

Step1. Open the data/ bmw-browsers.arff Dataset

Load the data file bmw-browsers.arff into WEKA

Weka Explorer

Preprocess | Classify | Cluster | Associate | Select attributes | Visualize | Forecast

Open fil... | Open U... | Open D... | Generat... | Undo | Edit... | Save...

Filter: Choose None Apply

Current relation
Relation: car-browsers
Instances: 100
Attributes: 8
Sum of weights: 100

Attributes
All | None | Invert | Pattern

No.	Name
1	Dealership
2	Showroom
3	ComputerSearch
4	M5
5	3Series
6	Z4
7	Financing
8	Purchase

Remove

Selected attribute
Name: Dealership
Missing: 0 (0%)
Distinct: 2
Type: Numeric
Unique: 0 (0%)

Statistic	Value
Minimum	0
Maximum	1
Mean	0.6
StdDev	0.492

Class: Purchase (Num) Visualize All

40 60
0 0.5 1

Status: OK Log x 0

Step2. Creating the Clusters with WEKA

1. click on the **Cluster** tab
2. Click **Choose** and select **SimpleKMeans** from the choices that appear.
3. Adjusting the **numClusters** field from 2 to 5.
4. Click Start button to create clusters.

Weka Explorer

Preprocess **Classify** **Cluster** Associate Select attributes Visualize Forecast

Clusterer
 SimpleKMeans -init 0 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t1 -1.25 -t2 -1.0 -N 5 -f

Cluster mode
☒ Use training set
☐ Supplied test set Set...
☐ Percentage split % 66
☐ Classes to clusters evaluation
 (Num) Purchase
☒ Store clusters for visualization

Ignore attributes

Start Stop

Result list (right-click for options)
 15:09:22 - SimpleKMeans

Clusterer output

Dealership	0.6	0.9615	0.6667	1
Showroom	0.72	0.6923	0.6667	0
ComputerSearch	0.43	0.6538	0	1
M5	0.53	0.4615	0.963	1
3Series	0.55	0.3846	0.4444	0.8
Z4	0.45	0.5385	0	0.8
Financing	0.61	0.4615	0.6296	0.8
Purchase	0.39	0	0.5185	0.4

Time taken to build model (full training data) : 0.03 seconds

=== Model and evaluation on training set ===

Clustered Instances

0	26 (26%)
1	27 (27%)
2	5 (5%)
3	14 (14%)
4	28 (28%)

Status
OK

Log x 0

Step3: Interpreting the Clusters

The output is telling us how each cluster comes together, with a "1" meaning everyone in that cluster shares the same value of one, and a "0" meaning everyone in that cluster has a value of zero for that attribute. Numbers are the average value of everyone in the cluster. Each cluster shows us a type of behavior in our customers, from which we can begin to draw some conclusions:

- **Cluster 0:** This group we can call the "Dreamers," as they appear to wander around the dealership, looking at cars parked outside on the lots, but trail off when it comes to coming into the dealership, and worst of all, they don't purchase anything.
- **Cluster 1:** We'll call this group the "M5 Lovers" because they tend to walk straight to the M5s, ignoring the 3-series cars and the Z4. However, they don't have a high purchase rate only 52 percent. This is a potential problem and could be a focus for improvement for the dealership, perhaps by sending more salespeople to the M5 section.
- **Cluster 2** This group is so small we can call them the "Throw-Aways" because they aren't statistically relevant, and we can't draw any good conclusions from their behavior. (This happens sometimes with clusters and may indicate that you should reduce the number of clusters you've created).

- **Cluster 3:** This group we'll call the "BMW Babies" because they always end up purchasing a car and always end up financing it. Here's where the data shows us some interesting things: It appears they walk around the lot looking at cars, then turn to the computer search available at the dealership. Ultimately, they tend to buy M5s or Z4s (but never 3-series). This cluster tells the dealership that it should consider making its search computers more prominent around the lots (outdoor search computers?), and perhaps making the M5 or Z4 much more prominent in the search results. Once the customer has made up his mind to purchase the vehicle, he always qualifies for financing and completes the purchase.
- **Cluster 4:** This group we'll call the "Starting Out With BMW" because they always look at the 3-series and never look at the much more expensive M5. They walk right into the showroom, choosing not to walk around the lot and tend to ignore the computer search terminals. While 50 percent get to the financing stage, only 32 percent ultimately finish the transaction. The dealership could draw the conclusion that these customers looking to buy their first BMWs know exactly what kind of car they want (the 3-series entry-level model) and are hoping to qualify for financing to be able to afford it. The dealership could possibly increase sales to this group by relaxing their financing standards or by reducing the 3-series prices.

Step4: Cluster visual inspection

- To do this, you should right-click on the **Result List** section of the **Cluster** tab ,Choose one of the options from this pop-up menu is **Visualize Cluster Assignments**.

