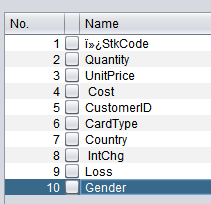
I am doing a comparison between two countries (India and Australia) and the credit card usage in the counties as per the gender.

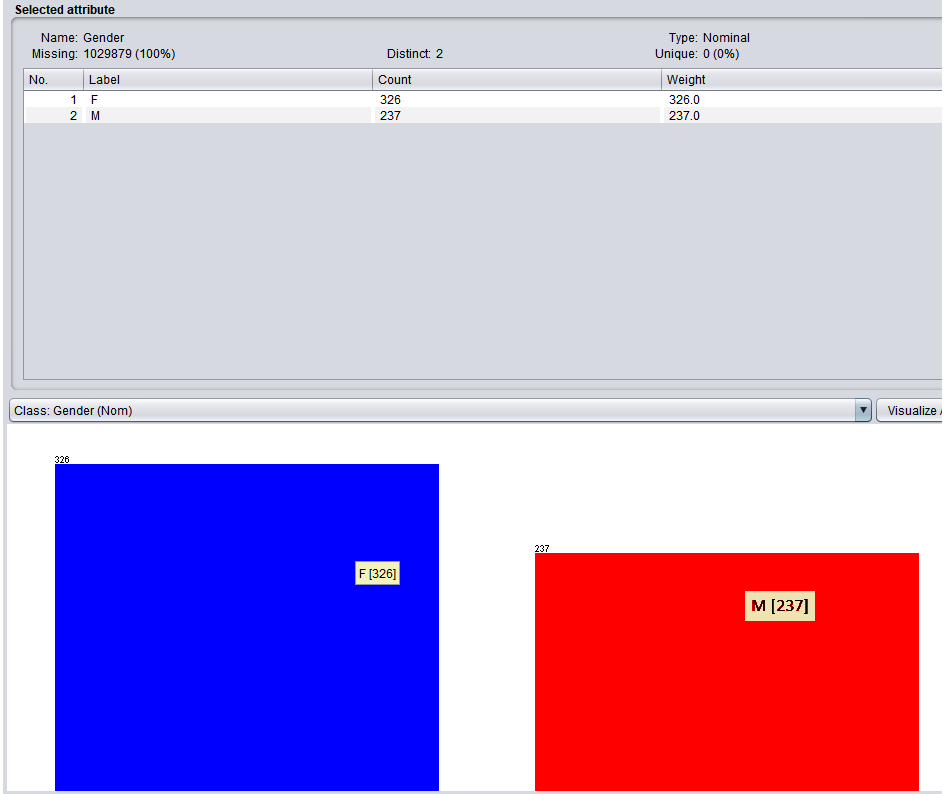
India Data

Data visualization on the following columns: Stk Code is the description of the items purchased.

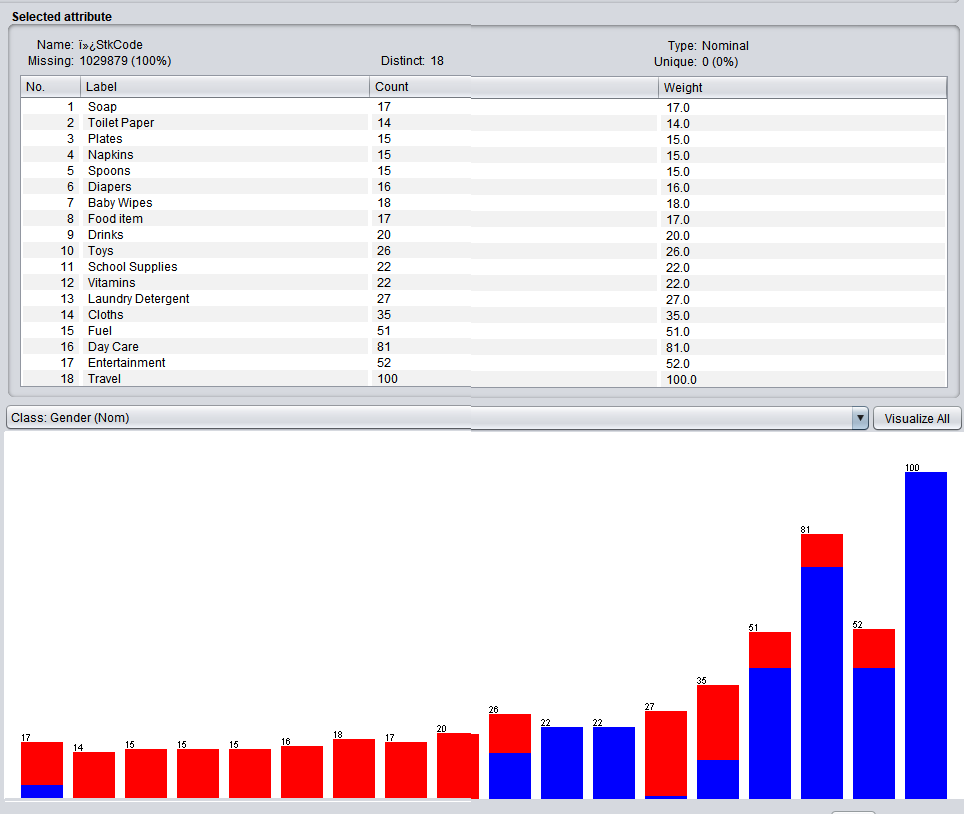


Below is the data for India

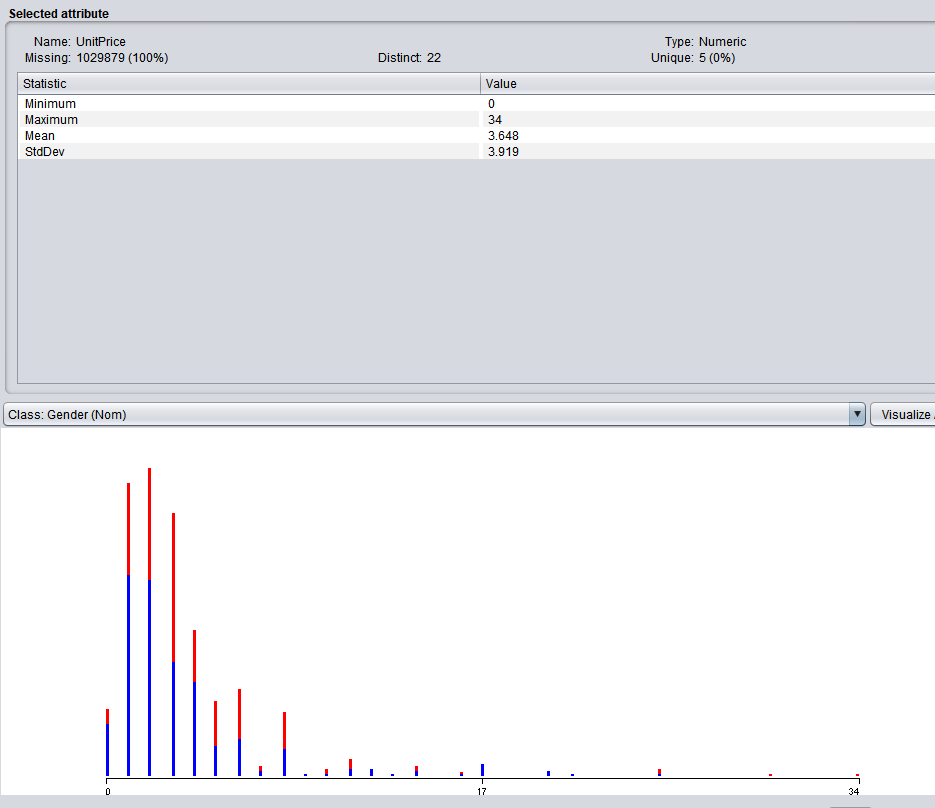
Credit cards used by Male and Female Population



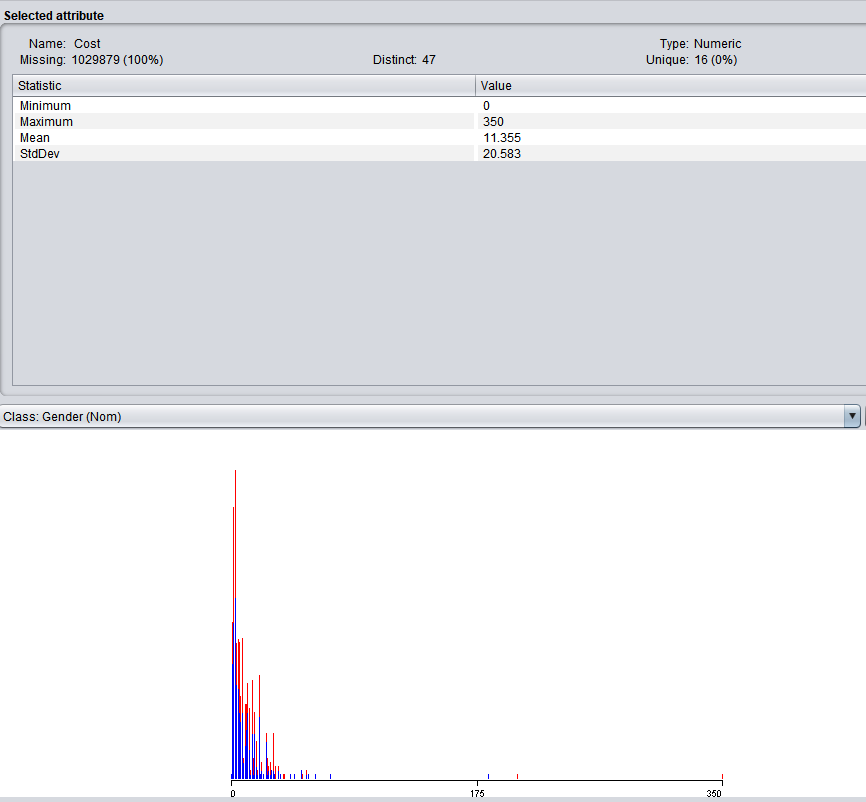
This screen shot show what type of purchase was the credit card was used for by which gender.



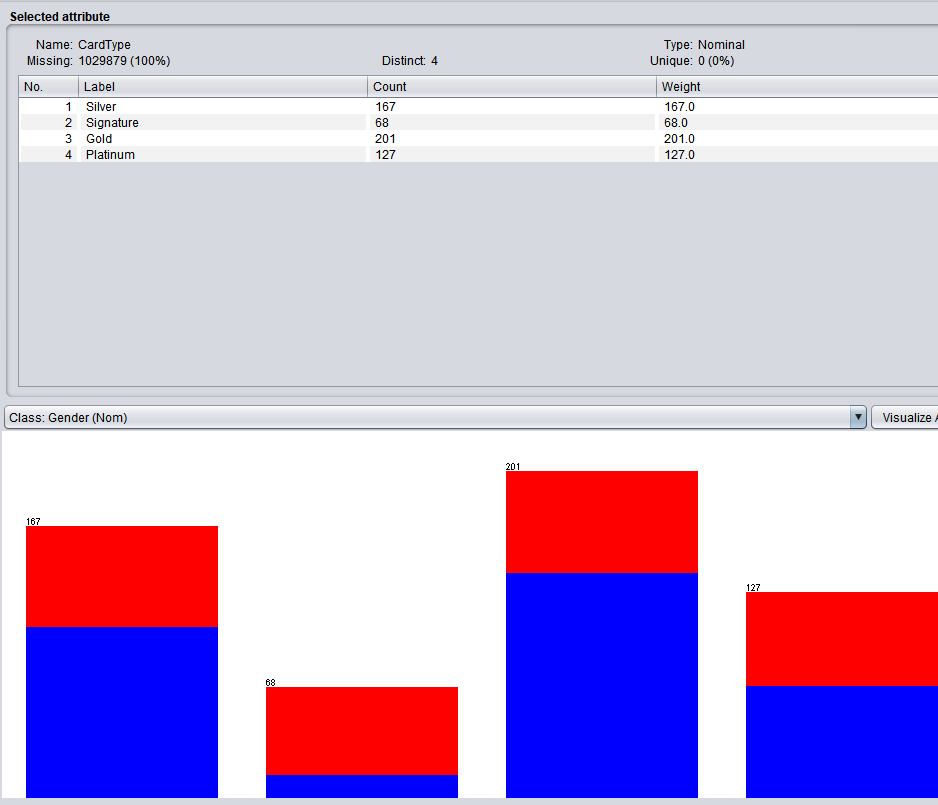
Unit price of items purchased (shows – max, min and standard deviation)



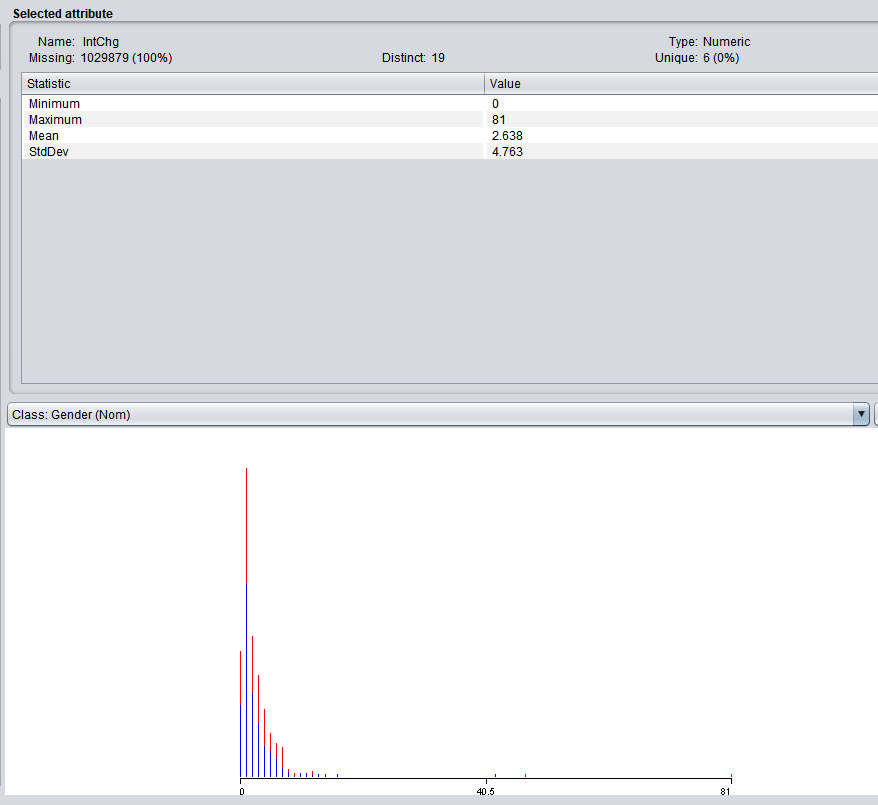
Item Cost is shown below with min, max and std Deviation.



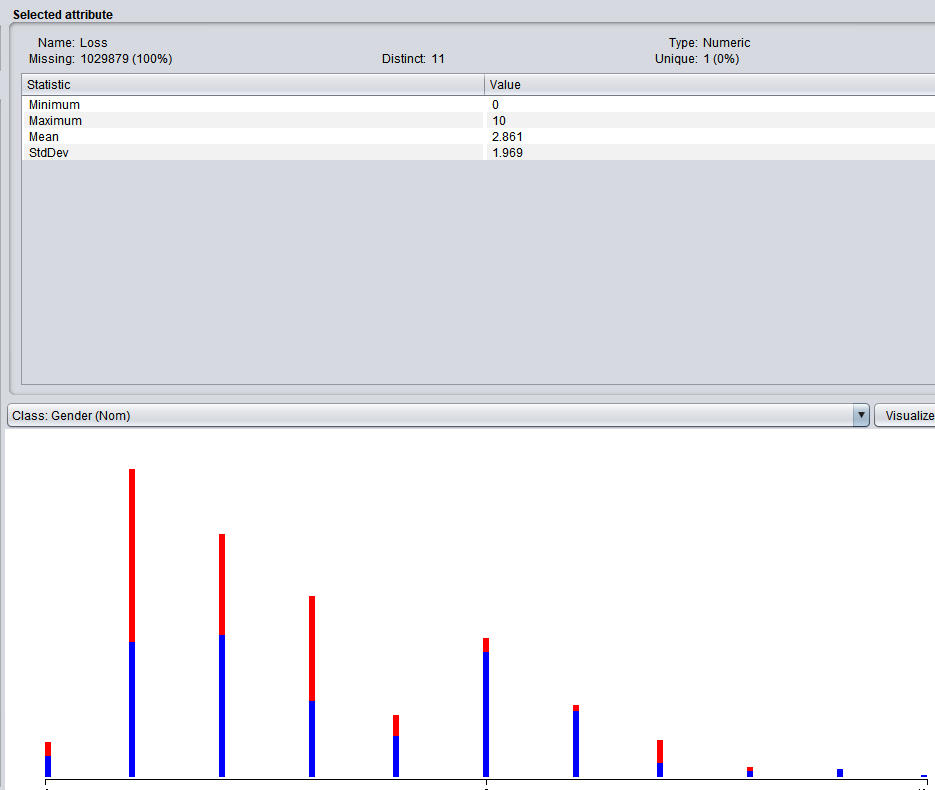
Card Usage by card type is shown below where blue is the Female and “gold” has the most usages.



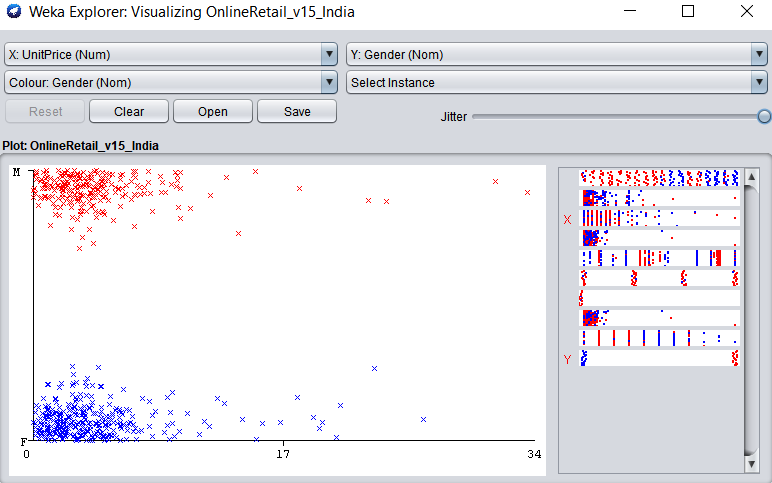
Interest Charged were shown here



The screen shot below shows the cc losses



Unit price by gender visualization



Classifier SubSet

=== Run information ===

Evaluator: weka.attributeSelection.ClassifierSubsetEval -B weka.classifiers.rules.ZeroR -T -H "Click to set hold out or test instances" -E DEFAULT

Search: weka.attributeSelection.BestFirst -D 1 -N 5

Relation: Online Retail\_v2

Instances: 205386

Attributes: 11

ï»¿InvoieNo

StockCode

Quantity

UnitPrice

Cost

CustomerID

CardType

Country

IntChg

Loss

Gender

Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

number of folds (%) attribute

0( 0 %) 1 ï»¿InvoieNo

0( 0 %) 2 StockCode

0( 0 %) 3 Quantity

0( 0 %) 4 UnitPrice

0( 0 %) 5 Cost

0( 0 %) 6 CustomerID

0( 0 %) 7 CardType

0( 0 %) 8 Country

0( 0 %) 9 IntChg

0( 0 %) 10 Loss

=== Classifier Run information ===

Scheme: weka.classifiers.rules.ZeroR

Relation: OnlineRetail\_v1\_IndiaAustralia

Instances: 1030357

Attributes: 10

ï»¿StockCode

Quantity

UnitPrice

Cost

CustomerID

CardType

Country

IntChg

Loss

Gender

Test mode: evaluate on training data

=== Classifier model (full training set) ===

ZeroR predicts class value: F

Time taken to build model: 0.05 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.4 seconds

=== Summary ===

Correctly Classified Instances 722 64.2349 %

Incorrectly Classified Instances 402 35.7651 %

Kappa statistic 0

Mean absolute error 0.4595

Root mean squared error 0.4793

Relative absolute error 100 %

Root relative squared error 100 %

Total Number of Instances 1124

Ignored Class Unknown Instances 1029233

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure MCC ROC Area PRC Area Class

1.000 1.000 0.642 1.000 0.782 ? 0.500 0.001 F

0.000 0.000 ? 0.000 ? ? 0.500 0.000 M

Weighted Avg. 0.642 0.642 ? 0.642 ? ? 0.500 0.001

=== Confusion Matrix ===

a b <-- classified as

722 0 | a = F

402 0 | b = M

Cluster

=== Run information ===

Scheme: weka.clusterers.FilteredClusterer -F "weka.filters.AllFilter " -W weka.clusterers.SimpleKMeans -- -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 2 -A "weka.core.EuclideanDistance -R first-last" -I 500 -num-slots 1 -S 10

Relation: Online Retail\_v2

Instances: 205386

Attributes: 11

ï»¿InvoieNo

StockCode

UnitPrice

Cost

CustomerID

CardType

Country

IntChg

Loss

Gender

Ignored:

Quantity

Test mode: evaluate on training data

Selected Attributes

== Run information ===

Evaluator: weka.attributeSelection.PrincipalComponents -R 0.95 -A 5

Search: weka.attributeSelection.Ranker -T -1.7976931348623157E308 -N -1

Relation: Online Retail\_v2

Instances: 205386

Attributes: 11

ï»¿InvoieNo

StockCode

Quantity

UnitPrice

Cost

CustomerID

CardType

Country

IntChg

Loss

Gender

Evaluation mode: evaluate on all training data