Final Report (Output)

**----------------------Statistical derivations----------------------**

Max Min Mean Median Standard Deviation

Length 0.815 0.075 0.523992 0.545 0.120093

Diameter 0.65 0.055 0.407881 0.425 0.0992399

Height 1.13 0 0.139516 0.14 0.0418271

Whole weight 2.8255 0.002 0.828742 0.7995 0.490389

Shucked weight 1.488 0.001 0.359367 0.336 0.221963

Viscera weight 0.76 0.0005 0.180594 0.171 0.109614

Shell weight 1.005 0.0015 0.238831 0.234 0.139203

Rings 29 1 9.93368 9 3.22417

Age 30.5 2.5 11.4337 10.5 3.22417

-------------------------------------------------------------------

**-----------------Correlation between different attributes------------------**

Length Diameter Height Whole weight Shucked weight \

Length 1.000000 0.986812 0.827554 0.925261 0.897914

Diameter 0.986812 1.000000 0.833684 0.925452 0.893162

Height 0.827554 0.833684 1.000000 0.819221 0.774972

Whole weight 0.925261 0.925452 0.819221 1.000000 0.969405

Shucked weight 0.897914 0.893162 0.774972 0.969405 1.000000

Viscera weight 0.903018 0.899724 0.798319 0.966375 0.931961

Shell weight 0.897706 0.905330 0.817338 0.955355 0.882617

Rings 0.556720 0.574660 0.557467 0.540390 0.420884

Age 0.556720 0.574660 0.557467 0.540390 0.420884

Viscera weight Shell weight Rings Age

Length 0.903018 0.897706 0.556720 0.556720

Diameter 0.899724 0.905330 0.574660 0.574660

Height 0.798319 0.817338 0.557467 0.557467

Whole weight 0.966375 0.955355 0.540390 0.540390

Shucked weight 0.931961 0.882617 0.420884 0.420884

Viscera weight 1.000000 0.907656 0.503819 0.503819

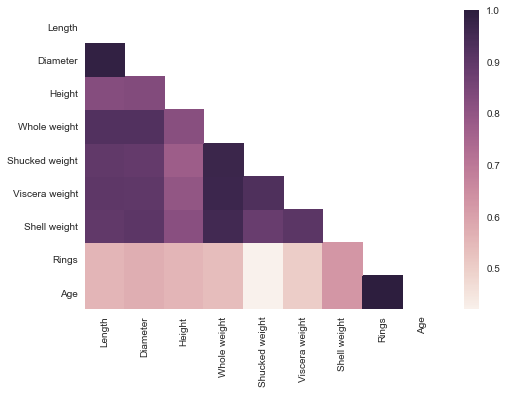
Shell weight 0.907656 1.000000 0.627574 0.627574

Rings 0.503819 0.627574 1.000000 1.000000

Age 0.503819 0.627574 1.000000 1.000000

---------------------------------------------------------------------------

**------------------Correlation Heatmap-------------------**



--------------------------------------------------------

**-------------Outlier Detection using IQR---------------**

**Outlier Detection on the basis of Height**

Outlier Index = [236, 237, 238, 239, 306, 694, 718, 719, 720, 1174, 1257, 1417, 1428, 1429, 1763, 1987, 2051, 2114, 2169, 2171, 2172, 2179, 2381, 2711, 3190, 3837, 3899, 3902, 3996]

Outlier Value = [0.01, 0.03, 0.03, 0.035, 0.03, 0.02, 0.035, 0.025, 0.025, 0.015, 0.0, 0.515, 0.25, 0.035, 0.25, 0.025, 1.13, 0.035, 0.015, 0.03, 0.03, 0.25, 0.025, 0.03, 0.025, 0.035, 0.035, 0.02, 0.0]

**Outlier Detection on the basis of Diameter**

Outlier Index = [43, 44, 148, 149, 236, 237, 238, 239, 305, 306, 321, 465, 520, 523, 525, 526, 611, 694, 696, 718, 719, 720, 892, 898, 1053, 1054, 1055, 1429, 1824, 1986, 1987, 1988, 2114, 2115, 2169, 2171, 2172, 2343, 2371, 2380, 2381, 2458, 2545, 2711, 2712, 3141, 3143, 3190, 3318, 3380, 3472, 3473, 3521, 3600, 3716, 3837, 3899, 3902, 3994]

Outlier Value = [0.15, 0.15, 0.13, 0.13, 0.055, 0.1, 0.09, 0.12, 0.145, 0.12, 0.145, 0.125, 0.15, 0.14, 0.125, 0.11, 0.145, 0.11, 0.105, 0.125, 0.1, 0.11, 0.14, 0.12, 0.13, 0.15, 0.135, 0.105, 0.135, 0.13, 0.11, 0.15, 0.095, 0.13, 0.115, 0.13, 0.15, 0.125, 0.125, 0.135, 0.115, 0.145, 0.15, 0.14, 0.15, 0.135, 0.15, 0.145, 0.125, 0.13, 0.12, 0.15, 0.15, 0.125, 0.13, 0.105, 0.105, 0.12, 0.135]

**Outlier Detection on the basis of Length**

Outlier Index = [148, 149, 236, 237, 238, 239, 305, 306, 321, 465, 523, 525, 526, 611, 694, 696, 718, 719, 720, 1053, 1054, 1055, 1056, 1210, 1429, 1824, 1986, 1987, 2114, 2115, 2169, 2171, 2343, 2371, 2380, 2381, 2458, 2711, 3141, 3143, 3190, 3318, 3380, 3472, 3600, 3837, 3899, 3902, 3994]

Outlier Value = [0.175, 0.17, 0.075, 0.13, 0.11, 0.16, 0.2, 0.165, 0.19, 0.175, 0.2, 0.175, 0.155, 0.195, 0.165, 0.155, 0.18, 0.15, 0.16, 0.185, 0.195, 0.195, 0.2, 0.185, 0.14, 0.185, 0.135, 0.16, 0.13, 0.18, 0.165, 0.19, 0.17, 0.18, 0.175, 0.155, 0.2, 0.19, 0.18, 0.2, 0.2, 0.165, 0.19, 0.165, 0.175, 0.17, 0.14, 0.16, 0.185]

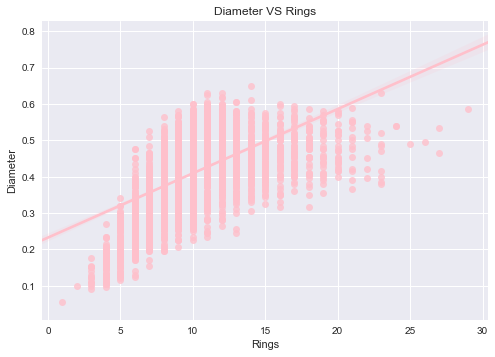
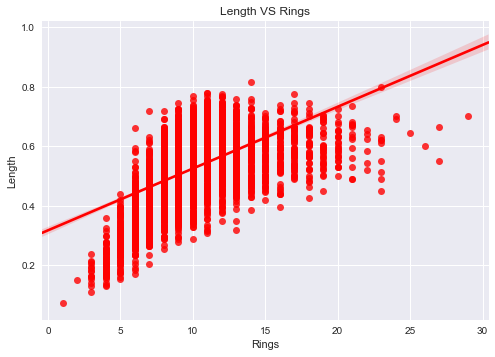
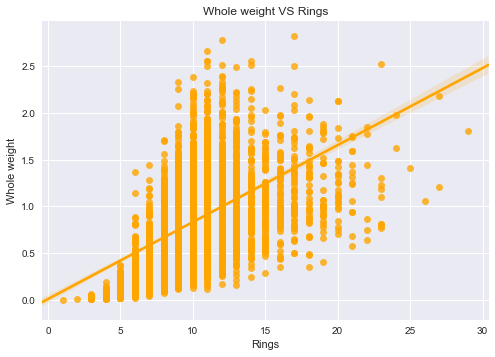
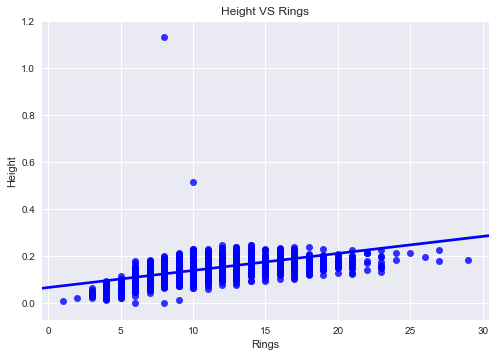
**Outlier Detection on the basis of Whole weight**

Outlier Index = [165, 358, 891, 1051, 1052, 1193, 1206, 1207, 1209, 1426, 1427, 1428, 1761, 1762, 1763, 2265, 2334, 2623, 2624, 2811, 2862, 2863, 3007, 3008, 3188, 3427, 3599, 3715, 3800, 3993]

Outlier Value = [2.55, 2.499, 2.8255, 2.555, 2.302, 2.273, 2.2359999999999998, 2.505, 2.657, 2.25, 2.5085, 2.255, 2.2205, 2.5155, 2.7795, 2.226, 2.526, 2.381, 2.398, 2.3305, 2.2355, 2.3235, 2.2305, 2.2635, 2.2695, 2.4925, 2.333, 2.548, 2.381, 2.2385]

--------------------------------------------------------------------

**----------Scatter plots with Regression line-----------**



--------------------------------------------------------

**-------------------------Encoding of categorical features------------------**

**Original data:-**

Sex Length Diameter Height Whole weight Shucked weight Viscera weight \

0 M 0.455 0.365 0.095 0.5140 0.2245 0.1010

1 M 0.350 0.265 0.090 0.2255 0.0995 0.0485

2 F 0.530 0.420 0.135 0.6770 0.2565 0.1415

3 M 0.440 0.365 0.125 0.5160 0.2155 0.1140

4 I 0.330 0.255 0.080 0.2050 0.0895 0.0395

Shell weight Rings Age

0 0.150 15 16.5

1 0.070 7 8.5

2 0.210 9 10.5

3 0.155 10 11.5

4 0.055 7 8.5

**Encoded data with the Sex attibute encoded (categorical to numeric feature):-**

Sex Length Diameter Height Whole weight Shucked weight \

0 2 0.455 0.365 0.095 0.5140 0.2245

1 2 0.350 0.265 0.090 0.2255 0.0995

2 0 0.530 0.420 0.135 0.6770 0.2565

3 2 0.440 0.365 0.125 0.5160 0.2155

4 1 0.330 0.255 0.080 0.2050 0.0895

Viscera weight Shell weight Rings

0 0.1010 0.150 15

1 0.0485 0.070 7

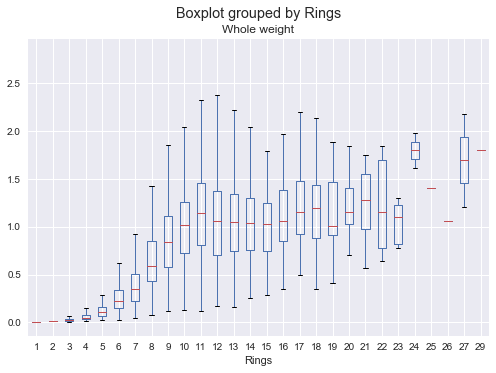
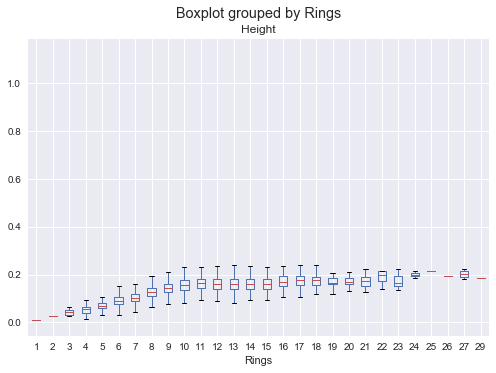
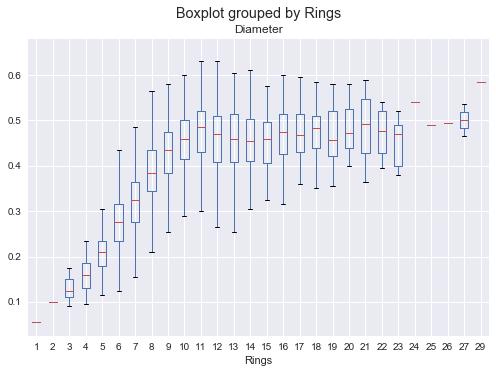
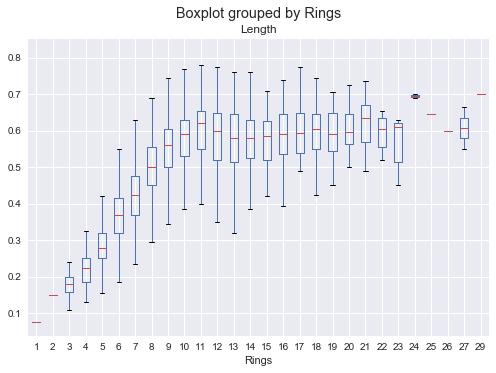
2 0.1415 0.210 9

3 0.1140 0.155 10

4 0.0395 0.055 7

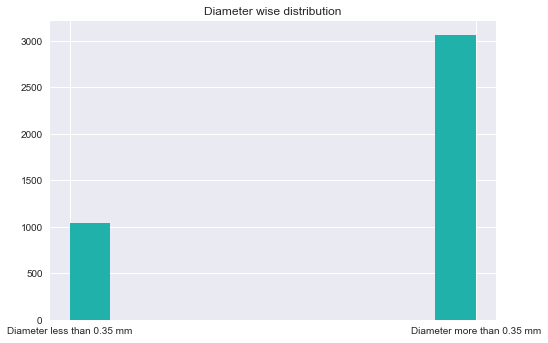
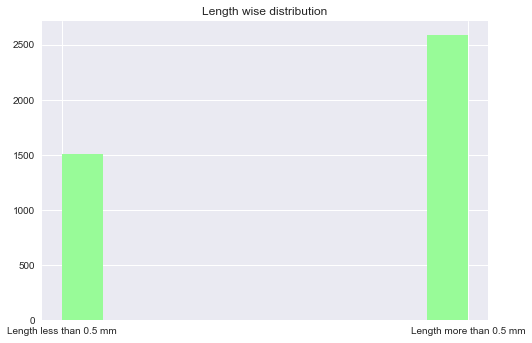
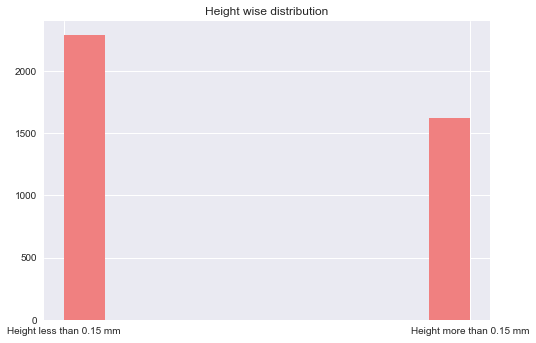
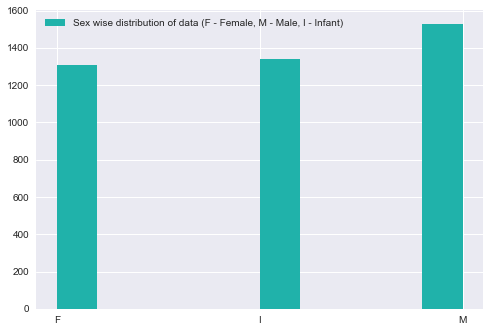
---------------------------------------------------------------------------

**-------------------------Box plots--------------------------**



------------------------------------------------------------

**-------------------------Histograms--------------------------**



------------------------------------------------------------