Out[1]: The raw code for this IPython notebook is by default hidden for easier reading. To toggle on/off the raw code, click here.

Exercice 1 - Linear and logistic regression

In this exercise, we are going to apply linear and logistic regressionto two datasets and study the results. The datasets are artificial and represent the salary (target variable) according to employee experience.

1. Start by loading the dataset, Salary_Data.csv.
2. Using the corrcoef library of numpy, compute the correlation matrix
3. Display the data on a graphe.
4. Create a linear regression model from the datausing the scikit-learn LinearRegression package.
5. What is the error ?
6. Plot the linear regression curve on a graph, in addition to the data in the dataset
7. Do again the steps 1 to 5, but Salary_Data2.csv dataset. Any conclusion ?
8. Transform your dataset to create additional columns x^0, X^2, et X^3 and do again linear regression.
9. Plot the linear regression curve on a graph, along with the **initial (Salary_Data2.csv)** dataset data