Final Project proposal for ELE: 581

Name: Muhammad Enayetur Rahman
Student ID: 100635221
Data set: [Airfoil Self-Noise Data Set]

Total 6 attributes are:

1. Frequency, in Hertz
2. Angle of arrack, in degrees
3. Chord length in meters
4. Free-stream velocity, in meters per second
5. Suction side displacement thickness, in meters

And the final output is:

6. Scaled sound pressure level, in meters

Here is the basic statistical summary for each attribute. Attribute 'Sound level' would be my final output regression attribute.

Summary of each attribute:

frequency	
Min.	: 200
1st Qu.	: 800
Median	: 1600
Mean	: 2886
3rd Qu.	: 4000
Max.	: 20000

angle_of_attack	
Min.	:0.000
1st Qu.	:2.000
Median	:5.400
Mean	:6.782
3rd Qu.	:9.900
Max.	:22.200

chord_length	
Min.	: 0.0254
1st Qu.	: 0.0508
Median	: 0.1016
Mean	: 0.1365
3rd Qu.	: 0.2286
Max.	: 0.3048

free_stream_velocity	
Min.	: 31.70
1st Qu.	: 39.60
Median	: 39.60
Mean	: 50.86
3rd Qu.	: 71.30
Max.	: 71.30

displacement_thickness	
Min.	: 0.0004007
1st Qu.	: 0.0025351
Median	: 0.0049574
Mean	: 0.0111399
3rd Qu.	: 0.0155759
Max.	: 0.0584113

sound_level	
Min.	: 103.4
1st Qu.	: 120.2
Median	: 125.7
Mean	: 124.8
3rd Qu.	: 130.0
Max.	: 141.0

Histogram of each independent variable:

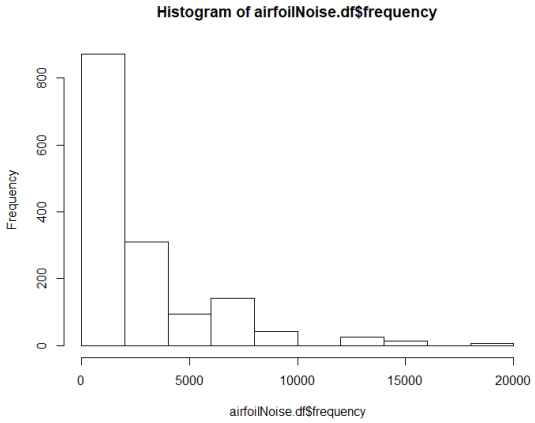


Figure 1: Histogram of Attribute Frequency

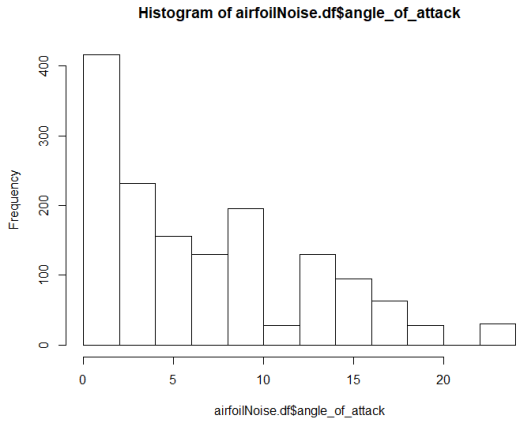


Figure 2: Histogram of Attribute Angle of Attack

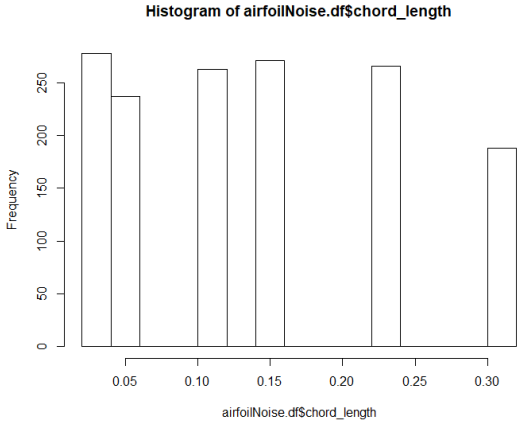


Figure 3: Histogram of Attribute Chord Length

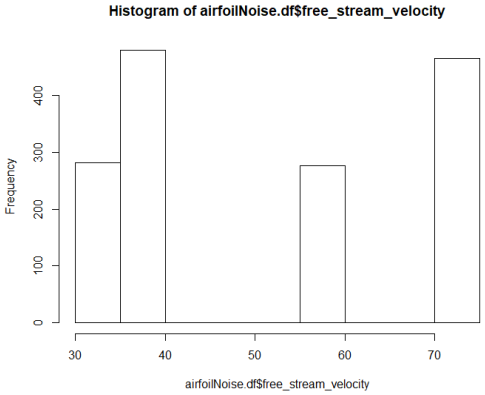


Figure 4: Histogram of Attribute Free stream velocity

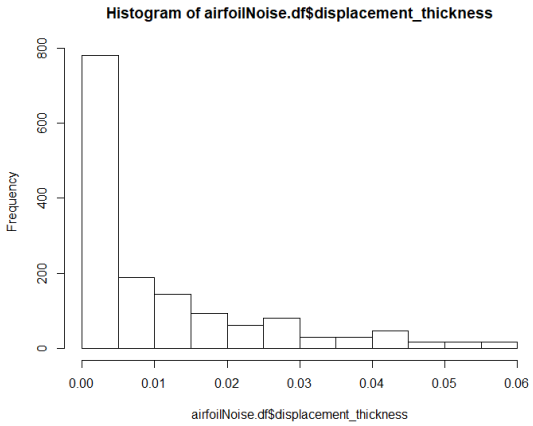


Figure 5: Histogram of Attribute Displacement Thickness

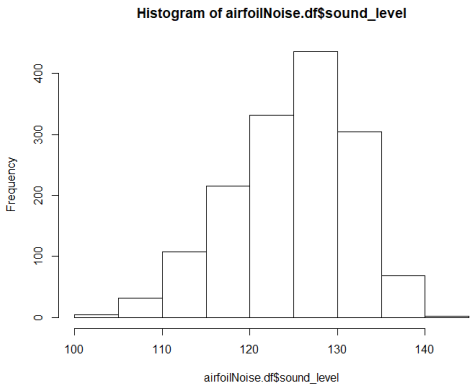
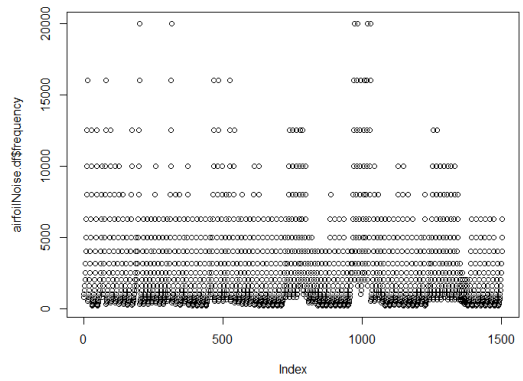


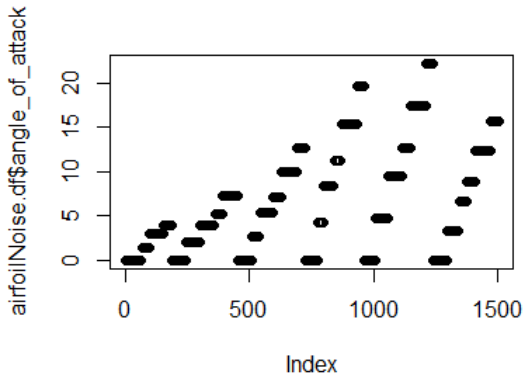
Figure 6: Histogram of Attribute Sound Level

Graphs of each independent variable:

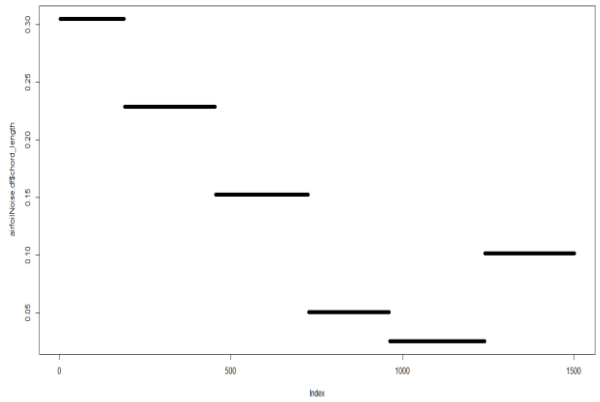
Frequency



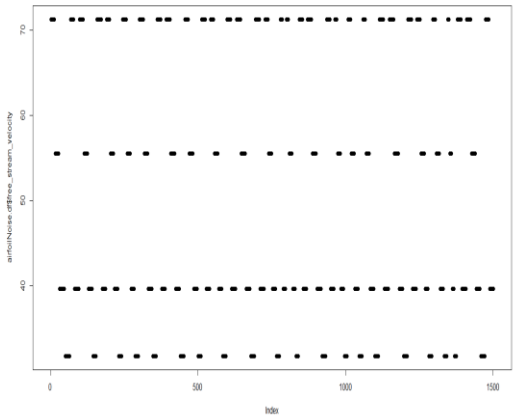
Angle of Attack



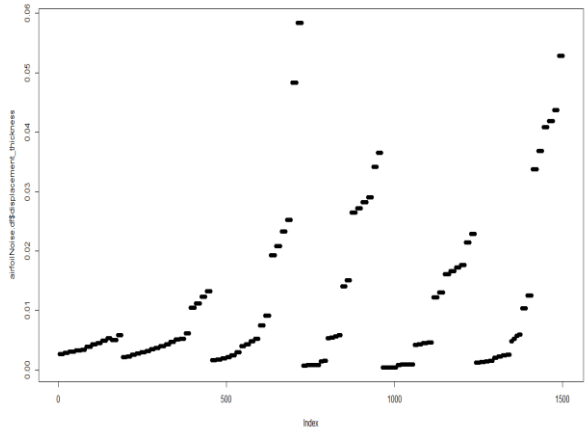
Chord Length



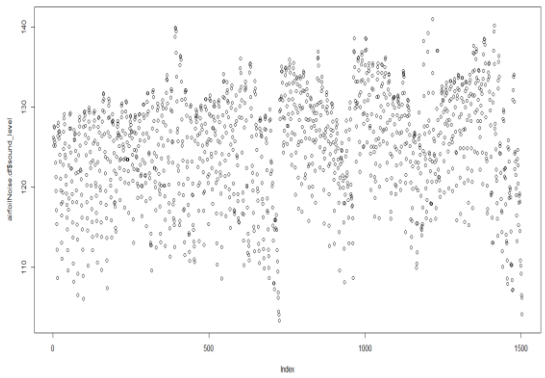
Free stream velocity



Displacement Thickness



Sound level



Why using this data set?

In the final project, I will choose option 1 that is, to build and evaluate support vector machine regression models. For this reason, I chose Airfoil Self-Noise Data Set.
This data set provides the airfoil self-noise which is an important component of the total airframe noise, which is due to the interaction between an airfoil blade and the turbulence produced in its own boundary layer and near wake. This data set provides different size NACA 0012 airfoils at various wind tunnel speeds and angles of attack. In this data set total instances are (rows) are 1503 and total attributes are (columns) are 6.
From the histograms it is seen that, most of the data set contains low frequencies and the target or, output variable has well distributed histogram.