**Lab 10**

|  |  |
| --- | --- |
| Name: | 陳柏霖 |
| Student ID: | B09611007 |
| Total Score: |  |

1. **Multiple Choice (35 points, 5 points each question)**

|  |  |  |  |
| --- | --- | --- | --- |
| # | Answer | Explanation (Optional) | Score |
| 1 | A |  |  |
| 2 | C |  |  |
| 3 | B |  |  |
| 4 | D |  |  |
| 5 | C |  |  |
| 6 | C |  |  |
| 7 | E |  |  |

1. **Simple Linear Regression (30 points, 10 points each question)**

|  |  |  |
| --- | --- | --- |
| # | Description | Score |
| 1 |  |  |
| 2 | Yes. The values of the slope and the intercept are positive. |  |
| 3 | Yes. The plotted lines are the same. |  |

1. **Multiple and Polynomial Regression (35 points)**

|  |  |  |
| --- | --- | --- |
| # | Description | Score |
| 1 | 1. Load the csv file, and add “age^2”, “log\_pages” to the dataframe.  2. Split the data with reasonable train/test = 0.8/0.2  3. Plot violin plots to validate the appropriate distributions of the training data and testing data |  |
| 2 | 1) Training R^2 = 0.69, testing R^2 = 0.61. MSE of this model is 60.2.  2) The coefficients of age(6.7) and big club(9.5) show that they are crucial for the market value, while age^2(-0.13) and log2(page\_views) are not.  3) The player who wants to improve his market values should join the big club. If the player wants to increase his market value by 10, he should reach (current page views)^(10/1.911) pages views.  4) The model performs well on testing data. I consider it fits well. |  |