**DATA ANALYSIS WITH PYTHON**

**MODULE 1: INTRODUCTION**

**İBRAHİM HALİL BAYAT**

**İSTANBUL TECHNICAL UNIVERSITY**

***Learning the Data***

The data is about Automobiles in year 1987. The data has 205 samples and 26 attributes/features. The features are:

1. *Symboling ( “symboling”)*

This feature is about risk rating. This feature can be -3, -2, -1, 0, 1, 2, and 3 where 3 indicates that the auto is risky based on the insurance and the price it indicates. –3 indicates that it is probably pretty safe. **Type: ‘int64’**

1. *Normalized Losses (“normalized-losses”)*

The sample’s normalized loss in use as compared to other cars. This feature is a continuos variable from 65 to 256. **Type: ‘object’**

1. *Make (“make”)*

Brand of the sample. This feature can be as follows and the **Type: ‘object’**

* alfa-romoro
* audi
* bmw
* chevrolet
* dodge
* hondo
* isuzu
* jaguar
* mazda
* mercedes-benz
* mercury
* mitsubishi
* nissan
* peugot
* plymouth
* porsche
* renault
* saab
* subaru
* toyota
* volkswagen
* volvo

1. *Fuel Type (“fuel-type”)*

This feature can be wheter “diesel” or “gas”. **Type: ‘object’**

1. *Aspiration (“aspiration”)*

This feature can be whether “std” or “turbo”. **Type: ‘object’**

1. *Number of Doors (“num-of-doors”)*

This feature can be whether “four” or “two”. **Type: ‘object’**

1. *Body Style (“body-style”)*

This feature can be as follows and **Type: ‘object’**

* hardtop
* wagon
* sedan
* hatchback
* convertible

1. *Drive Wheels (“drive-wheels”)*

This feature can be “4wd”, “fwd” or “rwd”. **Type: ‘object’**

1. *Engine Location (“engine-location”)*

This feature can be whether “front” or “rear”. **Type: ‘object’**

1. *Wheel Base (“wheel-base”)*

This feature is continus from 86.6 to 120.9. **Type: ‘float64’**

1. *Length (“length”)*

This feature is continuos from 141.1 to 208.1. **Type: ‘float64’**

1. Width (“width”)

This feature is continuos from 60.3 to 72.3. **Type: ‘float64’**

1. *Height (“height”)*

This feature is continuos from 47.8 to 59.8. **Type: ‘float64’**

1. *Curb Weight (“curb-weight”)*

This feature is continuos from 1488 to 4066. **Type: ‘int64’**

1. *Engine Type (“engine-type”)*

This feature can be “dohc”, “dohcv”, “I”, “ohc”, “ohcf”, “ohcv”, and “rotor”. **Type: ‘object’**

1. *Number of Cylnders (“num-of-cylinders”)*

This feature can be “eight”, “five”, “four”, “six”, “three”, “twelve”, and “two”. **Type: ‘object’**

*17. Engine Size (“engine-size”)*

This feature is continuous from 61 to 326. **Type: ‘int64’**

*18. Fuel System (“fuel-system”)*

This feature can be “1bbl”, “2bbl”, “4bbl”, “idi”, “mfi”, “mpfi”, “spdi”, and “spfi”. **Type: ‘object’**

*19. Bore (“bore”)*

This feature is continuos from 2.54 to 3.94. **Type: ‘object’**

*20. Stroke (“stroke”)*

Continuos from 2.07 to 4.17. **Type: ‘object’**

*21. Compression Ration (“compression-ration”)*

Cthis feature is continuos from 7 to 23. **Type: ‘float64’**

*22. Horse Power (“horsepower”)*

This feature is continuos from 48 to 288. **Type: ‘object’**

*23. Peak RPM (“peak-rpm”)*

This feature is continuos from 4150 to 6600. **Type: ‘object’**

*24. City MPG (“city-mpg”)*

This feature is continuos from 13 to 49. **Type: ‘int64’**

*25. Highway MPG (“highway-mpg”)*

This feature is contnuous from 16 to 54. **Type: ‘int64’**

*26. Price (“price”)*

This feature is continuos from 5118 to 45400. **Type: ‘object’**