

E-Commerce Sales Analysis

Dzikri Nur Rochim

Table of Contents

- Background of Project
- Business Questions
- Analysis Approach
- Key Findings
- Recommendations
- Conclusions

Background of Project

This data is sourced from the Kaggle website. The dataset covers a sales period of 4 years, from 2011 to 2014. The data encompasses 22 columns with 9994 rows, focusing on sales transactions for furniture, office supplies, and technology products. There are 219,868 entries with key columns including order ID, product name, category, customer name, quantity, sales, and profit. The objective of this data analysis is to enhance sales and customer satisfaction. Analysis techniques employed will involve data visualization and descriptive statistics

By continuing this data analysis, it is expected to gain insights into product performance, customer segmentation, and factors influencing sales, allowing for price optimization. The data used is ensured to have undergone a data cleaning process to ensure valid and accountable analysis results.

Business Questions

Here are a few potential smart method questions for analyzing the sales data:

1. Which products had the highest total sales revenue?
2. Which customers spent the most on furniture?
3. Which product categories had the lowest average discount percent?
4. Which regions had the highest total sales for Technology?
5. What was the best selling product?
6. What product categories make up the largest percentage of overall revenue?
7. Which products have the highest profit margins?
8. Which customers have the highest lifetime value?
9. How do sales and profits vary by region?
10. How do sales fluctuate seasonally?

Analysis Approach

Q.1 Which products had the highest total sales revenue?

1. Using Excel to filter the columns to be used
Creating a pivot table to facilitate finding answers to the questions
2. The rows column is filled with the category column, and the values column is filled with the sales column, calculated as the sum of sales
3. After that, the values are sorted from largest to smallest, and the highest sales value based on the product is obtained.

Analysis Approach

Q.2 Which customers spent the most on furniture?

1. In the Columns area, fill it with the category column, the Rows area is filled with the Rows ID column, and the Values area is filled with the Sales column displayed as the Sum of Sales. Then, simply sort the Furniture column with a descending filter by the sum of sales.
2. The result is obtaining the customer ID with the highest spending for furniture products.

Analysis Approach

Q.3 Which product categories had the lowest average discount percent?

1. Creating a pivot table with the rows section filled with categories, the columns section as values, and the values section filled with 2 data columns, namely discount and sales. Formatting the discount column as the average discount and the sales column as sum of sales.

2. Then, you will obtain a table of category vs average of discount and sum of sales. Perform a filter with `More Sort Options` >> `ascending (A to Z) by` >> choose the average of discount column.

Analysis Approach

Q.4 Which regions had the highest total sales for Technology?

1. In the Rows area, fill it with the category column and Region, and the Values area is filled with the Sales column displayed as the Sum of Sales. Then, simply click on 'More Sort Options,' choose 'Descending (Z to A) by sum of sales' to sort the category rows from highest to lowest sum of sales.

Analysis Approach

Q.5 What was the best selling product?

1. What should be noted is that the term 'best-selling product' can refer to either the quantity of items sold or the high revenue generated.
2. In the Columns section, you'll want to input values. For the Rows section, go ahead and populate it with the 'Product Name' column. Now, in the Values section, include three columns: quantity, price, and sales. Make sure these three columns are set in the sum format, and you can name them 'Sum of Quantity,' 'Sum of Price,' and 'Sum of Sales.'
3. We filter using 'value filters,' select 'top 10' to choose the top 10 products. In the 'top 10 items by' section, choose 'sum of sales' to determine the top 10 sales products or choose 'sum of quantity' based on the quantity of items sold. Then, sort using 'more sort options' > 'descending (Z to A) by' > select 'sum of sales' or 'sum of quantity.' This refers to best-selling by quantity or sales.

Key Findings

1. Which products had the highest total sales revenue?

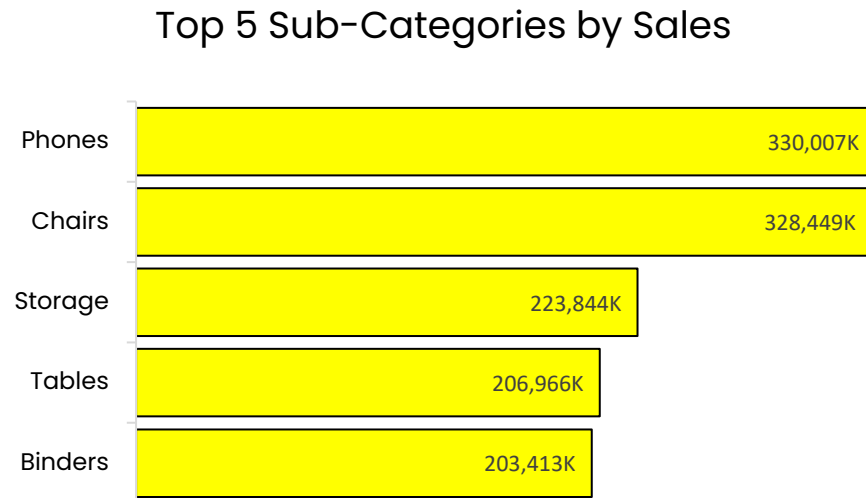


Figure 1. Bar chart displaying category vs sales.

From Figure 1, which shows products/sub-categories against sales, the highest sales were achieved by the "phones" product with a total sales of around 330 thousand USD. Followed by the sales of "chairs" in the second position with a total sales of around 328 thousand USD.

Key Findings

2. Which customers spent the most on furniture?

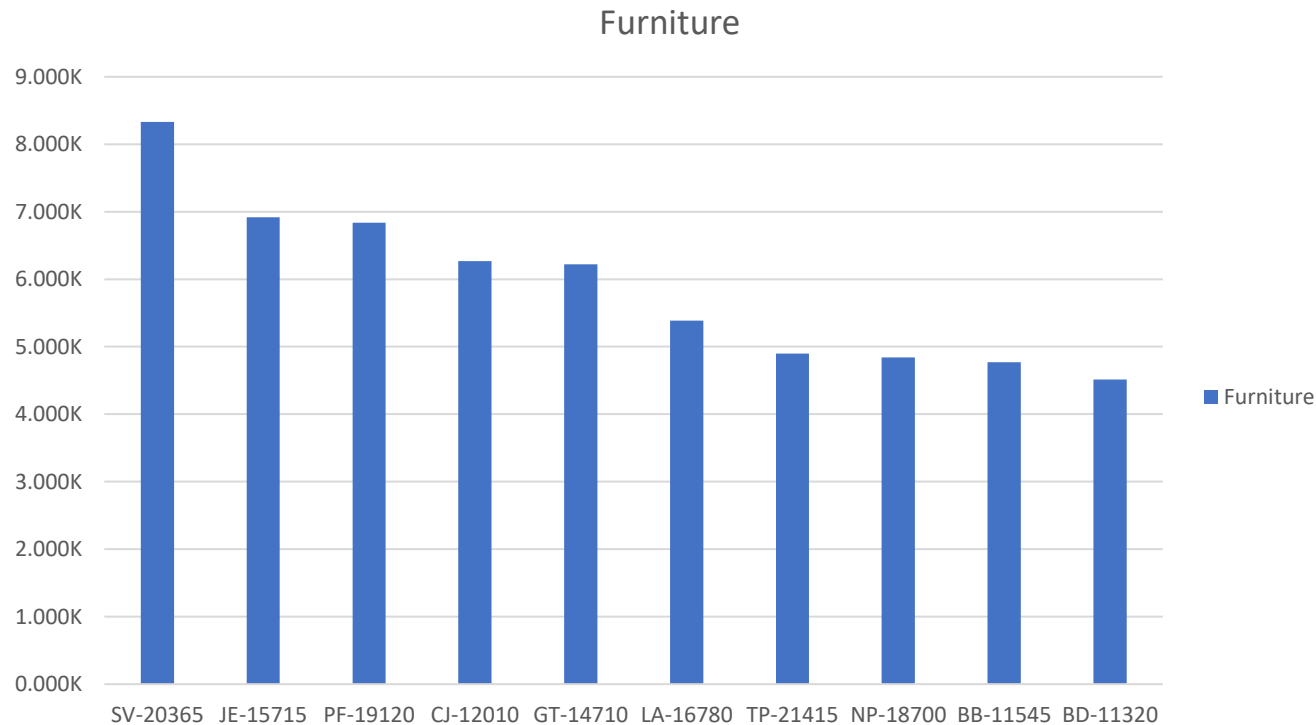


Figure 2. Column chart displaying sales vs customer ID.

In Figure 2, it illustrates the amount of money spent against customer ID. It can be observed that customer ID SV-20365 spent the most, totaling around 8.3 thousand USD. This is followed by a spending of approximately 6.9 thousand USD by customer ID JE-15715.

Key Findings

3. Which product categories had the lowest average discount percent?

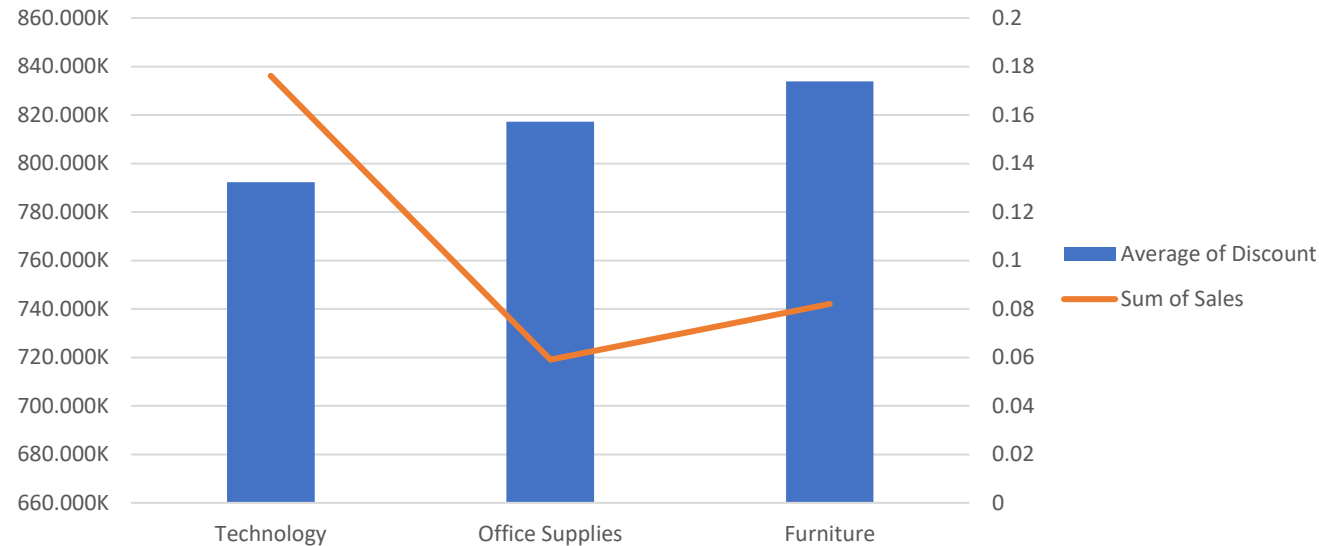


Figure 3. Combo chart displaying sales vs customer ID.

In Figure 3, it shows a combo chart depicting a bar chart for the average discount for each product category combined with sales for each product category visualized with a line chart. It can be seen that the Technology product category has the lowest average discount at 13.2%, followed by the Office Supplies category at 15.7%, and the Furniture category with the highest discount at 17.4%. Interestingly, the Technology product category with the lowest discount provides the highest sales revenue compared to the other two categories. Typically, the category with the highest discount would be expected to have the lowest sales, but in this case, it is higher than the Office Supplies category.

Key Findings

4. Which regions/state had the highest total sales for Technology?

Table 1. The table consists of columns for Category and Sales.

Category	Sales
Office Supplies	719.047K
Central	167.026K
East	205.516K
South	125.651K
West	220.853K
Furniture	742.000K
Central	163.797K
East	208.291K
South	117.299K
West	252.613K
Technology	836.154K
Central	170.416K
East	264.974K
South	148.772K
West	251.992K
Grand Total	2,297.201K

Table 1 is a table that contains columns for sales and rows for categories and regions. For the category with the highest revenue, Technology contributes 836.1k USD, accounting for around 36% of the total sales. The region with the highest sales in the Technology category is the West region, with a total sales of 251.9k USD, followed by the East region with sales reaching 264.9k USD. The South region is the lowest-performing area for sales, with a total sales value of only 148.7k USD.

Key Findings

5. What was the best selling product? Bisa dari kuantitas/ total sales

Table 2. The Sales Table contains columns for Product Name/Items, Quantity, Price, and Sales, sorted by the highest quantity.

Items	Quantity	Price	Sales
Staples	876	8.000228311	7008.2
KI Adjustable-Height Table	74	61.52217568	4552.641
Storex Dura Pro Binders	71	3.923746479	278.586
Avery Non-Stick Binders	71	3.060788732	217.316
GBC Premium Transparent Covers with Diagonal Lined Pattern	67	11.64859701	780.456
Situations Contoured Folding Chairs, 4/Set	64	56.89490625	3641.274
Chromcraft Round Conference Tables	61	134.5747377	8209.059
Eldon Wave Desk Accessories	61	4.372622951	266.73
Wilson Jones Turn Tabs Binder Tool for Ring Binders	59	3.594576271	212.08
Global Wood Trimmed Manager's Task Chair, Khaki	59	70.62515254	4166.884
Grand Total	1463	20.05005195	29333.226

Table 2 shows the best-selling products in terms of quantity. It is evident that the Staples product takes the lead in sales with a quantity of 876, followed by the KI Adjustable-Height Table with a quantity of 74. The top 10 best-selling products have a total quantity sold reaching 1463, contributing to 14.6% of the total products sold.

Table 3. The Sales Table contains columns for Product Name/Items, Quantity, Price, and Sales, sorted by the highest sales.

Items	Quantity	Price	Sales
Canon imageCLASS 2200 Advanced Copier	20	3079.9912	61599.824
Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind	31	885.5930323	27453.384
Cisco TelePresence System EX90 Videoconferencing Unit	6	3773.08	22638.48
HON 5400 Series Task Chairs for Big and Tall	39	560.784	21870.576
GBC DocuBind TL300 Electric Binding System	37	535.7697027	19823.479
GBC Ibimaster 500 Manual ProClick Binding System	48	396.34375	19024.5
Hewlett Packard LaserJet 3310 Copier	38	495.7812105	18839.686
HP Designjet T520 Inkjet Large Format Printer - 24" Color	12	1531.24125	18374.895
GBC DocuBind P400 Electric Binding System	27	665.3728889	17965.068
High Speed Automatic Electric Letter Opener	11	1548.210182	17030.312
Grand Total	269	909.3687881	244620.204

Table 3 displays the top-selling products in terms of sales. It is evident that the Canon imageCLASS 2200 Advanced Copier is the best-selling product with sales of 61.5K USD, followed by the Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind, totaling 27.4K USD in sales.

The top 10 best-selling products have a combined total sales of 244.62K USD, contributing to around 10% of the total sales during the period from 2011 to 2014.

Recommendations

1. Which products had the highest total sales revenue?

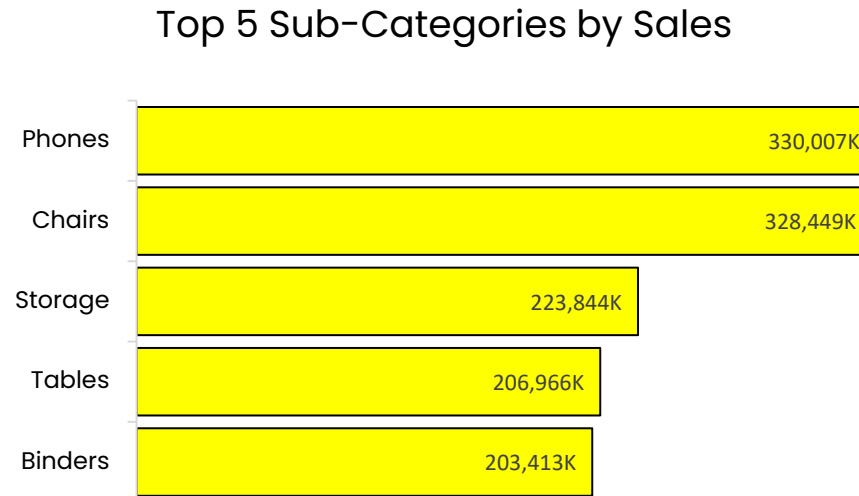


Figure 1. Bar chart displaying category vs sales.

Recommendation: We can focus on the two sub-categories with the best sales, namely phones and chairs. Considering this, there are two actionable things. The first is to invest in advertising and promote both sub-categories to enhance visibility and sales. The second is to continue regular promotions for all products and expand promotional and marketing efforts for the sub-category with the lowest sales to boost product purchasing power.

Recommendations

2. Which customers spent the most on furniture?

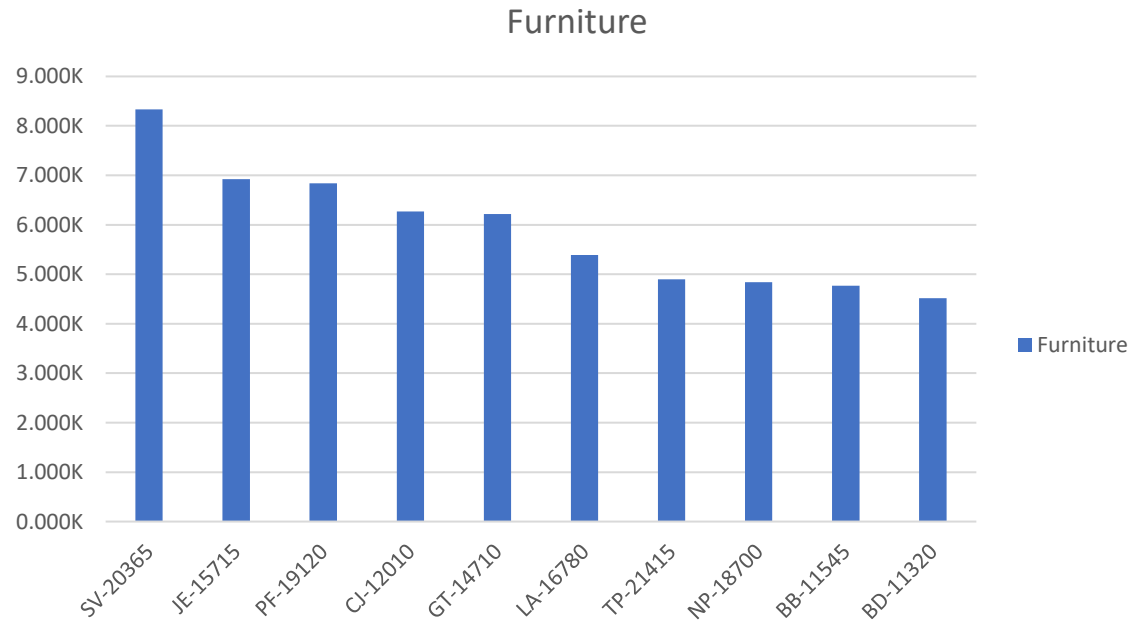


Figure 2. Column chart displaying sales vs customer ID.

Recommendation: Consider creating personalized loyalty programs or incentives for high-spending customers like SV-20365 and JE-15715. Rewarding their loyalty with exclusive benefits can encourage them to continue their high spending. Create targeted marketing campaigns specifically designed to attract and retain customers with similar profiles to SV-20365 and JE-15715. These campaigns can focus on the products or services that resonate with this customer segment.

Recommendations

3. Which product categories had the lowest average discount percent?

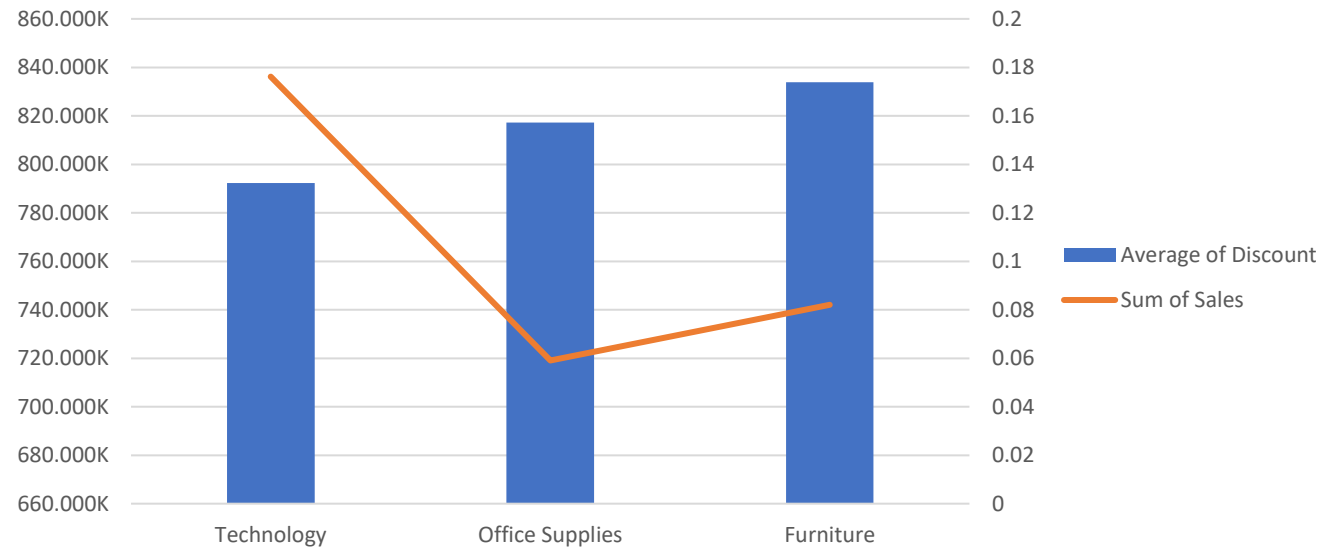


Figure 3. Combo chart displaying sales vs customer ID.

Recommendation: Since the Technology category has the lowest average discount but the highest sales, the company should focus more marketing and promotions on this category to drive even higher sales. Technology products tend to have high demand, so they can still sell well even with lower discounts. For the Furniture category which has the highest average discount but lower sales than Technology, the company could consider decreasing the discounts a bit. Furniture is more elastic in demand, so reducing discounts may optimize profitability without too much reduction in sales volume. For Office Supplies, the average discount is in the middle but it has the lowest sales. The company could increase discounts here to drive higher sales volume, or revamp the product offerings to focus on faster-selling office items. Beyond discounts, the company can also boost sales through marketing campaigns, loyalty programs, bundle deals, and offering value-added services. A comprehensive strategy is needed

Recommendations

4. Which regions/state had the highest total sales for Technology?

Table 1. The table consists of columns for Category and Sales.

Category	Sales
Office Supplies	719.047K
Central	167.026K
East	205.516K
South	125.651K
West	220.853K
Furniture	742.000K
Central	163.797K
East	208.291K
South	117.299K
West	252.613K
Technology	836.154K
Central	170.416K
East	264.974K
South	148.772K
West	251.992K
Grand Total	2,297.201K

Recommendation: Focus marketing and promotions for Office Supplies more in the West region, since it has the highest sales for this category. Central region has the lowest Office Supplies sales, so the company could consider reducing inventory/selection here. For Furniture, the West region again has significantly higher sales. Additional marketing, competitive pricing, and maintaining good inventory levels could further boost Furniture sales in the West. For Technology, sales are more evenly distributed by region, with East having the highest share. The company should ensure Technology products are well-stocked across regions to align with demand. Consider running regional promotions or discounts periodically for categories that are underperforming in certain regions relative to their category potential. This can help boost sales.

Recommendations

5. What was the best selling product? Bisa dari kuantitas/ total sales

Table 2. The Sales Table contains columns for Product Name/Items, Quantity, Price, and Sales, sorted by the highest quantity.

Items	Quantity	Price	Sales
Staples	876	8.000228311	7008.2
KI Adjustable-Height Table	74	61.52217568	4552.641
Storex Dura Pro Binders	71	3.923746479	278.586
Avery Non-Stick Binders	71	3.060788732	217.316
GBC Premium Transparent Covers with Diagonal Lined Pattern	67	11.64859701	780.456
Situations Contoured Folding Chairs, 4/Set	64	56.89490625	3641.274
Chromcraft Round Conference Tables	61	134.5747377	8209.059
Eldon Wave Desk Accessories	61	4.372622951	266.73
Wilson Jones Turn Tabs Binder Tool for Ring Binders	59	3.594576271	212.08
Global Wood Trimmed Manager's Task Chair, Khaki	59	70.62515254	4166.884
Grand Total	1463	20.05005195	29333.226

Table 3. The Sales Table contains columns for Product Name/Items, Quantity, Price, and Sales, sorted by the highest sales.

Items	Quantity	Price	Sales
Canon imageCLASS 2200 Advanced Copier	20	3079.9912	61599.824
Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind	31	885.5930323	27453.384
Cisco TelePresence System EX90 Videoconferencing Unit	6	3773.08	22638.48
HON 5400 Series Task Chairs for Big and Tall	39	560.784	21870.576
GBC DocuBind TL300 Electric Binding System	37	535.7697027	19823.479
GBC Ibimaster 500 Manual ProClick Binding System	48	396.34375	19024.5
Hewlett Packard LaserJet 3310 Copier	38	495.7812105	18839.686
HP Designjet T520 Inkjet Large Format Printer - 24" Color	12	1531.24125	18374.895
GBC DocuBind P400 Electric Binding System	27	665.3728889	17965.068
High Speed Automatic Electric Letter Opener	11	1548.210182	17030.312
Grand Total	269	909.3687881	244620.204

Recommendation: Focus marketing and promotions on the top selling products like Staples, Canon copiers, and Fellowes binding machines to drive more sales of already popular items. Analyze what is making lower selling products less successful and identify opportunities to boost their sales through pricing adjustments, bundling, and targeted promotions. For lower volume but higher priced products like the video conferencing system, explore opportunities to drive business sales and partnerships to increase volumes. Assess inventory levels regularly and ensure top products are well-stocked to align with demand. Avoid overstocking of slower items.

Conclusion

The analysis of the e-commerce sales data from 2011 to 2014 provides valuable insights into product performance, customer segments, and sales trends. The top selling sub-categories are phones and chairs, indicating strong demand in those product lines. Customers like SV-20365 and JE-15715 have high lifetime value and spending levels, presenting opportunities for personalized loyalty programs.

Technology has the lowest average discount but the highest sales volume, so discounts could potentially be decreased without reducing demand. Meanwhile, sales of Office Supplies lag the other categories, so expanded promotions and discounts may help improve performance. Geographically, the West region leads sales for Office Supplies and Furniture, while demand for Technology is more uniform across regions.

At the product level, office staples, Canon copiers, and Fellowes binding machines are the top sellers based on quantity and revenue. Maintaining availability of leading products can help match supply to high demand. Slow moving items may require pricing adjustments or bundling strategies to increase sales.

Overall, the analysis reveals areas where targeted marketing, competitive pricing, inventory optimization, and customer loyalty programs could boost revenues. By leveraging data-driven insights on product, customer, and geographic performance, the company can refine its e-commerce strategy to maximize sales growth and profitability.

Background of Project

Data ini berasal dari website **kaggle**. Data mencakup periode penjualan selama **4 tahun dari 2011 hingga 2014**. Ruang lingkup data meliputi **22 kolom** dengan **9994 baris**. Transaksi penjualan untuk produk **furniture, office supplies, dan technology**. Terdapat **219,868 data** dengan **kolom utama** meliputi order ID, product name, category, customer name, quantity, sales, dan profit. **Tujuan analisis** data ini adalah untuk meningkatkan penjualan dan kepuasan pelanggan. **Teknik analisis** yang akan digunakan mencakup visualisasi data, statistik deskriptif, regresi, dan clustering.

Dengan melanjutkan analisis data ini, **diharapkan dapat diperoleh** wawasan tentang performa produk, segmentasi pelanggan, serta faktor-faktor yang mempengaruhi penjualan sehingga dapat dilakukan optimasi harga. Data yang digunakan dipastikan telah lolos proses pembersihan data agar hasil analisis valid dan dapat dipertanggungjawabkan.

Conclusion

- Smart questions and data analysis provide business insights
- Identified top performing products, customers, regions
- Provided data-driven recommendations to support strategy

Recommendations

1. Kita dapat focus terhadap 2 sub-kategori dengan penjualan terbaik, yaitu telepon dan kursi. Mempertimbangkan hal itu, terdapat 2 actionable things, yaitu invest iklan dan promoting kedua sub-kategori produk tersebut. Yang kedua adalah tetap melakukan promosi seperti biasa dan expanding promoting and marketing untuk sub-kategori dengan penjualan terendah untuk meningkatkan daya beli produk.
2. Consider creating personalized loyalty programs or incentives for high-spending customers like SV-20365 and JE-15715. Rewarding their loyalty with exclusive benefits can encourage them to continue their high spending. Create targeted marketing campaigns specifically designed to attract and retain customers with similar profiles to SV-20365 and JE-15715. These campaigns can focus on the products or services that resonate with this customer segment.
3. Since the Technology category has the lowest average discount but the highest sales, the company should focus more marketing and promotions on this category to drive even higher sales. Technology products tend to have high demand, so they can still sell well even with lower discounts. For the Furniture category which has the highest average discount but lower sales than Technology, the company could consider decreasing the discounts a bit. Furniture is more elastic in demand, so reducing discounts may optimize profitability without too much reduction in sales volume. For Office Supplies, the average discount is in the middle but it has the lowest sales. The company could increase discounts here to drive higher sales volume, or revamp the product offerings to focus on faster-selling office items. Beyond discounts, the company can also boost sales through marketing campaigns, loyalty programs, bundle deals, and offering value-added services. A comprehensive strategy is needed to optimize sales and profits.

Recommendations

4. Focus marketing and promotions for Office Supplies more in the West region, since it has the highest sales for this category. Central region has the lowest Office Supplies sales, so the company could consider reducing inventory/selection here. For Furniture, the West region again has significantly higher sales. Additional marketing, competitive pricing, and maintaining good inventory levels could further boost Furniture sales in the West. For Technology, sales are more evenly distributed by region, with East having the highest share. The company should ensure Technology products are well-stocked across regions to align with demand. Consider running regional promotions or discounts periodically for categories that are underperforming in certain regions relative to their category potential. This can help boost sales.

5. Focus marketing and promotions on the top selling products like Staples, Canon copiers, and Fellowes binding machines to drive more sales of already popular items. Analyze what is making lower selling products less successful and identify opportunities to boost their sales through pricing adjustments, bundling, and targeted promotions. For lower volume but higher priced products like the video conferencing system, explore opportunities to drive business sales and partnerships to increase volumes. Assess inventory levels regularly and ensure top products are well-stocked to align with demand. Avoid overstocking of slower items.

Analysis Approach

- For the analysis approach section of a data presentation, you should describe the specific methods and analytical techniques used to answer each business question.
- Rather than listing all the methods you used for the entire analysis, you want to break it down by each question:
- For example:
- Business Question 1: What are the top selling products?
- Analysis Approach: Used Excel to filter the data for only product sales records. Sorted the filtered data by sales amount in descending order to reveal the products with highest sales at the top.
- Business Question 2: Which customers have the highest lifetime value?
- Analysis Approach: Used Excel pivot tables to group data by customer name and sum the total sales per customer. Applied a sort on the total sales column to show customers with highest totals.
- Business Question 3: How do sales vary by region?
- Analysis Approach: Built a pivot table grouped by region. Calculated total sales per region and created a chart showing sales by region for comparison.

Analysis Approach

Here are a few potential smart method questions for analyzing the sales data:

1. Which products had the highest total sales revenue?
2. Which customers spent the most on furniture?
3. Which product categories had the lowest average discount percent?
4. Which regions had the highest total sales for Technology?
5. What was the best selling product?

Analysis approach Q.1 Which products had the highest total sales revenue?

1. Menggunakan excel untuk menyaring kolom yang akan digunakan
2. Membuat pivot table untuk memudahkan menemukan jawaban dari pertanyaan tertentu
3. Kolom rows diisi dengan kolom category dan kolom values diisi kolom sales yang dihitung menjadi sum of sales
4. Setelah itu diurutkan nilainya dari terbesar ke terkecil dan didapatkan nilai penjualan tertinggi berdasarkan produk

Analysis approach Q.2

1. Menggunakan pivot table pada Q1, kemudian tinggal mengurutkan kolom Furniture dengan filter descending
2. Didapatkan customer-id dengan spent tertinggi untuk produk furniture

Analysis Approach

Here are a few potential smart method questions for analyzing the sales data:

6. What product categories make up the largest percentage of overall revenue?
 7. Which products have the highest profit margins?
 8. Which customers have the highest lifetime value?
 9. How do sales and profits vary by region?
 10. How do sales fluctuate seasonally?
6. Which products have increasing or decreasing sales momentum?
7. How do discounts and promotions impact profits?

Analysis Approach

Here are a few potential smart method questions for analyzing the sales data:

1. Which products had the highest total sales revenue?
2. Which customers spent the most on furniture?
3. Which product categories had the lowest average discount percent?
4. Which regions had the highest total sales for Technology?
5. What was the best selling product?

Analysis approach Q.1 Which products had the highest total sales revenue?

1. Menggunakan excel untuk menyaring kolom yang akan digunakan
2. Membuat pivot table untuk memudahkan menemukan jawaban dari pertanyaan tertentu
3. Kolom rows diisi dengan kolom category dan kolom values diisi kolom sales yang dihitung menjadi sum of sales
4. Setelah itu diurutkan nilainya dari terbesar ke terkecil dan didapatkan nilai penjualan tertinggi berdasarkan produk

Analysis approach Q.2

1. Menggunakan pivot table pada Q1, kemudian tinggal mengurutkan kolom Furniture dengan filter descending
2. Didapatkan customer-id dengan spent tertinggi untuk produk furniture