

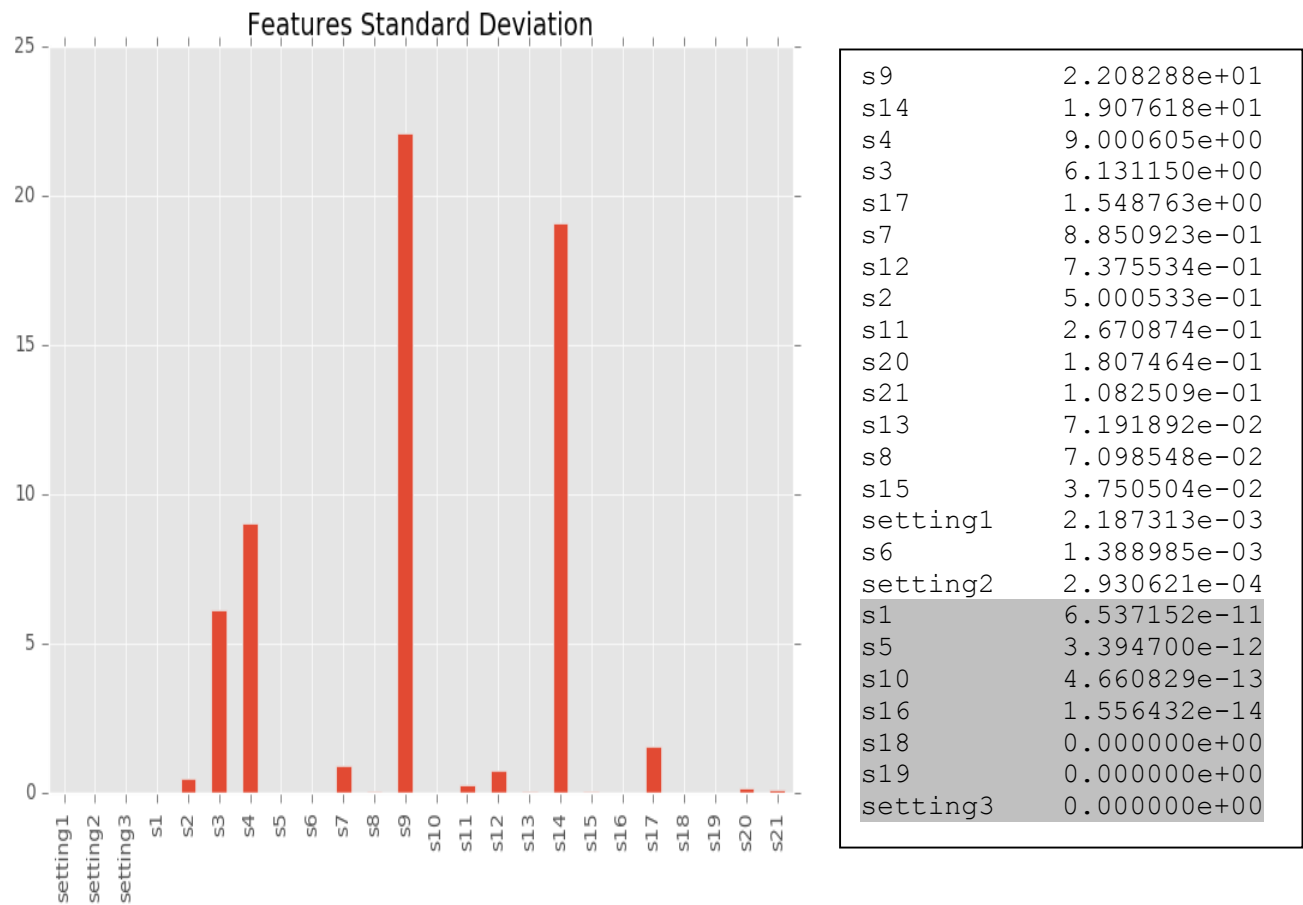
Project Inferential Statistics Report

Introduction:

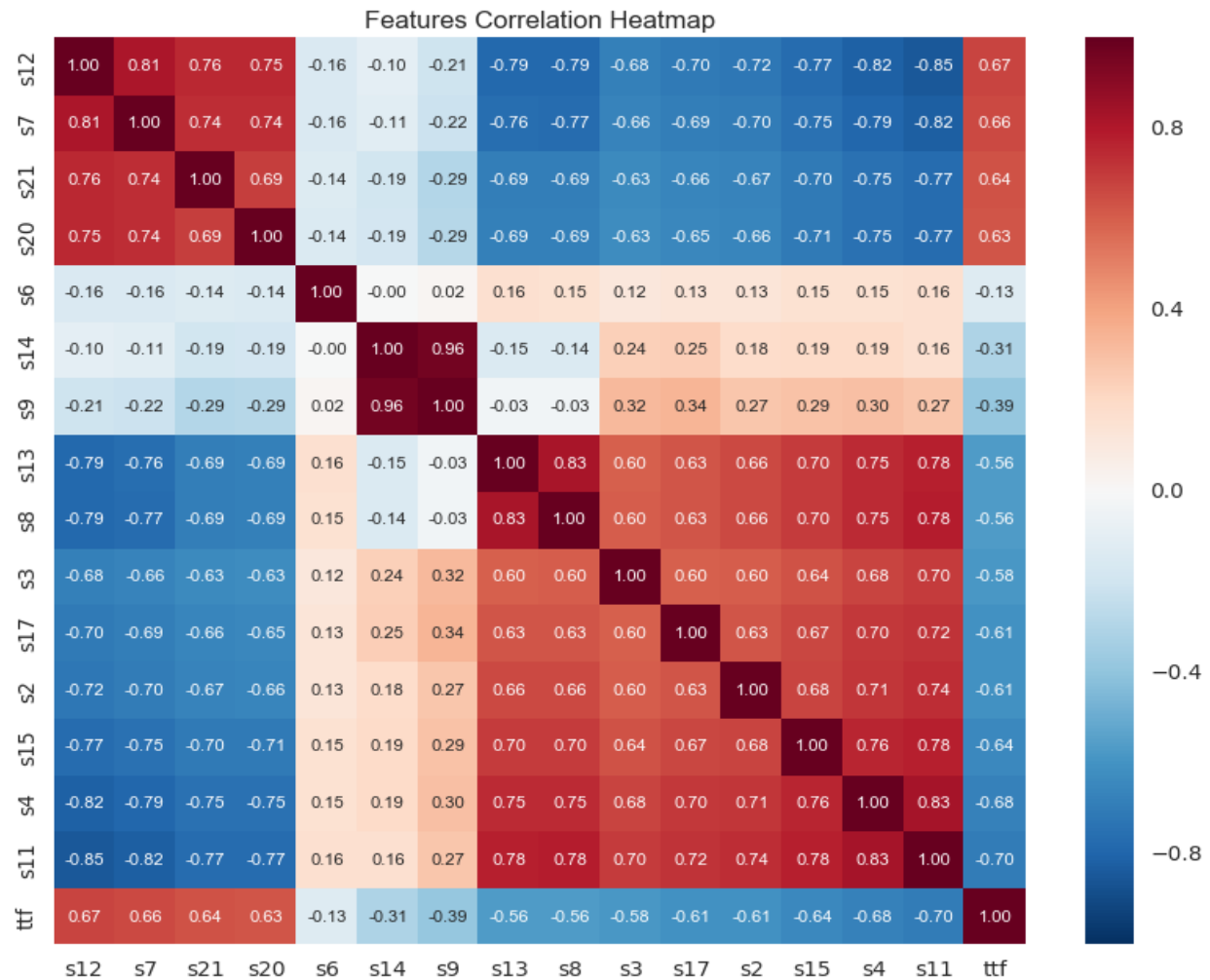
This report presents some statistical inferences done on the project dataset through answering some key question about the training data.

Are there variables that are particularly significant in terms of explaining the answer to your project question?

First step is check the features variability, and remove near zero variance predictors (or predictors which have one value only) since they are considered to have less predictive power. We can consider the other features for further analysis with the assumption they have significant effect on the dependent variable.



Are there strong correlations between pairs of independent variable?



As displayed in the above features correlation heatmap, there is a very high correlation (> 0.8) between the following pairs of independent variable:

(s14, s9), (s11, s4), (s11, s7), (s11, s12), (s4, s12), (s8, s13), (s7, s12)

Are there strong correlations between an independent and a dependent variable?

s20, s21, s7, and s12 have positive correlation with the dependent variable (ttf) in the range 0.63 to 0.67

s17, s2, s15, s4, s11 have a negative correlation in with the dependent variable (ttf) in the range 0.61 to 0.70