Tae Coding

Introduction to Data Science: CS61

Summer 2018

Class Exercise#5

Date Given: June 26, 2018 Due Date:

=======================================================================

There are 4 problems in this assignment.

Please add the following table on the first page of your submission.

|  |  |
| --- | --- |
| **Problem Number** | **Answer** |
| 1 | Answers: |
| 2 | Answers: |
| 3 | Answers: |
| 4 | Answers: |

This assignment comprises 6 problems.

----------------------------------------------------------------------------------------------------------------------------

**Problem#1**

Linear Regression

The values of 2 variables X and Y are given below. Here X is the predictor variable and Y is the response variable.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| X | 25 | 30 | 35 | 40 | 45 |
| Y | 5 | 260 | 480 | 745 | 1100 |

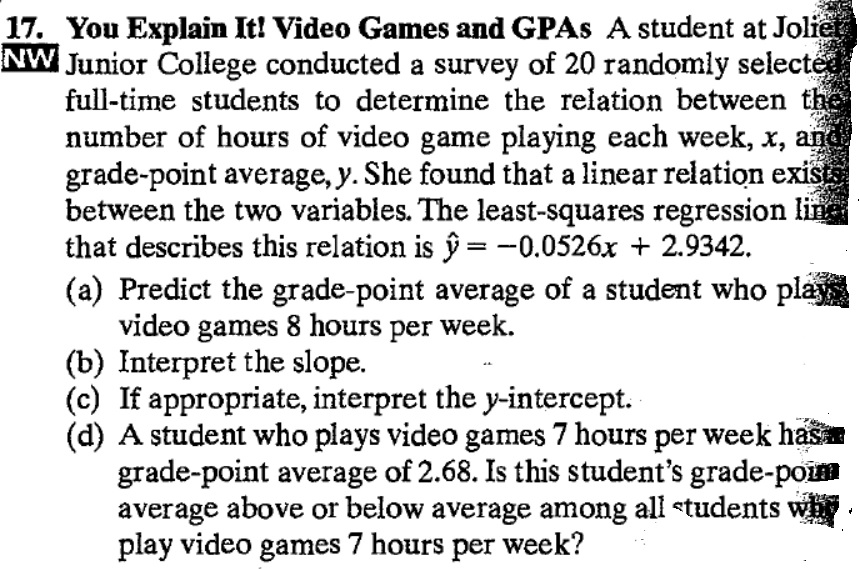
Build your regression model using the following 3 methods.

1. Closed form solution – using only the mean of ‘x’, ‘y’, ‘x\*y’, ‘x2’ variables.
2. Closed form solution – using the correlation coefficient between ‘x’ and ‘y’ variables and the standard deviation of both variables.
3. Closed form solution – using matrix (use Python for matrix manipulation)
4. Python: Scikit-Learn

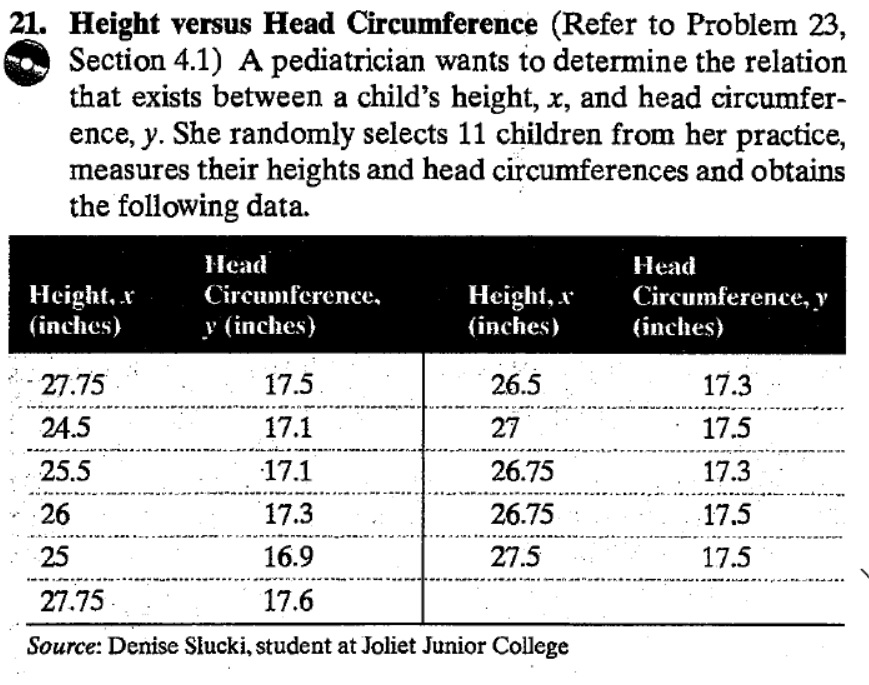
Make sure that your answers are the same using all the 3 methods.

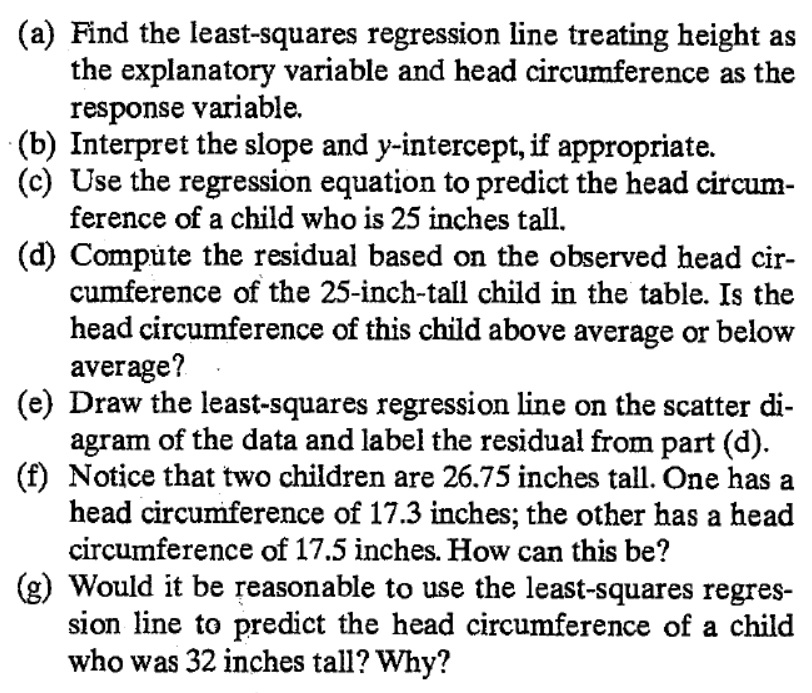
Solve Problem#2 + #3 + #4 in Python.

**Problem#2**



**Problem#3**





**Problem#4**

