Module 5: Pandas Case Study

Problem Statement:

You work in XYZ Company as a Python developer. The company officials want you to build a Python program.

Tasks To Be Performed:

1. Write a function that takes start and end of a range returns a pandas series object containing numbers within that range.

   In case the user does not pass start or end or both they should default to 1 and 10 respectively. E.g:

   -> range_series() -> Should Return a pandas series from 1 to 10
   range_series(5) -> Should Return a pandas series from 5 to 10
   range_series(5, 10) -> Should Return a pandas series from 5 to 15

   Create a method that takes n NumPy arrays of the same dimensions, sums them and returns the answer.

2. Create a function that takes in two lists named keys and values as arguments

   Keys would be strings and contain n string values
   Values would be a list containing n lists
   The methods should return a new pandas DataFrame with keys as column names and values as their corresponding values, e.g:

   ->create_dataframe(["One", "Two"], [["X", "Y"], ["A", "B"]]) -> should return a data frame

   One Two
   0 X A
   1 Y B

3. Create a function that concatenates two DataFrames. Use a previously created function to create two DataFrames and pass them as parameters Make sure that the indexes are reset before returning.

4. Write code to load data from cars.csv into a dataframe and print its details. Details like: 'count', 'mean', 'std', 'min', '25%', '50%', '75%', 'max'.
5. Write a method that will take a column name as argument and return the name of the column with which the given column has the highest correlation.

   The data to be used is the cars dataset.
   The returned value should not the column named that was passed as the parameters, e.g. : get_max_correlated_column('mpg') -> should return 'drat