Walmart Sales & Customer Behavior Analysis

Using SQL + Python for Business Insights

Project Overview

As part of my continuing effort to develop my data analysis skills, I recently completed an analysis of Walmart sales data using PostgreSQL, Python, and VS Code.

This project follows the YouTube tutorial by [Zero Analyst](#), using the dataset provided [on GitHub](#). However, I extended the analysis by formulating new questions, applying statistical testing (e.g. t-tests in Python), and exploring operational patterns beyond the scope of the original guide. Here's a summary of the key findings:

Key Business Questions & Findings

1) Which payment methods are most used and most profitable?

Objective: To identify customer preferences for payment methods based on the number of transactions, items sold, and total sales value for each method. This analysis can inform strategies for payment optimization.

Payment Method	Number of Payments	Number of Items Sold	Total Sales Value
Credit card	4,256.00	9,567.00	488,821.02
E-wallet	3,881.00	8,932.00	457,316.07
Cash	1,832.00	4,984.00	263,589.29

Insight: Credit cards were the most frequently used payment method, accounting for the highest number of transactions, items sold, and total sales value. E-wallet followed as the second most popular method. Cash transactions were notably lower compared to electronic payment methods. This highlights the importance of prioritizing and optimizing processes for credit card and E-wallet payments.

2) Which product categories perform best, and which need attention? What is the total revenue and profit for each category?

Objective: To determine the category with the highest and lowest average customer rating, as well as the total revenue and profit for each category across all Walmart branches. This information can help Walmart recognize and promote popular categories, leading to increased customer satisfaction and higher return on investment (ROI).

Category	Times Ranked Highest	Times Ranked Lowest	Total Revenue	Profit
Electronic accessories	19	18	78,175.03	30,772.49
Fashion accessories	3	18	489,480.90	192,314.89
Food and beverages	26	13	53,471.28	21,552.86
Health and beauty	25	9	46,851.18	18,671.73
Home and lifestyle	2	31	489,250.06	192,213.64
Sports and travel	26	11	52,497.93	20,613.81

Insight: The four highest rated categories, Electronic Accessories, Food & Beverages, Health & Beauty, Sports & Travel, were the least profitable product categories. Conversely, Home & Lifestyle and Fashion Accessories had the lowest ratings, but contributed 80%+ of profit. As such, Walmart should improve customer satisfaction in these high profit but poorly rated categories to maximize ROI.

3) Which days are busiest and most profitable? What product category had the highest number of transactions each day?

Objective: To identify the busiest days and highest selling product category across all Walmart branches. This can help Walmart optimize its staffing and inventory management.

Day	Number of Transactions	Total Sales Value	Category with the Highest Number of Transactions
Monday	601	$155,096.36	Fashion accessories
Tuesday	668	$181,217.71	Home and lifestyle
Wednesday	683	$174,515.70	Fashion accessories
Thursday	663	$174,092.76	Home and lifestyle

Friday	654	$168,103.61	Fashion accessories
Saturday	652	$185,624.39	Fashion accessories
Sunday	662	$171,075.85	Home and lifestyle

Insight: Analysis shows mid-weekdays (Tuesday, Thursday, and Wednesday) have the highest number of transactions while the beginning and end of the week (Friday and Monday) have the fewest transactions. The total sales value by day shows the same pattern with the exception of Saturday, which has the highest number of sales. Home & Lifestyle and Fashion Accessories departments have the highest number of transactions on any given day.

4) Are any stores significantly above or below the national average in customer ratings?

Objective: To determine which, if any, Walmart stores have statistically significant higher or lower customer ratings compared to the national average.

City	Average Customer Rating	T-Value	P-Value	Significant	Group
Plano	5.338298	-4.514085	1.0077E-05	True	Below Average
Port Arthur	5.298745	-5.049786	8.8004E-07	True	Below Average
Richardson	5.403057	-3.874732	1.3955E-04	True	Below Average
Rockwall	5.356757	-4.159793	4.5608E-05	True	Below Average
Abilene	6.336111	2.599907	1.1336E-02	True	Above Average
Alamo	6.595588	3.450362	9.7333E-04	True	Above Average
Allen	6.389041	2.724514	8.0772E-03	True	Above Average
Amarillo	6.542308	2.793862	7.3187E-03	True	Above Average

Note: Only the first five below and above average stores are shown. For a complete list see appendix 1.

Insight: Python was used to run t-tests to compare each store's average rating against the national average. 23 stores have a significantly below-average customer satisfaction t-value, and 50 stores have significantly above-average customer satisfaction t-value. All these stores have a significant p-value (<0.05). This allows for data-driven branch performance evaluation.

5) Which branches saw the biggest revenue shifts from 2022 to 2023?

Objective: To analyze the change in the gross income percentage for various Walmart branches. This helps in understanding branch performance trends and identifying areas of growth or decline.

Branch	City	2022 Year Revenue	2023 Year Revenue	Revenue Change Percentage	Change Type
WALM006	El Paso	765	2.0880E+03	172.94	Increase
WALM010	Laredo	808	2.1140E+03	161.63	Increase
WALM091	Little Elm	962	2.3940E+03	148.86	Increase
WALM072	Lancaster	811	1.7850E+03	120.1	Increase
WALM077	Coppell	862	1.8840E+03	118.56	Increase
WALM081	Friendswood	1723	850	-50.67	Decrease
WALM033	Pearland	2099	931	-55.65	Decrease
WALM098	Mineral Wells	2446	1030	-57.89	Decrease
WALM047	Flower Mound	2581	1069	-58.58	Decrease
WALM045	Missouri City	1731	647	-62.62	Decrease

Note: only the first five stores with the highest increase and decrease in yearly revenue are shown. For a complete list see appendix 2.

Insight: 57 branches experienced an increase in gross income percentage ranging from 1.5% to 173.94%. However, 42 branches experienced a revenue decrease ranging from 0.16% to 62.62%. Only one branch had no change in revenue. This suggests that while overall performance is positive, some branches may require further investigation to understand the reasons behind the decline and implement corrective measures.

Appendices

Appendix 1:

Store customer satisfaction rating vs national average customer rating. Descending by average customer rating.

City	Average Customer Rating	T-Value	P-Value	Significant	Group
Plano	5.338298	-4.514085	1.0077E-05	TRUE	Below Average
Port Arthur	5.298745	-5.049786	8.8004E-07	TRUE	Below Average
Richardson	5.403057	-3.874732	1.3955E-04	TRUE	Below Average
Rockwall	5.356757	-4.159793	4.5608E-05	TRUE	Below Average
Rosenberg	5.183616	-5.573845	9.2217E-08	TRUE	Below Average
Round Rock	5.238265	-5.419118	1.7548E-07	TRUE	Below Average
Rowlett	4.989349	-7.043323	4.5990E-11	TRUE	Below Average
San Angelo	5.227746	-5.140097	7.4026E-07	TRUE	Below Average
San Antonio	5.37043	-3.906563	1.3116E-04	TRUE	Below Average
San Marcos	5.076064	-6.703408	2.3338E-10	TRUE	Below Average
Schertz	5.319512	-4.728936	4.2064E-06	TRUE	Below Average
Seguin	5.319632	-4.147167	5.4180E-05	TRUE	Below Average
Sherman	5.064571	-7.014892	4.9086E-11	TRUE	Below Average
Southlake	5.308791	-4.535776	1.0422E-05	TRUE	Below Average
Sugar Land	5.275661	-4.74448	4.1263E-06	TRUE	Below Average
Temple	5.247644	-4.952538	1.6127E-06	TRUE	Below Average
Texas City	5.042045	-7.364085	6.7196E-12	TRUE	Below Average
Tyler	5.268235	-4.816639	3.2276E-06	TRUE	Below Average
Victoria	5.119205	-6.049286	1.1103E-08	TRUE	Below Average
Waco	5.385	-3.215748	1.5762E-03	TRUE	Below Average
Waxahachie	5.215079	-8.017225	1.3734E-14	TRUE	Below Average

Weatherford	5.084242	-6.810867	1.7492E-10	TRUE	Below Average
Weslaco	5.180808	-8.046771	1.0054E-14	TRUE	Below Average
Abilene	6.336111	2.599907	1.1336E-02	TRUE	Above Average
Alamo	6.595588	3.450362	9.7333E-04	TRUE	Above Average
Allen	6.389041	2.724514	8.0772E-03	TRUE	Above Average
Amarillo	6.542308	2.793862	7.3187E-03	TRUE	Above Average
Angleton	6.361194	2.595982	1.1614E-02	TRUE	Above Average
Austin	7.001667	6.353958	3.3201E-08	TRUE	Above Average
Baytown	6.215	2.006601	4.8214E-02	TRUE	Above Average
Bedford	6.33125	2.076044	4.1976E-02	TRUE	Above Average
Big Spring	6.449398	3.55207	6.3675E-04	TRUE	Above Average
Brownsville	6.588571	3.703182	4.2495E-04	TRUE	Above Average
Brownwood	6.492105	3.553551	6.6159E-04	TRUE	Above Average
Bryan	6.609375	3.411744	1.1322E-03	TRUE	Above Average
Burleson	6.322034	2.253411	2.8029E-02	TRUE	Above Average
Carrollton	6.283077	2.197471	3.1613E-02	TRUE	Above Average
Cedar Park	6.569565	3.542114	7.2223E-04	TRUE	Above Average
Cleburne	6.468354	3.089472	2.7785E-03	TRUE	Above Average
College Station	6.671429	3.676289	5.3852E-04	TRUE	Above Average
Conroe	6.398649	2.800791	6.5206E-03	TRUE	Above Average
Coppell	6.569231	3.723066	4.1780E-04	TRUE	Above Average

Corpus Christi	6.388235	2.395954	1.9377E-02	TRUE	Above Average
DeSoto	6.241772	2.148724	3.4758E-02	TRUE	Above Average
Denton	6.678082	4.014886	1.4417E-04	TRUE	Above Average
Edinburg	6.415942	2.741317	7.8120E-03	TRUE	Above Average
El Paso	6.33662	2.307916	2.3964E-02	TRUE	Above Average
Euless	6.294366	2.289388	2.5077E-02	TRUE	Above Average
Flower Mound	6.469136	3.471448	8.3754E-04	TRUE	Above Average
Fort Worth	6.582143	3.860817	2.2288E-04	TRUE	Above Average
Friendswood	6.366176	2.243841	2.8152E-02	TRUE	Above Average
Frisco	6.559375	3.222958	2.0105E-03	TRUE	Above Average
Garland	6.39625	2.723157	7.9553E-03	TRUE	Above Average
Georgetown	6.304938	2.409781	1.8258E-02	TRUE	Above Average
Grapevine	6.294667	2.460687	1.6200E-02	TRUE	Above Average
Haltom City	6.382192	2.531428	1.3544E-02	TRUE	Above Average
Harlingen	6.506329	3.404851	1.0484E-03	TRUE	Above Average
Houston	6.3	2.038158	4.5158E-02	TRUE	Above Average
Huntsville	6.812162	5.159443	2.0528E-06	TRUE	Above Average
Hurst	6.404412	2.639452	1.0319E-02	TRUE	Above Average
Kerrville	6.545333	3.5401	6.9575E-04	TRUE	Above Average
Killeen	6.546053	3.430629	9.8236E-04	TRUE	Above Average
Lake Jackson	6.527451	2.879475	5.8491E-03	TRUE	Above Average

Laredo	6.643939	4.158195	9.5870E-05	TRUE	Above Average
Longview	6.450847	2.818621	6.5866E-03	TRUE	Above Average
Lufkin	6.311594	2.137948	3.6118E-02	TRUE	Above Average
Mansfield	6.397468	2.847309	5.6331E-03	TRUE	Above Average
Midland	6.421429	2.853272	5.7070E-03	TRUE	Above Average
Mineral Wells	6.315385	2.119267	3.7953E-02	TRUE	Above Average
Nacogdoches	6.397333	2.732437	7.8578E-03	TRUE	Above Average
New Braunfels	6.269118	2.0562	4.3662E-02	TRUE	Above Average
Pasadena	6.538028	3.435663	9.9796E-04	TRUE	Above Average
Pflugerville	6.732911	4.571897	1.7877E-05	TRUE	Above Average

Appendix 2:

Store revenue shifted from 2022 to 2024. Descending by revenue change ratio.

Branch	City	2022 Year Revenue	2023 Year Revenue	Revenue Change Ratio	Change Type
WALM006	El Paso	765	2.0880E+03	172.94	Increase
WALM010	Laredo	808	2.1140E+03	161.63	Increase
WALM091	Little Elm	962	2.3940E+03	148.86	Increase
WALM072	Lancaster	811	1.7850E+03	120.1	Increase
WALM077	Coppell	862	1.8840E+03	118.56	Increase
WALM014	Amarillo	939	2.0170E+03	114.8	Increase
WALM008	Corpus Christi	1141	2.4450E+03	114.29	Increase

WALM060	DeSoto	1291	2.5480E+03	97.37	Increase
WALM007	Arlington	1301	2.5470E+03	95.77	Increase
WALM039	Longview	868	1.6850E+03	94.12	Increase
WALM027	Abilene	1483	2.8730E+03	93.73	Increase
WALM068	Burleson	888	1.6960E+03	90.99	Increase
WALM057	Euless	893	1.5570E+03	74.36	Increase
WALM063	Georgetown	1294	2.1940E+03	69.55	Increase
WALM044	Pharr	968	1.6290E+03	68.29	Increase
WALM017	McKinney	1446	2.3320E+03	61.27	Increase
WALM040	Edinburg	1142	1.7860E+03	56.39	Increase
WALM066	Grapevine	1198	1.8360E+03	53.26	Increase
WALM038	Sugar Land	3864	5.9000E+03	52.69	Increase
WALM052	Mansfield	1192	1.8150E+03	52.27	Increase
WALM012	Garland	1481	2.1940E+03	48.14	Increase
WALM028	Odessa	1244	1.7990E+03	44.61	Increase
WALM036	Allen	989	1.4190E+03	43.48	Increase
WALM089	Southlake	3326	4.7440E+03	42.63	Increase
WALM090	Brownwood	1546	2.1810E+03	41.07	Increase
WALM021	McAllen	1162	1.6100E+03	38.55	Increase
WALM049	North Richland Hills	1539	2.1290E+03	38.34	Increase
WALM026	Denton	1476	2.0270E+03	37.33	Increase
WALM050	Victoria	3318	4.4990E+03	35.59	Increase
WALM100	Canyon	1156	1.5660E+03	35.47	Increase
WALM002	Dallas	1169	1.5720E+03	34.47	Increase
WALM054	Sherman	3484	4.6310E+03	32.92	Increase
WALM035	San Angelo	3986	5.2620E+03	32.01	Increase
WALM075	San Marcos	3391	4.4110E+03	30.08	Increase
WALM079	La Porte	1387	1.7700E+03	27.61	Increase

WALM011	Lubbock	1235	1.5570E+03	26.07	Increase
WALM042	Bryan	1073	1.3280E+03	23.77	Increase
WALM069	Rockwall	4439	5.4140E+03	21.96	Increase
WALM009	Plano	4876	5.9330E+03	21.68	Increase
WALM083	Farmers Branch	1352	1.6450E+03	21.67	Increase
WALM055	Waxahachie	4022	4.8600E+03	20.84	Increase
WALM074	Weslaco	4257	5.1250E+03	20.39	Increase
WALM061	Cedar Park	1368	1.5850E+03	15.86	Increase
WALM082	Weslaco	4299	4.9740E+03	15.7	Increase
WALM022	Mesquite	1684	1.8900E+03	12.23	Increase
WALM070	Hurst	1264	1.4130E+03	11.79	Increase
WALM023	Midland	1222	1.3660E+03	11.78	Increase
WALM019	Pasadena	1370	1.5220E+03	11.09	Increase
WALM067	Haltom City	1850	2.0460E+03	10.59	Increase
WALM071	Lufkin	1280	1.3990E+03	9.3	Increase
WALM015	Grand Prairie	2,388.00	2.6000E+03	8.88	Increase
WALM051	New Braunfels	1416	1.5360E+03	8.47	Increase
WALM005	Fort Worth	1994	2.1520E+03	7.92	Increase
WALM080	Nacogdoches	1524	1.6410E+03	7.68	Increase
WALM003	San Antonio	4436	4.7410E+03	6.88	Increase
WALM058	Port Arthur	5227	5.4050E+03	3.41	Increase
WALM020	Killeen	1067	1.0830E+03	1.5	Increase
WALM048	**Harlingen**	**1546**	**1.5460E+03**	**0**	**No Change**
WALM056	Rowlett	3736	3.7300E+03	-0.16	Decrease
WALM032	Tyler	3860	3.8330E+03	-0.7	Decrease
WALM087	Waxahachie	4056	4.0050E+03	-1.26	Decrease
WALM096	Eagle Pass	2288	2.2550E+03	-1.44	Decrease

WALM095	Big Spring	1725	1.6990E+03	-1.51	Decrease
WALM001	Houston	1488	1.4630E+03	-1.68	Decrease
WALM031	Lewisville	1013	9.8300E+02	-2.96	Decrease
WALM065	Texas City	4966	4.6810E+03	-5.74	Decrease
WALM062	Galveston	1500	1.3960E+03	-6.93	Decrease
WALM025	Waco	3262	3.0280E+03	-7.17	Decrease
WALM092	Lake Jackson	1130	1.0260E+03	-9.2	Decrease
WALM024	Carrollton	1496	1.3490E+03	-9.83	Decrease
WALM041	Mission	1417	1.2310E+03	-13.13	Decrease
WALM073	Seguin	4446	3.7770E+03	-15.05	Decrease
WALM016	Brownsville	1969	1.6670E+03	-15.34	Decrease
WALM086	Rosenberg	4091	3.4270E+03	-16.23	Decrease
WALM013	Irving	1308	1.0940E+03	-16.36	Decrease
WALM053	Conroe	1594	1319	-17.25	Decrease
WALM076	Huntsville	1854	1520	-18.02	Decrease
WALM094	Alamo	1623	1323	-18.48	Decrease
WALM099	Weatherford	4662	3712	-20.38	Decrease
WALM018	Frisco	1746	1376	-21.19	Decrease
WALM043	Baytown	1899	1470	-22.59	Decrease
WALM093	Angleton	1640	1263	-22.99	Decrease
WALM030	Richardson	5622	4130	-26.54	Decrease
WALM037	League City	1748	1255	-28.2	Decrease
WALM004	Austin	1337	958	-28.35	Decrease
WALM046	Temple	5074	3528	-30.47	Decrease
WALM029	Round Rock	5400	3750	-30.56	Decrease
WALM034	College Station	1645	1131	-31.25	Decrease
WALM088	Cleburne	1693	1126	-33.49	Decrease
WALM084	Schertz	6067	4011	-33.89	Decrease

WALM078	Del Rio	2298	1421	-38.16	Decrease
WALM085	Kerrville	1921	1179	-38.63	Decrease
WALM097	Alice	1446	814	-43.71	Decrease
WALM064	Bedford	1786	980	-45.13	Decrease
WALM059	Pflugerville	2060	1123	-45.49	Decrease
WALM081	Friendswood	1723	850	-50.67	Decrease
WALM033	Pearland	2099	931	-55.65	Decrease
WALM098	Mineral Wells	2446	1030	-57.89	Decrease
WALM047	Flower Mound	2581	1069	-58.58	Decrease
WALM045	Missouri City	1731	647	-62.62	Decrease

References

Tutorial video:

https://youtu.be/49C3Mkmhskg?si=EwnajZu_FxV-Y3-z

GitHub dataset:

https://github.com/najirh/Walmart_SQL_Python