图示

描述已自动生成

In **single imputation**, each missing value is replaced by a reasonable alternative value (i.e., a plausible guess).

1. In **simple imputation,** the missing values in a variable are replaced with a single value (for example, the mean, median, or mode). Using mean substitution, you could replace each missing value on the Dreamvariable with 1.972 (the mean of the non-missing values).
2. **K-nearest neighbor imputation**: For an observation with one or more missing values, find the observations that are most similar but have the values, and use these cases to make the imputation.
3. missForest: For larger datasets, **random forests** (chapter 17) can be used to impute missing values.

**Multipleimputation(**MI) provides an approach to missing values that’s based on repeated simulations. In MI, a set of complete datasets (typically 3 to 10) is generated from an existing dataset that has missing values.

图示

描述已自动生成