ACTU PS5841 Data Science in Finance & Insurance - Spring 2022 (Y. Wang)
Assignment - 1
Assigned 1/20/22, Due 1/27/22 (Thur)

Problem 1. Base Python, Method of Least Squares
Let **base python** refer to the python distribution from www.python.org. Please use only base python to complete the following tasks and submit a functioning python script, SLR.py.

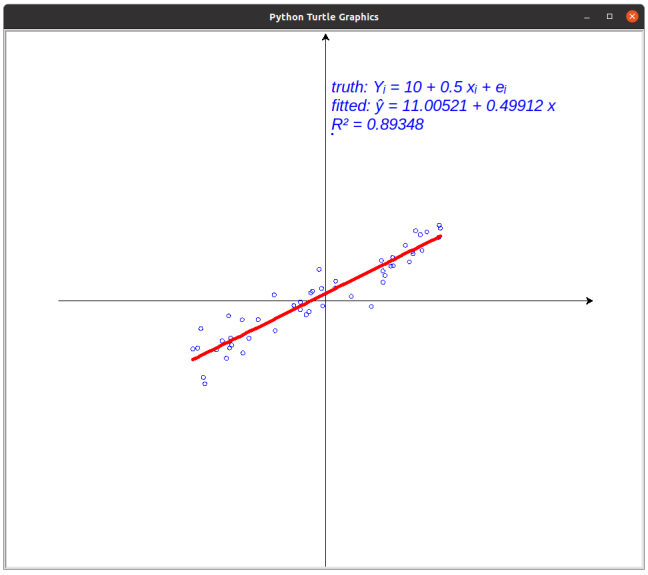
[1] Generate a training set containing 50 observations according to
$Y_i = 10 + 0.5x_i + \epsilon_i$ or $\boldsymbol{y} = 10 + 0.5\boldsymbol{x} + \boldsymbol{\epsilon}$
where $x_i \in (-200, 200)$ and $E(\epsilon_i) = 0, Var(\epsilon_i) = 400$ and ϵ_i are independent random variables.

[2] Adopt the model $E(Y_i) = \beta_0 + \beta_1 x_i$ and apply the *method of least squares* to the training set to estimate the model parameters and produce the fitted values:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

[3] Calculate the coefficient of determination R^2 of the fitted model.

[4] Use turtle graphics to present a scatter plot of the training set, the fitted regression line, and the related annotation, similar to the one below.



You may find the following resources useful.

random — Generate pseudo-random numbers
https://docs.python.org/release/3.10.2/library/random.html
random.uniform(a, b)
random.gauss(mu, sigma)

turtle — Turtle graphics
https://docs.python.org/release/3.10.2/library/turtle.html
turtle.write()
ClassFolder/U1-BasePy/Lib-standard-turtle.py

data structures
ClassFolder/U1-BasePy/BasePython-DataStructures.html

program control
ClassFolder/U1-BasePy/BasePython-ControlFlow.html

unicode character table
https://unicode-table.com/en/