Worksheet: Calories Burned vs Walking & Running

Code with Plot:

```
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Generate random data for walking and running durations
walking = np.random.uniform(10, 60, 50)  # 50 values between 10 and 60 minutes
running = np.random.uniform(5, 30, 50)  # 50 values between 5 and 30 minutes

# Calculate calories burned based on walking and running
calories = 4 * walking + 10 * running + np.random.normal(0, 20, 50)

# Combine walking and running into a 2D array for regression
X3 = np.column_stack((walking, running))
y3 = calories

# Fit a linear regression model
model3 = LinearRegression().fit(X3, y3)

# Plot the data
plt.scatter(walking, y3, color='orange', label='Walking')
plt.scatter(running, y3, color='brown', label='Running')
plt.xlabel('Minutes')
plt.ylabel('Calories')
plt.title('Calories Burned vs Walking & Running')
plt.legend()
plt.show()
```

Explanation:

This code estimates calories burned based on walking and running durations.
We generate synthetic data for walking and running using `np.random.uniform`.
Calories are calculated using a weighted formula, with noise added to simulate real-world variation.
`np.column_stack` combines the features for regression.
We train a linear regression model to learn the relationship.
The plot visualizes how calories burned vary with walking and running times.

[Space for answers]

Questions:

1. Which activity burns more calories per minute?
2. Why do we use different weights for walking and running?
3. What could be other factors affecting calories burned?