**Pandas Quiz**

**Section 2**

1. What is the keyboard shortcut to run a cell in Jupyter Notebook?
2. **SHIFT+ENTER**
3. CTRL+SPACE
4. ALT+ENTER
5. SHIFT+TAB

**Section 2**

1. Suppose that

dict1 = {‘a’ : 1 , ’b’ : 2 , ’c’ : 3 , ’d’ : 4 , ’e’ : 5}

s1 = pd.Series(dict1 , index = [ ‘a’ , ‘b’ , ‘f’ ]

what is the output of the below statement?

print( s1[‘f’] )

* 1. Error
  2. **None**
  3. 6
  4. 0

1. Suppose that : df = pd.DataFrame([“Alex” , “C++”, 97], [“Bob” , “Java”, 93]]) ,

what is the output of the following statement?

print(df.shape)

1. 6
2. 2
3. (3,2)
4. **(2,3)**
5. Which of these methods let us know the number non-nan value in each column as well as data type for each column?
   1. **info()**
   2. is\_null()
   3. head()
   4. tail()
6. Assume that df is a DataFrame containing 10 rows and 8 columns. What is the output of the following statement?

print(len(df.head()))

1. **5**
2. 10
3. 8
4. We need to know what is the data inside df.
5. When your csv file does not contain labels for each column, which of the following should be added to the read\_csv method?
   1. Skiprows = 1
   2. **Header = None**
   3. Header = 1
   4. comment = ‘#’
6. Suppose that our DataFrame has 5 columns including ID, Name, Gender, City and Score. Which of the following statements return the Gender column?
   1. df.iloc[:,2]
   2. df.loc[:,' Gender']
   3. df.Gender
   4. **All of the above**
7. Suppose that you have a DataFrame with two columns, Movie name and comment. Each row represents a comment for a movie. How we get the number of comments for Titanic Movie that has the word romantic?
   1. len((Df[‘Comment’]==’Romantic’) & (df[‘Movie’]==’Titanic’))
   2. sum((Df[‘Comment’]==’Romantic’) & (df[‘Movie’]==’Titanic’))
   3. len(df.loc[(df[‘Comment’]==’Romantic’) & (df[‘Movie’]==’Titanic’))
   4. **len(df.loc[(df[‘Comment’].str.contains(’Romantic’)) & (df[‘Movie’]==’Titanic’))**

**Section 3:**

1. What is the purpose of the following statement?

Df.dropna(threshold = .20 \* len(df) , axis = 1)

* 1. Removes the columns that at least 20 percent of the column is Null
  2. **Keeps the columns that at least 20 percent of the columns are not Null**
  3. Removes the first 20 percent of rows
  4. Removes 20 percent of the first columns

1. Which of the following is incorrect?
   1. **Categorical data type helps represent string features with lots of options.**
   2. Categorical data types can decrease memory usage and make features comparable.
   3. We can convert numerical data to categorical data types.
   4. Some Categorical objects are ordered, and some of them are not ordered.

**Section 4:**

1. What is the purpose of the corr method for DataFrame?
   1. It computes correlation between all possible pairs of numerical variables.
   2. It can remove highly correlated features.
   3. It is a Series method, not a method for DataFrame.
   4. **a and b**

**Section 5:**

1. Which of the following graph is suitable for showing the general trend over time?
   1. **Line graph**
   2. Scatter plot
   3. Bar
   4. Histogram
2. Suppose that your DataFrame has a column named Hair color. Which of the following creates a valuable plot for comparing the number of people with different hair colors?
   1. **Df[‘HairColor’].value\_counts().plot(‘bar’)**
   2. Df.plot(y = ‘HairColor’)
   3. Df[‘HairColor’].plot(kind = ‘hist’)
   4. Df.plot(X = ‘HairColor’ , kind = ‘scatter’)