**Airline Delay Prediction**

Use this [dataset of airline information](http://stat-computing.org/dataexpo/2009/the-data.html) to predict how late flights will be. A flight only counts as late if it is more than 30 minutes late.

1. The project should include the introduction, describing the data, visualizations (scatter plot, bar graph, histogram, boxplot, subplot, pie chart, heatmap, etc., each graph should have an explained analysis followed), filtering data to different categories, analysis (include t-test or Mann- Whitney for Hypothesis testing).
2. Apply models in Naïve Bayes, Logistic Regression, Decision Tree, Random Forest, Gradient Boosting and SVM.
3. Apply PCA, SelectKBest and RFE for feature selections.
4. Tuning models to obtain best parameters.
5. Compare performances among models, write up analysis why the model is good or bad in the algorithmic approach (explain why the algorithm is good or bad for the dataset structure, can you do something to improve the model?)
6. Include the conclusions.