'''

Lesson 6 - Day 2 - Pandas DataFrame

\* 'pandas' is a python library/module

\* 'dataframe' is a data-structure provided by pandas.

It is just like excel sheet

\* A dataframe as 'indexes' and 'columns'

'''

# %%

# import pandas module into a variable pd

import pandas as pd

# %%

# create a pandas dataframe from list of lists

dataArr = [

    ['PersonA', 12, 'male'],    ['PersonB', 95, 'male'],

    ['PersonC', 84, 'female']

]

df = pd.DataFrame(dataArr,

                  columns=['name', 'age', 'sexuality'],

                  index=[114,546,987])

# %%

# read dataframe from a csv file using 'read\_csv' function

df = pd.read\_csv('data/personsDb.csv')

# %%

# get all the column names of the dataframe using the 'columns' attribute

colNames = df.columns.tolist()

print('The dataframe columns are {0}'.format(colNames))

# %%

# get all the names from the name column

names = df['name'].values.tolist()

print('All the names are {0}'.format(names))

# %%

# filtering the rows of a dataframe based upon its column values

# get the names of the people who are atleast 50 years

df2 = df[df['age']>=50]

# %%

# filter people who are male and atleast 50 years

df3 = df[(df['age']>=50) & (df['sexuality']=='Male')]

# %%

# get the names of only male people

df4 = df[df['sexuality'] == 'Male']

maleNames = df4['name'].values.tolist()

print('The male people are {0}'.format(', '.join(maleNames)))

# %%

**### initialize a git repository in a folder**

\* In the folder right click->open git bash here

\* use command 'git init'

\* use command 'git status' to see the status of files

\* use command 'git add .' to add all untracked files and changes for staging to get committed

Git \* use command git commit -am 'commit comment' to commit the changes of your code to git

**### pushing code to remote git repositories like GitHub**

\* Create an account in sites like GitHub, Bitbucket

\* Create a repository in your account

\* In the folder right click and select 'Git bash here' to open git command prompt

\* add remote repo url to local git using the command ```git remote add origin <your\_repo\_url>```

\* Pull remote code to local git folder using the command ```git pull origin master```

For unrelated histories issue use command ```git pull origin master --allow-unrelated-histories```

\* Push code to remote repository using the command ```git push origin master```

'utf-8' codec can't decode bytes in position 0-1: invalid continuation byte

Error tokenizing data. C error: Expected 2 fields in line 8, saw 3

**AttributeError**: partially initialized module 'pandas' has no attribute 'DataFrame' (most likely due to a circular import)