



Case Study Report

For the Subject

**Data Warehousing Data Mining
(Trimester VIII)**

Submitted by

Shivansh Kushwaha

S1032180745

Sourabh Gavhane

S1032180909

Pratik Singh

S1032181204

Saish Malunde

S1032181324

Under the Guidance of

Prof. Sheetal Girase

**School of Computer Engineering and Technology
MIT World Peace University, Kothrud,
Pune 411 038, Maharashtra - India
2020-2021**

List of Figures

Figure 1.1: File widget dropped on the Canvas.

Figure 2.1: Workflow on canvas.

Figure 3.1: Python Script widget on canvas.

Figure 3.2: Preprocess widget on canvas

Figure 3.3: Python Script widget on canvas, running clustering algorithm.

Figure 4.1: Animal Rescue Centre Dataset.

Figure 4.2: Iris Dataset.

Figure 6.1: Scatter Plot.

Figure 6.2: Bar Plot (Date Intake).

Figure 6.3: Bar Plot (Intake Type).

Figure 6.4: Bar Plot (Stay Length).

Figure 6.5: Bar Plot (Animal Type).

Figure 6.6: Bar Plot (Gender).

Figure 6.7: Bar Plot (Age).

Figure 6.8: Histogram (Date Intake).

Contents

List of Figures	I
-----------------------	---

1	Introduction of the Tool: Orange		
	1.1	Features	
	1.2	Applications	
2	How to perform Supervised/Unsupervised Learning		
	2.1	Supervised Learning	
	2.2	Unsupervised Learning	
	2.3	Semi-supervised Learning	
3	How to perform Pre-processing/ Clustering.		
	3.1	Pre-Processing	
	3.2	Clustering	
4	Description of the Dataset Used		
		Animal Rescue Dataset	
		Iris Dataset	
5	Implementation		
6	Visualization		
7	Conclusion		
		Strengths & Weakness of the tool	
8	References		

Chapter 1

Introduction of the Tool

Orange data mining software was originally developed by scientists at the University of Ljubljana (Slovenia) in 1997 using the Python, Cython, C++ and C programming languages. The software's graphical environment and interfaces have been developed using the Python and Qt3 libraries (Demsar et al.,2013). This software, with the latest version presents on March 6, 2017 with Orange 3.4.0 has a simple interface on which users create a data analysis workflow by placing graphical components(widgets).

Orange is an open-source data mining tool with very strong data visualization capabilities. It allows you to use a GUI (Orange Canvas) to drag and drop modules and connect them to evaluate and test various machine learning algorithms on your data. This hands-on tutorial will go through setting up Orange and getting familiar with its GUI components. We do this by exploring a sample data set with some visualization widgets included with Orange.

1.1 Features:

1.1.1 **Canvas:** Graphical front-end for data analysis.

1.1.2 **Widgets:**

- **Data:** Widgets for data input, data filtering, sampling, imputation, feature manipulation.
- **Visualize:** Widgets for common visualization and multivariate visualization.
- **Classify:** A set of supervised machines learning algorithms.
- **Regression:** A set of supervised machines learning algorithms.
- **Evaluate:** Cross-validation, sampling-based procedures, reliability estimation and scoring of prediction methods.
- **Unsupervised:** Unsupervised learning algorithms for clustering and data projection techniques.
- **Add-ons:**
 - **Associate:** Widgets for mining frequent item sets and association rule learning.

- **Bioinformatics:** Widgets for gene set analysis, enrichment and access to pathway libraries.
- **Data fusion:** Widgets for fusing different data sets, collective matrix factorization of latent factors.
- **Educational:** Widgets for teaching machine learning concepts.
- **Geo:** Widgets for working with geospatial data.
- **Image analytics:** Widgets for working with images and ImageNet embeddings.
- **Network:** Widgets for graph and network analysis.
- **Text mining:** Widgets for natural language processing (NLP) and text mining.
- **Spectroscopy:** Widgets for analyzing and visualization of spectral datasets.

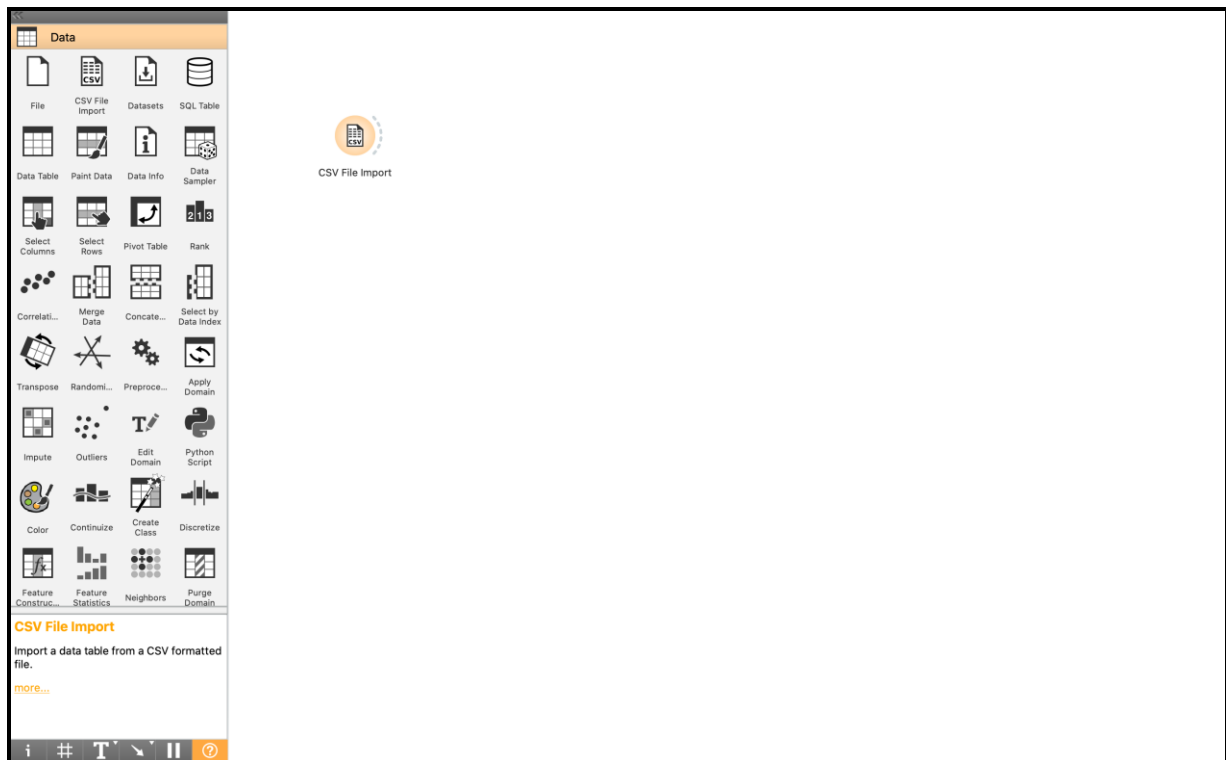


Figure 1.1: File widget dropped on the Canvas.

Chapter 2

How To perform Supervised/Unsupervised Learning

2.1 Supervised learning:

If you're learning a task under supervision, someone is present judging whether you're getting the right answer. Similarly, in supervised learning, that means having a full set of labeled data while training an algorithm.

Fully labeled means that each example in the training dataset is tagged with the answer the algorithm should come up with on its own. So, a labeled dataset of flower images would tell the model which photos were of roses, daisies and daffodils. When shown a new image, the model compares it to the training examples to predict the correct label.

There are two main areas where supervised learning is useful: classification problems and regression problems.

2.1.1 Supervised learning in Orange:

- **Classification:** Much of Orange is devoted to machine learning methods for classification, or supervised data mining. These methods rely on data with class-labeled instances, like that of senate voting.

Orange implements functions for construction of classification models, their evaluation and scoring.

For supervised learning, Orange uses learners. These are objects that receive the data and return classifiers. Learners are passed to evaluation routines

Classification Models:

- Learners and Classifiers
- Probabilistic Classification
- Logistic Regression
- Random Forrest
- Cross-Validation
- Handful of Classifiers
- k-Nearest Neighbors
- Naive Bayes
- Support Vector Machines
- Linear Support Vector Machines
- Nu-Support Vector Machines
- Classification Tree
- Simple Tree

- Majority Classifier
 - Neural Network
 - CN2 Rule Induction
 - Calibration and threshold optimization
- **Regression:** Regression in Orange is, from the interface, very similar to classification. These both require class-labeled data. Just like in classification, regression is implemented with learners and regression models (regressors). Regression learners are objects that accept data and return regressors. Regression models are given data items to predict the value of continuous class.
Regression Models:
 - Linear Regression
 - Polynomial
 - Mean
 - Random Forest
 - Simple Random Forest
 - Regression Tree
 - Neural Network
 - Handful of Regressors
 - Cross Validation

2.2 Unsupervised learning:

Clean, perfectly labeled datasets aren't easy to come by. And sometimes, researchers are asking the algorithm questions they don't know the answer to. That's where unsupervised learning comes in.

In unsupervised learning, a deep learning model is handed a dataset without explicit instructions on what to do with it. The training dataset is a collection of examples without a specific desired outcome or correct answer. The neural network then attempts to automatically find structure in the data by extracting useful features and analyzing its structure.

2.2.1 Unsupervised learning in Orange:

- **Clustering:** Without being an expert ornithologist, it's possible to look at a collection of bird photos and separate them roughly by species, relying on cues like feather color, size or beak shape. That's how the most common application for unsupervised learning, clustering, works: the deep learning model looks for training data that are similar to each other and groups them together.

- **Association:** Fill an online shopping cart with diapers, applesauce and sippy cups and the site just may recommend that you add a bib and a baby monitor to your order. This is an example of association, where certain features of a data sample correlate with other features. By looking at a couple key attributes of a data point, an unsupervised learning model can predict the other attributes with which they're commonly associated.
- **Autoencoders:** Autoencoders take input data, compress it into a code, then try to recreate the input data from that summarized code. It's like starting with Moby Dick, creating a SparkNotes version and then trying to rewrite the original story using only SparkNotes for reference. While a neat deep learning trick, there are fewer real-world cases where a simple autocoder is useful. But add a layer of complexity and the possibilities multiply: by using both noisy and clean versions of an image during training, autoencoders can remove noise from visual data like images, video or medical scans to improve picture quality.
- **Anomaly Detection:** Banks detect fraudulent transactions by looking for unusual patterns in customer's purchasing behavior. For instance, if the same credit card is used in California and Denmark within the same day, that's cause for suspicion. Similarly, unsupervised learning can be used to flag outliers in a dataset.

2.3 Semi-Supervised learning:

Semi-supervised learning is, for the most part, just what it sounds like: a training dataset with both labeled and unlabeled data. This method is particularly useful when extracting relevant features from the data is difficult, and labeling examples is a time-intensive task for experts.

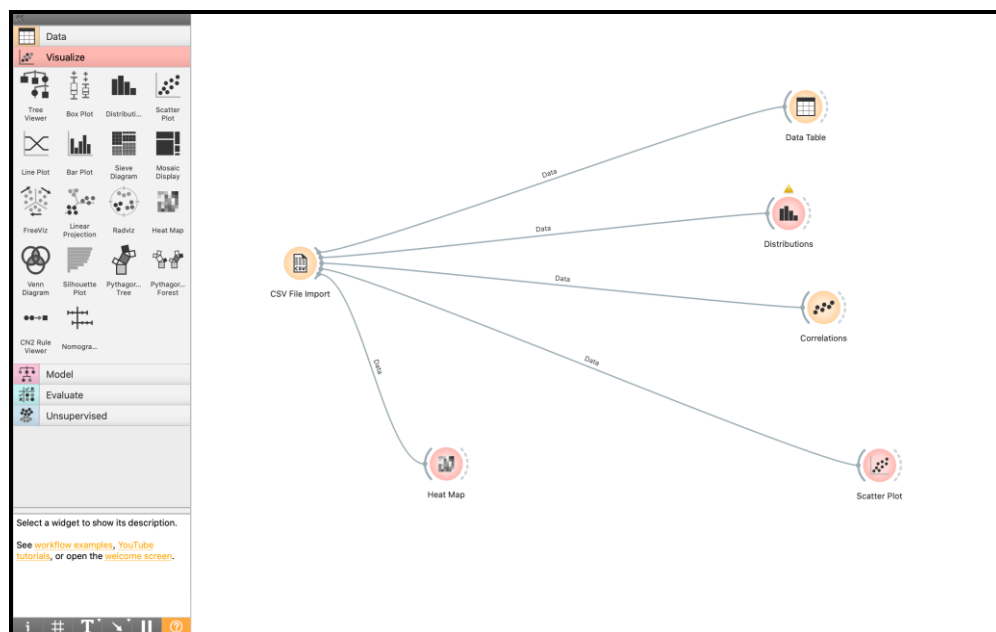


Figure 2.1: Workflow on canvas.

Chapter 3

How to perform Pre-processing/ Clustering.

3.1 Pre-processing of data-set:

- **Pre-processing using Python Script:**

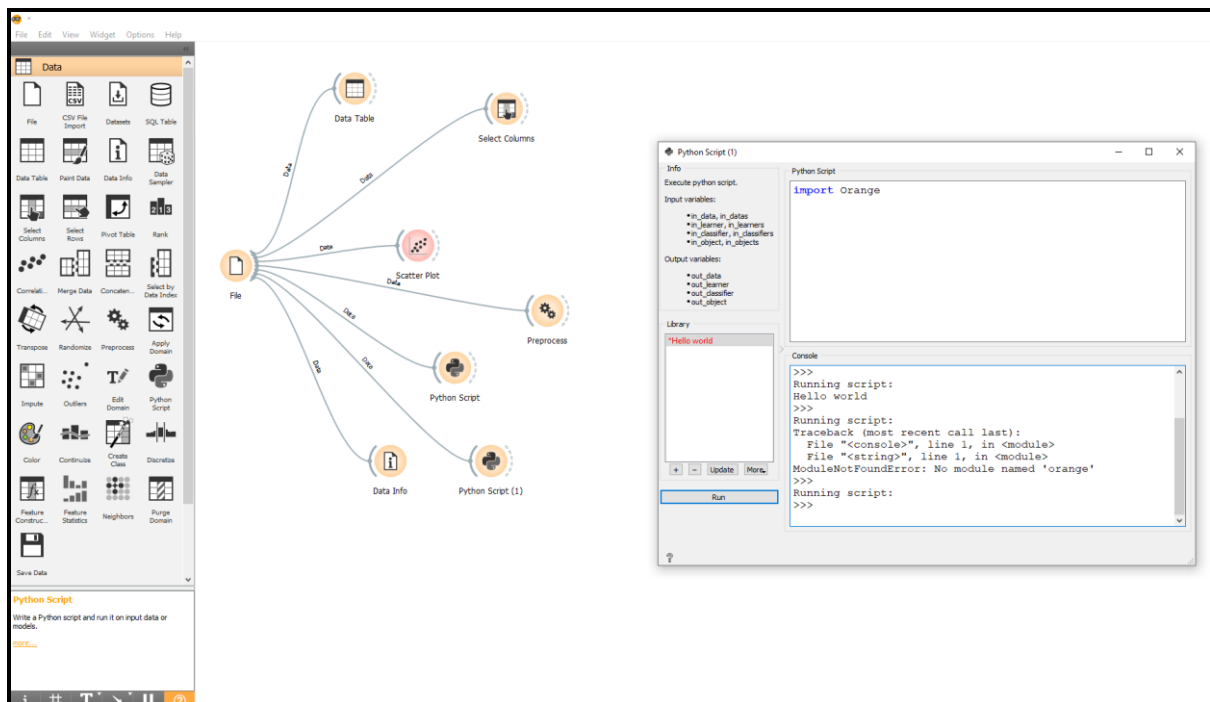


Figure 3.1: Python Script widget on canvas.

Python Script provided by the Orange console can be used to write all the functions or pre-processing such as eliminating null values from the dataset etc.

This can be achieved either by using the provided widget i.e., Python Script or on Anaconda by using “**import Orange**”.

- **Pre-processing using Preprocess widget.**

Orange console also provides a widget called **Preprocess**. Using this widget, we can select the process we want to perform on our data-set such as:

- Impute Missing Values
- Normalize Features
- Randomize

- Remove Sparse Features
- Select Relevant Features
- Select Random Features
- Continuize Discrete Variables

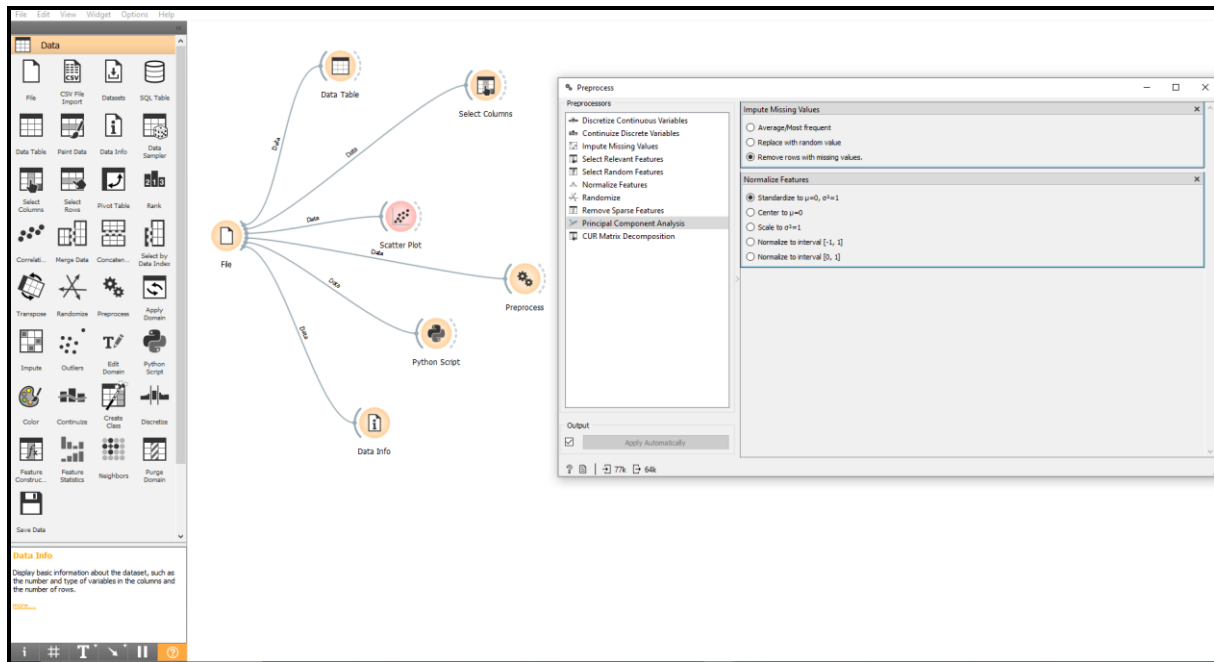


Figure 3.2: Preprocess widget on canvas.

3.2 Clustering:

Clustering is the process of making a group of abstract objects into classes of similar objects.

Points to Remember

- A cluster of data objects can be treated as one group.
- While doing cluster analysis, we first partition the set of data into groups based on data similarity and then assign the labels to the groups.
- The main advantage of clustering over classification is that, it is adaptable to changes and helps single out useful features that distinguish different groups.

Clustering can be executed on console's Python script using the following algorithm of hierarchical clustering.

Chapter 4

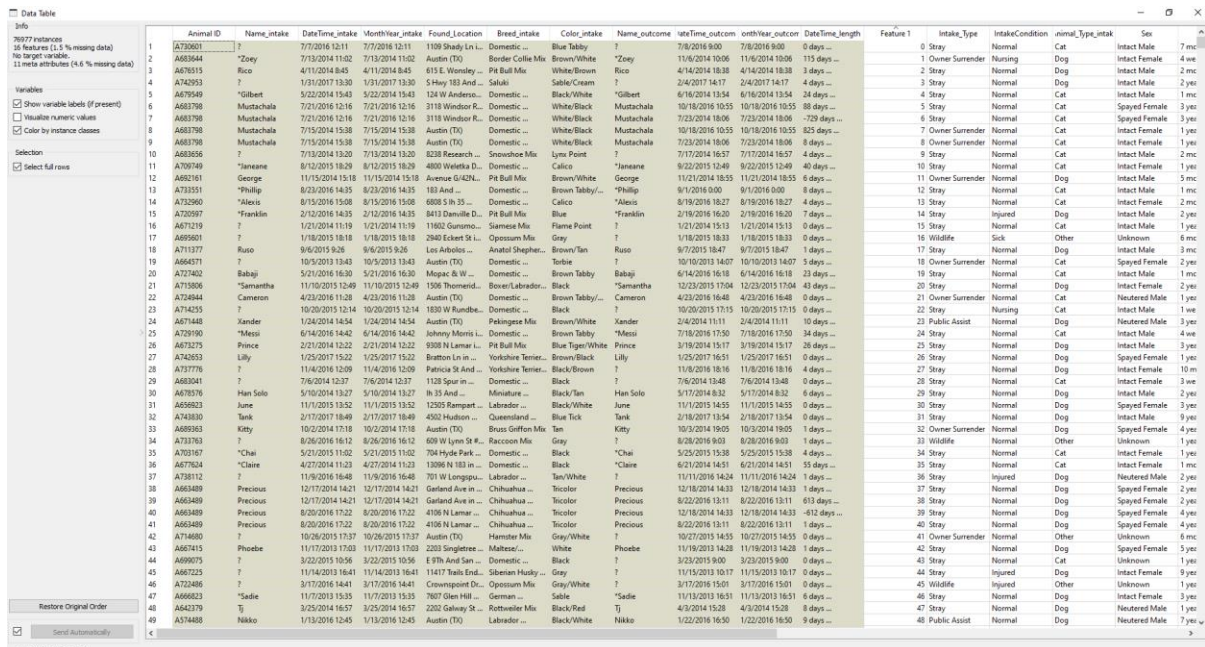
Description of the Dataset Used

There are two data sets used to discover and utilise various features of Orange Data Mining Tool.

- Animal rescue centre dataset.

This dataset is recorded by Austin City Animal Shelter, Texas, USA. It features all the details of the rescued animal of all sorts and their outcome.

It helped in keeping a record of stray animals and lost animals which later on provided the establishment of probability on adoption of animals according to their feature.



Animal ID	Name_intake	DateTime_intake	MonthYear_intake	Found_Location	Breed_intake	Color_intake	Name_outcome	IntakeTime_outcome	MonthYear_outcome	DateTime_length	Feature 1	Intake_Type	IntakeCondition	animal_Type_intak	Sex
A730021	Zoe	7/7/2016 12:11	7/7/2016 12:11	1109 Shady Ln.	Domestic ...	Blue Tabby	Zoe	7/8/2016 9:00	7/8/2016 9:00	0 days ...	0	Stray	Normal	Cat	Intact Male
A683644	Zoe	7/7/2016 11:02	7/13/2016 11:02	Austin (TX)	Border Collie Mix	Brown/White	Zoe	11/6/2014 10:06	11/6/2014 10:06	115 days ...	1	Owner Surrender	Nursing	Dog	Intact Female
A676153	Rico	4/11/2014 8:45	4/11/2014 8:45	611 E. Wombley	Pk Bull Mix	White/Brown	Rico	4/14/2014 18:38	4/14/2014 18:38	3 days ...	2	Stray	Normal	Dog	Intact Male
A742953	?	1/31/2017 13:30	1/31/2017 13:30	5 Hwy 183 And...	Sabalo/Cream	?	?	2/4/2017 14:17	2/4/2017 14:17	4 days ...	3	Stray	Normal	Dog	Intact Male
A679549	Gilbert	5/22/2014 15:43	5/22/2014 15:43	124 W Anderson...	Domestic ...	Black/White	Gilbert	6/16/2014 13:54	6/16/2014 13:54	24 days ...	4	Stray	Normal	Cat	Intact Male
A683798	Mustachula	7/21/2016 12:16	7/21/2016 12:16	3118 Windsor R...	Domestic ...	White/Black	Mustachula	10/18/2016 10:55	10/18/2016 10:55	88 days ...	5	Stray	Normal	Cat	Spayed Female
A683798	Mustachula	7/21/2016 12:16	7/21/2016 12:16	3118 Windsor R...	Domestic ...	White/Black	Mustachula	7/23/2014 18:06	7/23/2014 18:06	-729 days ...	6	Stray	Normal	Cat	Spayed Female
A683798	Mustachula	7/15/2014 13:38	7/15/2014 13:38	Austin (TX)	Domestic ...	White/Black	Mustachula	10/18/2016 10:55	10/18/2016 10:55	825 days ...	7	Owner Surrender	Normal	Cat	Intact Female
A683798	Mustachula	7/15/2014 13:38	7/15/2014 13:38	Austin (TX)	Domestic ...	White/Black	Mustachula	7/23/2014 18:06	7/23/2014 18:06	8 days ...	8	Owner Surrender	Normal	Cat	Intact Female
A683656	?	7/13/2014 13:20	7/13/2014 13:20	8238 Research ...	Snowshoe Mix	Lynx Point	?	7/17/2014 16:57	7/17/2014 16:57	4 days ...	9	Stray	Normal	Cat	Intact Male
A700749	Janeane	8/12/2015 18:29	8/12/2015 18:29	4800 Waterloo D...	Domestic ...	Calico	Janeane	8/22/2015 12:49	8/22/2015 12:49	40 days ...	10	Stray	Normal	Cat	Intact Female
A682161	George	11/15/2014 15:18	11/15/2014 15:18	Avenue G/42N...	Pk Bull Mix	Brown/White	George	11/21/2014 18:55	11/21/2014 18:55	6 days ...	11	Owner Surrender	Normal	Dog	Intact Male
A733551	Phillip	8/23/2016 14:35	8/23/2016 14:35	183 And ...	Domestic ...	Brown Tabby/...	Phillip	9/1/2016 0:00	9/1/2016 0:00	8 days ...	12	Stray	Normal	Cat	Intact Male
A732600	Alexis	8/15/2016 15:08	8/15/2016 15:08	6808 S H 35 ...	Domestic ...	Calico	Alexis	8/19/2016 18:27	8/19/2016 18:27	4 days ...	13	Stray	Normal	Cat	Intact Female
A720597	Franklin	2/12/2016 14:35	2/12/2016 14:35	8413 Danville D...	Pk Bull Mix	Blue	Franklin	2/19/2016 16:20	2/19/2016 16:20	7 days ...	14	Stray	Injured	Dog	Intact Male
A671219	?	1/21/2014 11:19	1/21/2014 11:19	11802 Gurnee...	Siamese Mix	Flame Point	?	1/21/2014 15:13	1/21/2014 15:13	0 days ...	15	Stray	Normal	Cat	Intact Male
A695601	?	1/18/2015 18:18	1/18/2015 18:18	2940 Eckert St L...	Opussum Mix	Gray	?	1/18/2015 18:33	1/18/2015 18:33	0 days ...	16	Wildlife	Sick	Other	Unknown
A711377	Ruso	9/6/2015 9:26	9/6/2015 9:26	Los Arboles ...	Anatoli Shepher...	Brown/Tan	Ruso	9/7/2015 18:47	9/7/2015 18:47	1 days ...	17	Stray	Normal	Dog	Intact Male
A664571	?	10/9/2013 13:43	10/9/2013 13:43	Austin (TX)	Domestic ...	Tortue	?	10/10/2013 14:07	10/10/2013 14:07	3 days ...	18	Owner Surrender	Normal	Cat	Spayed Female
A727462	Babaji	5/21/2016 16:30	5/21/2016 16:30	Mexpe & W ...	Domestic ...	Brown Tabby	Babaji	6/14/2016 16:18	6/14/2016 16:18	23 days ...	19	Stray	Normal	Cat	Intact Male
A715066	Samantha	11/10/2015 12:49	11/10/2015 12:49	1506 Thorneid...	Bowser/Labrador...	Black	Samantha	12/23/2015 17:04	12/23/2015 17:04	43 days ...	20	Stray	Normal	Dog	Intact Female
A724844	Cameron	4/23/2016 11:28	4/23/2016 11:28	Austin (TX)	Domestic ...	Brown Tabby/...	Cameron	4/23/2016 16:48	4/23/2016 16:48	0 days ...	21	Owner Surrender	Normal	Cat	Neutered Male
A714255	?	10/20/2015 12:14	10/20/2015 12:14	1830 W Runbel...	Domestic ...	Black	?	10/20/2015 17:15	10/20/2015 17:15	0 days ...	22	Stray	Nursing	Cat	Intact Male
A671446	Xander	1/24/2014 14:54	1/24/2014 14:54	Austin (TX)	Peangone Mix	Brown/White	Xander	2/4/2014 13:11	2/4/2014 13:11	19 days ...	23	Public Assist	Normal	Dog	Neutered Male
A729190	Messi	6/14/2016 14:42	6/14/2016 14:42	Johnny Morris L...	Domestic ...	Brown Tabby	Messi	7/18/2016 17:50	7/18/2016 17:50	34 days ...	24	Stray	Normal	Cat	Intact Male
A673275	Prince	2/21/2014 12:22	2/21/2014 12:22	9308 N Lamar L...	Pk Bull Mix	Blue Tiger/White	Prince	3/18/2014 15:17	3/18/2014 15:17	26 days ...	25	Stray	Normal	Dog	Intact Male
A742653	Lilly	1/25/2017 15:22	1/25/2017 15:22	Barton Ln in ...	Yorkshire Terrier...	Brown/Black	Lilly	1/25/2017 16:51	1/25/2017 16:51	0 days ...	26	Stray	Normal	Dog	Spayed Female
A737776	?	11/4/2016 12:09	11/4/2016 12:09	Patrice St And...	Yorkshire Terrier...	Black/Brown	?	11/8/2016 18:16	11/8/2016 18:16	4 days ...	27	Stray	Normal	Dog	Intact Female
A683041	?	7/6/2014 12:37	7/6/2014 12:37	1128 Spur in ...	Domestic ...	Black	?	7/16/2014 13:48	7/16/2014 13:48	0 days ...	28	Stray	Normal	Cat	Intact Female
A678576	Han Solo	5/10/2014 13:27	5/10/2014 13:27	th 35 And ...	Miniature ...	Black/Tan	Han Solo	5/17/2014 8:32	5/17/2014 8:32	6 days ...	29	Stray	Normal	Dog	Intact Male
A686823	June	11/1/2015 13:52	11/1/2015 13:52	12305 Rampart ...	Labrador ...	Black/White	June	11/1/2015 14:55	11/1/2015 14:55	0 days ...	30	Stray	Normal	Dog	Spayed Female
A743830	Tank	2/17/2017 18:49	2/17/2017 18:49	4502 Hutton ...	Queeneland ...	Blue Tick	Tank	2/18/2017 13:54	2/18/2017 13:54	0 days ...	31	Stray	Normal	Dog	Intact Male
A689363	Kitty	10/2/2014 17:18	10/2/2014 17:18	Austin (TX)	Bruss Griffon Mix	Tan	Kitty	10/3/2014 19:05	10/3/2014 19:05	1 days ...	32	Owner Surrender	Normal	Dog	Spayed Female
A733763	?	8/26/2016 16:12	8/26/2016 16:12	609 W Lynn St #...	Raccoon Mix	Gray	?	8/28/2016 9:03	8/28/2016 9:03	1 days ...	33	Wildlife	Normal	Other	Unknown
A701167	*Chi	5/21/2015 15:02	5/21/2015 15:02	794 Hyde Park ...	Domestic ...	Black	*Chi	5/25/2015 15:38	5/25/2015 15:38	4 days ...	34	Stray	Normal	Cat	Intact Female
A677924	*Claire	4/27/2014 11:23	4/27/2014 11:23	13096 N 183 in ...	Domestic ...	Black	*Claire	6/21/2014 14:51	6/21/2014 14:51	55 days ...	35	Stray	Normal	Cat	Intact Male
A738112	?	11/9/2016 16:48	11/9/2016 16:48	701 W Longgip...	Labrador ...	Tan/White	?	11/11/2016 14:24	11/11/2016 14:24	1 days ...	36	Stray	Injured	Dog	Neutered Male
A663489	Precious	12/17/2014 14:21	12/17/2014 14:21	Garland Ave in ...	Chihuahua ...	Ticlor	Precious	12/18/2014 14:33	12/18/2014 14:33	1 days ...	37	Stray	Normal	Dog	Spayed Female
A663489	Precious	12/17/2014 14:21	12/17/2014 14:21	Garland Ave in ...	Chihuahua ...	Ticlor	Precious	8/22/2016 13:11	8/22/2016 13:11	613 days ...	38	Stray	Normal	Dog	Spayed Female
A663489	Precious	8/20/2016 17:32	8/20/2016 17:32	4106 N Lamar ...	Chihuahua ...	Ticlor	Precious	12/18/2014 14:33	12/18/2014 14:33	-612 days ...	39	Stray	Normal	Dog	Spayed Female
A663489	Precious	8/20/2016 17:32	8/20/2016 17:32	4106 N Lamar ...	Chihuahua ...	Ticlor	Precious	8/22/2016 13:11	8/22/2016 13:11	1 days ...	40	Stray	Normal	Dog	Spayed Female
A714800	?	10/26/2015 17:37	10/26/2015 17:37	Austin (TX)	Hamster Mix	Gray/White	?	10/27/2015 14:55	10/27/2015 14:55	0 days ...	41	Owner Surrender	Normal	Other	Unknown
A687415	Phoebe	11/17/2013 17:03	11/17/2013 17:03	2203 Singlere...	Matured ...	White	Phoebe	11/19/2013 14:28	11/19/2013 14:28	1 days ...	42	Stray	Normal	Dog	Spayed Female
A699075	?	3/12/2013 15:56	3/12/2013 15:56	61th And San ...	Domestic ...	Gray	?	3/23/2013 9:00	3/23/2013 9:00	0 days ...	43	Stray	Normal	Cat	Unknown
A687225	?	11/14/2013 16:41	11/14/2013 16:41	11417 Trails End...	Siberian Husky ...	Black	?	11/15/2013 10:17	11/15/2013 10:17	0 days ...	44	Stray	Injured	Dog	Intact Female
A723486	?	3/17/2016 14:41	3/17/2016 14:41	Crownpoint Dr...	Opussum Mix	Gray/White	?	3/17/2016 15:01	3/17/2016 15:01	0 days ...	45	Wildlife	Injured	Other	Unknown
A666623	*Sadie	11/7/2013 13:35	11/7/2013 13:35	7807 Glen Hwy ...	German ...	Sable	*Sadie	11/13/2013 16:51	11/13/2013 16:51	6 days ...	46	Stray	Normal	Dog	Intact Female
A642379	TJ	3/25/2014 16:57	3/25/2014 16:57	2202 Galway St ...	Rottweiler Mix	Black/Red	TJ	4/3/2014 15:28	4/3/2014 15:28	8 days ...	47	Stray	Normal	Dog	Neutered Male
A734488	Nikko	1/13/2016 12:45	1/13/2016 12:45	Austin (TX)	Labrador ...	Black/White	Nikko	1/22/2016 16:30	1/22/2016 16:30	9 days ...	48	Public Assist	Normal	Dog	Neutered Male

Figure 4.1: Animal Rescue Centre Dataset.

- Iris dataset (sample dataset of the console).

This dataset is provided by the console as a sample to help in learning the variety of features on the software.

Info						
150 instances (no missing data)						
4 features						
Target with 3 values						
No meta attributes						
Variables						
<input checked="" type="checkbox"/> Show variable labels (if present)						
<input type="checkbox"/> Visualize numeric values						
<input checked="" type="checkbox"/> Color by instance classes						
Selection						
<input checked="" type="checkbox"/> Select full rows						
Restore Original Order						
<input checked="" type="checkbox"/> Send Automatically						
	iris	sepal length	sepal width	petal length	petal width	
1	Iris-setosa	5.1	3.5	1.4	0.2	
2	Iris-setosa	4.9	3.0	1.4	0.2	
3	Iris-setosa	4.7	3.2	1.3	0.2	
4	Iris-setosa	4.6	3.1	1.5	0.2	
5	Iris-setosa	5.0	3.6	1.4	0.2	
6	Iris-setosa	5.4	3.9	1.7	0.4	
7	Iris-setosa	4.6	3.4	1.4	0.3	
8	Iris-setosa	5.0	3.4	1.5	0.2	
9	Iris-setosa	4.4	2.9	1.4	0.2	
10	Iris-setosa	4.9	3.1	1.5	0.1	
11	Iris-setosa	5.4	3.7	1.5	0.2	
12	Iris-setosa	4.8	3.4	1.6	0.2	
13	Iris-setosa	4.8	3.0	1.4	0.1	
14	Iris-setosa	4.3	3.0	1.1	0.1	
15	Iris-setosa	5.8	4.0	1.2	0.2	
16	Iris-setosa	5.7	4.4	1.5	0.4	
17	Iris-setosa	5.4	3.9	1.3	0.4	
18	Iris-setosa	5.1	3.5	1.4	0.3	
19	Iris-setosa	5.7	3.8	1.7	0.3	
20	Iris-setosa	5.1	3.8	1.5	0.3	
21	Iris-setosa	5.4	3.4	1.7	0.2	
22	Iris-setosa	5.1	3.7	1.5	0.4	
23	Iris-setosa	4.6	3.6	1.0	0.2	
24	Iris-setosa	5.1	3.3	1.7	0.5	
25	Iris-setosa	4.8	3.4	1.9	0.2	
26	Iris-setosa	5.0	3.0	1.6	0.2	
27	Iris-setosa	5.0	3.4	1.6	0.4	
28	Iris-setosa	5.2	3.5	1.5	0.2	
29	Iris-setosa	5.2	3.4	1.4	0.2	
30	Iris-setosa	4.7	3.2	1.6	0.2	
31	Iris-setosa	4.8	3.1	1.6	0.2	
32	Iris-setosa	5.4	3.4	1.5	0.4	
33	Iris-setosa	5.2	4.1	1.5	0.1	
34	Iris-setosa	5.5	4.2	1.4	0.2	
35	Iris-setosa	4.9	3.1	1.5	0.1	
36	Iris-setosa	5.0	3.2	1.2	0.2	
37	Iris-setosa	5.5	3.5	1.3	0.2	
38	Iris-setosa	4.9	3.1	1.5	0.1	
39	Iris-setosa	4.4	3.0	1.3	0.2	
40	Iris-setosa	5.1	3.4	1.5	0.2	
41	Iris-setosa	5.0	3.5	1.3	0.3	

Figure 4.2: Iris Dataset.

We used two datasets as the Animal rescue centre dataset is non-categorical and Isis being categorical, visualization was effectively executed.

Chapter 5

Implementation

- **Data Input:** Orange can read files in native tab-delimited format, or can load data from any of the major standard spreadsheet file types, like CSV and Excel. Native format starts with a header row with feature (column) names. The second header row gives the attribute type, which can be continuous, discrete, time, or string. The third header line contains meta information to identify dependent features (class), irrelevant features (ignore) or meta features (meta).

```
import Orange
data = Orange.data.Table("grp_data")
```

- **Saving Data:** Data objects can be saved to a file:

```
data.save("new_data.tab")
```

- **Exploring data:** Data table stores information on data instances as well as on data domain. Domain holds the names of attributes, optional classes, their types and, and if categorical, the value names.

```
import Orange

data = Orange.data.Table("grp_data")
n = len(data.domain.attributes)
n_cont = sum(1 for a in data.domain.attributes if a.is_continuous)
n_disc = sum(1 for a in data.domain.attributes if a.is_discrete)
print("%d attributes: %d continuous, %d discrete" % (n, n_cont, n_disc))

print(
    "First three attributes:",
    ", ".join(data.domain.attributes[i].name for i in range(3)),
)

print("Class:", data.domain.class_var.name)
```

- **Missing Values:** The particular data instance included missing data (represented with '?') for the first and the fourth attribute. In the original dataset file, the missing values are, by default, represented with a blank space.

```
>>> import numpy as np
>>> data = Orange.data.Table("iris")
>>> data[2]
```

```
[?, y, y, ?, y, ... | sepal]
>>> np.isnan(data[2][0])
True
>>> np.isnan(data[2][1])
False
```

- **Data Instances:** Data table stores data instances (or examples). These can be indexed or traversed as any Python list. Data instances can be considered as vectors, accessed through element index, or through feature name.

```
import Orange

data = Orange.data.Table("iris")
print("First three data instances:")
for d in data[:3]:
    print(d)

print("25-th data instance:")
print(data[24])

name = "sepal width"
print("Value of '%s' for the first instance:" % name, data[0][name])
print("The 3rd value of the 25th data instance:", data[24][2])
```

- **Mean:** The Iris dataset we have used above has four continuous attributes. Here's a script that computes their mean:

```
average = lambda x: sum(x) / len(x)

data = Orange.data.Table("iris")
print("%-15s %s" % ("Feature", "Mean"))
for x in data.domain.attributes:
    print("%-15s %.2f" % (x.name, average([d[x] for d in data])))
```

- **Data Selection and Sampling:** Besides the name of the data file, `Orange.data.Table` can accept the data domain and a list of data items and returns a new dataset. This is useful for any data subsetting:

```
data = Orange.data.Table("iris.tab")
print("Dataset instances:", len(data))
subset = Orange.data.Table(data.domain, [d for d in data if d["petal
length"] > 3.0])
print("Subset size:", len(subset))
```


- The code outputs: (Data Selection and Sampling).

Dataset instances: 150
Subset size: 99

- Classification Tree:** Orange includes three implemenations of classification trees. TreeLearner is home-grown and properly handles multinominal and missing values. The one from scikit-learn, SkITreeLearner, is faster. Another home-grown, SimpleTreeLearner, is simpler and still faster.

```
>>> import Orange
>>> iris = Orange.data.Table('iris')
>>> tr = Orange.classification.TreeLearner()
>>> classifier = tr(data)
>>> printed_tree = classifier.print_tree()
>>> for i in printed_tree.split('\n'):
>>>     print(i)
[50.  0.  0.] petal length ≤ 1.9
[ 0. 50. 50.] petal length > 1.9
[ 0. 49.  5.]     petal width ≤ 1.7
[ 0. 47.  1.]     petal length ≤ 4.9
[ 0.  2.  4.]     petal length > 4.9
[ 0.  0.  3.]     petal width ≤ 1.5
[ 0.  2.  1.]     petal width > 1.5
[ 0.  2.  0.]     sepal length ≤ 6.7
[ 0.  0.  1.]     sepal length > 6.7
[ 0.  1. 45.]     petal width > 1.7
```

Data table																	
Info																	
9977 instances 36 features (1.5 % missing data) No target variable 1 meta attribute (4.6 % missing data)																	
Variables																	
<input type="checkbox"/> Show variable labels (if present) <input type="checkbox"/> Visualize numeric values <input type="checkbox"/> Color by instance classes																	
Selection																	
<input checked="" type="checkbox"/> Select follows																	
ID	Animal ID	Name_intake	DateTime_intake	MonthYear_intake	Found_Location	Breed_intake	Color_intake	Name_outcome	DateTime_outcome	MonthYear_outcome	DateTime_length	Feature 1	Intake_Type	IntakeCondition	animal_Type_intake	Sex	
1	A720601	?	7/7/2016 12:11	7/7/2016 12:11	1109 Shady Ln.	Domestic ...	Blue Tabby	?	7/8/2016 9:00	7/8/2016 9:00	0 days ...	0	Stay	Normal	Cat	Intact Male	7 mo
2	A683444	*Zoey	7/13/2014 11:02	7/13/2014 11:02	Austin (TX)	Border Collie Mix	Brown/White	*Zoey	11/6/2014 10:06	11/6/2014 10:06	115 days ...	1	Owner Surrender	Nursing	Dog	Intact Female	4 mo
3	A676153	Rico	4/11/2014 8:43	4/11/2014 8:43	815 E. Worsley	Pit Bull Mix	White/Brown	Rico	4/14/2014 18:38	4/14/2014 18:38	3 days ...	2	Stay	Normal	Dog	Intact Male	2 mo
4	A742953	?	1/31/2017 13:30	1/31/2017 13:30	5 Hay 183 And.	Sable	Sable/Cream	?	2/4/2017 14:17	2/4/2017 14:17	4 days ...	3	Stay	Normal	Dog	Intact Male	2 yrs
5	A679549	*Gilbert	5/22/2014 15:43	5/22/2014 15:43	124 W Anderson	Domestic ...	Black/White	*Gilbert	6/16/2014 13:54	6/16/2014 13:54	24 days ...	4	Stay	Normal	Cat	Intact Male	1 mo
6	A683798	Mustachula	7/21/2016 12:16	7/21/2016 12:16	3118 Windsor R.	Domestic ...	White/Black	Mustachula	10/18/2016 10:55	10/18/2016 10:55	88 days ...	5	Stay	Normal	Cat	Spayed Female	3 yrs
7	A683798	Mustachula	7/21/2016 12:16	7/21/2016 12:16	3118 Windsor R.	Domestic ...	White/Black	Mustachula	7/23/2014 18:06	7/23/2014 18:06	-729 days ...	6	Stay	Normal	Cat	Spayed Female	3 yrs
8	A683798	Mustachula	7/15/2014 15:38	7/15/2014 15:38	Austin (TX)	Domestic ...	White/Black	Mustachula	10/18/2016 10:55	10/18/2016 10:55	823 days ...	7	Owner Surrender	Normal	Cat	Intact Female	1 yrs
9	A683798	Mustachula	7/15/2014 15:38	7/15/2014 15:38	Austin (TX)	Domestic ...	White/Black	Mustachula	7/23/2014 18:06	7/23/2014 18:06	8 days ...	8	Owner Surrender	Normal	Cat	Intact Female	1 yrs
10	A683556	?	7/13/2014 13:20	7/13/2014 13:20	8238 Research	Domestic ...	Lyne Point	?	7/17/2014 16:57	7/17/2014 16:57	4 days ...	9	Stay	Normal	Cat	Intact Male	2 mo
11	A707149	*Janeane	8/12/2015 18:29	8/12/2015 18:29	4800 Virelita D.	Domestic ...	Calico	*Janeane	9/22/2015 12:49	9/22/2015 12:49	40 days ...	10	Stay	Normal	Cat	Intact Female	1 yrs
12	A693161	George	11/15/2014 15:18	11/15/2014 15:18	Avenue G12LN.	Pit Bull Mix	Brown/White	George	11/21/2014 18:59	11/21/2014 18:59	6 days ...	11	Owner Surrender	Normal	Dog	Intact Male	5 mo
13	A733551	*Phillip	8/23/2016 14:35	8/23/2016 14:35	183 And.	Domestic ...	Calico	*Phillip	9/1/2016 9:00	9/1/2016 9:00	8 days ...	12	Stay	Normal	Cat	Intact Male	1 mo
14	A732960	*Alexis	8/15/2016 15:08	8/15/2016 15:08	6805 S In 35	Domestic ...	Calico	*Alexis	8/18/2016 18:27	8/18/2016 18:27	4 days ...	13	Stay	Normal	Cat	Intact Female	2 mo
15	A720597	*Franklin	2/12/2016 14:35	2/12/2016 14:35	8413 Danville D.	Pit Bull Mix	Blue	*Franklin	2/19/2016 16:20	2/19/2016 16:20	7 days ...	14	Stay	Injured	Dog	Intact Male	2 yrs
16	A671219	?	1/21/2014 11:19	1/21/2014 11:19	11602 Gunsmo.	Samoyed Mix	Flame Point	?	1/21/2014 19:13	1/21/2014 19:13	0 days ...	15	Stay	Normal	Cat	Intact Male	1 yrs
17	A695501	?	1/18/2015 18:18	1/18/2015 18:18	2840 Eckert St.	Opussum Mix	Gray	?	1/18/2015 18:33	1/18/2015 18:33	0 days ...	16	Wildlife	Sick	Other	Unknown	6 mo
18	A711377	Ruso	9/6/2015 9:26	9/6/2015 9:26	Les Arboles	Border Shepher.	Brown/Tan	Ruso	9/7/2015 18:47	9/7/2015 18:47	1 days ...	17	Stay	Normal	Dog	Intact Male	3 mo
19	A684371	?	10/5/2013 13:43	10/5/2013 13:43	Austin (TX)	Domestic ...	Tobie	?	10/10/2013 14:07	10/10/2013 14:07	5 days ...	18	Owner Surrender	Normal	Cat	Spayed Female	2 yrs
20	A723482	Babaji	5/21/2016 16:39	5/21/2016 16:39	Maple & W.	Domestic ...	Brown Tabby	Babaji	6/14/2016 16:18	6/14/2016 16:18	23 days ...	19	Stay	Normal	Cat	Intact Male	1 mo
21	A713906	*Samantha	11/10/2013 12:49	11/10/2013 12:49	1506 Thornhill	Boxer/Labrador	Black	*Samantha	12/23/2013 17:04	12/23/2013 17:04	43 days ...	20	Stay	Normal	Dog	Intact Female	2 yrs
22	A724844	Cameron	4/23/2016 11:28	4/23/2016 11:28	Austin (TX)	Domestic ...	Brown Tabby...	Cameron	4/23/2016 16:48	4/23/2016 16:48	0 days ...	21	Owner Surrender	Normal	Cat	Neutered Male	1 yrs
23	A714233	?	10/20/2013 12:14	10/20/2013 12:14	1830 W Rundick	Domestic ...	Black	?	10/20/2013 17:15	10/20/2013 17:15	0 days ...	22	Stay	Nursing	Cat	Intact Male	1 mo
24	A674448	Xander	1/24/2014 14:54	1/24/2014 14:54	Austin (TX)	Pekingese Mix	Brown/White	Xander	2/4/2014 11:11	2/4/2014 11:11	10 days ...	23	Public Assist	Normal	Dog	Neutered Male	3 yrs
25	A729190	*Messi	6/14/2016 14:42	6/14/2016 14:42	Johnny Morris L.	Domestic ...	Brown Tabby	*Messi	7/18/2016 17:50	7/18/2016 17:50	34 days ...	24	Stay	Normal	Cat	Intact Male	4 mo
26	A672735	Prince	2/21/2014 12:22	2/21/2014 12:22	9308 N Lamar	Pit Bull Mix	Blue Tiger/White	Prince	3/18/2014 15:17	3/18/2014 15:17	26 days ...	25	Stay	Normal	Dog	Intact Male	3 yrs
27	A724593	Lilly	1/25/2017 15:22	1/25/2017 15:22	Bottom Ln in	Yorkshire Terrier	Brown/Black	Lilly	1/25/2017 16:51	1/25/2017 16:51	0 days ...	26	Stay	Normal	Dog	Spayed Female	1 mo
28	A737776	?	11/4/2016 12:09	11/4/2016 12:09	Patricia St And.	Yorkshire Terrier	Black/Brown	?	11/9/2016 18:16	11/9/2016 18:16	4 days ...	27	Stay	Normal	Dog	Intact Female	10 mo
29	A683241	?	7/6/2014 12:37	7/6/2014 12:37	1128 Spur in	Domestic ...	Black	?	7/6/2014 13:48	7/6/2014 13:48	0 days ...	28	Stay	Normal	Cat	Intact Female	3 mo
30	A678776	Han Solo	5/10/2014 13:27	5/10/2014 13:27	In 35 And.	Miniature ...	Black/Tan	Han Solo	5/17/2014 8:32	5/17/2014 8:32	6 days ...	29	Stay	Normal	Dog	Intact Male	2 yrs
31	A686833	June	11/1/2013 13:52	11/1/2013 13:52	12505 Rampart	Labrador	Black/White	June	11/1/2013 14:53	11/1/2013 14:53	0 days ...	30	Stay	Normal	Dog	Spayed Female	3 yrs
32	A724830	Tank	2/17/2017 18:49	2/17/2017 18:49	4802 Hudson	Queensland	Blue Tick	Tank	2/18/2017 13:54	2/18/2017 13:54	0 days ...	31	Stay	Normal	Dog	Intact Male	8 mo
33	A689363	Kitty	10/2/2014 17:18	10/2/2014 17:18	Austin (TX)	Bruss Griffon Mix	Tan	Kitty	10/3/2014 19:05	10/3/2014 19:05	1 days ...	32	Owner Surrender	Normal	Dog	Spayed Female	4 yrs
34	A737163	?	8/26/2016 16:12	8/26/2016 16:12	609 W Lynn St	Raccoon Mix	Gray	?	8/28/2016 9:03	8/28/2016 9:03	1 days ...	33	Wildlife	Normal	Other	Unknown	1 yrs
35	A703187	*Chai	5/21/2013 11:02	5/21/2013 11:02	794 Hyde Park	Domestic ...	Black	*Chai	5/25/2013 15:38	5/25/2013 15:38	4 days ...	34	Stay	Normal	Cat	Intact Female	1 yrs
36	A677534	*Claire	4/27/2014 11:23	4/27/2014 11:23	12086 N 181 in	Domestic ...	Black	*Claire	6/1/2014 14:51	6/1/2014 14:51	55 days ...	35	Stay	Normal	Cat	Intact Female	1 mo
37	A738112	?	11/9/2016 16:48	11/9/2016 16:48	701 W Longspur	Labrador	Tan/White	?	11/11/2016 14:24	11/11/2016 14:24	1 days ...	36	Stay	Injured	Dog	Neutered Male	2 yrs
38	A663489	Precious	12/17/2014 14:21	12/17/2014 14:21	Garland Ave in	Chihuahua	Tricolor	Precious	12/18/2014 14:33	12/18/2014 14:33	1 days ...	37	Stay	Normal	Dog	Spayed Female	2 yrs
39	A663489	Precious	10/17/2014 14:21	10/17/2014 14:21	Garland Ave in	Chihuahua	Tricolor	Precious	8/22/2016 16:11	8/22/2016 16:11	613 days ...	38	Stay	Normal	Dog	Spayed Female	2 yrs
40	A663489	Precious	8/20/2016 17:22	8/20/2016 17:22	4106 N Lamar	Chihuahua	Tricolor	Precious	12/18/2014 14:33	12/18/2014 14:33	-413 days ...	39	Stay	Normal	Dog	Spayed Female	4 yrs
41	A663489	Precious	8/20/2016 17:22	8/20/2016 17:22	4106 N Lamar	Chihuahua	Tricolor	Precious	8/22/2016 13:11	8/22/2016 13:11	1 days ...	40	Stay	Normal	Dog	Spayed Female	4 yrs
42	A714800	?	10/26/2013 17:37	10/26/2013 17:37	Austin (TX)	Husky Mix	Gray/White	?	10/27/2013 14:55	10/27/2013 14:55	0 days ...	41	Owner Surrender	Normal	Other	Unknown	6 mo
43	A687815	Phoebe	11/17/2013 17:03	11/17/2013 17:03	2203 Singletree	White	White	Phoebe	11/19/2013 14:28	11/19/2013 14:28	1 days ...	42	Stay	Normal	Dog	Spayed Female	5 mo
44	A699075	?	3/22/2015 10:56	3/22/2015 10:56	E 9th And San	Domestic ...	Black	?	3/23/2015 9:00	3/23/2015 9:00	0 days ...	43	Stay	Normal	Cat	Unknown	1 mo
45	A686723	?	11/14/2013 16:41	11/14/2013 16:41	11417 Trail End	Siberian Husky	Gray	?	11/15/2013 10:17	11/15/2013 10:17	0 days ...	44	Stay	Injured	Dog	Intact Female	9 yrs
46	A722486	?	3/17/2016 14:41	3/17/2016 14:41	Crowpoint Dr.	Opussum Mix	Gray/White	?	3/17/2016 15:01	3/17/2016 15:01	0 days ...	45	Wildlife	Injured	Other	Unknown	1 yrs
47	A666823	*Sedie	11/17/2013 15:35	11/17/2013 15:35	7807 Glen Hill	German ...	Sable	*Sedie	11/13/2013 16:51	11/13/2013 16:51	6 days ...	46	Stay	Normal	Dog	Intact Female	3 yrs
48	A642179	?	3/25/2014 16:57	3/25/2014 16:57	2202 Galveston St	Rottweiler Mix	Black/Red	?	4/3/2014 15:38	4/3/2014 15:38	8 days ...	47	Stay	Normal	Dog	Neutered Male	1 yrs
49	A574488	Nikko	1/13/2016 12:45	1/13/2016 12:45	Austin (TX)	Labrador	Black/White	Nikko	1/22/2016 16:50	1/22/2016 16:50	9 days ...	48	Public Assist	Normal	Dog	Neutered Male	7 yrs

Info

150 instances (no missing data)

4 features

Target with 3 values

No meta attributes

Variables

Show variable labels (if present)

Visualize numeric values

Color by instance classes

Selection

Select full rows

Restore Original Order

Send Automatically

	iris	sepal length	sepal width	petal length	petal width
1	Iris-setosa	5.1	3.5	1.4	0.2
2	Iris-setosa	4.9	3.0	1.4	0.2
3	Iris-setosa	4.7	3.2	1.3	0.2
4	Iris-setosa	4.6	3.1	1.5	0.2
5	Iris-setosa	5.0	3.6	1.4	0.2
6	Iris-setosa	5.4	3.9	1.7	0.4
7	Iris-setosa	4.6	3.4	1.4	0.3
8	Iris-setosa	5.0	3.4	1.5	0.2
9	Iris-setosa	4.4	2.9	1.4	0.2
10	Iris-setosa	4.9	3.1	1.5	0.1
11	Iris-setosa	5.4	3.7	1.5	0.2
12	Iris-setosa	4.8	3.4	1.6	0.2
13	Iris-setosa	4.8	3.0	1.4	0.1
14	Iris-setosa	4.3	3.0	1.1	0.1
15	Iris-setosa	5.8	4.0	1.2	0.2
16	Iris-setosa	5.7	4.4	1.5	0.4
17	Iris-setosa	5.4	3.9	1.3	0.4
18	Iris-setosa	5.1	3.5	1.4	0.3
19	Iris-setosa	5.7	3.8	1.7	0.3
20	Iris-setosa	5.1	3.8	1.5	0.3
21	Iris-setosa	5.4	3.4	1.7	0.2
22	Iris-setosa	5.1	3.7	1.5	0.4
23	Iris-setosa	4.6	3.6	1.0	0.2
24	Iris-setosa	5.1	3.3	1.7	0.5
25	Iris-setosa	4.8	3.4	1.9	0.2
26	Iris-setosa	5.0	3.0	1.6	0.2
27	Iris-setosa	5.0	3.4	1.6	0.4
28	Iris-setosa	5.2	3.5	1.5	0.2
29	Iris-setosa	5.2	3.4	1.4	0.2
30	Iris-setosa	4.7	3.2	1.6	0.2
31	Iris-setosa	4.8	3.1	1.6	0.2
32	Iris-setosa	5.4	3.4	1.5	0.4
33	Iris-setosa	5.2	4.1	1.5	0.1
34	Iris-setosa	5.5	4.2	1.4	0.2
35	Iris-setosa	4.9	3.1	1.5	0.1
36	Iris-setosa	5.0	3.2	1.2	0.2
37	Iris-setosa	5.5	3.5	1.3	0.2
38	Iris-setosa	4.9	3.1	1.5	0.1
39	Iris-setosa	4.4	3.0	1.3	0.2
40	Iris-setosa	5.1	3.4	1.5	0.2
41	Iris-setosa	5.0	3.5	1.3	0.3

Chapter 6

Visualisation

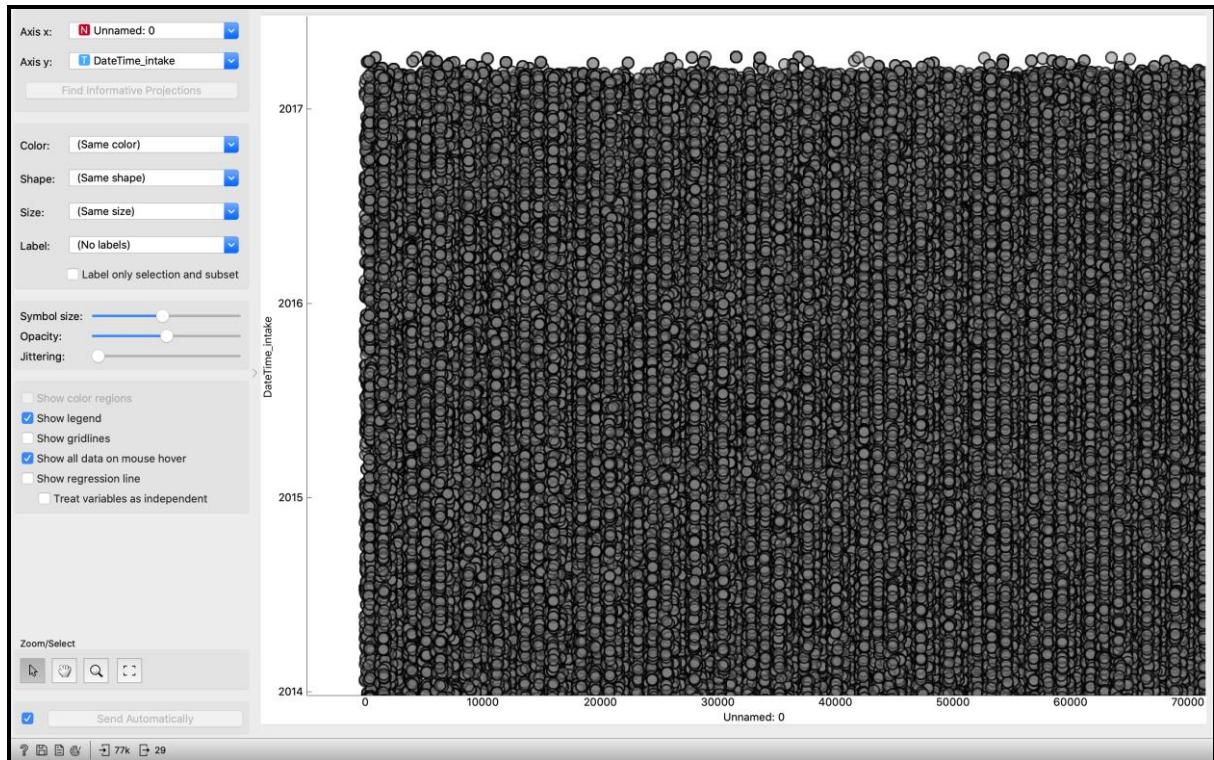


Figure 6.1: Scatter Plot.

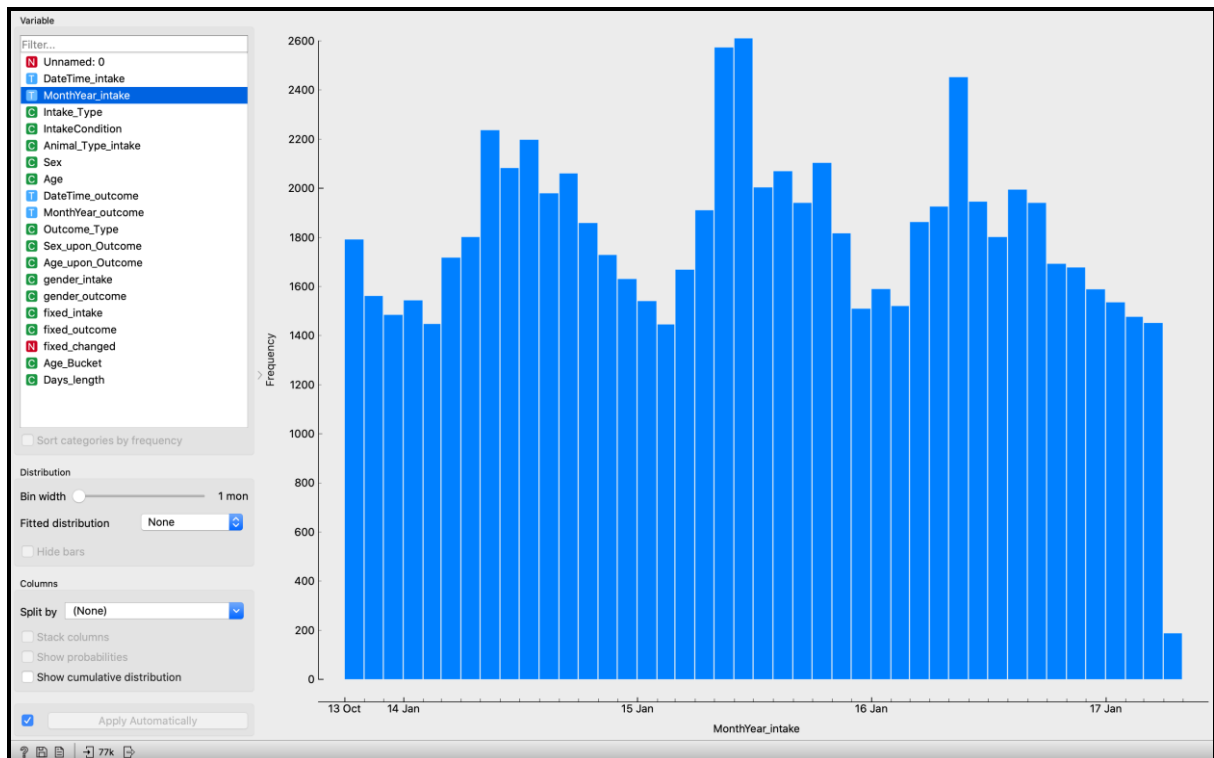


Figure 6.2: Bar Plot (Date Intake).

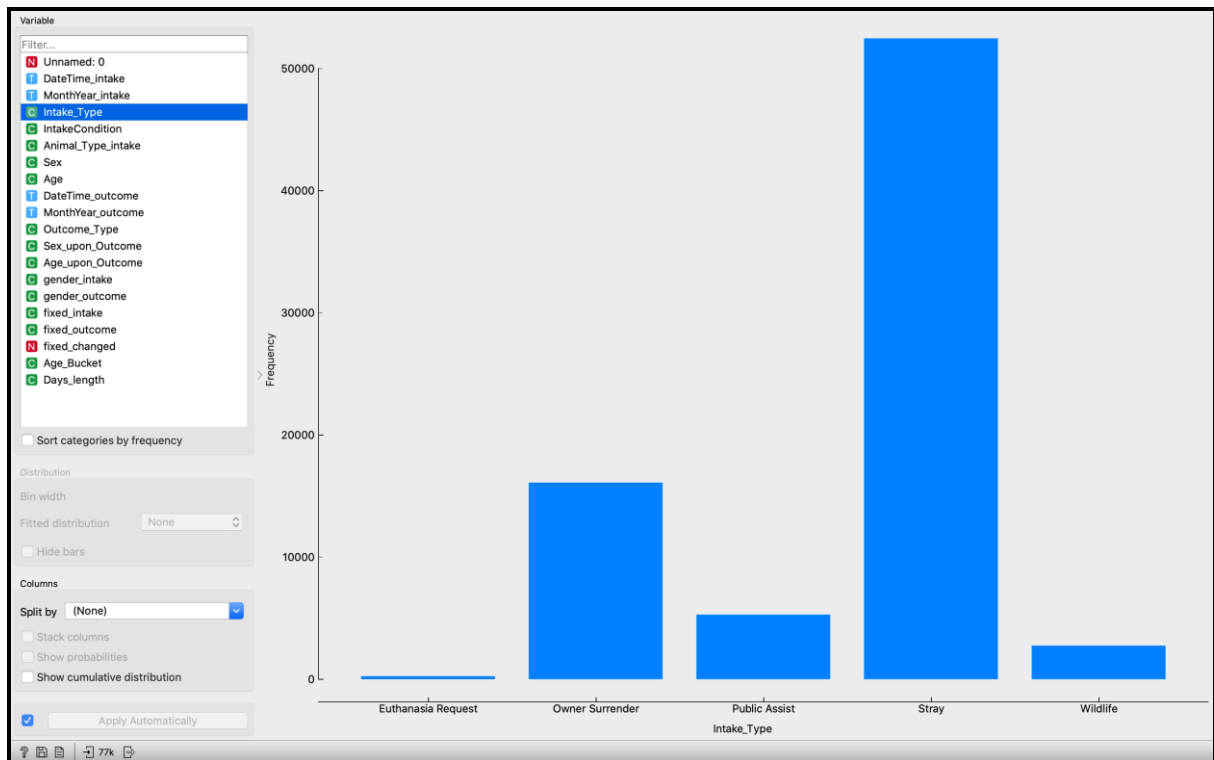


Figure 6.3: Bar Plot (Intake Type).

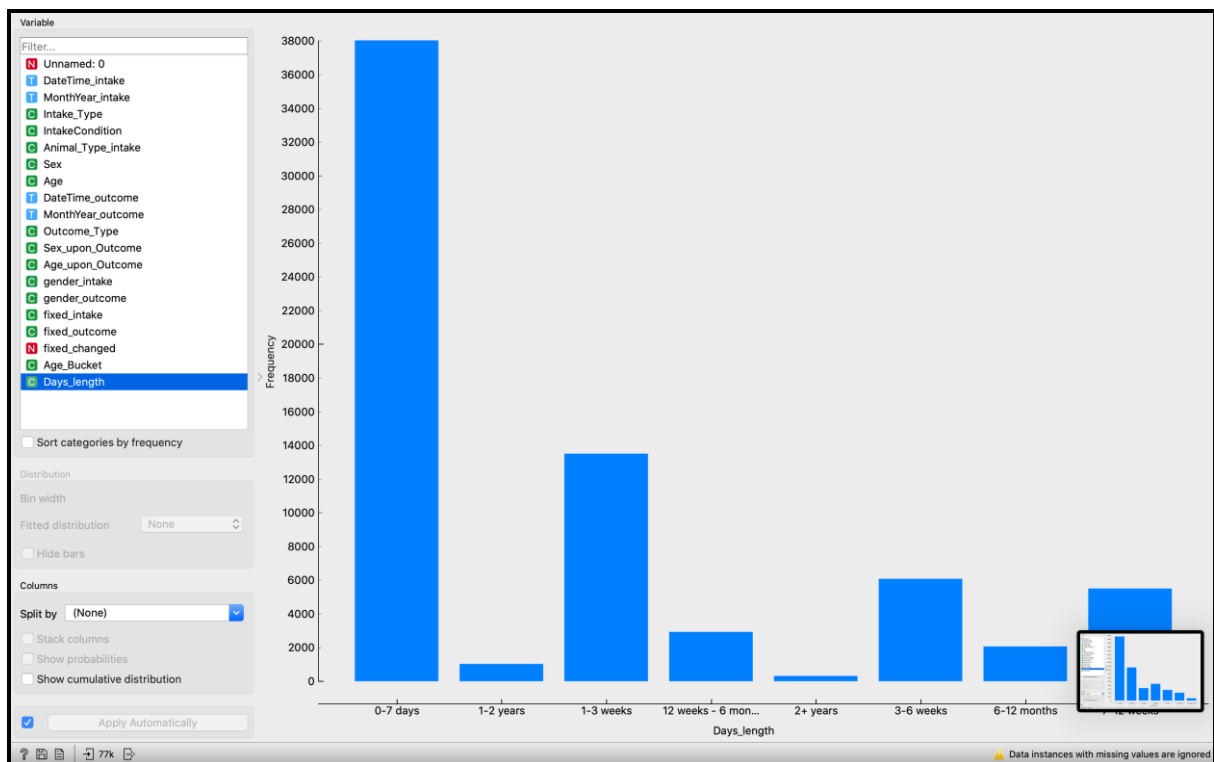


Figure 6.4: Bar Plot (Stay Length).

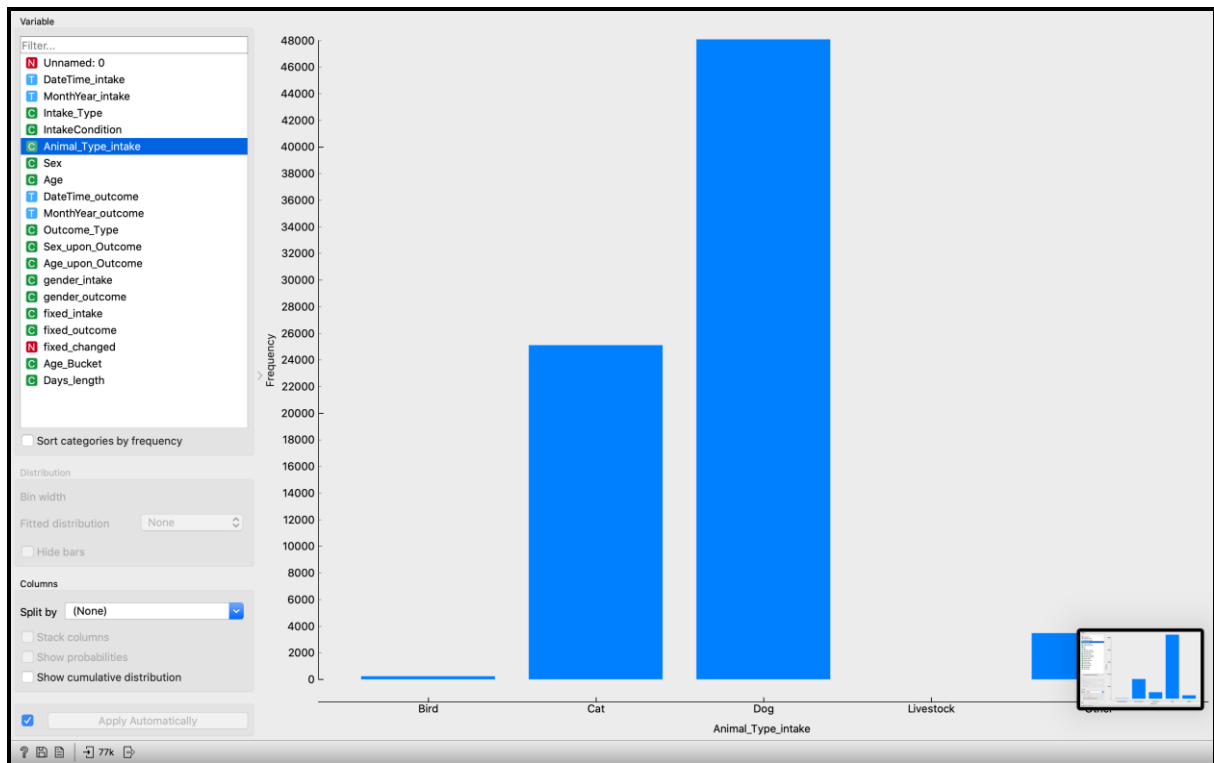


Figure 6.5: Bar Plot (Animal Type).

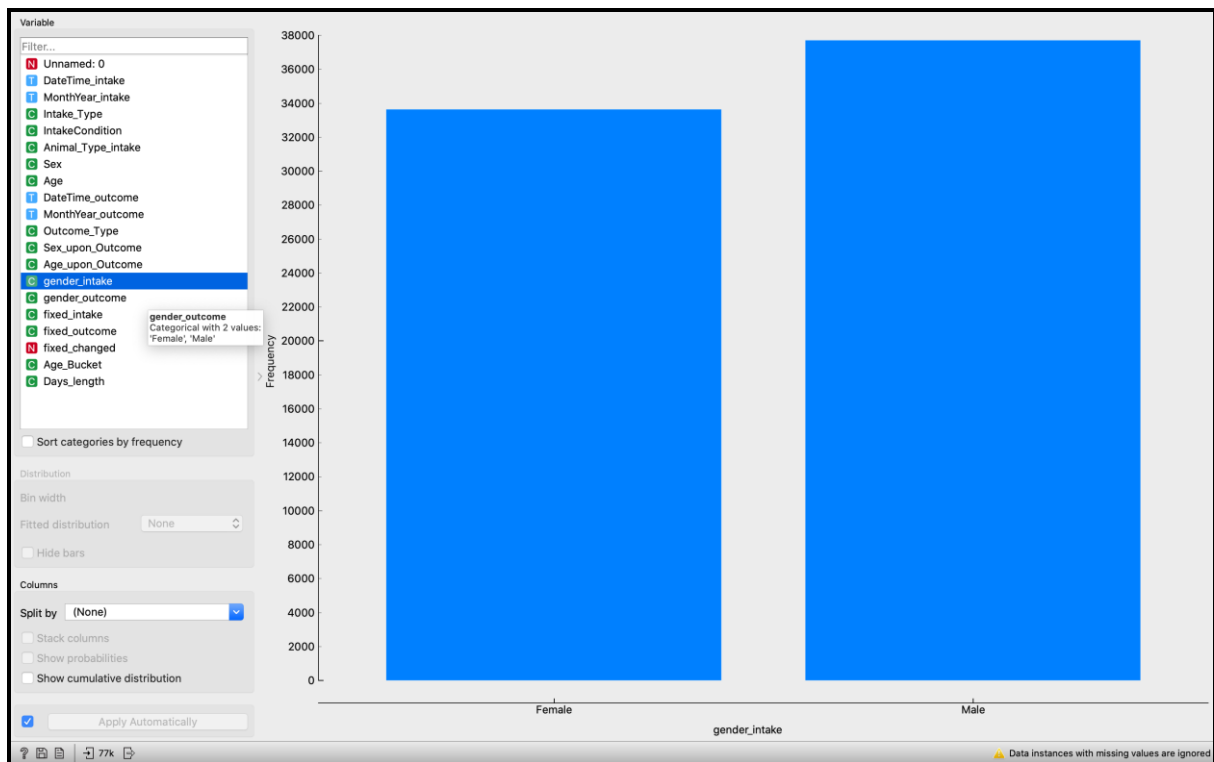


Figure 6.6: Bar Plot (Gender).

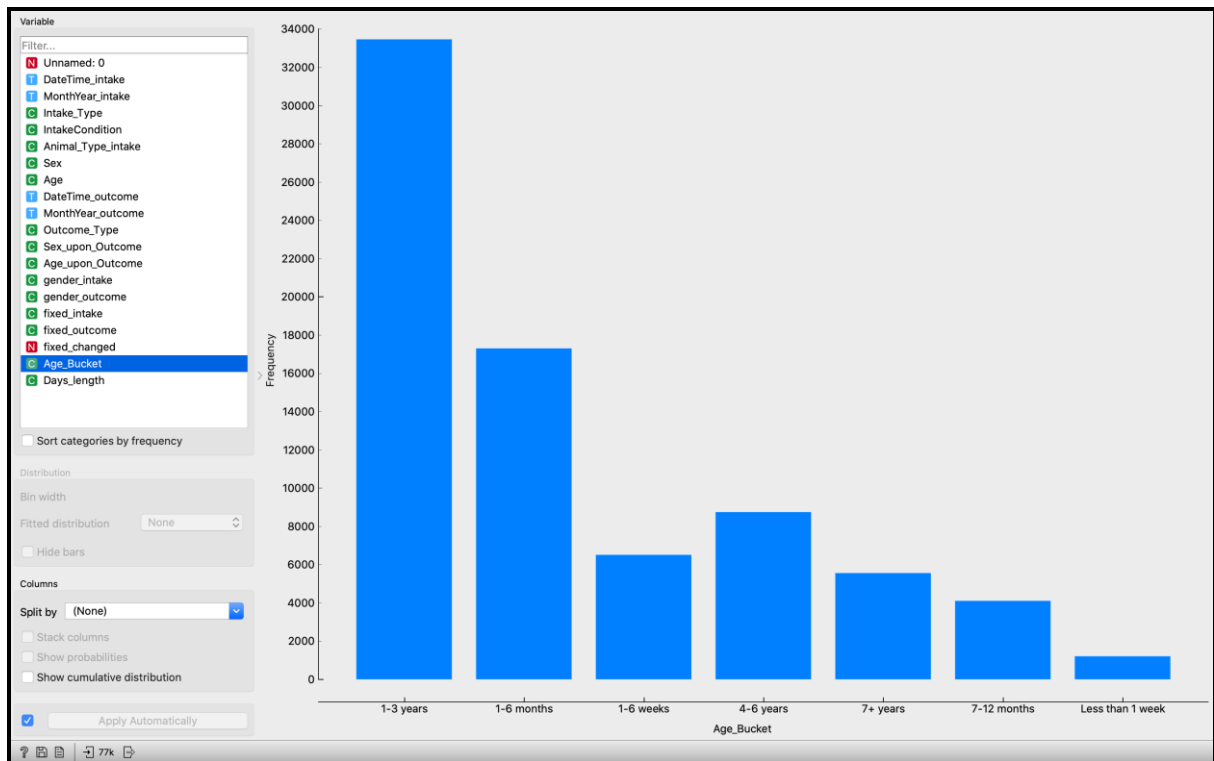


Figure 6.7: Bar Plot (Age).

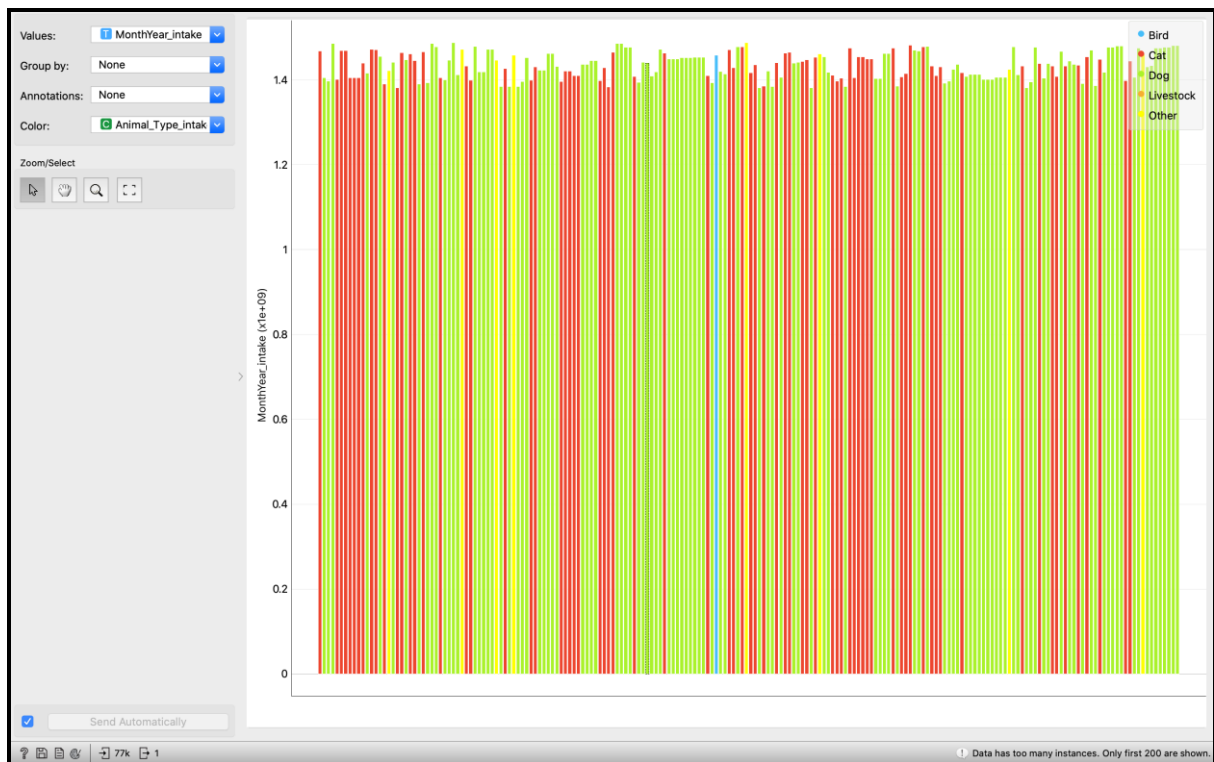


Figure 6.8: Histogram (Date Intake).

Chapter 6

Conclusion

Orange is a C++ core object and routines library that incorporates a huge variety of standard and non-standard machine learning and data mining algorithms. It is an open-source data visualization, data mining, and machine learning tool. Orange is a scriptable environment for quick prototyping of the latest algorithms and testing patterns. It is a group of python-based modules that exist in the core library. It implements some functionalities for which execution time is not essential, and that is done in Python.



orange
DATA MINING

It incorporates a variety of tasks such as pretty-print of decision trees, bagging and boosting, attribute subset, and many more. Orange is a set of graphical widgets that utilizes strategies from the core library and orange modules and gives a decent user interface. The widget supports digital-based communication and can be gathered together into an application by a visual programming tool called an orange canvas.

All these together make an orange an exclusive component-based algorithm for data mining and machine learning. Orange is proposed for both experienced users and analysts in data mining and machine learning who want to create and test their own algorithms while reusing as much of the code as possible, and for those simply entering the field who can either write short python contents for data analysis.

- **Strengths of using Orange as data mining tool:**
 - Open Source
 - Available on different prominent OS.
 - Python supported.
 - Python scripting inbuilt.
 - Appealing GUI.
 - Easy accessibility.
 - Various features of visualization and data processing available.
 - Support many Machine Learning Algorithms.

- **Weakness of using Orange data mining tool:**
 - No scope for developing ML programming skills. As usually we can connect objects on the canvas.
 - Not repeatable. You'll have to start again for different parameter intake.
 - Big Installation.
 - Limited Reporting capabilities.

References

- [1] Manisha Chandak, Sheetal Girase, Introducing hybrid technique for optimization of book recommender system, International Conference on Advanced Computing Technologies and Applications (ICACTA-2015), Procedia Computer Science page nos- 23 – 31, ELSEVIER
- [2] Semantic Scholar: <https://www.semanticscholar.org/paper/Comparative-Study-of-Data-Mining-Tools-Rangra-Bansal/1597c7a98a9915dcde0f6e4aefda15cd032d6b13>
- [3] Clustering: https://www.tutorialspoint.com/data_mining/dm_cluster_analysis.htm#:~:text=Clustering%20is%20the%20process%20of,the%20labels%20to%20the%20groups.
- [4] Orange Data Mining Library: <https://orange3.readthedocs.io/projects/orange-data-mining-library/en/latest/tutorial/data.html>
- [5] Nvidia Blogs: <https://blogs.nvidia.com/blog/2018/08/02/supervised-unsupervised-learning/#:~:text=In%20a%20supervised%20learning%20model,and%20patterns%20on%20its%20own.>