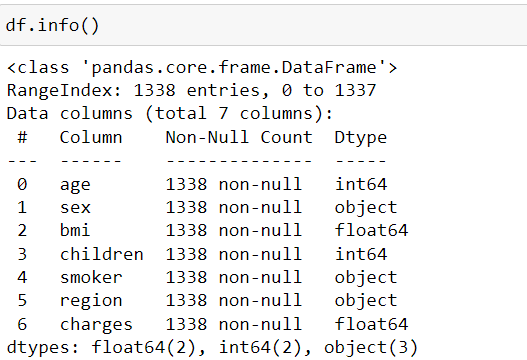
Q1. Explain the data. Check It is supervised or Unsupervised. Demonstrate?



Based on the description of the insurance data set, it appears to be a supervised learning data as it has both the independent variables (age, sex, bmi, children, smoker, region) and the dependent variable (charges). The goal would be to train a model to predict the charges based on the other variables.

Q2. Estimate the minimum sample size n to get the 99% accurate predictions. (precision = 0.02)

To estimate the minimum sample size needed for 99% accurate predictions with a precision of 0.02, the following formula can be used:

n = (z^2 \* p \* (1-p)) / e^2

where:

n is the minimum sample size

z is the standard normal deviate for the desired level of confidence (for 99% confidence, z = 2.576)

p is the proportion of successes in the population (for a precision of 0.02, p = 0.02)

e is the margin of error (for a precision of 0.02, e = 0.02)

Substituting in the values: n = (2.576^2 \* 0.02 \* (1-0.02)) / 0.02^2

n = 665.34

So, the minimum sample size required to get 99% accurate predictions with a precision of 0.02 is 665