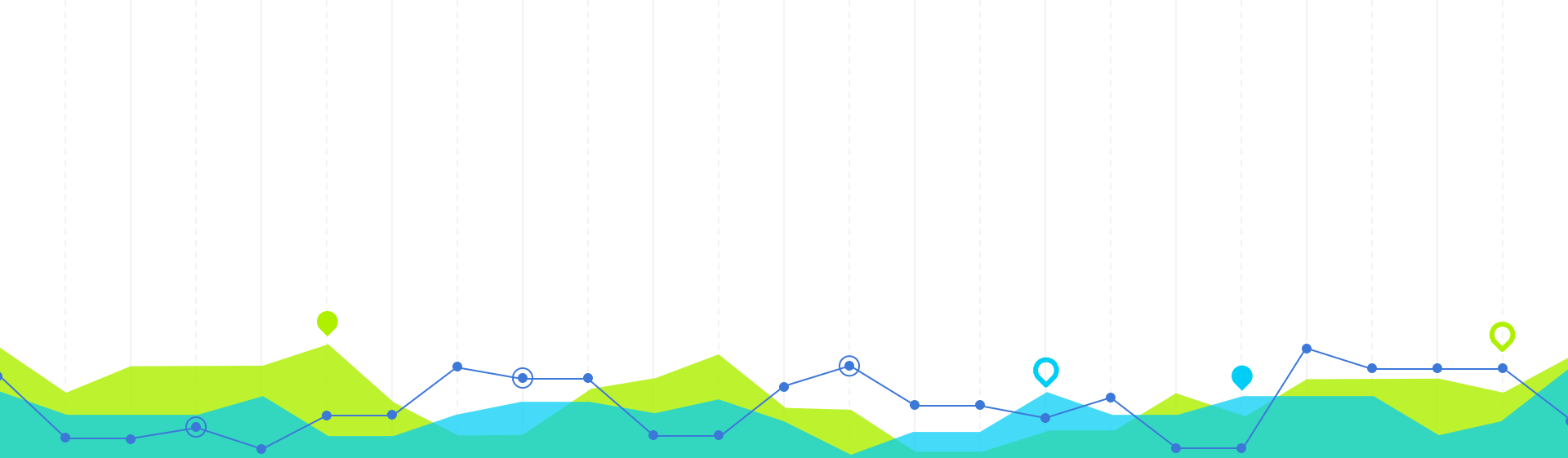


Predict outcome of individual items located in groups to classify iris species.



What does this dataset contains?

TASKS TO PERFORM

- Load the Iris Dataset
- Exploring the dataset
- Build a Decision Tree model for data analysis
- Analyze the rules
- Plot the Decision Tree
- Evaluation of Model: Confusion Matrix

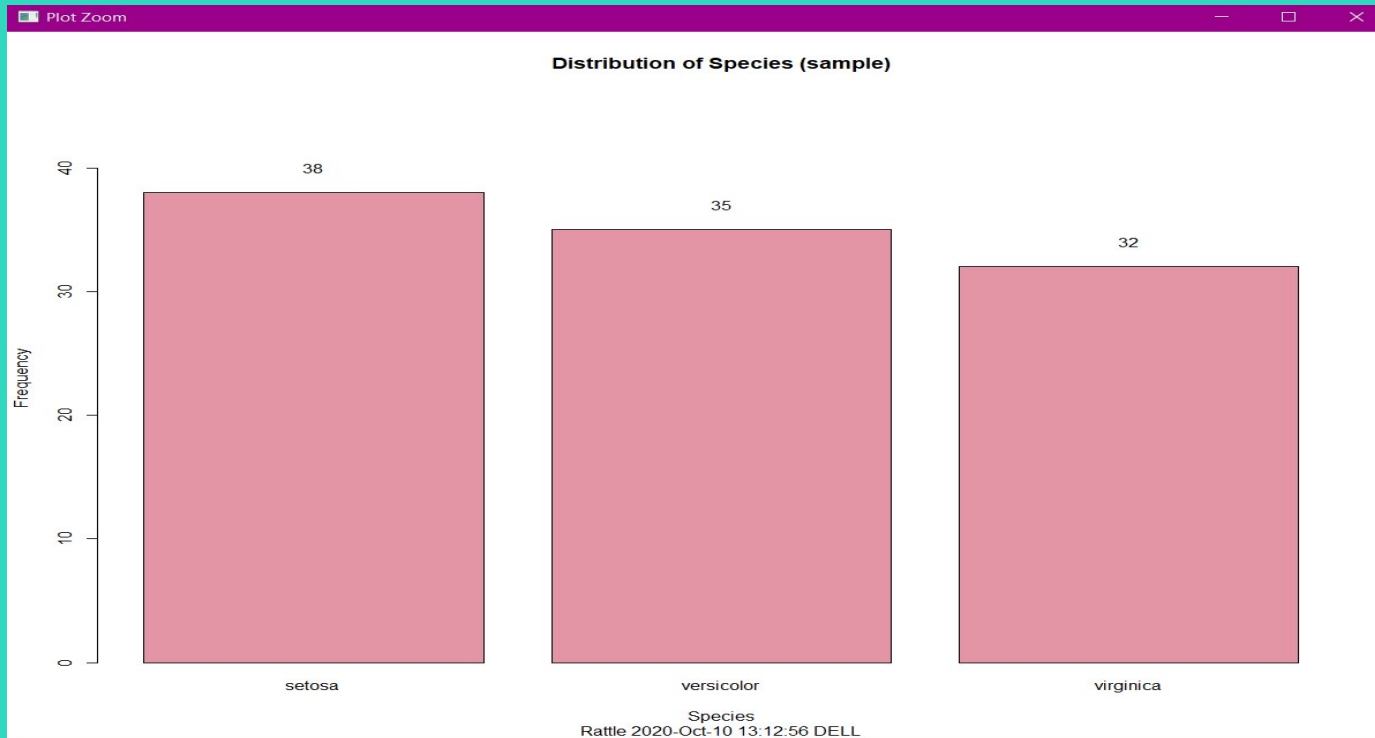
Source: ☐ File ☐ ARFF ☐ ODBC ☐ R Dataset ☐ RData File ☒ Library ☐ Corpus ☐ Script

Data Name: iris:datasets:Edgar Anderson's Iris Data

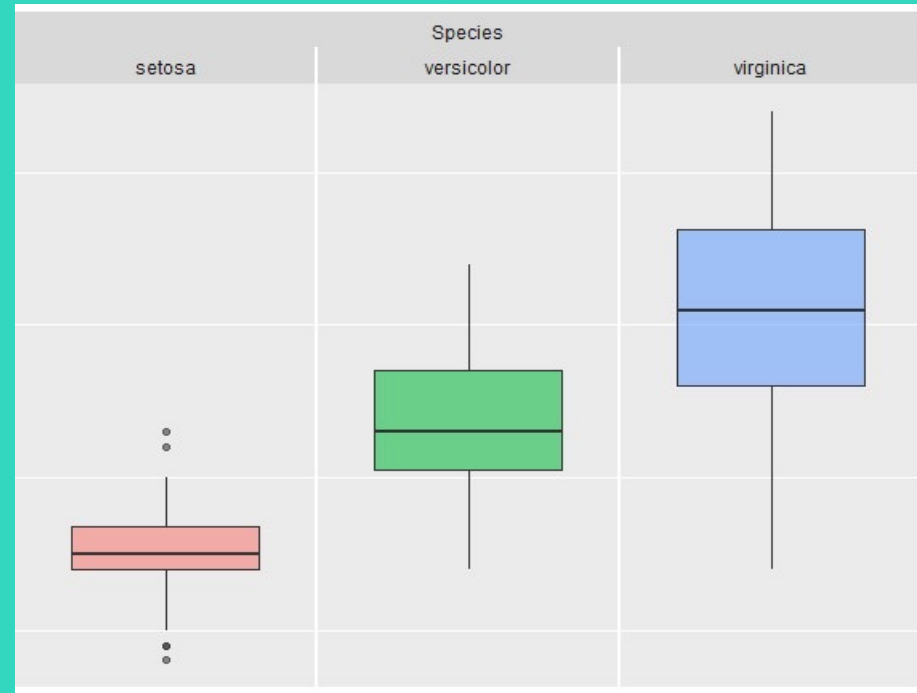
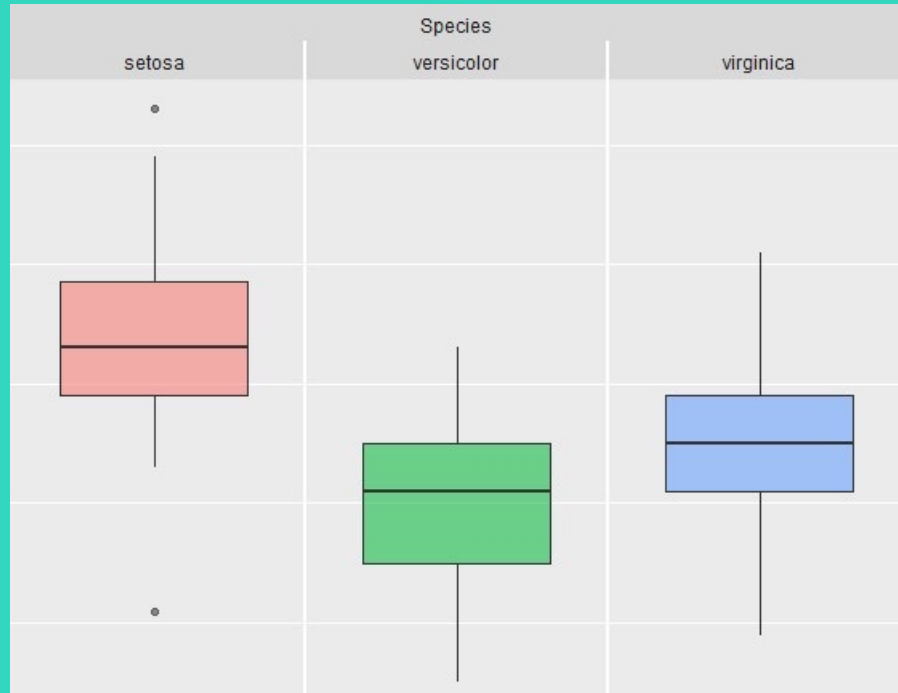
☒ Partition 70/15/15 Seed: 42 View EditInput ☒ Ignore Weight Calculator: Target Data Type: ☒ Auto ☐ Categorical ☐ Numeric ☐ Survival

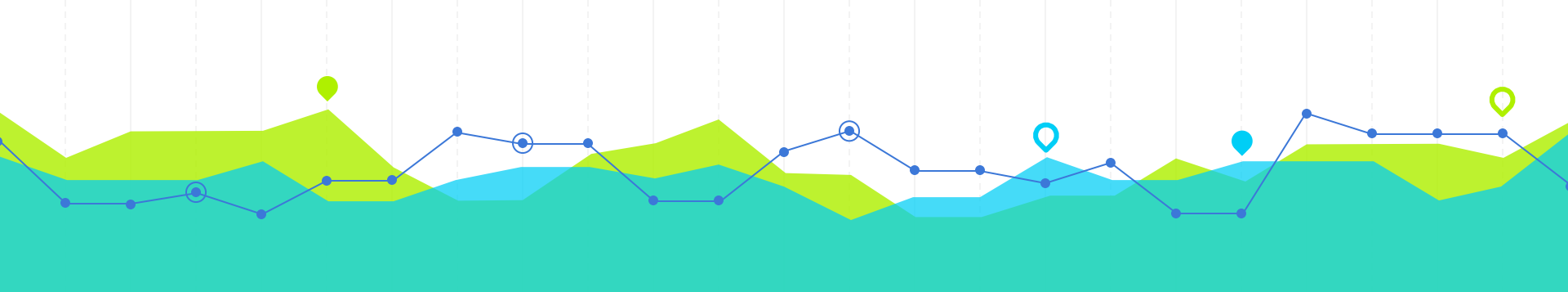
No.	Variable	Data Type	Input	Target	Risk	Ident	Ignore	Weight	Comment
1	Sepal.Length	Numeric	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Unique: 35
2	Sepal.Width	Numeric	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Unique: 23
3	Petal.Length	Numeric	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Unique: 43
4	Petal.Width	Numeric	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Unique: 22
5	Species	Categorical	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Unique: 3

LOADING THE DATASET

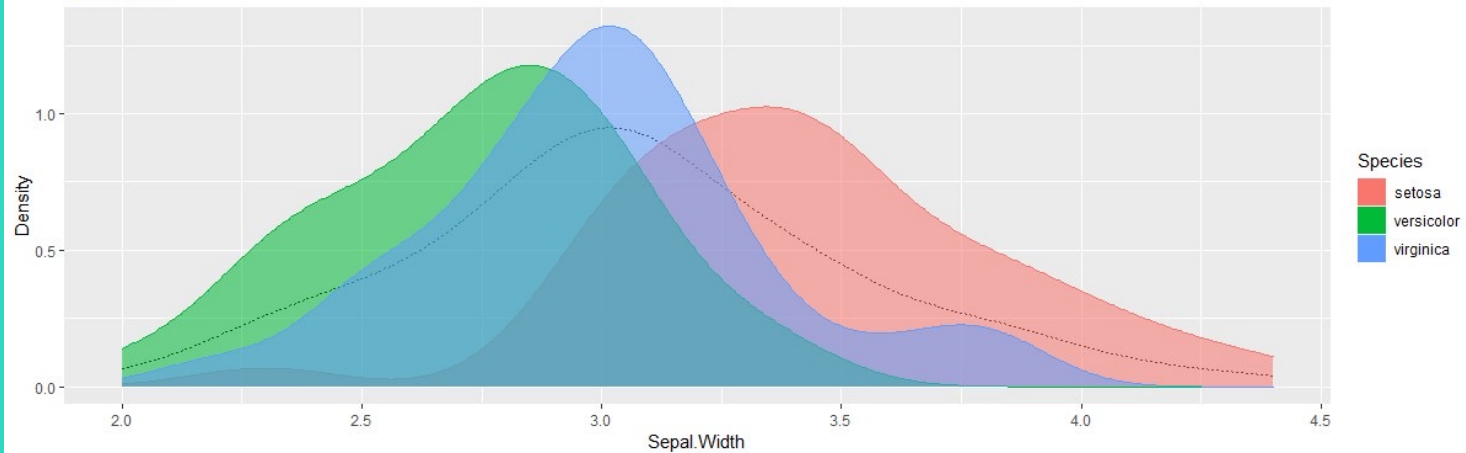


EXPLORING THE DATASET.....

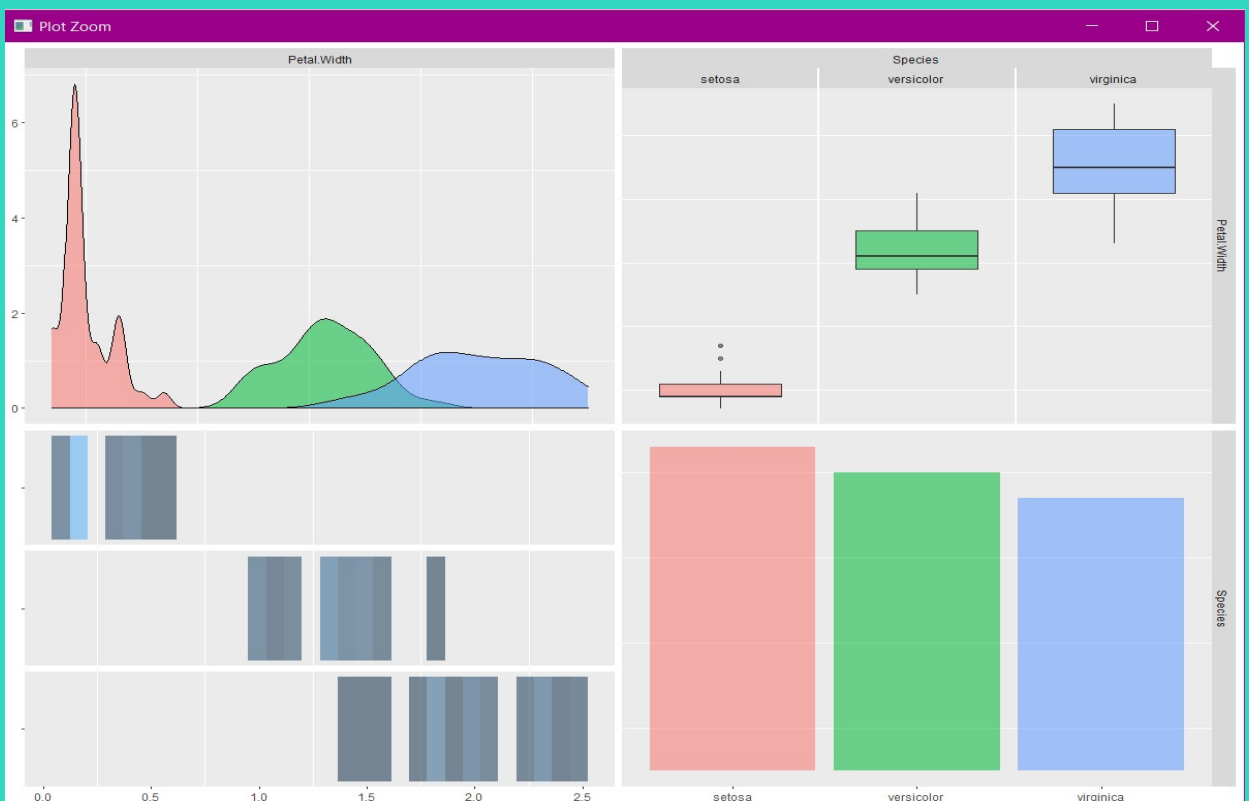
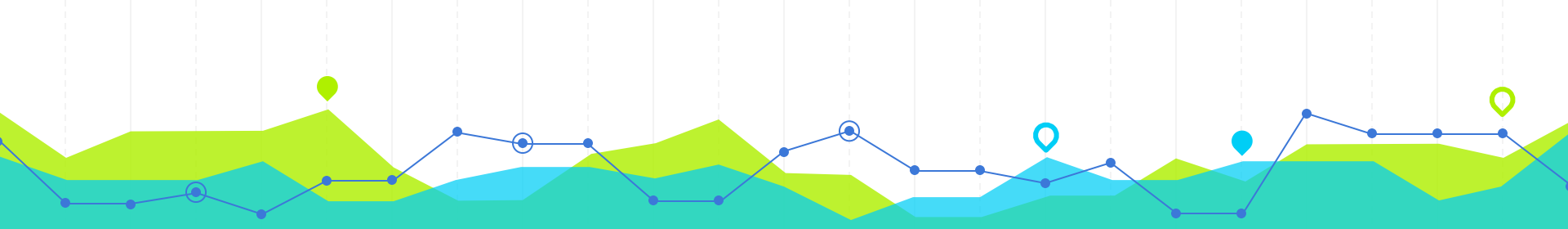




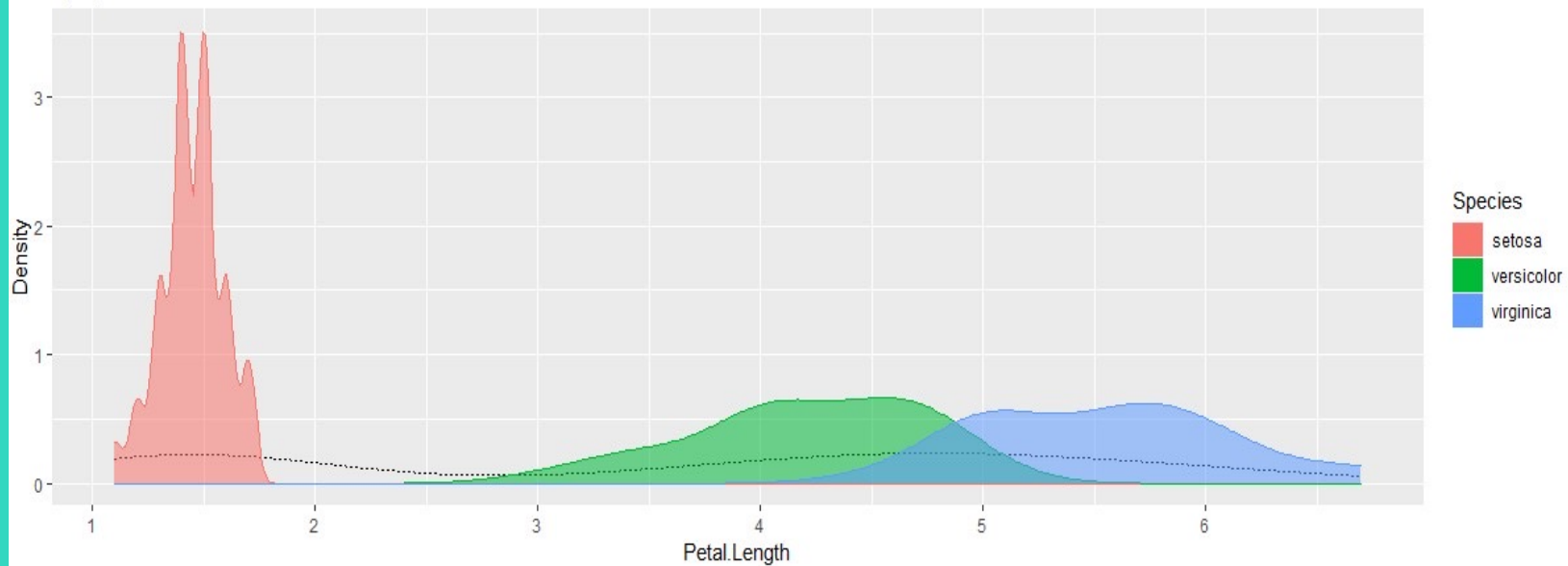
Distribution of Sepal.Width (sample)
by Species



Rattle 2020-Oct-09 16:12:31 DELL



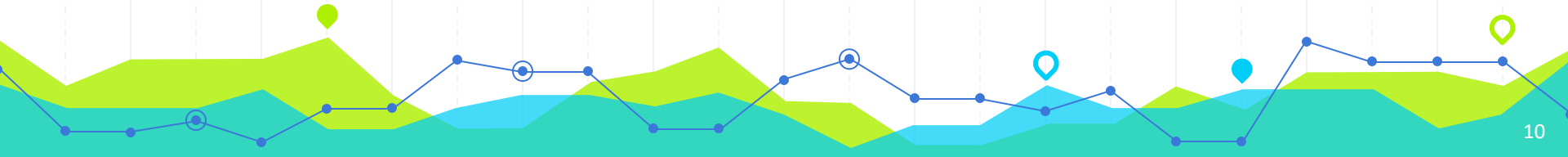
Distribution of Petal.Length (sample)
by Species



Rattle 2020-Oct-10 02:04:36 DELL



ALGORITHM





Data Explore Test Transform Cluster Associate Model Evaluate Log

Type: ☒ Tree ☐ Forest ☐ Boost ☐ SVM ☐ Linear ☐ Neural Net ☐ Survival ☐ All

Target: Species Algorithm: ☒ Traditional ☐ Conditional

Min Split: 20 Max Depth: 30 Priors:

Min Bucket: 7 Complexity: 0.0100 Loss Matrix:

Summary of the Decision Tree model for Classification (built using 'rpart'):

n= 105

```
node), split, n, loss, yval, (yprob)
* denotes terminal node
```

```
1) root 105 67 setosa (0.36190476 0.33333333 0.30476190)
2) Petal.Length< 2.35 38 0 setosa (1.00000000 0.00000000 0.00000000) *
3) Petal.Length>=2.35 67 32 versicolor (0.00000000 0.52238806 0.47761194)
6) Petal.Width< 1.65 37 3 versicolor (0.00000000 0.91891892 0.08108108) *
7) Petal.Width>=1.65 30 1 virginica (0.00000000 0.03333333 0.96666667) *
```

Classification tree:

```
rpart(formula = Species ~ ., data = crs$dataset[crs$train, c(crs$input,
crs$target)], method = "class", model = TRUE, parms = list(split = "information"),
control = rpart.control(usesurrogate = 0, maxsurrogate = 0))
```

Variables actually used in tree construction:

[1] Petal.Length Petal.Width

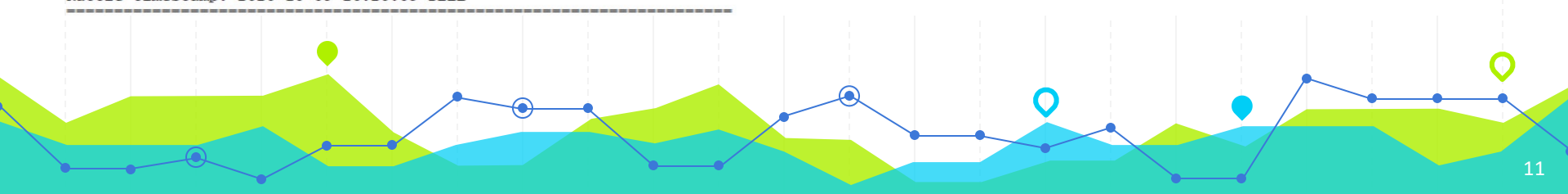
Root node error: 67/105 = 0.6381

n= 105

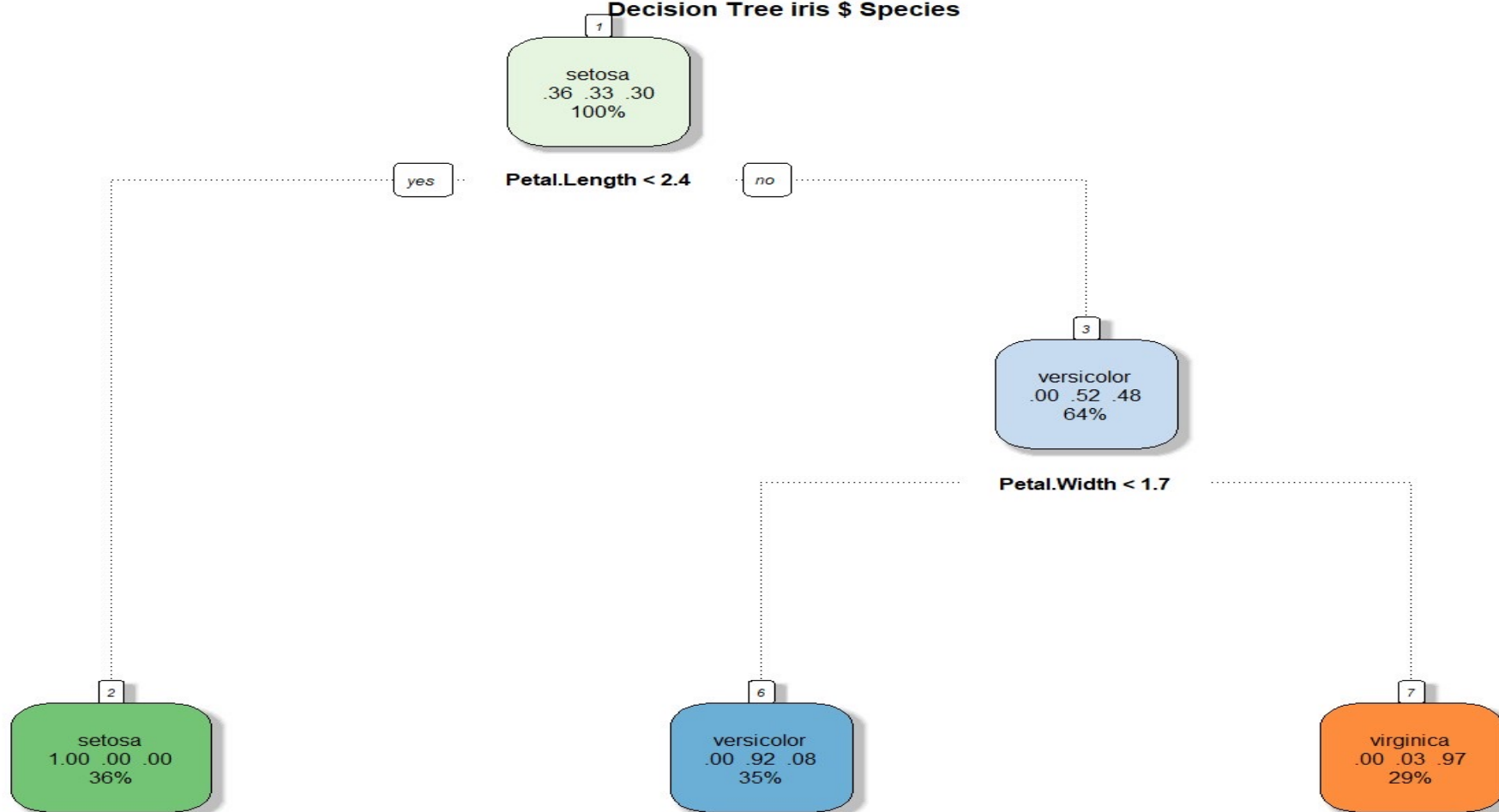
	CP	nsplit	rel error	xerror	xstd
1	0.52239	0	1.000000	1.05970	0.071565
2	0.41791	1	0.477612	0.53731	0.072595
3	0.01000	2	0.059701	0.11940	0.040575

Time taken: 0.03 secs

Rattle timestamp: 2020-10-09 16:20:09 DELL



Decision Tree iris \$ Species





EVALUATION OF MODEL: CONFUSION MATRIX



Type: ☒ Error Matrix ☐ Risk ☐ Cost Curve ☐ Hand ☐ Lift ☐ ROC ☐ Precision ☐ Sensitivity ☐ Pr v Ob ☐ ScoreModel: ☒ Tree ☐ Boost ☐ Forest ☐ SVM ☐ Linear ☐ Neural Net ☐ Survival ☐ KMeans ☐ HClustData: ☐ Training ☐ Validation ☐ Testing ☒ Full ☐ Enter ☐ CSV File ☐ R DatasetRisk Variable: Report: ☒ Class ☐ Probability Include: ☒ Identifiers ☐ All

Error matrix for the Decision Tree model on iris (counts):

Actual	Predicted			Error
	setosa	versicolor	virginica	
setosa	50	0	0	0
versicolor	0	48	2	4
virginica	0	4	46	8

Error matrix for the Decision Tree model on iris (proportions):

Actual	Predicted			Error
	setosa	versicolor	virginica	
setosa	33.3	0.0	0.0	0
versicolor	0.0	32.0	1.3	4
virginica	0.0	2.7	30.7	8

Overall error: 4%, Averaged class error: 4%

Rattle timestamp: 2020-10-10 13:33:16 DELL