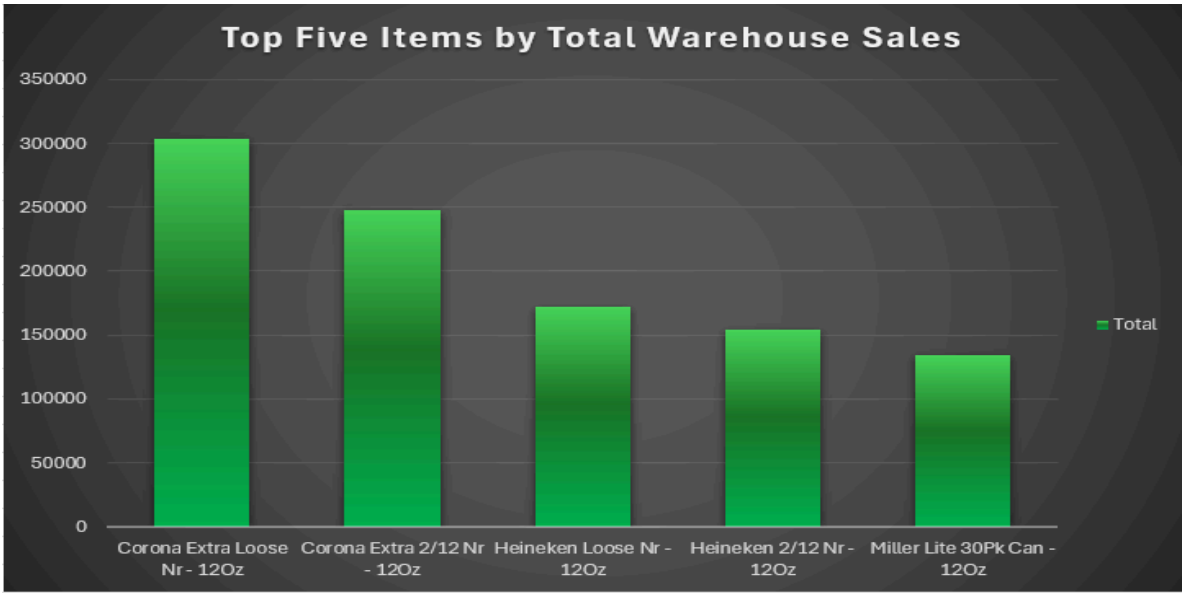
- The top five suppliers based on total retail sales are: E & J Gallo Winery, Diageo North America Inc, Constellation Brands, Anheuser Busch Inc, and Jim Beam Brands Co.

- The top five suppliers based on total warehouse sales are: Crown Imports, Miller Brewing Company, Anheuser Busch Inc, Heineken Usa, and E & J Gallo Winery.

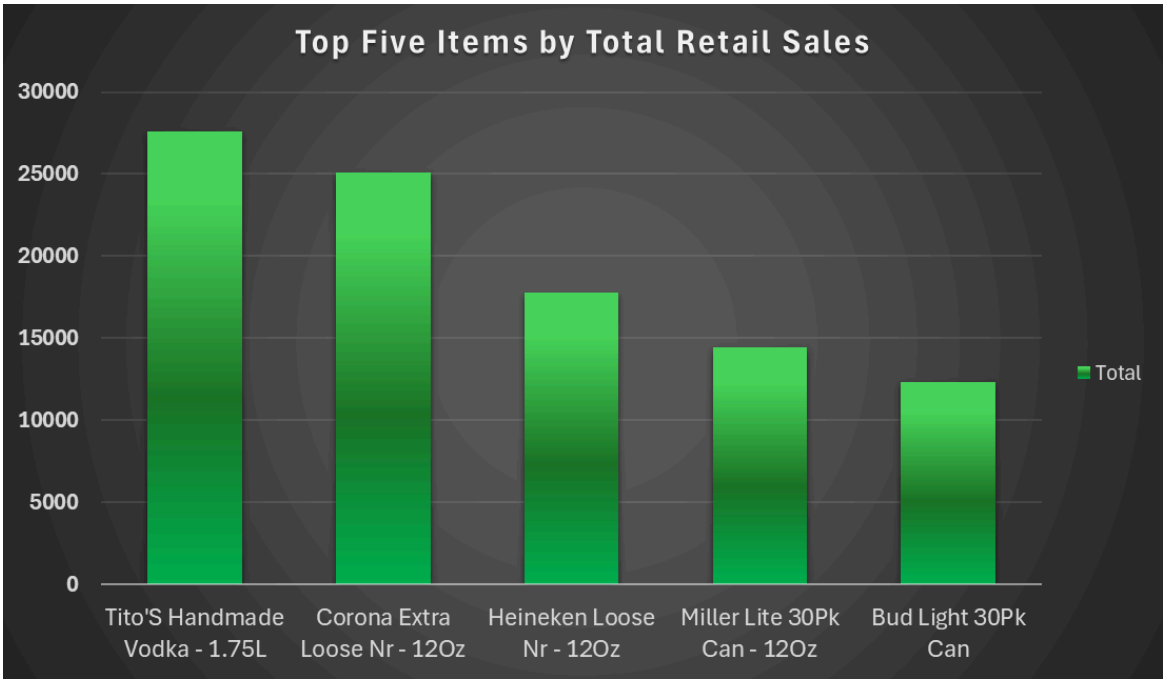
Top Items by Total Warehouse Sales and Total Retail Sales

Moving onto item performance:

- The top five total warehouse sale items were Corona Extra Loose NR 12 oz (where 'loose' refers to individual bottles and 'NR' means non-returnable), Corona Extra 2/12 Nr 12 oz (2/12 means two 12-packs), Heineken Loose Nr 12 oz, Heineken 2/12 Nr 12 oz, and Miller Lite 30 pk can 12 oz.



- The top five total retail sale items were Tito's Handmade Vodka 1.75 L, Corona Extra Loose Nr 12 oz, Heineken Loose Nr 12 oz, Miller Lite 30 pk can 12 oz, and Bud Light 30 pk can.



Conclusions

The first main insight is the dominance of warehouse sales for beer and wine compared to retail sales and transfers. This makes sense as the purchasers in this case are restaurants, bars, and stadiums, which need to make bulk purchases to meet customer demand.

The next insight was how retail sales and transfers were the dominant force for Liquor and Non-alcohol products. While the liquor is a surprise, it may be due to the fact that liquor isn't like beer which can be sold in cases 12/24/36 etc. Furthermore, liquor can potentially serve as a bigger bang for the buck, given the initial cost. The non-alcohol products do make sense in this, especially considering non-alcoholic beer, drink mix, and ice.

Keg sales were completely dominated in the warehouse, which does make sense as restaurants/bars may like to serve beer/wine straight from the keg to their consumers. Furthermore, there is a small negative value for retail transfers, indicating that broken kegs likely occurred during transportation.

Dunnage and ref products were very negative for warehouse sales. Given that dunnage includes empty beer and wine kegs, this could be due to kegs arriving broken, requiring customer reimbursement. For the ref category, this includes store special wines and beers, along with beer and wine credit, indicating that the warehouse provided credit for defective products.

While there are some hints for seasonality, especially a decrease from November/December to January in warehouse sales and retail sales/transfers, the gaps in monthly data for 2018 and 2019 limit a comprehensive assessment of seasonal patterns.

Although the sales are consistently greater for warehouse sales, they are more volatile than retail sales, which are more consistent. However, due to warehouse sales being greater regardless, if desired, it still may be beneficial to make sure both sales are positive.

Recommendations

My first recommendation is to further investigate why warehouse sales were negative for products in the dunnage/ref item categories. While I can hypothesize the reasoning for this, an inquiry could provide information that could help us limit this in the future.

The next recommendation is to maintain consistent supply of Corona, Heineken, Modelo, Bud Light, Miller Lite, and Stella Artois, as they have all been in the top five for warehouse and retail sales. Furthermore, Miller Lite, Stella Artois, Yuengling Lager, Bud Light, and Blue Moon (all ½ kegs) dominated total warehouse sales, indicating these types of beers and size of kegs are most popular. Additionally, Tito's, Montezuma Tequila, Triple Sec, Jameson, and Jack Daniel's are all fairly popular in both warehouse and retail sales.

Potential Future Work / Limitations

While this was a good starting point, it would be helpful to revisit in the future for a couple of reasons.

As this data covers 2017 to 2020 (with some gaps), there is a chance that patterns have shifted significantly. For example, newer products could have completely taken over sales in recent years, but since we are looking at data from 2020 and earlier, recent patterns may not be visible. We would likely see different beer preferences following the Bud Light controversy that occurred after this data period.

Another thing that could impact the validity is that there were some gaps, especially in 2018 as only two months were recorded. Although perfect datasets are rare in practice, expanding coverage in future iterations would enhance reliability.

Additionally, analyzing poorly performing items could help us identify not only which items need strong supply chains, but also which items to avoid purchasing. Last but not least, this data was specifically located in the Maryland area, so the results may not be transferable to other regions in the United States.

Dashboards

Dashboard 1 displays Monthly/Yearly Total Sales and Transfers as a line chart and Item Type Sales and Transfers as a column chart. Both charts can be customized with the Year and Item Type slicers, to view how well Beer sold in 2019 or how well Wine sold in 2017, etc.

Dashboard 2 displays Total Warehouse Sales by Supplier and Total Retail Sales by Supplier, with customizable Year and Item Type slicers. This helps identify which suppliers have the best-selling products, indicating potential profitability.

Dashboard 3 displays the Top Five items by Warehouse Sales and the Top Five Items by Retail Sales, while also having a slicer for Year and Item Type. This allows us to determine which beer sold best in retail in 2018 or which wine sold best in warehouse sales in 2019.

A video of each dashboard in action, along with the analysis and data page in Excel, can be viewed at the following video : Overview of Excel Workbook Mini-Project

Citations

[1] Warehouse and Retail Sales | Open Data Portal. (2025, July 5). https://data.montgomerycountymd.gov/Community-Recreation/Warehouse-and-Retail-Sales/v76h-r7br/about_data