HOMEWORK – 5

NAME: RUCHIN PATEL
USC ID: 9520665364

Q-1 (MEAN ABSOLUTE ERROR)(0 points WEKA)

ANS: 1.5835

****

Q-2 (EQUATION FOR NUM_RINGS USING ALL ATTRIBUTES)
(3 points WEKA)

ANS: num_rings =   0.8607 * sex=M,F +
10.5383 * diameter +
10.7251 * height +
8.9743 * whole_weight +
-19.769 * shucked_weight +
-10.6481 * viscera_weight +
8.7497 * shell_weight +
3.0551

FIGURE:

```
Linear Regression Model

num_rings =

      0.8607 * sex=M,F +
     10.5383 * diameter +
     10.7251 * height +
      8.9743 * whole_weight +
    -19.769  * shucked_weight +
    -10.6481 * viscera_weight +
      8.7497 * shell_weight +
      3.0551

Time taken to build model: 0.04 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                 0.7268
Mean absolute error                     1.5835
Root mean squared error                 2.2147
Relative absolute error                67.023  %
Root relative squared error            68.686  %
Total Number of Instances            4177
```

****

Q-3 (EQUATION FOR NUM_RINGS USING length,diameter,whole_weight,num_rings)*(1 point WEKA)*

ANS: num_rings = -11.8042 * length +
29.8645 * diameter +
0.6345 * whole_weight +
3.412

FIGURE:

```
num_rings =

    -11.8042 * length +
     29.8645 * diameter +
      0.6345 * whole_weight +
      3.412


Time taken to build model: 0.02 seconds

=== Cross-validation ===
=== Summary ===

Correlation coefficient                    0.5785
Mean absolute error                        1.9117
Root mean squared error                    2.6295
Relative absolute error                   80.9118 %
Root relative squared error               81.5515 %
Total Number of Instances             4177
```

****

Q-4 (EQUATION FOR NUM_RINGS USING ALL ATTRIBUTES AND COMPARE WITH WEKA OUTPUT) *(2+1=3 points KNIME)*
ANS:

- **Linear regression equation:**

$$num\_rings = (-0.8249 * sex=I) + (0.0577 * sex=M) +$$
$$-0.4583 * length +$$
$$11.0751 * diameter +$$
$$10.7615 * height +$$
$$8.9754 * whole\_weight +$$
$$-19.7869 * shucked\_weight +$$
$$-10.5818 * viscera\_weight +$$
$$8.7418 * shell\_weight +$$
$$3.8946$$

- **Comparing to Weka's output that have parameters that differ at-most by 0.5**

  1. **length :** Differs by 0.4583
  2. **height :** Differs by 0.0364
  3. **whole_weight :** Differs by 0.0011
  4. **shucked_weight :** Differs by 0.0179
  5. **viscera_weight :** Differs by 0.0663
  6. **shell_weight :** Differs by 0.0079

FIGURE: Design:
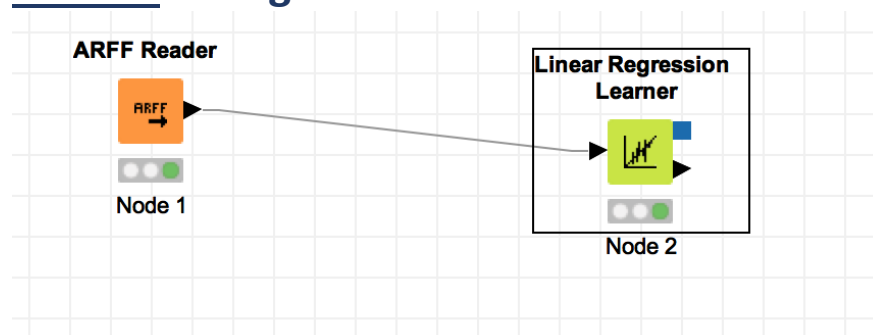
Result:

Statistics on Linear Regression

Variable	Coeff.	Std. Err.	t-value	P>|t|
sex=I	-0.8249	0.1024	-8.0558	1.11E-15
sex=M	0.0577	0.0833	0.6925	0.4887
length	-0.4583	1.8091	-0.2533	0.8
diameter	11.0751	2.2273	4.9725	6.88E-7
height	10.7615	1.5362	7.0053	2.86E-12
whole_weight	8.9754	0.7254	12.373	0.0
shucked_weight	-19.7869	0.8174	-24.2086	0.0
viscera_weight	-10.5818	1.2937	-8.1792	4.44E-16
shell_weight	8.7418	1.1247	7.7723	9.55E-15
Intercept	3.8946	0.2916	13.3576	0.0

Multiple R-Squared: 0.5379
Adjusted R-Squared: 0.5369

****

Q-5 (DECISION TREE LEARNER)(1 point KNIME)
ANS:

SCREENSHOT:

FIGURE: Design:

****
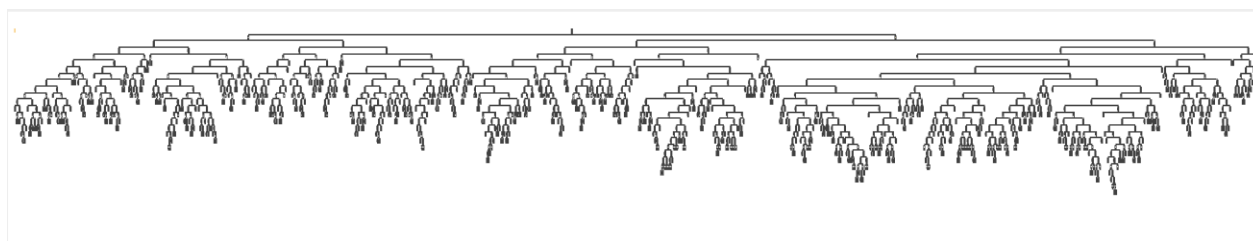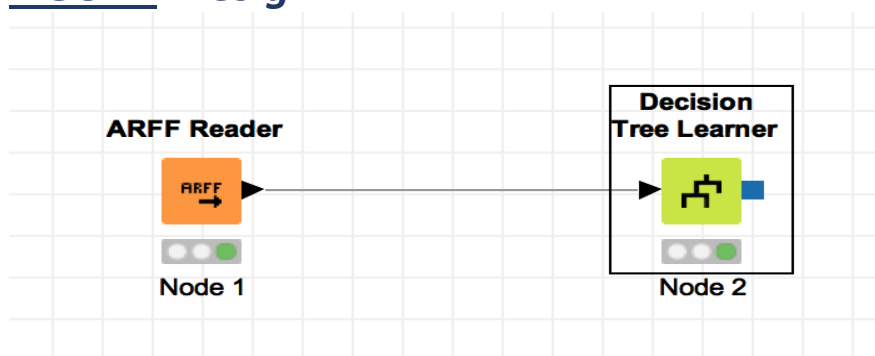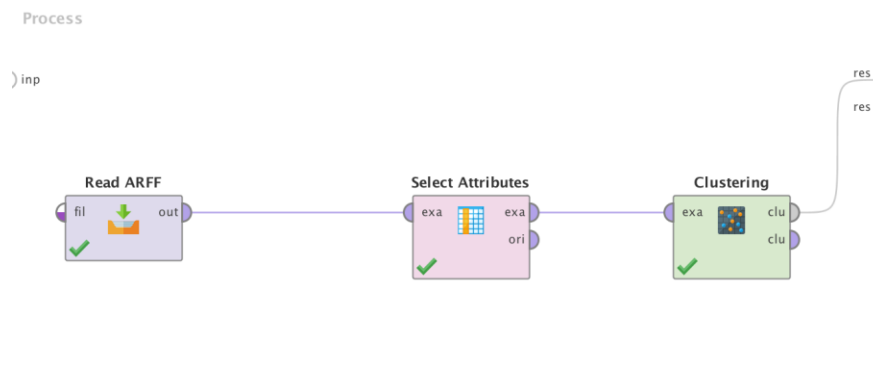
BONUS QUESTION:

1. HOW MANY POINTS IN EACH CLUSTER

ANS: **Cluster 0: 634 items**
        Cluster 1: 754 items
        Cluster 2: 499 items
        Cluster 3: 1257 items
        Cluster 4: 194 items
        Cluster 5: 839 items
        Total number of items: 4177

FIGURES: Design:

Result:

Cluster Model

```
Cluster 0: 634 items
Cluster 1: 754 items
Cluster 2: 499 items
Cluster 3: 1257 items
Cluster 4: 194 items
Cluster 5: 839 items
Total number of items: 4177
```

2. Linear Regression to predict num_rings using length diameter and height

ANS: num_rings = -11.933 * length +
                    25.766 * diameter +
                    20.358 * height +
                    2.836

FIGURE: Design:

Read ARFF → Select Attributes → Set Role → Linear Regression
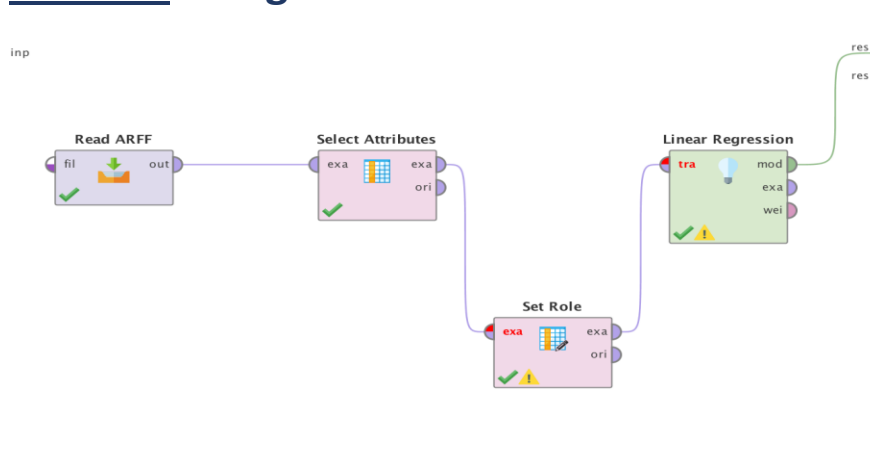
Result:

Attribute	Coeffici...	Std. Error	Std. Coe...	Tolerance	t-Stat	p-Value	Code
length	-11.933	2.064	-0.444	0.078	-5.781	0.000	****
diameter	25.766	2.539	0.793	0.094	10.147	0	****
height	20.358	1.737	0.264	0.319	11.719	0	****
(Intercept)	2.836	0.186	?	?	15.243	0	****

***********************END*************************
****