

Lab 11 (Due: Nov 11)

PYTHON PROGRAMMING FOR DATA SCIENCE - COSC 3360

Department of Computer Science and Electrical Engineering

Fall Semester, 2022

Exercises

Create a **New Project** for every exercise. Take a screenshot of the source code along with its output and place the **source code** and the **screenshot** in a **zipped folder** named **LastNameFirstName_Lab11**

Exercise 1

Given the dataset: (-5, -10), (-1, -3), (3, 5), (7, 8), (5, 7), (9, 9), find the **regression line** using **Matrix Algebra Least Squares** (slides 323-325). Find the **parameters** of the line (slope and y-intercept) and **plot** the data points along with the regression line

Note 1: You can use the **x.transpose()** method to *transpose* a matrix, the **np.matmul(x, y)** method to multiply matrices *x* and *y*, and the **np.linalg.inv(x)** to find the *inverse* of a matrix

Note 2: Your algorithm should be able to work with any dataset not just a dataset of 5 data points

Exercise 2

Based on the dataset from Ex. 1, find the **regression line** using **Ordinary Least Squares** (slides 326-328). You can use the **mean()** built-in function

Exercise 3

Based on the dataset from Ex. 1, find the **regression line** using the built-in **linregress()** method (slides 331-334). **Plot** the data points along with the regression line and **print:** *slope, y-intercept, r (correlation coefficient), p-value, standard error*

Exercise 4

Based on the dataset from Ex. 1, compute the: **SST, SSR, SSE** (slide 336) and the: **r, R²** (slide 337). You can use the **sum()** and **mean()** built-in functions

Exercise 5

Given the **grades.csv** file, where the first column is *hours studied* and the second column is *grades*, **read** its data and place them into two **numpy** arrays. Find the **regression line** using the built-in **linregress()** method (slides 331-334). **Plot** the data points along with the regression line and **print:** *slope, y-intercept, r (correlation coefficient), p-value, standard error*. **Predict** a student's grade given **10.5 hours of study**

Note: Submit through **Canvas**