

TABLE OF CONTENTS

- The problem of degrees of freedom
- Missing Value Patterns
- A Shared Sense of identity (Essence of KNN algorithm)
- Distance calculation in the presence of missing values
- Imputation Approach With KNN Impute

THE PROBLEM OF DEGREESS OF FREEDOM:

Missing values in a data set can be a hornet nest for any data scientist. Variable with missing values can be non-trivial problem as there is no easy way out to deal with them.

Generally if the proportion of missing observation in data is small relative to the number of observation we can simply remove the observation. Deleting the rows containing missing values may lead to parting away with useful information or patterns.

In Statistical term this leads to degree of freedom as the number of independent values piece of information goes down.