COMPUTING SUBJECT: Machine Learning

TYPE: WORK ASSIGNMENT

IDENTIFICATION: Panda Introduction

COPYRIGHT: *Michael Claudius*

DEGREE OF DIFFICULTY: Easy

TIME CONSUMPTION: 1 hours

EXTENT: < 50 lines

OBJECTIVE: Get familiar with Spyder and Jupyter

COMMANDS:

IDENTIFICATION: Panda Introduction

The Mission

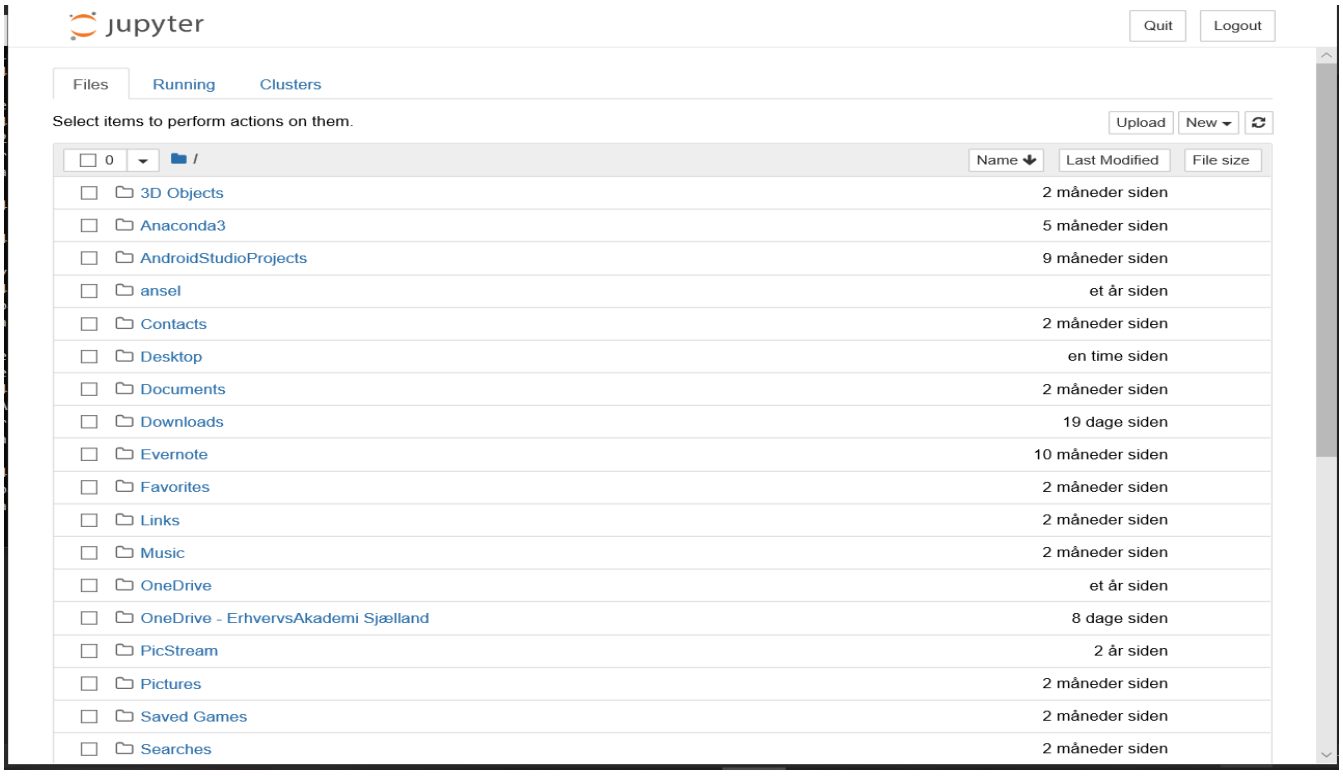
In this assignment, you are to become familiar with use of Panda.
To print out and to plot data from a 2 dimensional array (matrix) using Panda.
Remark
The Pandas library is part of the Data Science / Machine Learning ecosystem in Python. It provides the programmer with an efficient tool for modelling, storing and outputting data.

The problem

In this assignment you are to import the Pandas library, initialize a DataFrame (Pandas data model) and plot the data.

Assignment 1 Create a small program Jupyter

Start Anaconda and open the Jupyter NoteBook.
Your main page should look like this. Find you folder and directory by browsing your folders list.



Then create a new Jupyter notebook

Import the Pandas library simply by typing:

```
import pandas as pd
import matplotlib.pyplot as plt
```

In the next cell initialize a matrix as dictionary containing two lists; i.e. a two-dimensional array.

```
grades = {"Student":["Ole", "Richo", "Jens", "Lars"], "Grade": [2, 7, 12, 2]}
```

View the content

```
grades
```

Now, create a DataFrame that takes the dictionary as argument and view it on the screen.

```
df = pd.DataFrame(grades)
df
```

Your output should look like this:

	Student	Grade
0	Ole	2
1	Richo	7
2	Jens	12
3	Lars	2

Finally, try to plot the data in 3 different ways one by one:

```
df.plot()
df.plot(x='Student', y='Grade')
plt.scatter(x=df.Student, y=df.Grade)
```

What did you get, explain.

At last, save the file in your repository Machine Learning/Solutions/PandaIntro

In [1]:
```
import pandas as pd
```

In [2]:
```
grades = {"Student":["Ole", "Richo", "Jens", "Lars"], "Grade": [2, 7, 12, 2]}
```

In [3]:
```
df = pd.DataFrame(grades)
```

In [4]:
```
print(df)
```