

**>install.packages("package name")**

# 1 Pre Modeling Stage

- 1. Data Visualization ggplot2, googleVis
- 2. Data Transformation plyr, data.table
- 3. Missing Value Imputations MissForest, MissMDA
- 4. Outlier Detection Outliers, EVIR
- 5. Feature Selection Features, RRF
- 6. Dimension Reduction FactoMineR, CCP



# Data Analysis Useful Libraries in



# 2 Modeling Stage

- 1. Continuous Regression car, randomforest
- 2. Ordinal Regression RMiner, CoreLearn
- 3. Classification Caret, BigRF
- 4. Clustering CBA, RankCluster
- 5. Time Series forecast, LTSA
- 6. Survival survival, Basta

# 3 Post Modeling Stage

- 1. General Model Validation LSMeans, Comparison
- 2. Regression Validation RegTest, ACD
- 3. Classification Validation BinomTools, DAIM
- 4. Clustering Validation ClustEval, SigClust
- 5. ROC Analysis PROC, TimeROC



## Other Libraries

A. Improve performance Rcpp, parallel

B. Work with web XML, jsonlite, httr

C. Report results shiny, RMarkdown

D. Text Mining tm, twitterR

E. Database sqldf, RODBC, RMongo

F. Miscellaneous swirl, reshape2, qcc