Simple Linear Regression: -Any two given three Problems

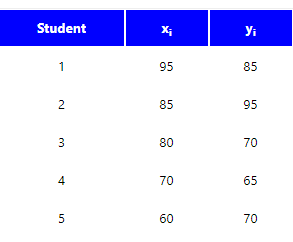
Use OLS and Matrix method

1. Problem 1

Last year, five randomly selected students took a math aptitude test before they began their statistics course. The Statistics Department has three questions.

* What linear regression equation best predicts statistics performance, based on math aptitude scores?
* If a student made an 80 on the aptitude test, what grade would we expect her to make in statistics?
* How well does the regression equation fit the data?

In the table below, the xi column shows scores on the aptitude test. Similarly, the yi column shows statistics grades.



1. Problem 2

The sales of a company (in million dollars) for each year are shown in the table below.

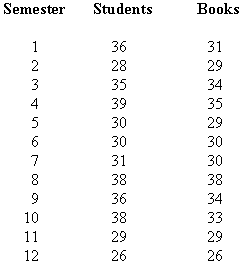
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| x (year) | 2005 | 2006 | 2007 | 2008 | 2009 |
| y (sales) | 12 | 19 | 29 | 37 | 45 |

a) Find the least square regression line y = a x + b.  
b) Use the least squares regression line as a model to estimate the sales of the company in 2012.

c) How well does the regression equation fit the data?

1. Problem 3

A college bookstore must order books two months before each semester starts. They believe that the number of books that will ultimately be sold for any course is related to the number of students registered for the course when the books are ordered. They would like to develop a linear regression equation to help plan how many books to order. From past records, the bookstore obtains the number of students registered, X, and the number of books sold for a course, Y, for 12 different semesters. These data are below.

****

a) Find the least square regression line y = a x + b.

b) How well does the regression equation fit the data?

**Reference:**

<http://faculty.cas.usf.edu/mbrannick/regression/Part3/Reg2.html>

## Estimating R2

## Testing the Significance of R2

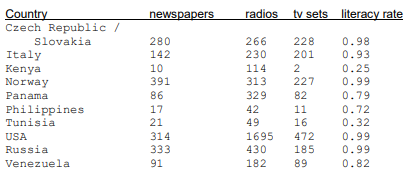
## Tests of Regression Coefficients

Multiple Linear Regression:

Use OLS and Matrix method

1. Problem 4

Literacy rate is a reflection of the educational facilities and quality of education available in a country, and mass communication plays a large part in the educational process. In an effort to relate the literacy rate of a country to various mass communication outlets, a demographer has proposed to relate literacy rate to the following variables: number of daily newspaper copies (per 1000 population), number of radios (per 1000 population), and number of TV sets (per 1000 population). Here are the data for a sample of 10 countries:



1. What is the response variable? What are the explanatory variables?
2. Write the least-squares regression equation for this problem. Explain what each term in the regression equation represents in terms of the problem.
3. How well does the regression equation fit the data?

**Note:** Those who wants to do **research** on this please follow all the steps given in the below link and approach me for clarification:

https://courses.ecampus.oregonstate.edu/st352/lesson21-mlranalysis-w.pdf

Before 22.08.2021lab:

Python implementation

Hard Code and Gradient Descent

1. Python Implementation for Problem 3 and 4
2. Practice the following dataset using Linear Regression and apply Z-Score standardization if necessary.

https://cdn.scribbr.com/wp-content/uploads//2020/02/heart.data\_.zip

1. What is the response variable? What are the explanatory variables?
2. Write the least-squares regression equation for this problem. Explain what each term in the regression equation represents in terms of the problem.
3. How well does the regression equation fit the data?