Test 2 (WQD7005 Data Mining)

1. Write a Pandas program to select the 'name' and 'score' columns from the following DataFrame and order by score (highest to lowest).  
     
   Sample DataFrame: (i)  
   exam\_data = {'name': ['Ali', 'Abu', 'Katherine', 'Site', 'Emily', 'Michael', 'Matthew', 'Laura', 'Kevin', 'James'],  
   'score': [12.5, 9, 16.5, np.nan, 9, 20, 14.5, np.nan, 8, 19],  
   'attempts': [1, 3, 2, 3, 2, 3, 1, 1, 2, 1],  
   'qualify': ['yes', 'no', 'yes', 'no', 'no', 'yes', 'yes', 'no', 'no', 'yes']}
2. Select 'name' , ‘attempts’ and 'score' columns in rows 1, 3, 5, 6 from the above DataFrame (i).
3. Write a Pandas program to select the rows where the number of attempts in the examination is greater than 2 from the above DataFrame(i).
4. Write a Pandas program to count the number of rows and columns of the above DataFrame(i).
5. Write a Pandas program to select the rows where the score is missing, i.e. is NaN.
6. Write a Pandas program to select the rows the score is between 12 and 20 (inclusive).
7. Write a Pandas program to select the rows where number of attempts in the examination is less than 2 and score greater than 10.
8. Write a Pandas program to change the score in row '3' to 11.5.
9. Write a Pandas program to calculate the sum of the examination attempts by the students.
10. Write a Pandas program to calculate the mean score for each different student in DataFrame(i) and list out student(s) name and score(s) who is/are more than mean score.

* End -