

1. Which of the following methods should you use to replace a missing value of an attribute with continuous values?

1 point

- ☐ Use the mean square error of the other data in the column
- ☐ Use the difference between the minimum and maximum values of the other data in the column
- ☐ Use an educated guess
- ☒ Use the average of the other values in the column

2. Which of the following helps you decide on bin values when pre-processing data?

1 point

- ☐ Divide the average by the standard deviation
- ☒ Visualize the distribution using a histogram
- ☐ Use the interquartile range
- ☐ Convert objects to ints

3. Which of the following data types should numbers with decimals be if you want to use them as input for training a statistical model?

1 point

666, 1.1, 232, 23.12

- ☐ data frame
- ☒ float
- ☐ object
- ☐ int

4. Which of the following is the primary purpose of simple feature scaling?

1 point

- ☒ It brings data into a common standard of expression
- ☐ To make comparing and analyzing values easier.
- ☐ So all the variables have a similar influence on the models you build
- ☐ To get rid of “not a number” or NaN values

5. Which of the following is the primary purpose of the `get_dummies()` method?

1 point

- ☐ Converts the data's data type
- ☐ Converts numerical values into categorical ones
- ☒ Converts categorical values into numerical ones
- ☐ To help you group your data into bins