

# CS381 Data Analytic Regression Practice Quizz

---

Question 1: How could you include a categorical variable in Linear or Logistic Regression?

Answer: Create additional dummy variables where each variable corresponds to one possible/allowed value of that categorical variable. Each of these dummy variable will be encoded as 0 or 1.

Question 2: If a categorical variable has 6 possible values, and if we use the `get_dummies` method to transform it into a set of dummy variable where each variable will be either 0 or 1, how many these additional dummy variables will be created? In other words, how many additional features that one categorical variable will turn into?

Answer: 5. In general, it will  $N - 1$  for categorical variable that can take  $N$  possible values

Question 3: What is feature scaling? What are two of the common methods in feature scaling?

Answer: Feature scaling is a method use to normalize the data when we have different units (ie scale) among the different features. z-score and min-max normalization are two of the most common normalization method.