Data Science amb Python

Sprint 19

S19 T01: Interactive display with ElasticSearch Stack

Cristiane de Souza da Silva

July 2021

Description

We are beginning to become familiar with the interactive display of graphics using Kibana.

Nivel 1

- Exercise 1

Download Kibana and display some graphs using Example datasets.

Here are some examples from sample data n Kibana

Nivel 2

- Exercise 1

Implement a dashboard that interactively displays the data you choose.

The chosen dataset comes from the site Kaggle and contains information about a test written by some students.

First, I will check if the dataset needs some cleaning.

In [1]:
```
#import libraries

import pandas as pd

#Load the dataset

df_scores = pd.read_csv('test_scores.csv')
df_scores.head()
```

Out[1]:

	school	school_setting	school_type	classroom	teaching_method	n_student	student_i
0	ANKYI	Urban	Non-public	6OL	Standard	20.0	2FHT
1	ANKYI	Urban	Non-public	6OL	Standard	20.0	3JIVI
2	ANKYI	Urban	Non-public	6OL	Standard	20.0	3XOW
3	ANKYI	Urban	Non-public	6OL	Standard	20.0	556O
4	ANKYI	Urban	Non-public	6OL	Standard	20.0	74LO

In [3]:
```
df_scores.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2133 entries, 0 to 2132
Data columns (total 11 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   school           2133 non-null   object 
 1   school_setting   2133 non-null   object 
 2   school_type      2133 non-null   object 
 3   classroom        2133 non-null   object 
 4   teaching_method  2133 non-null   object 
 5   n_student        2133 non-null   float64
 6   student_id       2133 non-null   object 
 7   gender           2133 non-null   object 
 8   lunch            2133 non-null   object 
 9   pretest          2133 non-null   float64
 10  posttest         2133 non-null   float64
dtypes: float64(3), object(8)
memory usage: 183.4+ KB
```

In [4]:
```
df_scores.isnull().sum()
```

Out[4]:
```
school             0
school_setting     0
school_type        0
classroom          0
teaching_method    0
n_student          0
student_id         0
gender             0
lunch              0
pretest            0
posttest           0
dtype: int64
```

In [6]:
```
df_scores['school_setting'].value_counts()
```

Out[6]:
```
Urban       906
Suburban    717
Rural       510
Name: school_setting, dtype: int64
```

We can see there is no null values in the dataset.

Next, we can see some graphs from this dataset done on Kibana.

Nivel 3

- Exercise 1

Implement a dashboard that generates graphics using each of the different types of graphics that Kibana offers.

Now , we can see the dashboard done.

The version I downloaded did not allowed save the dashboard in pdf. we can see the link of dashboard here.

In [ ]: