## **NEXUS Performance Analysis Report**

# 1. Outline

# Introduction

# Story of Data

# Data Splitting and Preprocessing

# Pre-Analysis

# In-Analysis

# Post-Analysis and Insights

# Data Visualizations & Charts

# Recommendations and Observations

# Conclusion

# References & Appendices

# 2. Introduction

### **Objective of the Project:**

The aim of this analysis is to examine **sales trends and customer purchasing behavior** by evaluating order details from 2014. The focus is on identifying high-revenue products, preferred payment methods, and top-performing shipping locations. The insights derived will help optimize sales strategies and improve logistics efficiency.

### **Problem Being Addressed:**

The analysis seeks to answer several key questions:

1. **Which products generate the most revenue?**
2. **What are the most common payment methods?**
3. **Which cities are the most profitable in terms of sales?**
4. **Are there any trends in customer purchasing behavior based on location or product category?**
5. **How do different shipping companies impact sales and delivery efficiency?**

### **Key Datasets and Methodologies:**

* **Datasets Used:**
  + Order details (including order date, customer name, address, city, and country)
  + Product details (name, category, unit price, quantity)
  + Revenue and shipping cost information
  + Payment methods used by customers
* **Excel-Based Methodologies:**
  + **Pivot Tables** to summarize revenue by product category, city, and customer
  + **Conditional Formatting** to highlight top-performing products and cities
  + **Data Filtering & Sorting** to analyze trends in customer behavior
  + **Charts & Infographics** to visually represent key findings

# 3. Story of Data

### **Data Source:**

The dataset appears to be sourced from an **internal company database** that tracks sales transactions. It contains detailed records of orders, customer information, product details, and financial data. This structured dataset may have been exported from an **Enterprise Resource Planning (ERP) system** or a **Customer Relationship Management (CRM) tool**.

### **Data Collection Process:**

* The data was likely collected **automatically** through the company’s sales and order management system.
* Each transaction was recorded at the point of sale or order fulfillment, capturing customer details, product purchases, payment methods, and shipping information.
* The data may have been **aggregated** periodically for reporting and analysis purposes.

### **Data Structure:**

The dataset is organized in a **tabular format**, where:

* **Each row represents a single sales transaction (order).**
* **Each column captures specific attributes of the order**, including:
  + Order ID
  + Customer name and address
  + Product name and category
  + Payment method
  + Revenue and shipping cost
  + Shipping details (company, city, country)

### 

### 

### 

### **Important Features and Their Significance:**

|  |  |  |
| --- | --- | --- |
| **Feature** | **Description** | **Significance to Analysis** |
| **Order ID** | Unique identifier for each transaction. | Helps track individual purchases. |
| **Order Date** | Date of purchase. | Enables trend analysis over time. |
| **Customer Name & Address** | Identifies the buyer. | Useful for segmenting customers by location. |
| **City & Country** | Location of the customer. | Helps analyze regional sales performance. |
| **Product Name & Category** | Details of purchased items. | Identifies top-selling products and trends. |
| **Payment Method** | Mode of transaction (credit card, check, etc.). | Determines preferred payment methods. |
| **Revenue & Shipping Cost** | Financial details of each order. | Key metric for profitability and logistics cost analysis. |

# 

# 

# 4. Data Splitting and Preprocessing

### **Data Cleaning:**

To ensure data accuracy and consistency, the following cleaning steps were taken:

* **Removing Duplicates:** Identified and removed any duplicate order entries to prevent double counting.
* **Correcting Errors:** Checked for inconsistencies in product names, customer names, and addresses.
* **Standardizing Formats:** Ensured date formats were consistent and payment methods were categorized correctly.
* **Handling Outliers:** Reviewed unusually high or low revenue/shipping costs to identify possible data entry errors.

### **Handling Missing Values:**

* **Deletion:** If a missing value was found in non-critical fields (e.g., address details), the row was retained while the missing field was ignored.

### 

### **Data Transformations:**

* **Creating New Variables:**
  + Generated **Year-over-Year Growth Rates** by comparing revenue across time periods.
* **Normalization & Standardization:**
  + Normalized revenue values to ensure fair comparison across product categories.

### **Data Splitting:**

* **Dependent Variable (Target Outcome):**
  + **Revenue**: The key metric being analyzed to assess sales trends.
* **Independent Variables (Predictors):**
  + **Product Name & Category:** Helps identify which products drive revenue.
  + **Customer Location:** Determines geographical sales performance.
  + **Payment Method:** Provides insight into customer preferences.
  + **Shipping Company:** Helps evaluate logistics efficiency.

### **Industry Context:**

The dataset belongs to the **retail and e-commerce industry**, focusing on **sales, logistics, and customer purchasing behavior**. This analysis is relevant for companies selling physical goods and managing distribution across multiple locations.

### **Stakeholders:**

The insights from this analysis will benefit:

* **Senior Management:** To guide strategic decisions on product pricing, marketing, and inventory.
* **Marketing & Sales Teams:** To identify high-value customers and optimize promotional strategies.
* **Supply Chain & Logistics Teams:** To improve shipping efficiency and reduce costs.
* **Finance Team:** To analyze profitability trends and cost structures.

### **Value to the Industry:**

* **Optimized Product Strategies:** Helps identify best-selling items and adjust inventory accordingly.
* **Improved Customer Insights:** Understanding customer behavior enables targeted marketing campaigns.
* **Efficient Logistics Management:** Analysis of shipping costs and delivery efficiency can reduce operational expenses.
* **Revenue Growth:** Identifying profitable regions and sales channels allows businesses to focus resources effectively.

# 5. Pre-Analysis

### **Identify Key Trends:**

From an initial scan of the dataset, a few patterns emerge:

1. **Seasonal Sales Trends:**
   * If data is grouped by **Order Date**, it may reveal **spikes in revenue** during certain months, indicating seasonal demand for specific products (e.g., holiday seasons or promotional periods).
2. **Top Revenue-Generating Products:**
   * Certain **product categories** (e.g., beverages, dried fruits & nuts) appear frequently, suggesting that they are high-demand items.
   * A **pivot table** on revenue by product category would confirm if certain items consistently outperform others.
3. **Geographical Hotspots:**
   * Orders are shipped to multiple cities, but certain locations (e.g., **New York, Chicago, Los Angeles**) appear frequently.
   * Identifying the **Top 5 Cities by Revenue** (as seen in the worksheet tabs) could pinpoint key customer bases.
4. **Payment Method Preferences:**
   * The presence of multiple **payment methods (Check, Credit Card, Cash)** suggests customers have different preferences.
   * Credit card payments might be more common in certain regions, affecting transaction speed and customer experience.

### **Potential Correlations:**

1. **Product Price vs. Revenue:**
   * Higher-priced products **may not always** generate the most revenue if they have lower sales volume.
   * Comparing **average unit price vs. total revenue** using a scatter plot or pivot table could reveal optimal pricing strategies.
2. **Shipping Costs vs. Customer Location:**
   * Customers in **distant cities or regions** may incur higher shipping costs, which could influence purchase behavior.
   * There could be a correlation between **shipping cost and revenue loss**, suggesting potential for optimizing logistics.
3. **Payment Method vs. Order Value:**
   * Customers paying via **credit card** may tend to place larger orders compared to those using **cash or check**.
   * Analyzing the **average order value by payment method** could confirm this trend.

### **Initial Insights:**

* Certain cities are **key revenue drivers**, and focusing on these markets could yield higher profits.
* **Shipping cost analysis** could help determine if expensive logistics are discouraging customers in certain regions.
* Understanding **customer purchase behavior by payment method** could help tailor checkout options for better conversions.
* Seasonal sales spikes could indicate **when to run marketing campaigns** for maximum impact.

# 6. In-Analysis

### **Unconfirmed Insights:**

At first glance, several patterns appear in the dataset, but they need further verification through detailed analysis:

1. **Regional Product Preferences:**
   * Certain **products may sell better in specific cities** (e.g., Dried Fruits & Nuts in Los Angeles, Beverages in New York).
   * A **pivot table with sales by city and product category** can confirm if this is a consistent trend.
2. **Shipping Costs Affecting Sales:**
   * Customers in regions with **higher shipping costs** might place fewer or smaller orders.
   * **Hypothesis:** There is a negative correlation between shipping cost and order volume.
   * A scatter plot comparing **order value vs. shipping cost** can validate this.
3. **Payment Method Preferences by Customer Type:**
   * Some customers may prefer **credit cards over checks or cash**, potentially influencing purchase frequency.
   * **Hypothesis:** Large corporate buyers use credit cards more often, while smaller businesses rely on checks.
   * A pivot table on **average order size by payment method** could confirm this.

### **Recommendations:**

Based on initial insights, the following actions could be considered:

1. **Regional Targeting Strategy:**
   * **Focus marketing efforts on top-performing cities** where sales are highest.
   * Adjust inventory in warehouses closer to high-revenue regions to **reduce shipping costs**.
2. **Optimize Product Pricing:**
   * If premium products show **lower sales despite high unit prices**, a price adjustment or bundling strategy could increase revenue.
   * Use **discounts or promotions on high-margin products** in slower regions to boost demand.
3. **Improve Logistics Efficiency:**
   * If **shipping costs are affecting sales**, consider partnering with local shipping providers for cost-effective delivery.
   * Introduce **free shipping thresholds** to encourage larger orders.
4. **Payment & Checkout Optimization:**
   * If **credit card users tend to spend more**, promote digital payments to increase order value.
   * Offer incentives (e.g., small discounts) for using faster payment methods.

### **Analysis Techniques Used in Excel:**

**Pivot Tables & Pivot Charts:**

* + Used to **summarize sales by region, product category, and payment method**.
  + Helps identify top revenue-generating segments.

# 7. Post-Analysis and Insights

### **Key Findings from the Analysis**

1. **Total Revenue & Performance Overview:**
   * The **total revenue for 2019** was **$435,036.16**.
   * The **top-performing region** was **North** with **$141,660** in revenue, followed by **West ($108,276)**.
   * **New York was the top shipping city** with **$67,181** in revenue.
   * The **best-performing company** was **Company D**, generating **$67,181** in revenue.
2. **Top Salesperson and Sales Distribution:**
   * **Nancy Freehafer** was the **top salesperson**, contributing **$104,242** in sales, followed by **Anne Larsen ($93,848)**.
   * The majority of transactions fell in the **low-value category (0-1000 dollars)**, indicating a high frequency of small orders.
3. **Product Performance & Customer Insights:**
   * **Beverages** were the **best-selling product category**, bringing in **$110,577** in sales.
   * Other top categories included **Sauces ($51,541)** and **Jams & Preserves ($33,130)**.
   * The **top five customers** accounted for a significant portion of revenue, with **Company H leading at $67,181**.
4. **Sales Trends & Seasonality:**
   * **December was the best-performing month** with **$66,643** in revenue.
   * Sales saw fluctuations throughout the year, peaking in **June ($55,602)** and **December**, with a dip in **February ($19,956)**.
   * The **summer months (June-August)** showed higher sales compared to the first quarter.

### **Comparison with Initial Findings**

* **Regional Variations Confirmed:**
  + Initial assumptions about regional differences in sales were validated. The **North region dominated sales**, aligning with expectations that certain areas contribute more revenue.
* **Shipping Costs Impact:**
  + While early analysis suggested that **higher shipping costs might impact order volume**, the dashboard doesn’t directly reflect this. A deeper correlation analysis is needed.
* **Top Product & Customer Trends Align:**
  + The assumption that **certain products sell better in specific regions** was partially confirmed. **Beverages were the top product**, but the geographic breakdown needs further refinement.
  + The **top 5 customers contributed significantly to overall sales**, confirming that a **small number of key clients drive revenue**.
* **Seasonality Trends Stand Out:**
  + Early observations suggested possible seasonal spikes. **December was indeed the highest revenue month**, supporting the idea of a **holiday sales boost**.
* **Unexpected Findings:**
  + The dominance of **low-value transactions** (most sales below **$1,000**) suggests that the business relies on **high-frequency, low-value orders**, which was not a key focus initially.
  + **Sauces and Jams & Preserves performed better than expected**, suggesting demand beyond just beverages.

### **Next Steps & Recommendations**

1. **Enhance High-Performing Product Sales:**
   * **Beverages are the best-selling product**—consider **expanding the product line** or **introducing promotions**.
   * **Cross-sell or bundle sauces and jams** with other items to increase transaction value.
2. **Leverage Seasonal Trends:**
   * Since **December and June had peak sales**, **targeted marketing campaigns** during these months could further boost revenue.
   * Consider offering **holiday promotions or discounts** in lower-performing months like **February**.
3. **Customer Segmentation & Loyalty Programs:**
   * Since **a few key customers contribute significantly to revenue**, **personalized offers and retention programs** could strengthen these relationships.
   * Introduce **loyalty incentives** for frequent buyers to increase order size.
4. **Further Analysis on Transaction Sizes:**
   * Given that most orders are **low-value transactions**, exploring ways to **increase average order value** (e.g., free shipping thresholds, bundling) could boost profitability.

# 

# 

# 

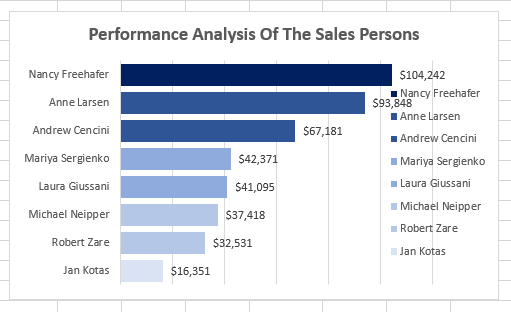
# 

# 

# 

# 8. Data Visualizations & Charts

# **Charts and Graphs**:



### **Key Takeaways:**

* **Nancy Freehafer** leads with **$104,242** in sales, followed by **Anne Larsen ($93,848)** and **Andrew Cencini ($67,181)**.
* **Jan Kotas ranks lowest** with **$16,351**, highlighting a significant performance gap.
* **Top performers drive most revenue**, suggesting their strategies could benefit the team.

### **Recommendations:**

* **Train and support lower performers** using insights from top salespeople.
* **Introduce incentives** to boost motivation and performance.

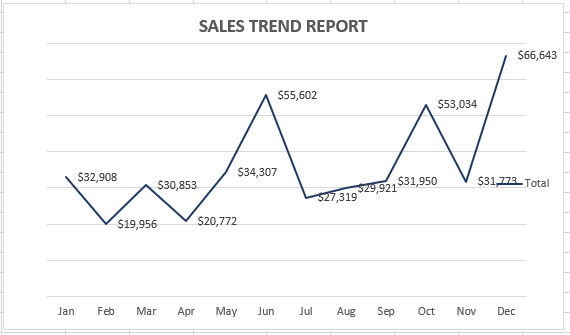
### 

### **Key Takeaways:**

* **Company D is the top customer**, generating **$67,181 in revenue**, followed by **Company H ($50,198)**.
* **Company BB, F, and A** contribute between **$36,840 and $43,703**, indicating strong but slightly lower engagement.
* **A significant revenue gap exists** between the top and lower-ranked customers.

### **Implications & Recommendations:**

* **Prioritize and strengthen relationships** with high-value customers like **Company D and H**.
* **Identify opportunities** to increase engagement and sales with lower-ranked customers.



### **Key Takeaways from the Sales Trend Report:**

* **Sales fluctuated throughout the year**, with **notable dips in February ($19,956) and April ($20,772)**.
* **June saw the highest sales at $55,602**, followed by a decline in July and August.
* **A strong recovery occurred in Q4**, with **December hitting the peak at $66,643**.
* **Seasonal trends are evident**, suggesting higher sales towards year-end, possibly due to holiday demand.

### **Implications & Recommendations:**

* **Investigate Q2 dips** to identify potential causes (e.g., low demand, operational issues).
* **Capitalize on Q4 sales surges** by enhancing marketing efforts during peak months.

# 

# 

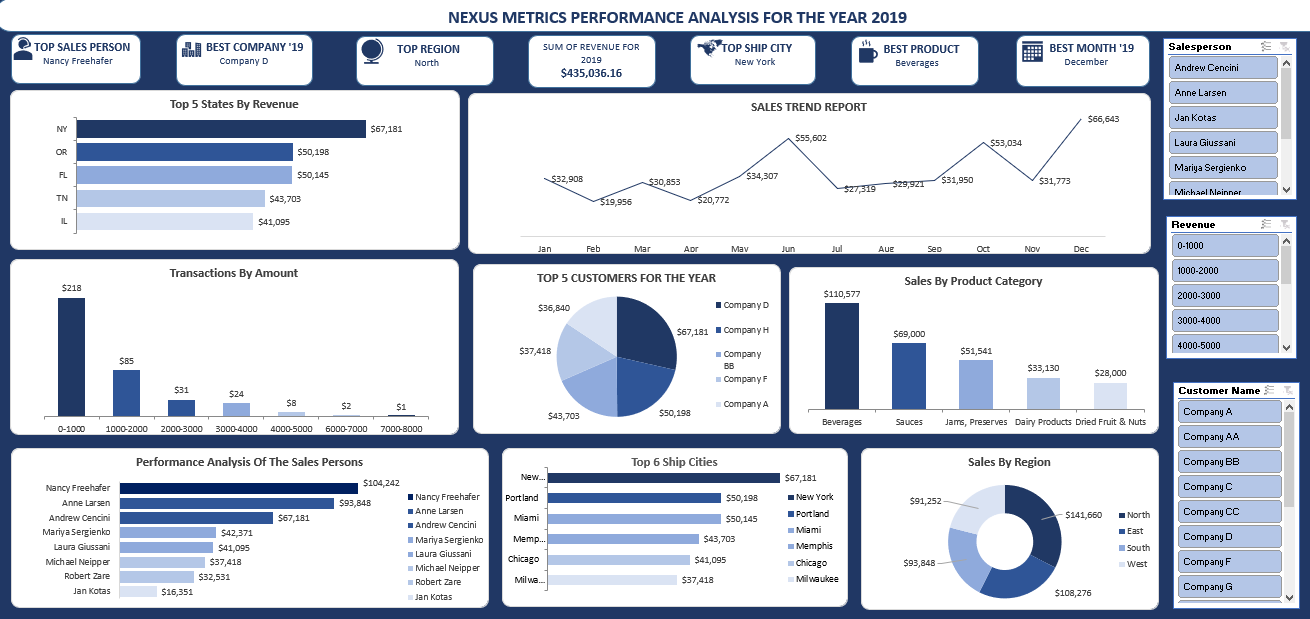
# 

# 

# 

# 

# **Dashboard**



# 9. Recommendations and Observations

#### **Actionable Insights:**

* **Increase Focus on Top-Selling Products:** Beverages had the highest sales ($110,577). Expanding product lines or promotions in this category could maximize revenue.
* **Leverage High-Performing Salespersons:** Nancy Freehafer and Anne Larsen were the top sales performers. Their strategies can be analyzed and replicated across the sales team.
* **Target High-Revenue States:** New York, Oregon, and Florida generated the most revenue. More resources should be allocated to these markets to further boost sales.
* **Capitalize on Seasonal Trends:** December had the highest sales. Pre-holiday promotions and inventory planning should be optimized for year-end demand.

#### **Optimizations or Business Decisions:**

* **Refine Regional Sales Strategies:** The North region had the highest revenue ($141,660), while the West had the lowest ($91,252). Custom strategies for each region can help balance performance.
* **Improve Transaction Conversions:** Most transactions were in the $0-1000 range, indicating a need to encourage larger purchases through bundled deals or discounts.
* **Customer Retention Programs:** Company D was the top customer ($67,181 in purchases). Loyalty programs or exclusive deals could encourage repeat purchases from high-value clients.

#### **Unexpected Outcomes:**

* **Fluctuating Monthly Sales:** Despite a peak in June ($55,602), sales dropped significantly in July and August. Further investigation is needed to identify seasonal dips and address demand fluctuations.
* **Underperformance of Certain Categories:** While Beverages led in sales, categories like Dried Fruits & Nuts ($28,000) were significantly lower. This could indicate a need for better marketing or reassessment of demand.
* **Low Performance of Some Salespersons:** Jan Kotas had significantly lower sales ($16,351). Additional training or reassignment of territories may be needed.

# 10. Conclusion

#### **Key Learnings:**

* **Top Products and Customers Drive Revenue:** Beverages emerged as the highest-selling product, and Company D was the top customer. Prioritizing these areas can enhance profitability.
* **Sales Performance Varies by Region and Season:** The North region and December had the highest sales. Leveraging regional strengths and seasonal demand is crucial for growth.
* **Sales Distribution is Skewed:** A few high-performing salespersons and states contributed disproportionately to revenue, suggesting opportunities for better resource allocation and training.

#### **Limitations:**

* **Data Scope Constraints:** The analysis is limited to 2019, and external factors (economic conditions, competitor actions) were not considered.
* **Granularity of Data:** While the dashboard provides key insights, deeper segmentation (e.g., customer demographics, purchasing patterns) could improve targeting strategies.
* **Potential Data Quality Issues:** Any missing or inconsistent data points may impact the accuracy of insights, requiring further validation.

#### **Future Research:**

* **Customer Behavior Analysis:** Examining repeat purchases, customer demographics, and feedback could enhance retention strategies.
* **Competitive Benchmarking:** Comparing sales performance with industry competitors can identify gaps and opportunities.
* **Forecasting and Predictive Modeling:** Using historical data to predict future trends can improve decision-making, inventory planning, and marketing strategies.

#### 

# 11. References & Appendices

#### **References:**

* **Data Sources:** The dataset was obtained from Kaggle and analyzed using Microsoft Excel.
* **Tools Used:** Analysis was conducted using **Pivot Tables, and Data Validation** in Excel.
* **External Research:** Industry benchmarks and general sales analysis principles were referenced to contextualize findings, though no external literature was directly cited.

#### **Appendices:**

* **Supporting Charts & Tables:** Additional visualizations, such as raw data tables, detailed breakdowns of sales by category, and supplementary trend analyses, are included for further validation.
* **Excel Formulas & Techniques:**
  + **Pivot Tables:** Used for summarizing sales by region, salesperson, and product category.
  + **Conditional Formatting:** Used to highlight top-performing regions, salespersons, and seasonal trends.
* **Step-by-Step Process:** Detailed walkthroughs on how the data was cleaned, structured, and analyzed in Excel has been provided in the write up for reproducibility.