## Fictitious Names

## Introduction:

This time you will create a data again

Special thanks to Chris Albon for sharing the dataset and materials. All the credits to this exercise belongs to him.

In order to understand about it go here.

## Step 1. Import the necessary libraries

```
import pandas as pd
```

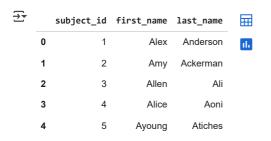
 ✓ Step 2. Create the 3 DataFrames based on the following raw data

Step 3. Assign each to a variable called data1, data2, data3

```
data1 = pd.DataFrame(raw_data_1)
data2 = pd.DataFrame(raw_data_2)
data3 = pd.DataFrame(raw_data_3)
```

Step 4. Join the two dataframes along rows and assign all\_data

```
all_data = pd.concat([data1, data2], ignore_index=True)
all data.head()
```



Step 5. Join the two dataframes along columns and assing to all\_data\_col

```
all_data_col = pd.concat([data1, data2], axis=1)
all_data_col.head()
```

New interactive sheet



## Step 6. Print data3

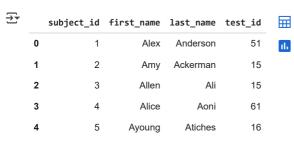
data3



View recommended plots

Step 7. Merge all\_data and data3 along the subject\_id value

merged\_data = pd.merge(all\_data, data3, on='subject\_id')
merged\_data.head()



Step 8. Merge only the data that has the same 'subject\_id' on both data1 and data2

merged\_same\_id = pd.merge(data1, data2, on='subject\_id', how='inner')
merged\_same\_id.head()



Step 9. Merge all values in data1 and data2, with matching records from both sides where available.

```
merged_all_values = pd.merge(data1, data2, on='subject_id', how='outer')
```

mer.Ren\_arr\_varnez·mean()

<del>∑</del> ▼	subject id	first name x	last name x	first name v	last name v	<b>≡</b>
				NaN		
	<b>0</b> 1	Alex	Anderson	inain	NaN	th.
	1 2	Amy	Ackerman	NaN	NaN	
	<b>2</b> 3	Allen	Ali	NaN	NaN	
	3 4	Alice	Aoni	Billy	Bonder	
	<b>4</b> 5	Ayoung	Atiches	Brian	Black	
Next steps: Generate code with merged_all_values View recommended plots New interactive sheet						