Data Analysis Practice Datasets

Sample Data for Analysis

Sales Data

Month, Product, Region, Sales, Expenses
January, Product_A, North, 50000, 20000
January, Product_B, North, 45000, 18000
January, Product_A, South, 55000, 22000
January, Product_B, South, 48000, 19000
February, Product_A, North, 52000, 21000
February, Product_B, North, 46000, 18500
February, Product_A, South, 58000, 23000
February, Product_B, South, 50000, 20000
March, Product_A, North, 51000, 20500
March, Product_B, North, 47000, 18800
March, Product_A, South, 56000, 22500
March, Product_B, South, 49000, 19500

Customer Data

CustomerID, Age, Gender, Income, SpendingScore
C001, 35, Male, 50000, 85
C002, 28, Female, 45000, 75
C003, 42, Male, 75000, 90
C004, 31, Female, 55000, 60
C005, 29, Male, 48000, 70
C006, 45, Female, 82000, 95
C007, 33, Male, 52000, 65
C008, 39, Female, 68000, 80
C009, 27, Male, 42000, 55
C010, 36, Female, 58000, 75

Analysis Template Code

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
# Load data
def load_data(file_path):
    return pd.read_csv(file_path)
# Basic statistics
def basic_stats(df):
    print("Dataset Shape:", df.shape)
    print("Basic Statistics:")
    print(df.describe())
    print("Missing Values:")
    print(df.isnull().sum())
# Create visualizations
def create_visualizations(df):
```