COSC 3337 - In-Class Assignment #1 Line Regression

The purpose of this in-class assignment is for you to get hands-on experience with linear regression.

We will be using the famous *Iris Dataset* from UC Irivine

Please follow closely using the ppt "4.COSC3337_Week2_lecture1_LineRegression", from slides 47-55,

Use these slides to fill in any missing lines of code and fill in the '???'

Reading the Data

Begin by importing the necessary libraries, and use these libraries only

```
In [1]:  import numpy as np
         import pandas as pd
         import matplotlib.pyplot as plt
```

Import the data using pandas and display the first 5 rows

```
In [2]:  iris_df = pd.read_csv('iris.csv')
```

```
In [3]:  ### WRITE YOUR CODE HERE, WHICH SHOULD REPRODUCE THE BELOW IF DONE CORRECTLY ###
```
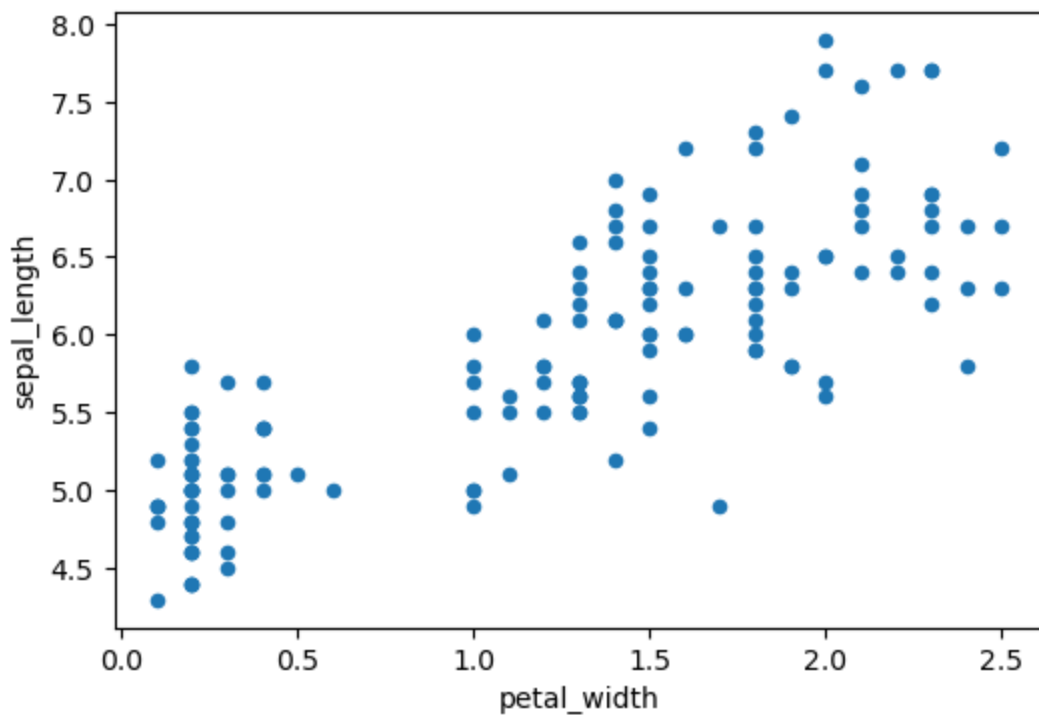
Out[3]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

Goal 1: Predict the sepal length of an iris, given that we know the iris' petal width

First, let's make a scatterplot using petal_width as the x and sepal_length as the y

```
In [4]:  iris_df.plot(kind = 'scatter', x = '???', y = '???', figsize=(6,4))
```

```
Out[4]:  <Axes: xlabel='petal_width', ylabel='sepal_length'>
```

Train a Linear Model

Let's set up the linear regression first, using petal_width as the x variable and sepal_length as the target variable

```
In [6]: petal_width = iris_df['petal_width']
        a = ???
        X = ???
        y = ???
```

Now, we are going to train the linear model, display theta_0 and theta_1

```
In [7]: theta = #???
        theta
```

Now let's graph your trained linear model

Your regression should look like the one displayed below

```
In [9]: iris_df.plot(kind = 'scatter', x = '???', y = '???', figsize=(6,4))
        plt.axis([0,3,4,10])
        X = np.linspace(0,3)
        plt.plot(???,???, "b") #plot the model here
        plt.text(0.5, 8, ???, fontsize=13, color="b") #display theta_0 here as shown below
        plt.text(0.5, 7, ???, fontsize=13, color="b") #display theta_1 here as shown below
        plt.show()
```

$\theta_0 = ???$

$\theta_1 = ???$
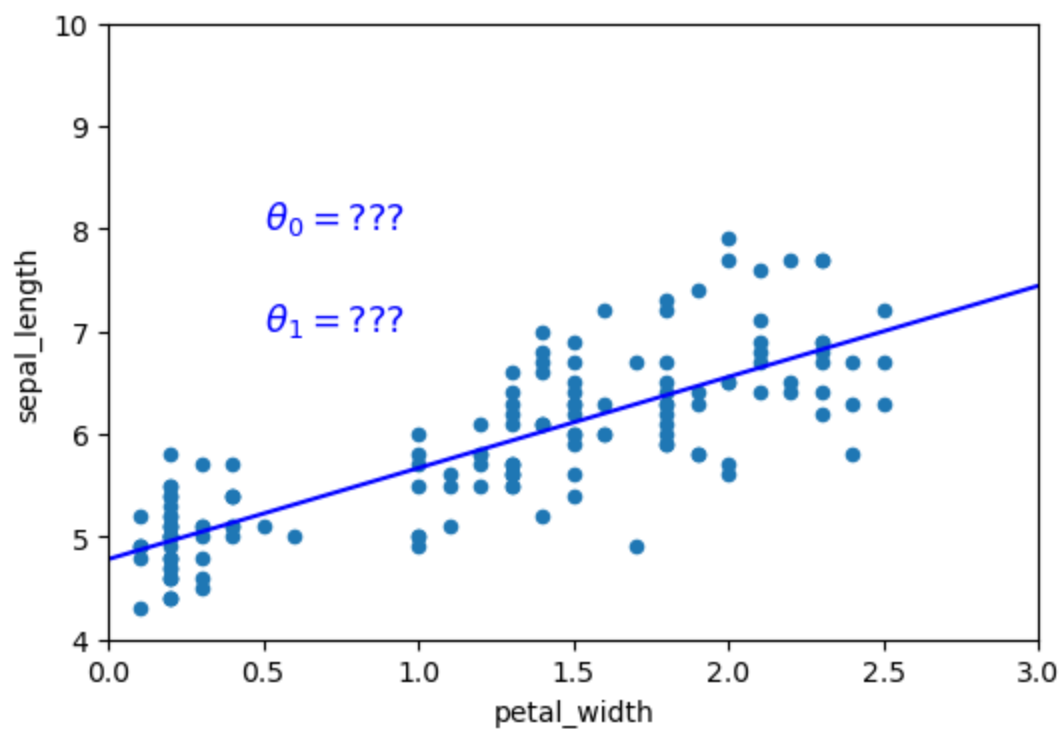
Make Predictions, Let's say we know the petal_width is 1.37, what would be the iris' sepal length?

```
In [10]:  example_petal_width = 1.37
          predicted_sepal_length = ???
          print('The predicted sepal_length is: ???')
```

```
The predicted sepal_length is: ???
```

Now let's plot our predicted sepal length onto our linear model

Your linear model should look like the one displayed below

```
In [11]:  iris_df.plot(kind = 'scatter', x = '???', y = '???', figsize=(6,4))
          plt.axis([0,3,0,10])
          X = np.linspace(0,3)
          plt.plot(???,???, "b") #plot the model here
          plt.text(0.5, 8, ???, fontsize=13, color="b") #display theta_0 here as shown below
          plt.text(0.5, 7, ???, fontsize=13, color="b") #display theta_1 here as shown below
          plt.plot([example_petal_width,example_petal_width],[0,predicted_sepal_length],"r--")
          plt.text(1.5, 4, ???, fontsize=13, color="b") #display your prediction here as shown bel
          plt.plot(???,???,"ro") #plot the petal width with respect to its predicted sepal length
          plt.show()
```