**Sales Forecasting and Analysis Report**

**Project Title**: Predicting Total Sales using Regression Modeling

Objective:

The primary goal of this report is to explore, clean, and analyze food sales data to build a regression model that accurately predicts total sales. This model supports business decisions, sales planning, and performance evaluation across different product categories, regions, and cities.

Dataset Summary:

* Source: Sample Data Food Sales
* Key Features: Region, City, Category, Product, Quantity, UnitPrice, OrderDate
* Engineered Feature: Total Sales = Quantity × Unit Price

Feature Engineering:

* Imported and explored the dataset with Pandas.
* Removed duplicates and fixed missing or inconsistent entries.
* I Created Total Sales as target variable
* Extracted Month, Weekday, and Day from OrderDate
* I used One-hot encoded categorical variables for modeling

Exploratory Analysis:

* I identified top selling products and cities
* Highlighted high performing regions and monthly sales patterns
* Visualized actual vs predicted trends over time

Regression Model:

* Model Used: Linear Regression (Scikit-learn)
* Inputs: Encoded Region, City, Category, Product, Quantity, UnitPrice
* Target: Total Sales
* Downloaded both actual and predicted sales for visualization.

Model Evaluation:

I evaluated the model using the following

* R² Score: which resulted in 0.91
* MAE(Mean Absolute error
* RMSE(Root mean square error)
* Created dataframe of features and coefficients and visualize top 10 features most important in predicting sales

**Power BI Dashboard:**

* KPI cards for Revenue, Total orders, Sales error, Best selling Region, MAE, RMSE, and R² etc.
* Absolute Residual by Product and Region with a bar chart
* Revenue by Region: A map visual displaying Region by Revenue with and a drill down button to help view Revenue generated by the different cities in both Regions.
* Revenue by City: A donut chart comparing Revenue generated across the 4 cities ( Los Angeles, Boston, San Diego and New York.
* Predicted vs Actual Sales: Line chart showing how actual sales align with the regression model’s forecast.
* Category-wise Performance: Pie chart tracking category performance by Revenue.
* Top 5 Products: Bar charts highlighting best selling Products.
* Monthly Sales Trends: Line chart visualizing sales trends , on a monthly basis across the years.

**Key Insights:**

* Top cities: Boston, New York (Having 39.81% And 24.78% of sales).
* Highest-selling product: Carrot.
* Best selling Region: East
* MAE(Mean Absolute error): ₦27.35
* RMSE(Root mean square error): ₦14.11
* Revenue: ₦33.3K
* Average Order: 63
* Total Order : ₦15.4K
* Regression model R² Score: 0.91

**Visual Highlights:**

* Category-wise performance showed Cookies leading over all by 51%.
* Predicted vs Actual Sales matched closely for most high-volume products.

**Recommendations:**

* Focus marketing efforts on top 3 cities(Boston, New York, Los Angeles)
* Increase inventory for best-selling products.
* Increase inventory for best-performing category.

**Conclusion:**

The regression model demonstrates reliable prediction of total sales and highlights important trends and performance patterns. This enables stakeholders to make data-driven decisions on product focus, regional strategy, and inventory planning.