# STATISTICAL ANALYSIS REPORT

## 1 Introduction

This project performs a comprehensive statistical business analysis on a retail sales dataset. The objective is to analyze sales performance, examine relationships between key variables such as price and quantity, conduct hypothesis testing, and derive meaningful business insights using statistical methods.

The study applies descriptive statistics, confidence intervals, correlation analysis, and regression modeling to evaluate business performance and support data-driven decision making.

## 2 Dataset Description

The dataset contains transactional sales data with the following variables:

- **Date** – Transaction date
- **Product** – Product category (Phone, Laptop, Headphones)
- **Quantity** – Number of units sold
- **Price** – Price per unit
- **Customer_ID** – Unique customer identifier
- **Region** – Sales region (East, West, North, etc.)
- **Total_Sales** – Total revenue generated (Quantity × Price)

The dataset represents daily retail sales activity and allows analysis of revenue behavior with respect to quantity and pricing.

## 3 Descriptive Statistics

Descriptive analysis was conducted on numerical variables including Quantity, Price, and Total Sales.

Key metrics calculated:

- Mean
- Median
- Standard Deviation
- Variance
- Minimum & Maximum values
- Quartiles

**Key Findings:**

- The **average total sales** indicate the central revenue trend.
- The **standard deviation** shows moderate variability in daily sales.
- Sales distribution shows variability depending on product type and quantity purchased.
- Higher quantities significantly increase total revenue.

Descriptive statistics provide a foundational understanding of sales performance and variability.

# 4 Confidence Interval (95%)

A 95% confidence interval was calculated for the mean Total Sales.

Interpretation:

We are 95% confident that the true average sales value lies within the computed interval range.

This provides statistical reliability and reduces uncertainty in estimating true population sales performance.

# 5 Hypothesis Testing Results

Three hypothesis tests were conducted:

### ◆ Test 1: One-Sample T-Test (Mean Sales)

$H_0$: Average Total Sales = Expected benchmark value
$H_1$: Average Total Sales $\neq$ Benchmark

Result:

- p-value $< 0.05$
- Null hypothesis rejected

Conclusion:
Average sales significantly differ from the benchmark value.

◆ **Test 2: Independent T-Test (High vs Low Quantity Sales)**

$H_0$: No difference in mean sales between high and low quantity groups
$H_1$: Sales differ between groups

Result:

- p-value $< 0.05$
- Null hypothesis rejected

Conclusion:
Higher quantity transactions significantly generate more revenue.

◆ **Test 3: Pearson Correlation Test**

$H_0$: No relationship between Price and Total Sales
$H_1$: Significant relationship exists

Result:

- Strong positive correlation
- p-value $< 0.05$

Conclusion:
Price significantly affects total revenue.

# 6 Correlation Analysis

Correlation analysis was performed between:

- Quantity and Total Sales
- Price and Total Sales

Findings:

- **Strong positive correlation between Quantity and Total Sales**
- **Moderate to strong positive correlation between Price and Total Sales**

This indicates that both selling more units and pricing strategy influence overall revenue.

# 7 Regression Analysis

A linear regression model was developed with:

Dependent Variable:
Total Sales

Independent Variable:
Quantity (or Price)

Model Results:

- $R^2$ value indicates strong explanatory power.
- Regression coefficients are statistically significant ($p < 0.05$).
- Positive slope shows increasing quantity leads to increasing sales revenue.

The regression confirms that quantity sold is a major driver of revenue.

# 8 Business Insights

From the statistical analysis:

1. Sales revenue is highly dependent on quantity sold.
2. Pricing strategy has significant impact on total revenue.
3. Increasing unit sales will substantially increase business revenue.
4. Variability in sales suggests opportunity for demand forecasting.
5. Focused marketing strategies could enhance high-performing regions/products.

Strategic Recommendation:

- Increase promotion of high-demand products.
- Optimize pricing for maximum revenue.
- Implement region-based sales targeting.

# 9 Conclusion

This statistical business analysis demonstrates that total sales are significantly influenced by both quantity sold and pricing strategy.

The hypothesis tests confirmed statistically significant relationships between key variables. Correlation and regression analysis showed strong positive associations, validating the impact of sales drivers.

Overall, statistical methods provide clear evidence-based insights for improving business performance and revenue planning.

Data-driven decision making can enhance profitability and operational efficiency.