

Module 4 Live Coding Assignment: Python Functions and Linear Regression Basics

Throughout this assignment, please remember to import all the necessary libraries at the beginning of every exercise.

Question 1

Define a function called “calculate_si_amount” that takes the parameters “principal”, “rate”, and “time” and computes the “interest” using the formula

$$\text{interest} = \text{principal} * \text{rate} * \text{time} / 100, \quad (1)$$

and returns the variable “total” computed by adding “principal” and “rate”. Next, call the function using the values principal = 1859, rate = 1.67, and time = 5.

Question 2

Define a function called “calculate_sum” that takes a variable number of parameters and computed their sum. Make sure you initialize your starting sum correctly and return the final result to a variable called “sum”. Finally, compute the same of the numbers 34, 73.8, 28 and 78.

Question 3

Import the public “diabetes” database from the sklearn library and save it to a dataframe called “diabetes”. Print to screen the shape of the dataset and the first ten lines of the dataset. Next, perform a simple linear regression using only the variable “Age” to predict the variable “Body mass index”. Save your model in a variable called “simple_prediction”. Print to screen the intercept and the coefficients obtained from your linear regression model and the score of your prediction.

Question 4

Import the public “diabetes” database from the sklearn library and save it to a dataframe called “diabetes”. Next, perform a multiple linear regression using the variables “Age”, “Body mass index” to predict the variable “Average blood pressure”. Save your model in a variable called “multiple_prediction”. Print to screen the intercept and the coefficients obtained from your linear regression model and the score of your prediction.

Question 5

Define the following dataset called “Housing”

Year	Income (K)	Bedrooms	Bathrooms
2017	67	2	1
2017	78	1	1
2017	69	1	2
2017	130	2	2
2018	95	2	1
2018	113	2	2
2018	99	3	1
2019	178	4	3
2019	156	3	2

Assign it to a dataframe called “df”, naming the columns accordingly to the definition of the dataset and print to screen the first 5 rows. Using the “statsmodel” library, compute a simple linear regression using the values in “Year” to predict “Bedrooms” and assign it to a variable called “model”. Print to screen the summary of your result.

Question 6

Consider the dataframe “df” created in the previous question. Using the “statsmodel” library, compute a simple linear regression using the values in “Year” and “Income (K)” to predict “Bedrooms” and assign it to a variable called “multiple_model”. Print to screen the summary of your result.