The reference and sample data set