**Notes for Q1.**

Subplots is used to analyse how different workout types affect calorie burn and heart rate based on weather and location.

**Key Steps in the Code**

1. **Importing Required Libraries:**
   * Pandas for data manipulation.
   * Matplotlib for data visualization.
   * NumPy for numerical operations.
2. **Reading the Dataset:**
   * pd.read\_csv() reads the file. The FileNotFoundError exception is handled appropriately.
3. **Data Preprocessing:**
   * The date column is converted to datetime format.
   * Numeric columns like sleep\_hours, active\_minutes, steps, etc., are coerced into numeric values.
   * dropna() ensures missing values are removed.
4. **Analysis Logic:**
   * **Grouped Analysis:** Groups data by workout\_type for each combination of weather and location.
   * **Visualization:**
     + Subplots are used for visualizing data per weather-location combination.
     + Bar plots compare average calories\_burned and heart\_rate\_avg.
5. **Overall Plot:**
   * Aggregates and visualizes the overall effects of workout types on calorie burn and heart rate.

Improved Version of the Code (Snippet)

|  |
| --- |
| # Check date format (adapted to match dataset format)  try:  df['date'] = pd.to\_datetime(df['date'], format='%Y-%m-%d', errors='coerce') # Adjust format if necessary  except ValueError:  print("Error with 'date' format. Adjust as needed.")  exit()  # Improved visualization (dynamic subplot handling)  fig, axes = plt.subplots(len(unique\_weather), len(unique\_locations), figsize=(15, 5 \* len(unique\_weather)))  axes = axes.flatten() # Flatten axes for simpler iteration  for idx, (weather, location) in enumerate([(w, l) for w in unique\_weather for l in unique\_locations]):  df\_filtered = df[(df['weather\_conditions'] == weather) & (df['location'] == location)]    if not df\_filtered.empty:  df\_grouped = df\_filtered.groupby('workout\_type')[['calories\_burned', 'heart\_rate\_avg']].mean()    x = np.arange(len(df\_grouped))  axes[idx].bar(x - width/2, df\_grouped['calories\_burned'], width, label='Calories Burned')  axes[idx].bar(x + width/2, df\_grouped['heart\_rate\_avg'], width, label='Heart Rate')    axes[idx].set\_xticks(x)  axes[idx].set\_xticklabels(df\_grouped.index, rotation=45)  axes[idx].set\_title(f"{weather}, {location}")  axes[idx].legend()  else:  axes[idx].axis('off') # Hide empty plots  plt.tight\_layout()  plt.show() |

 **Introduction:**

* Explain the goal: Understanding the effects of workout types on calorie burn and heart rate under varying weather and location conditions.
* Highlight the dataset's key features and variables.

 **Preprocessing Steps:**

* Describe the cleaning process (date conversion, numeric coercion, handling missing data).
* Emphasize error-handling strategies.

 **Analysis Approach:**

* Grouped data by weather, location, and workout type.
* Calculated average calories\_burned and heart\_rate\_avg.

 **Visualization:**

* Discuss the use of subplots to show granular insights per weather-location combination.
* Explain the overall summary plot for aggregated insights.

 **Key Insights:**

* Share patterns observed in the data (e.g., high calorie burn for running in clear weather at parks).

 **Improvements:**

* Suggest additional analyses, such as incorporating user demographics.