Lab 7

ITS265

Problem 1

Extract data columns from file Seattle2014.csv for rainfall and snowfall and get the average for the year for both. Also, extract out the first 31 days of data (January) and calculate the average rainfall and snowfall for January. Display the results.

Hint: The values in the file are in $1/10^{th}$ of a mm. Convert to inches when displaying the average. Also, you will need to check for any null values or negative values and either convert to 0 or do not use in calculation.

Problem 2

Write a Python code using NumPy to find the Linear regression equation for given data (lab2-data.txt). Use given equations to calculate b0 and b1 of linear regression and coefficient of determinant (R^2). Use useful NumPy functions, average and sum, for this question. Your code should display b_1, b_0, and R^2 values.

Hint: Extract out from the datafile an x and y column. Then average the extracted columns by using the average function. Use the numpy sum function to calculate the sum for b_1 and SS_{tot} and SS_{res}.

Equations:

$$y_{pred} = b_0 + b_1 * x$$

$$b_1 = \frac{\sum_{i=1}^{n}[(x_i - \bar{x})(y_i - \bar{y})]}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$b_0 = \bar{y} - b_1 * \bar{x}$$

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

$where$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$SS_{tot} = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$SS_{res} = \sum_{i=1}^{n}(y_i - y_{pred})^2$$

Problem 3

Find the weighted grade of each student. Use matrix formulation to calculate it. Final grade should use matrix multiplication. Student grade files are given (lab2-hw.txt, lab2-test.txt, lab2-quiz.txt, lab2-project.txt). Your code should display final grade for each student (A, B, C, D, E, and F).
Use following grading scale to calculate the final grade.

5	Homework	50 points each	30%
5	Quizzes	10 points each	10%
1	Project	100 points	20%
2	Tests	100 points each	40%

Useful function: astype, numpy colum_stack (use this to put together the averages of each grade category into array), numpy matmul (apply weights in array to the stacked grades).

Hint: each assessment has total score. You should covert them into 100 bases.
Final grade = G * W
        where G is grade matrix (rows are students' grades and columns are assessments, 6x4)
              W is weight matrix (weight for each assessment, 4x1)

Submission: Use **Zip** to compress all code and submit to Brightspace. Use "**ITS265-Lab7-**
*FirstnameLastname***.ZIP**" for the filename.