

Test -2

Please solve the below question accordingly using Pandas

- A. Imagine you work for an e-commerce company that has collected customer reviews for its products. The data is stored in a list of dictionaries, where each dictionary represents a customer review with the following structure

```
customer_reviews = [  
    {'product_id': 101, 'review': 'This product is amazing! I love it.'},  
    {'product_id': 102, 'review': 'The quality is not as expected. Disappointed.'},  
    {'product_id': 103, 'review': 'Great value for the price. Highly recommended.'},  
    {'product_id': 104, 'review': 'Not happy with the purchase. Will return it.'},  
    {'product_id': 105, 'review': 'Excellent service. Fast delivery.'}  
]
```

1. Identify and count the number of reviews that express a positive sentiment.
2. Identify and list the product IDs and reviews for the products with negative sentiments.
3. Calculate the average length (number of words) of the reviews
4. Determine the product that is mentioned the most in the reviews.
5. Create a distribution of sentiments (positive, negative, neutral) for the reviews.
6. Extract key words or phrases from the reviews that frequently appear.
7. Find and display the longest review along with its product ID.
8. Identify products that have both positive and negative reviews.
9. Change the sentiment of a selected review from positive to negative and vice versa.
10. Compare the similarity between two selected reviews without using any similarity metrics directly.

- B. Let's consider a scenario related to real estate data. Suppose you have a dataset containing information about real estate properties, and the data is stored in a list of dictionaries. Each dictionary represents a property with details such as the property type, size, location, number of bedrooms, and price. Here's the scenario and some questions:

```
real_estate_data = [  
    {'property_id': 1, 'property_type': 'Apartment', 'size_sqft': 1200, 'location':  
    'Downtown', 'bedrooms': 2, 'price_usd': 250000},  
  
    {'property_id': 2, 'property_type': 'House', 'size_sqft': 2000, 'location': 'Suburb',  
    'bedrooms': 3, 'price_usd': 350000},  
  
    {'property_id': 3, 'property_type': 'Apartment', 'size_sqft': 800, 'location': 'Uptown',  
    'bedrooms': 1, 'price_usd': 150000},  
  
    {'property_id': 4, 'property_type': 'Condo', 'size_sqft': 1500, 'location': 'Downtown',  
    'bedrooms': 2, 'price_usd': 300000},  
  
    {'property_id': 5, 'property_type': 'House', 'size_sqft': 1800, 'location': 'Suburb',  
    'bedrooms': 4, 'price_usd': 400000}  
]
```

Question:

1. Calculate the average size of properties in the dataset
2. Identify and list properties located in the downtown area.
3. Find properties with a price higher than \$300,000.
4. Determine the distribution of property types in the dataset.
5. Identify and list apartments with a price less than \$200,000.
6. Calculate the average price per square foot for all properties.
7. Identify and list houses with a size greater than 1,800 sqft.
8. Determine the distribution of the number of bedrooms in the dataset.
9. Find properties with 3 bedrooms and a price less than \$300,000.
10. Categorize properties into size ranges (e.g., Small, Medium, Large) based on their square footage.