PSYC 4330
Seminar in Statistics

Exercise #3

A researcher is interested in predicting depression from trait anxiety and agreeableness. The dataset is called **epi.bfi** and resides within the `psychTools` package. Depression was measured with the Beck Depression Inventory (bdi), trait anxiety was measured with the State-Trait Anxiety Inventory (traitanx), and agreeableness was measured with the Big 5 Inventory (bfagree). Use $\alpha = .05$ for any null hypothesis testing.

1) Use the **delete_MCAR** function from the `missMethods` package to randomly remove 15% of cases from each variable. Save this dataset with a new name.

2) Determine if the missing data can be classified as 'missing completely at random' (you should be able to predict the answer from Q1).

3) Run a general linear model predicting depression from trait anxiety and agreeableness using each of the following strategies:
    a. Complete Data
    b. Listwise Deletion
    c. Regression Imputation
    d. Stochastic Regression Imputation
    e. Expectation-Maximization
    f. Multiple Imputation

4) Explore the unstandardized regression coefficients resulting from each strategy in Q3. Are there any noticeable differences between strategies?

5) Explore the standard errors for the coefficients resulting from each strategy in Q3. Are there any noticeable differences between strategies? (We could/should have also looked at the p-values, standardized effect sizes, etc.)