Report for HW 2

Eren Özbek

In my homework, I investigated the relationship between Turkey’s ﻿Urban population (% of total population) and ﻿GDP per capita based on the data provided by World Bank. I used World Bank API for data scraping, NumPy and Pandas for data manipulation, and Matplotlib for data visualization. First of all, I defined a function that is able to make linear regression in matrix in my first py file. In my function, I defined the variables of beta, error term, variance, sigma square, and standard error based on the formulas provided in the HomeWork 2 file in Koç Python 2021 GitHub page. Next, I defined upper and lower values with the variables I defined earlier so that I was able to calculate the Confidence Interval. I worked in 95% C.I. and had the Z-score of 1.96. The last step in this code is to print the outputs in a dictionary, so that I am able to provide definitions for my data with the help of keys of dictionary. Then I used World Bank API ‘wbdata’ to obtain the data set needed for my analysis. I defined ‘tur’ variable and obtain Turkey’s data needed for my work. I saved my data to a csv file. I use my linear regression function to analyze the results. Here are my results:

﻿{'beta': array([0.0084878]), 'Confidence Interval': [array([0.00708351]), array([0.00989208])], 'SE': array([0.00071647])}

According to my results, the coefficient for my regression is 0.009.

The data found by me is different from the results of output I made by using STATA. The reason for this might be that the arguments my regression function takes are defined as arrays in my code. Another reason might be that there might be some complications with my way of using matmul function. I will investigate it further

Chart, line chart, histogram

Description automatically generated