

## TASK NO 5

**Digital Empowerment Network**  
**TASK submitted by:MubashirAbbas**

# Data Analysis Project: Sales Overview Using Python

## Objective:

Analyze sales data from a company's various product categories and regions. Create visual representations to understand sales distribution and gain insights using Python's Matplotlib and Seaborn libraries.

---

## Step 1: Import Necessary Libraries

We use **Pandas** to handle the data, and **Matplotlib** and **Seaborn** for visualizing it.

```
python
Copy code
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

---

## Step 2: Create a Simple Dataset

For this project, I'm using a sample dataset that covers sales of different products across various regions. This is a small dataset, so it's easier to understand and visualize.

```
python
Copy code
data = {
    'Product Category': ['Electronics', 'Furniture', 'Clothing', 'Toys',
    'Groceries'],
    'Units Sold': [150, 90, 200, 85, 250],
    'Sales (in $1000)': [120, 80, 150, 70, 200],
    'Region': ['North', 'South', 'East', 'West', 'Central']
}
df = pd.DataFrame(data)
```

---

## Step 3: Data Visualizations

Now, we create visualizations that summarize the data. We will use:

- **Bar Plot** to compare sales in different product categories.
- **Pie Chart** to show the proportion of sales in each region.
- **Dot Plot** to visualize units sold in each region.

### 1. Bar Plot: Sales by Product Category

```
python
```

```
plt.figure(figsize=(8, 6))
sns.barplot(x='Product Category', y='Sales (in $1000)', data=df,
palette='viridis')
plt.title('Sales by Product Category')
plt.xlabel('Product Category')
plt.ylabel('Sales (in $1000)')
plt.show()
```

This bar plot shows how much revenue each product category generated. You can easily see that **Groceries** bring in the highest sales, while **Toys** have the least.

---

## 2. Pie Chart: Proportion of Sales by Region

```
plt.figure(figsize=(6, 6))
plt.pie(df['Sales (in $1000)'], labels=df['Region'], autopct='%1.1f%%',
startangle=140, colors=sns.color_palette('Set2'))
plt.title('Sales Proportion by Region')
plt.show()
```

The pie chart shows the sales contribution from each region. **Central** is the top-performing region, while **West** has the lowest sales.

---

## 3. Dot Plot: Units Sold per Region

```
plt.figure(figsize=(8, 6))
sns.stripplot(x='Region', y='Units Sold', data=df, jitter=True, color='red',
size=10)
plt.title('Units Sold per Region')
plt.xlabel('Region')
plt.ylabel('Units Sold')
plt.show()
```

**Explanation:** The dot plot shows how many units were sold in each region. **Central** and **East** regions sold the most units, especially for **Groceries**.

---

## Step 4: Key Insights

Based on the visualizations, here are the important findings:

1. **Sales by Category:**
  - **Groceries** generate the highest sales, followed by **Clothing**.
  - **Toys** is the lowest-selling category, which may need some attention.
2. **Sales by Region:**
  - The **Central** region is the leader in sales, contributing the most to overall revenue.

- **West** region lags behind, contributing the least.
3. **Units Sold:**
- **Central** and **East** regions have the highest number of units sold, showing strong demand in these areas.
  - **Toys** have the fewest units sold, and efforts could be made to boost sales in this category.