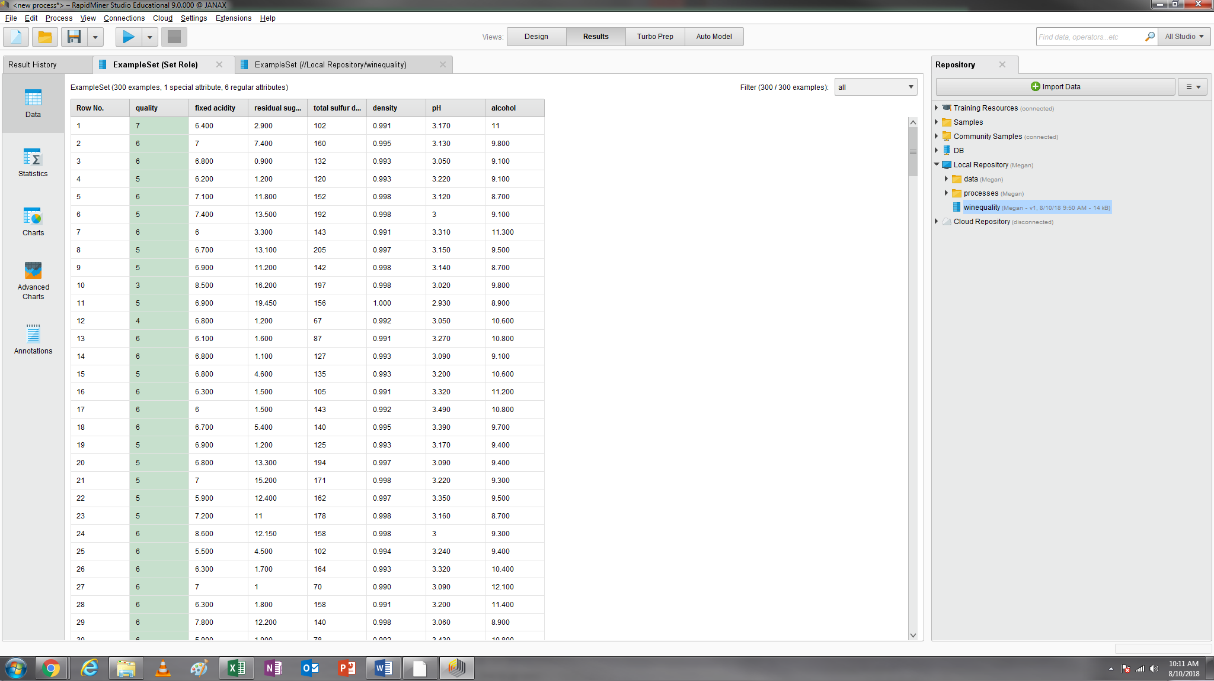
**Class 9 Lab & HW4 Screenshot Submission**

**(Due Midnight Sunday, August 12 2018)**

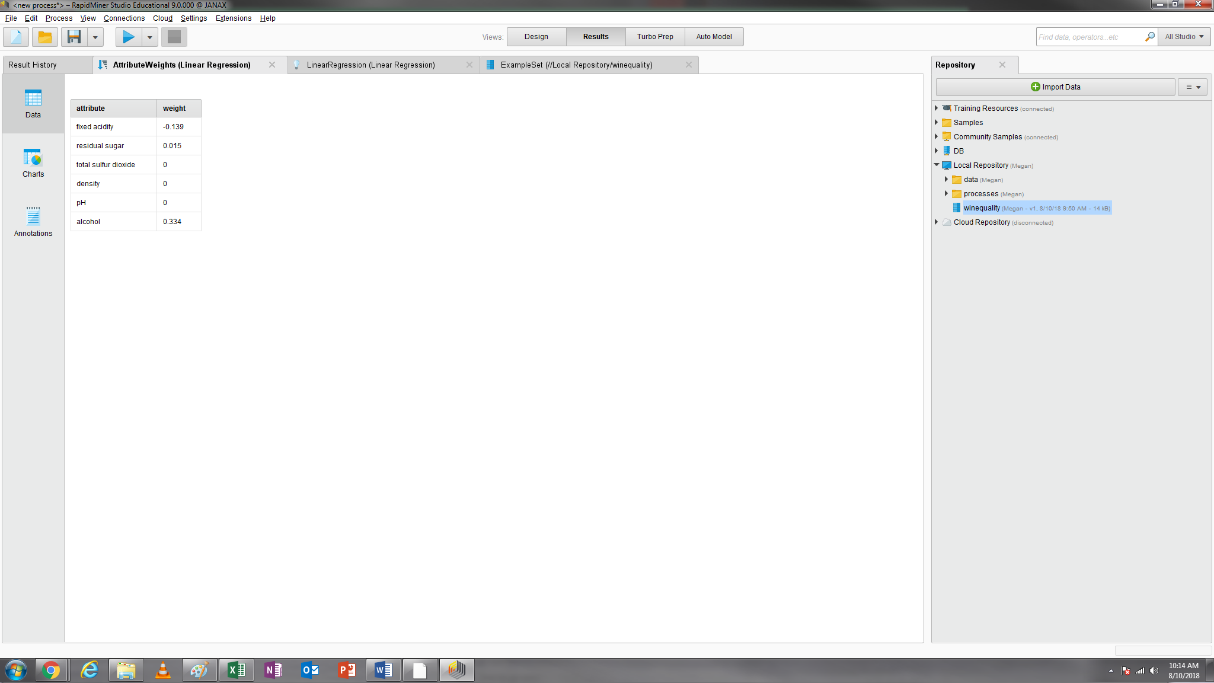
Your Name (First Last): Megan Moore Email Address: [mmmoore2@buffs.wtamu.edu](mailto:mmmoore2@buffs.wtamu.edu)

2.5 Please answer the question: how many special attributes can you find in your ExampleSet [typically, the one(s) highlighted in light green are special attribute(s)] (4 points) One

Screenshot 1



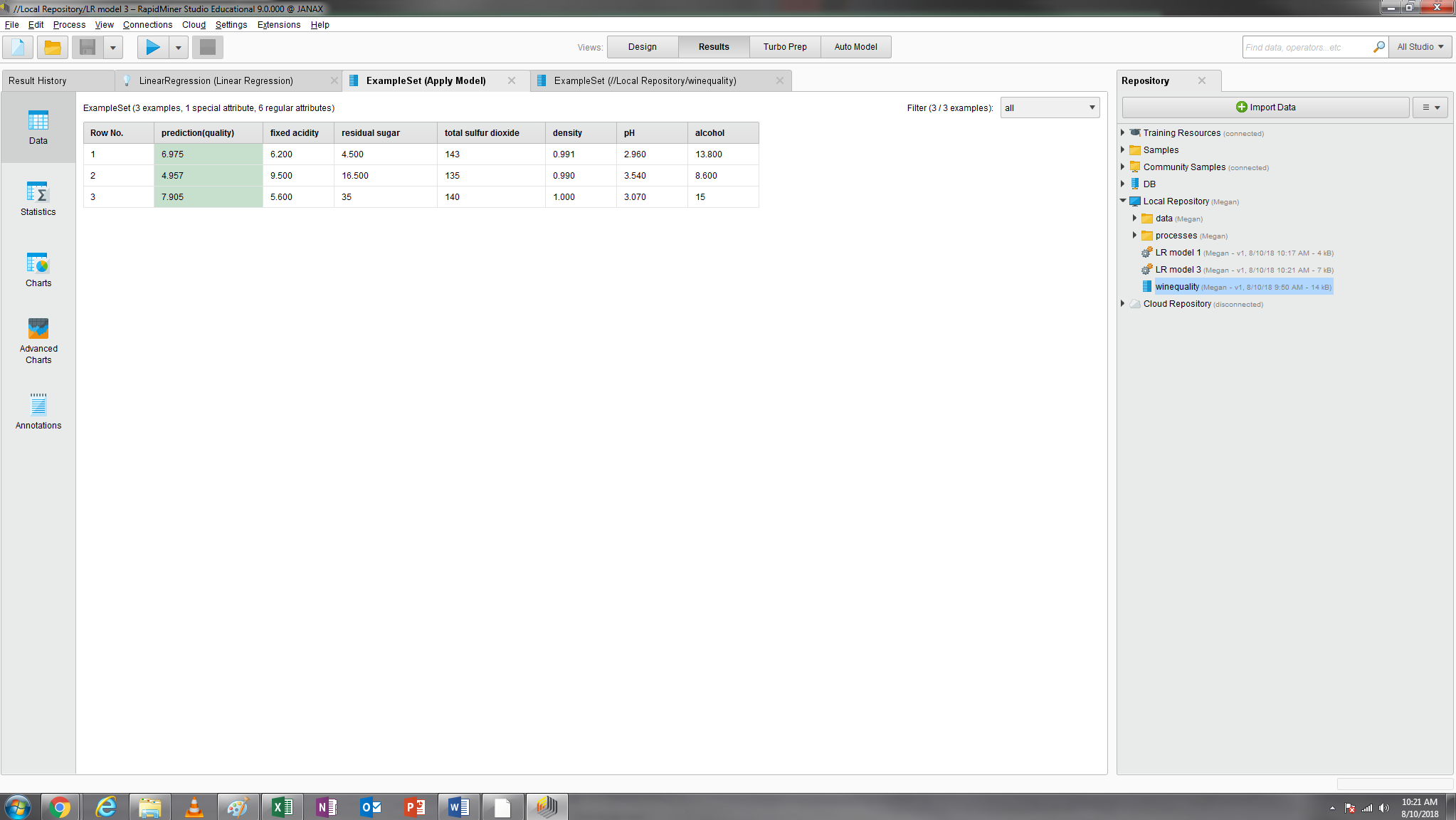
Screenshot 2



3.6.1 Which attribute is a significant and positive predictor of wine quality? Which one is a significant and negative predictor of wine quality? (4 points) Positive, alcohol. Negative, fixed acidity.

3.6.2 In the attribute weight result, please find the weight of the following two attributes: residual sugar and density? (4 points) Residual sugar, 0.015. Density, 0.

Screenshot 3



4.6.1 What is the predicted quality of the first new wine (with fixed acidity =6.2)? [round your answer to an integer] (4 points) 7

4.6.2 What is the predicted quality of the second new wine (with fixed acidity =9.5)? [round your answer to an integer] (4 points) 5

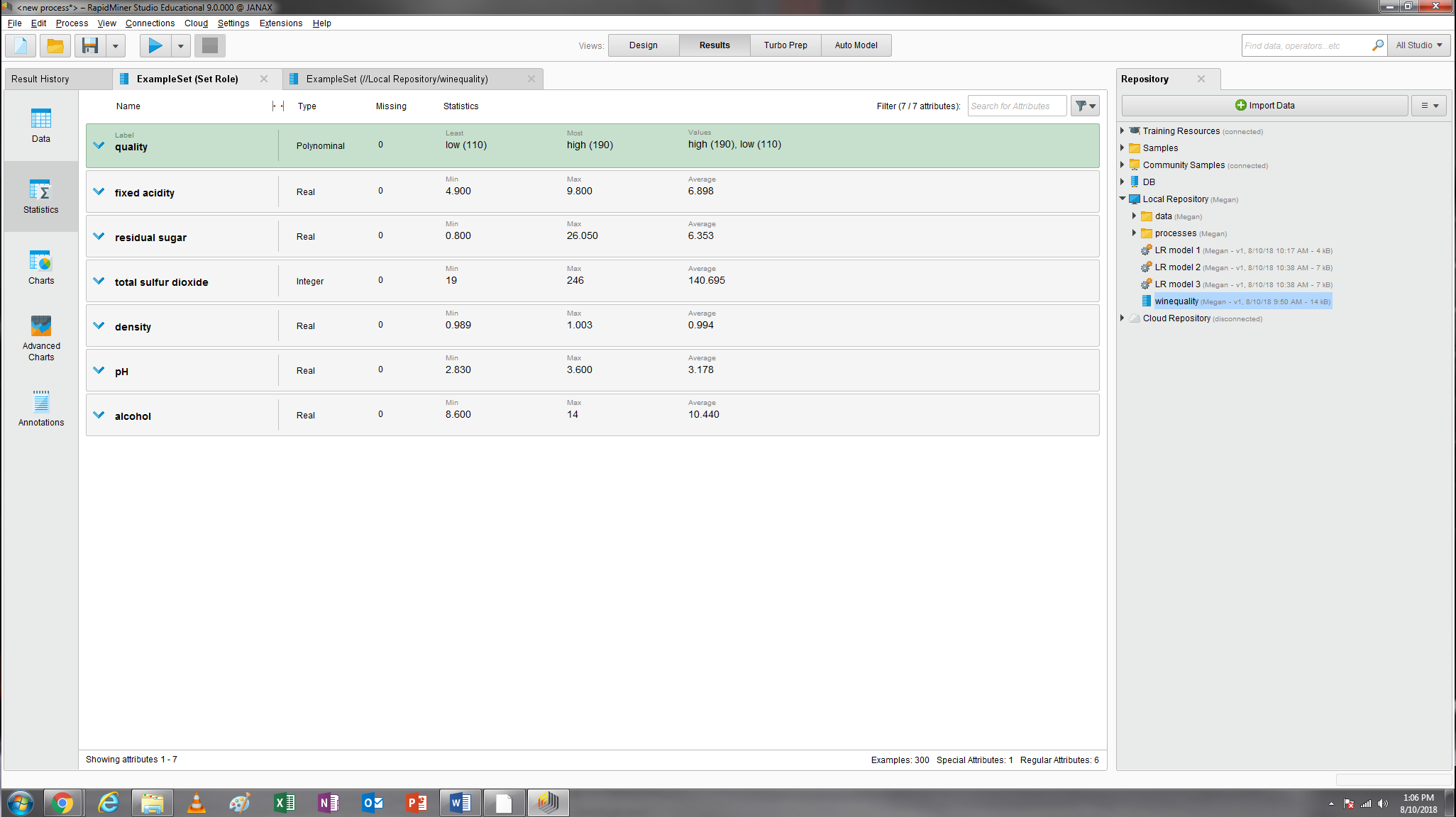
4.6.3 What is the predicted quality of the third new wine (with fixed acidity =5.6)? [round your answer to an integer] (4 points) 8

4.6.4 Is the linear regression model same with the one we generated in Step 3.5? Yes or No? (4 points) Yes, coefficients match up.

4.7 However, in this lab, the prediction dataset only includes three wines, so you can manually check how many records having their attributes outside the range of the training attributes. 0, 1, 2, or 3? (4 points) 1

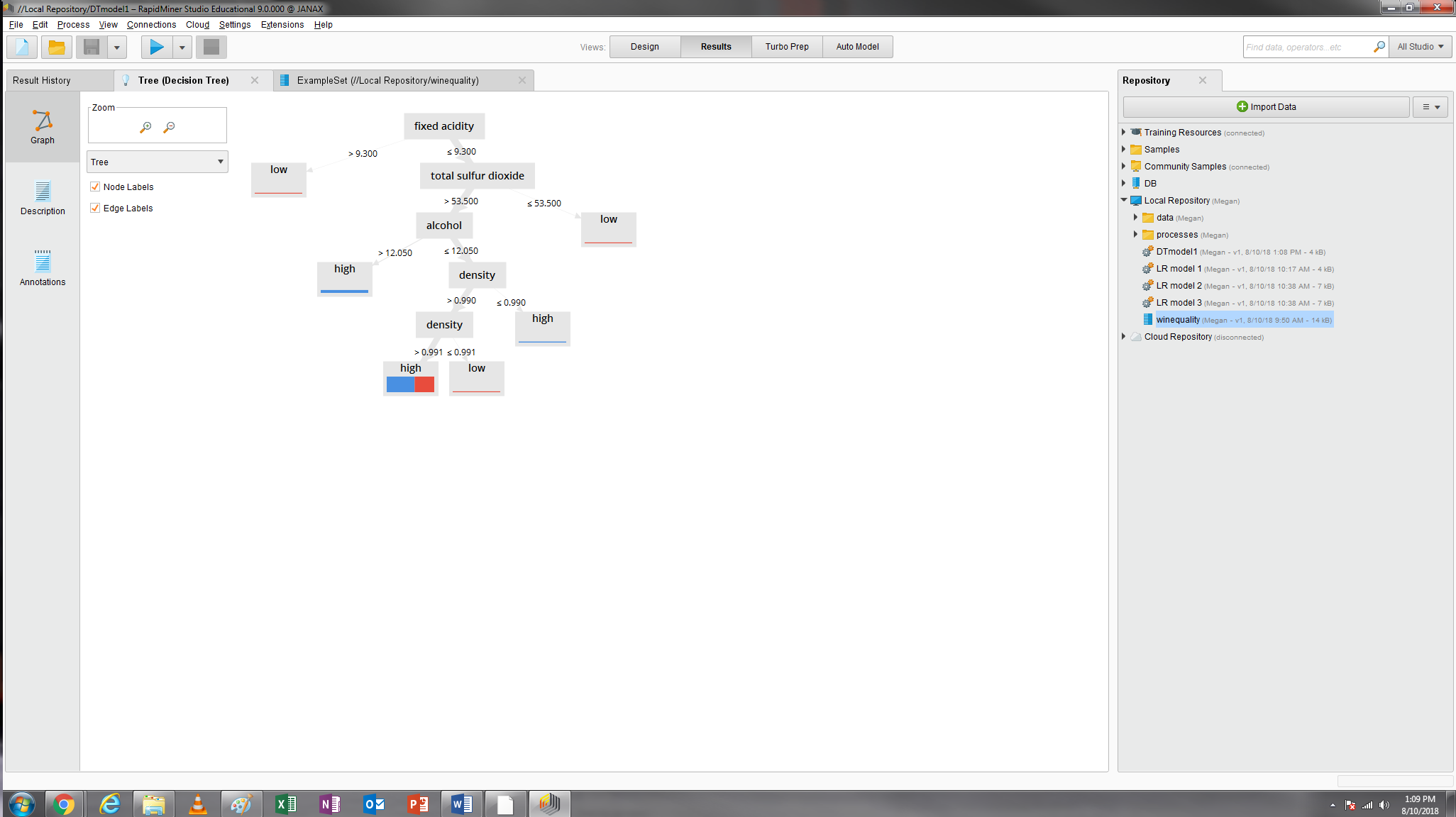
5.6 Run the process and in your ExampleSet click the Statistics button at the left sidebar. Review the results there and answer the question: among 300 observations, how many are rated as high quality? (4 points). 190

Screenshot 4



6.4 Save your process as DTmodel1 and then run this process and you will see a decision tree model and description (see the two screenshots at the next page). Take a screenshot of your decision tree model.

Screenshot 5



6.5.1 Which attribute is located at the root node in this tree? (4 points) fixed acidity

6.5.2 How many internal or split nodes in this tree? (4 points) 5

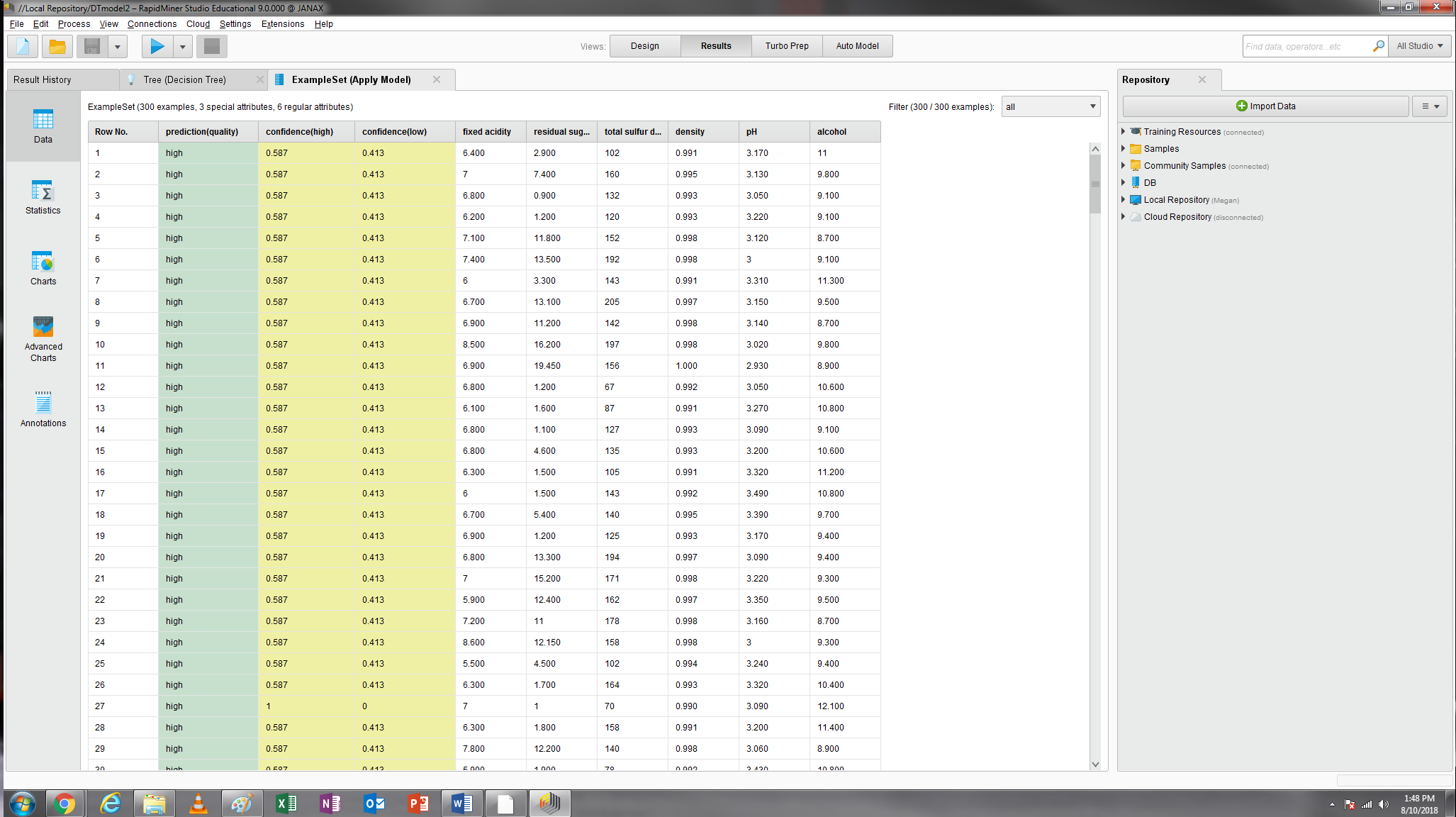
6.5.3 What is the size of the leaf node 4? (4 points) 5

6.5.4 Among the six leaf nodes, how many generate a pure class (i.e, single color in the leaf node)? (4 points) 5

6.5.5 Please use this decision tree to determine the quality of the following two wines (top-to-bottom): (4 points for each)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| fixed acidity | residual sugar | total sulfur dioxide | density | pH | alcohol | Quality (high or low) |
| 9.4 | 2.9 | 102 | 0.9912 | 3.17 | 11 | low |
| 7 | 6.4 | 160 | 0.9954 | 3.13 | 9.5 | high |

Screenshot 6

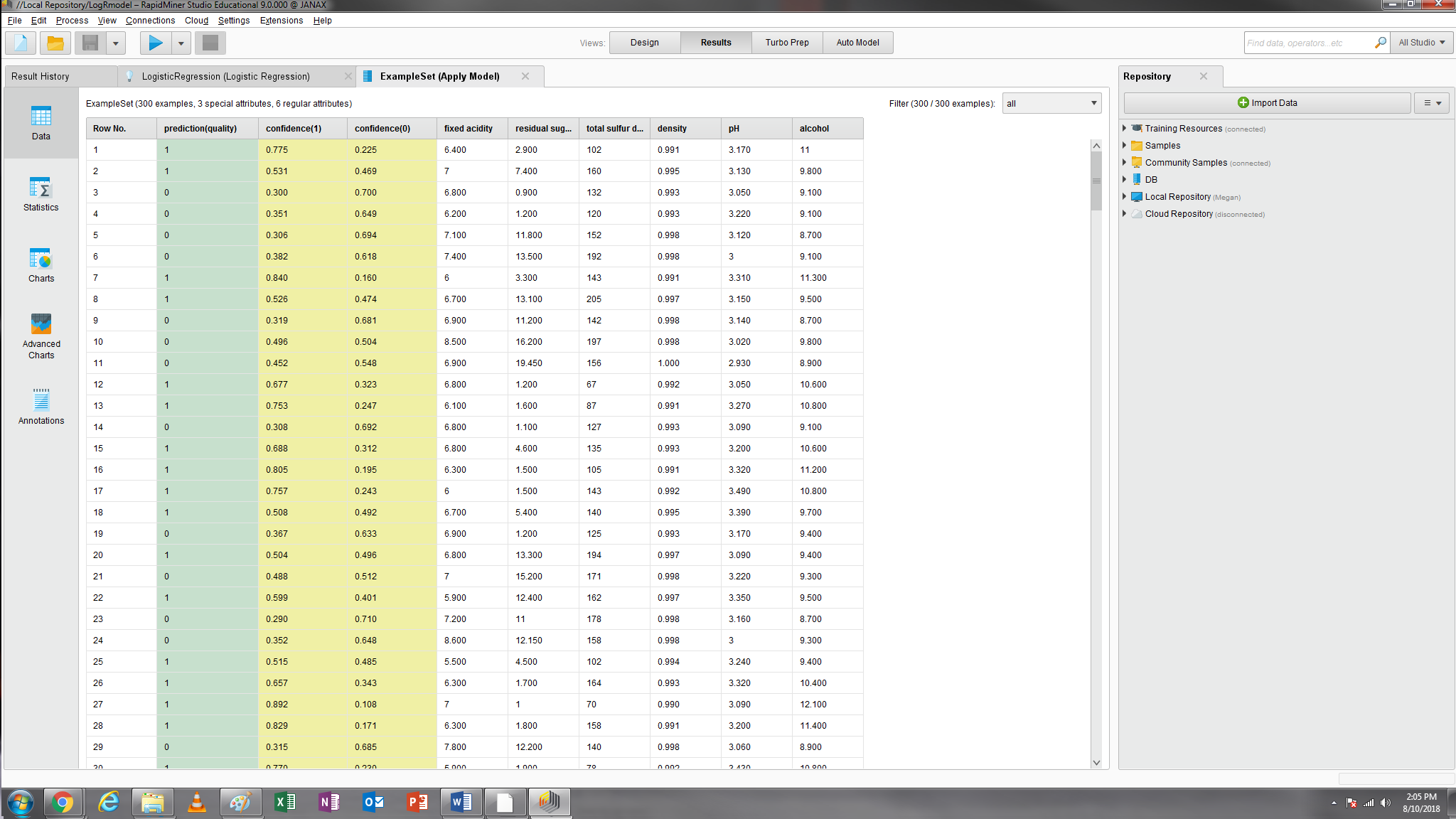


7.5.1 What is the predicted quality of the first new wine (with fixed acidity =6.2)? [low or high] high

7.5.2 What is the predicted quality of the second new wine (with fixed acidity =9.5)? [low or high] low

7.5.3 What is the predicted quality of the third new wine (with fixed acidity =5.6)? [low or high] high

Screenshot 7



9.6.1 What is the predicted quality of the first new wine (with fixed acidity =6.2)? [1 or 0] 1

9.6.2 What is the predicted quality of the second new wine (with fixed acidity =9.5)? [1 or 0] 0

9.6.3 What is the predicted quality of the third new wine (with fixed acidity =5.6)? [1 or 0] 1

10.3 The following table may help you answer this question.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Linear Regression | | Decision Tree | Logistic Regression | | Conclusion |
| The predicted quality of the first new wine (with fixed acidity =6.2) | | 7 | high | 1 | Same prediction or not | |
| The predicted quality of the second new wine (with fixed acidity =9.5) | | 5 | low | 0 | Same prediction or not | |
| The predicted quality of the third new wine (with fixed acidity =5.6) | | 8 | high | 1 | Same prediction or not | |