

Jun 2, 2022

- Data set selection
 - News article popularity
 - Regression
 - Target variable: number of shares
- Meeting schedule
 - Weekly? Twice per week?
 - Once per week for now
 - What day / time?
 - Thursday @ 8PM
- Technical Report sections:
 - Problem Statement
 - Data Description
 - Exploratory Data Analysis
 - Relationships b/w response variable and predictor variables
 - Data pre-processing
 - Zero-Variance
 - Missing values
 - Outliers
 - Skewed data
 - Correlated features
 - Data splitting
 - Model building strategies
 - Validation and testing (model tuning and evaluation)
 - Results and final model selection
 - Discussion and conclusions
 - Appendix (reproducible and clean R code)
- Action Items
 - Andrew Abeles to schedule recurring meeting
 - Dingyi Duan to create GitHub repository
 - All to use 100 as random seed
 - All to have started EDA, Pre-Processing, Data Splitting and reviewed each other's code by Thursday 6/9
 - All to have finished EDA, Pre-Processing, Data Splitting by Monday 6/13