IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/Python/Python38/program 16.py =
Error: File 'customer_reviews.csv' not found.
>>>
```

program 16.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 16.py (3.8.10)

File Edit Format Run Options Window Help

```
from collections import Counter
import re

# Load the dataset (add encoding for compatibility)
try:
    data = pd.read_csv('customer_reviews.csv', encoding='utf-8')
except FileNotFoundError:
    print("Error: File 'customer_reviews.csv' not found.")
    exit()

# Ensure 'review' column exists
if 'review' not in data.columns:
    print("Error: 'review' column not found in the dataset.")
    exit()

# Drop missing reviews
data = data.dropna(subset=['review'])

# Combine all reviews into a single string
all_reviews = ' '.join(data['review'].astype(str))

# Clean and split the reviews into words
words = re.findall(r'\w+', all_reviews.lower())

# Calculate the frequency distribution of words
word_freq = Counter(words)

# Convert to DataFrame for better visualization
word_freq_df = pd.DataFrame(word_freq.items(), columns=['Word', 'Frequency'])

# Sort the DataFrame by frequency
word_freq_df = word_freq_df.sort_values(by='Frequency', ascending=False)

# Display the top 10 most common words
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()"
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/P
Error: File 'customer_feedback.csv' not found.
>>>
```

program 17.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 17.py (3.8.10)

File Edit Format Run Options Window Help

```
except FileNotFoundError:
    print("Error: File 'customer_feedback.csv' not found.")
    exit()

# Ensure 'Rating' column exists
if 'Rating' not in data.columns:
    print("Error: 'Rating' column not found in the dataset.")
    exit()

# Drop rows with missing or NaN ratings
data = data.dropna(subset=['Rating'])

# Display the first few rows of the data
print(data.head())

# Analyze the feedback ratings
feedback_counts = data['Rating'].value_counts()

# Plot the feedback ratings
plt.figure(figsize=(10, 6))
feedback_counts.plot(kind='bar', color='skyblue')
plt.title('Customer Feedback Ratings')
plt.xlabel('Ratings')
plt.ylabel('Number of Feedbacks')
plt.xticks(rotation=0)
plt.grid(axis='y')
plt.show()

# Calculate the average rating
average_rating = data['Rating'].mean()
print(f'Average Customer Rating: {average_rating:.2f}')

# Save the analysis results to a new CSV file
analysis_results = pd.DataFrame({
    'Rating': feedback_counts.index,
    'Count': feedback_counts.values
})
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/Python/Python38/program 15.py =
Likes: 10, Frequency: 4
Likes: 20, Frequency: 3
Likes: 30, Frequency: 2
Likes: 40, Frequency: 1
>>>
```

program 15.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 15.py (3.8.10)

File Edit Format Run Options Window Help

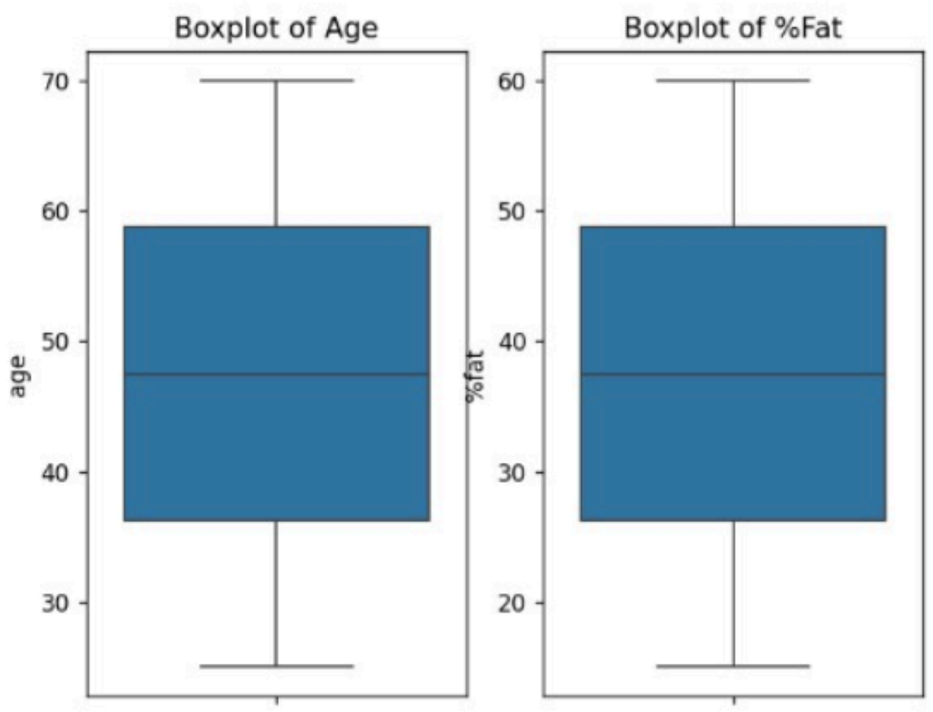
```
from collections import Counter

# Sample data: list of likes for each post
likes = [10, 20, 10, 30, 20, 10, 40, 30, 20, 10]

# Calculate frequency distribution
frequency_distribution = Counter(likes)

# Display the frequency distribution
for like, frequency in frequency_distribution.items():
    print(f'Likes: {like}, Frequency: {frequency}')
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [
Type "help", "copyright", "credits" or "license()" for more i
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/Python/Pytho
Mean Age: 47.5, Median Age: 47.5, Std Age: 15.138251770487457
Mean %Fat: 37.5, Median %Fat: 37.5, Std %Fat: 15.138251770487
```

Figure 1



program 18.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 18.py (3.8.10)

File Edit Format Run Options Window Help

```
# Calculate mean, median and standard deviation
mean_age = df['age'].mean()
median_age = df['age'].median()
std_age = df['age'].std()

mean_fat = df['%fat'].mean()
median_fat = df['%fat'].median()
std_fat = df['%fat'].std()

print(f'Mean Age: {mean_age}, Median Age: {median_age}, Std Age: {std_age}'
print(f'Mean %Fat: {mean_fat}, Median %Fat: {median_fat}, Std %Fat: {std_fa

# Boxplots
plt.figure(figsize=(12, 5))

plt.subplot(1, 2, 1)
sns.boxplot(y=df['age'])
plt.title('Boxplot of Age')

plt.subplot(1, 2, 2)
sns.boxplot(y=df['%fat'])
plt.title('Boxplot of %Fat')

plt.show()

# Scatter Plot
plt.figure(figsize=(8, 6))
plt.scatter(df['age'], df['%fat'])
plt.title('Scatter Plot of Age vs %Fat')
plt.xlabel('Age')
plt.ylabel('%Fat')
plt.show()

# Q-Q Plot
plt.figure(figsize=(8, 6))
stats.probplot(df['%fat'], dist="norm", plot=plt)
plt.title('Q-Q Plot of %Fat')
plt.show()
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/Python/Python38/program 20.py =
Error: 'customer_data.csv' not found.
>>>
```

program 20.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 20.py (3.8.10)

File Edit Format Run Options Window Help

```
import pandas as pd

# Load the customer data
try:
    df = pd.read_csv("customer_data.csv", encoding='utf-8')
except FileNotFoundError:
    print("Error: 'customer_data.csv' not found.")
    exit()

# Ensure required columns exist
required_columns = ['Total Spending', 'Age']
if not all(col in df.columns for col in required_columns):
    print(f"Error: Required columns missing. Found columns: {list(df.columns)}")
    exit()

# Segment customers based on total spending
def segment_spending(row):
    if row['Total Spending'] > 1000:
        return 'High Spenders'
    elif row['Total Spending'] > 500:
        return 'Medium Spenders'
    else:
        return 'Low Spenders'

df['Spending Segment'] = df.apply(segment_spending, axis=1)

# Calculate the average age of customers in each spending segment
average_age = df.groupby('Spending Segment')['Age'].mean()

print("Average Age per Spending Segment:\n")
print(average_age)
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs
Error: File 'sales_data.csv' not found.
>>>
```

program 19.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 19.py (3.8.10)

File Edit Format Run Options Window Help

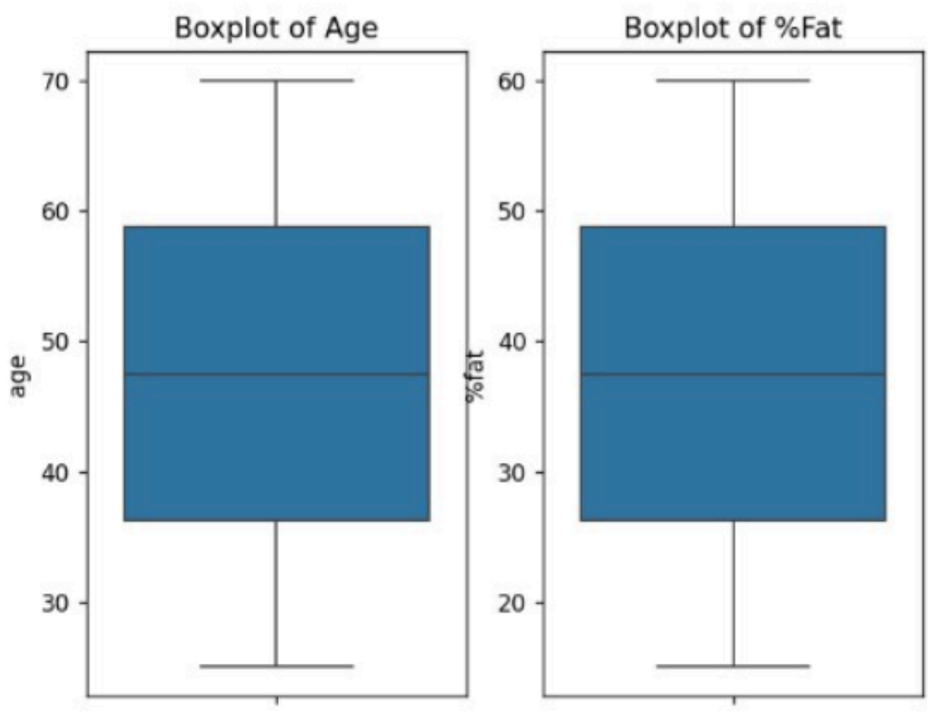
```
# Sales and Profit Analysis

import pandas as pd

# Load the sales data safely
try:
    df = pd.read_csv('sales_data.csv', encoding='utf-8')
except FileNotFoundError:
    print("Error: File 'sales_data.csv' not found.")
    exit()

# Check for necessary columns
required_columns = ['Product', 'Quantity Sold', 'Unit Price']
if not all(col in df.columns for col in required_columns):
    print(f"Error: Required columns missing. Found columns: {list(df.columns)}")
    exit()

# Create a new column for Total Sales
df['Total Sales'] = df['Quantity Sold'] * df['Unit Price']

# Calculate total sales for each product
total_sales_per_product = df.groupby('Product')['Total Sales'].sum()

# Calculate overall profit considering a 20% profit margin
profit_margin = 0.20
overall_profit = total_sales_per_product * profit_margin

# Display the top 5 most profitable products
top_profitable_products = overall_profit.sort_values(ascending=False).head(5)

print("Top 5 Most Profitable Products:")
print(top_profitable_products)
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [
Type "help", "copyright", "credits" or "license()" for more i
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/Python/Pytho
Mean Age: 47.5, Median Age: 47.5, Std Age: 15.138251770487457
Mean %Fat: 37.5, Median %Fat: 37.5, Std %Fat: 15.138251770487
```

Figure 1



program 18.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 18.py (3.8.10)

File Edit Format Run Options Window Help

```
# Calculate mean, median and standard deviation
mean_age = df['age'].mean()
median_age = df['age'].median()
std_age = df['age'].std()

mean_fat = df['%fat'].mean()
median_fat = df['%fat'].median()
std_fat = df['%fat'].std()

print(f'Mean Age: {mean_age}, Median Age: {median_age}, Std Age: {std_age}'
print(f'Mean %Fat: {mean_fat}, Median %Fat: {median_fat}, Std %Fat: {std_fa

# Boxplots
plt.figure(figsize=(12, 5))

plt.subplot(1, 2, 1)
sns.boxplot(y=df['age'])
plt.title('Boxplot of Age')

plt.subplot(1, 2, 2)
sns.boxplot(y=df['%fat'])
plt.title('Boxplot of %Fat')

plt.show()

# Scatter Plot
plt.figure(figsize=(8, 6))
plt.scatter(df['age'], df['%fat'])
plt.title('Scatter Plot of Age vs %Fat')
plt.xlabel('Age')
plt.ylabel('%Fat')
plt.show()

# Q-Q Plot
plt.figure(figsize=(8, 6))
stats.probplot(df['%fat'], dist="norm", plot=plt)
plt.title('Q-Q Plot of %Fat')
plt.show()
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs
Error: File 'sales_data.csv' not found.
>>>
```

program 19.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 19.py (3.8.10)

File Edit Format Run Options Window Help

```
# Sales and Profit Analysis

import pandas as pd

# Load the sales data safely
try:
    df = pd.read_csv('sales_data.csv', encoding='utf-8')
except FileNotFoundError:
    print("Error: File 'sales_data.csv' not found.")
    exit()

# Check for necessary columns
required_columns = ['Product', 'Quantity Sold', 'Unit Price']
if not all(col in df.columns for col in required_columns):
    print(f"Error: Required columns missing. Found columns: {list(df.columns)}")
    exit()

# Create a new column for Total Sales
df['Total Sales'] = df['Quantity Sold'] * df['Unit Price']

# Calculate total sales for each product
total_sales_per_product = df.groupby('Product')['Total Sales'].sum()

# Calculate overall profit considering a 20% profit margin
profit_margin = 0.20
overall_profit = total_sales_per_product * profit_margin

# Display the top 5 most profitable products
top_profitable_products = overall_profit.sort_values(ascending=False).head(5)

print("Top 5 Most Profitable Products:")
print(top_profitable_products)
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/Python/Python38/program 16.py =
Error: File 'customer_reviews.csv' not found.
>>>
```

program 16.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 16.py (3.8.10)

File Edit Format Run Options Window Help

```
from collections import Counter
import re

# Load the dataset (add encoding for compatibility)
try:
    data = pd.read_csv('customer_reviews.csv', encoding='utf-8')
except FileNotFoundError:
    print("Error: File 'customer_reviews.csv' not found.")
    exit()

# Ensure 'review' column exists
if 'review' not in data.columns:
    print("Error: 'review' column not found in the dataset.")
    exit()

# Drop missing reviews
data = data.dropna(subset=['review'])

# Combine all reviews into a single string
all_reviews = ' '.join(data['review'].astype(str))

# Clean and split the reviews into words
words = re.findall(r'\w+', all_reviews.lower())

# Calculate the frequency distribution of words
word_freq = Counter(words)

# Convert to DataFrame for better visualization
word_freq_df = pd.DataFrame(word_freq.items(), columns=['Word', 'Frequency'])

# Sort the DataFrame by frequency
word_freq_df = word_freq_df.sort_values(by='Frequency', ascending=False)

# Display the top 10 most common words
```

IDLE Shell 3.8.10

File Edit Shell Debug Options Window Help

```
Python 3.8.10 (tags/v3.8.10:3d8993a, May  3 2021, 11:48:03) [MSC v.1928 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()"
>>>
= RESTART: C:/Users/suman/AppData/Local/Programs/P
Error: File 'customer_feedback.csv' not found.
>>>
```

program 17.py - C:/Users/suman/AppData/Local/Programs/Python/Python38/program 17.py (3.8.10)

File Edit Format Run Options Window Help

```
except FileNotFoundError:
    print("Error: File 'customer_feedback.csv' not found.")
    exit()

# Ensure 'Rating' column exists
if 'Rating' not in data.columns:
    print("Error: 'Rating' column not found in the dataset.")
    exit()

# Drop rows with missing or NaN ratings
data = data.dropna(subset=['Rating'])

# Display the first few rows of the data
print(data.head())

# Analyze the feedback ratings
feedback_counts = data['Rating'].value_counts()

# Plot the feedback ratings
plt.figure(figsize=(10, 6))
feedback_counts.plot(kind='bar', color='skyblue')
plt.title('Customer Feedback Ratings')
plt.xlabel('Ratings')
plt.ylabel('Number of Feedbacks')
plt.xticks(rotation=0)
plt.grid(axis='y')
plt.show()

# Calculate the average rating
average_rating = data['Rating'].mean()
print(f'Average Customer Rating: {average_rating:.2f}')

# Save the analysis results to a new CSV file
analysis_results = pd.DataFrame({
    'Rating': feedback_counts.index,
    'Count': feedback_counts.values
})
```