Assignment-2

Problem Statement 1: Samantha has created a dataset named 'top50spotify.csv' of her top 50 songs from spotify.

Dataset Description:

top50spotify.csv - The dataset contains 14 features. Here's a brief description of a few columns in the dataset:

- SerialNo. - Serial number of songs
- Track.Name - Name of the track
- Artist.Name - Name of the artist
- Genre - Genre of the song
- Energy - Energy index of the song
- Length. - Length of the song
- Popularity - Popularity index of the song

Tasks to be performed:

1. Import the dataset as a DataFrame and drop the first column.
2. Save it as 'top50.csv'.
3. Find the average Energy and Length of first 10 songs.
4. Find the total length of songs, group by genre from top to bottom.
5. Print the artist name with the most number of tracks in one genre. **(Hint: Group by artist name and genre)**
6. Print the data of the tracks created by the artist from the previous question.

Note: Fetch and download the dataset using the following link
https://www.dropbox.com/s/2hg67jin2n852mz/top50spotify.csv

Problem Statement 2: Write a Python program to perform the following tasks-

1. Create a pandas series from the below dictionary where indices are subjects:

```
{'English':{'Sam':60,'Jackson':74,'Ahree':85},

'History':{'Gloria':83,'Sam':65,'Isla':78,'Aron':72,'Gray':61
},

'Geography':{'Jackson':92,'Gloria':95,'Isla':82,'Aron':75,'A
hree':76},
```

```
'Mathematics':{'Sam':99,'Gloria':74,'Jackson':89,'Ahree':85,
'Gray':95},

'Science':{'Sam':89,'Aron':82,'Gray':78,'Isla':93,'Ahree':87
}

}
```

2. Convert the created series into DataFrame and replace the null values with zeroes.
3. Transpose the DataFrame and create a new column 'Average' and fill the values in it by calculating the average of all subjects.

Problem Statement 3: Write a Python program to create a series from 1 to 1000 and select only numbers divisible by 7 and 17.

Problem Statement 4: Sylphia has a dataset of various cereals sold in the supermarket.

Dataset Description:

cereal.csv - The dataset contains 16 features. Here's a brief description of 3 columns in the dataset:

- **name** - Brand name of the cereals
- **MFR** - Manufacturer of the brands
- **rating** - Rating of the cereals

Syliphia wants to visualize the quality of cereals and determine which manufacturer delivers the best quality.

Tasks to be performed:

1. Import the dataset.
2. Plot ratings of different types of manufacturers.
3. Use xticks range form 0-100.
4. Change the style of the graph to seaborn

Note: Fetch and download the dataset using the following link
https://www.dropbox.com/s/idnul34dfo5cnke/cereal.csv