

The variables and operators that have the best effect on the predictive power of the model, also have a local performance indicator¹ attached (see previous figure). This indicator is a number between 50 and 100, where a lower number indicates lower predictive power. The total predictive power of this model is 79.04 (see top of model). The best individual variable is *ppersaut* (68.47). The performance significantly improves if this variable is combined with *minkgem* and *moshoofd*. Figure 2 shows the output of the operator “vex_cave” that combines these variables. The square-shaped segments in this figure represent the caravan-policy-interest² by their colour and height. The caravan-policy-interest increases with the colors ranging from green, yellow, red, purple to blue. Figure 2 shows that for *moshoofd* values 3,8,1,9 and especially 2 in combination with *minkgem* values 4 and 5, the model indicates an increased caravan-policy-interest. The nice thing about this approach is that it is not only based on probabilities for the subgroups but also on the number of cases that are in these subgroups. Statistically insignificant subgroups, no matter how high its probabilities, do not necessarily get high scores. Also, another nice thing is that other combinations in the model influence the combination of *moshoofd* and *minkgem*.

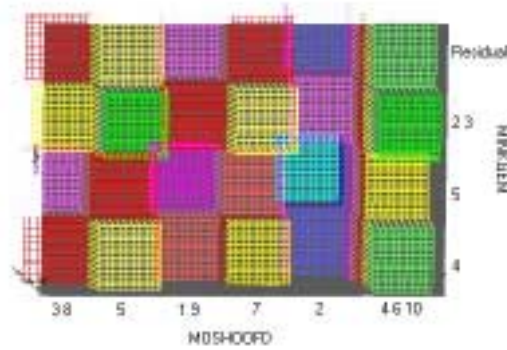


Figure 3. Combination of *moshoofd* and *minkgem*

By combining the output of the previously described variables via the operator “vex_cave” with *ppersaut*, the operator “quadv” generates scores locally in the model according to figure 3. This shows that for increase in “vex_cave” the local output also increases. Moreover, for *ppersaut* value 6, the local score is much higher.

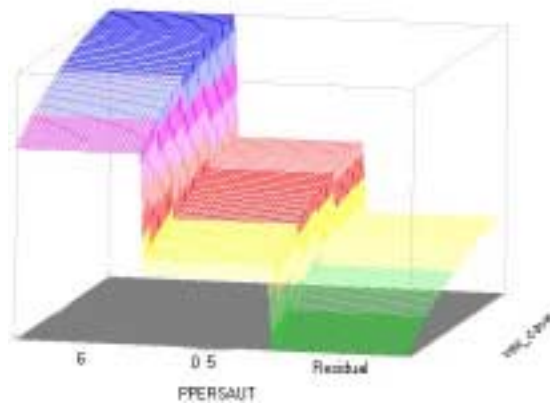


Figure 4. Combination of *ppersaut* and “vex_cave”

Therefore, clients with cars that are insured according to category 6 (fl.1000-fl.4999), have an increased caravan-policy-interest. The local performance has increased to 75.22.

From here, by moving up inside the model, the local performance increases. By taking other variables into account the model fine-tunes its predictions. For example, *pplazier* at

¹ This rankcorrelation indicator is related to Kendall’s tau [Kendall M.G., 1948, *Rank Correlation methods*, Griffin, London]

² The caravan-policy-interest is defined as the percentage of caravan policy holders in the selected group

a first glance, looks very important (if unequal 0, the corresponding caravan-policy-interest is about 40%), but discriminates over too few cases to be significant. Therefore, also in the model, its influence is little.

THIS MODEL HAS A RANKCORRELATION PERFORMANCE OF 79.04 ON THE TRAININGSET, WHICH CONSISTS OF AN 80% RANDOM SAMPLE OF ALL CASES. AT A MAILVOLUME OF 20%, THE CARAVAN-POLICY-INTEREST IS 19.4%. ON THE VALIDATION SET (THE REMAINING 20%), THE MODEL HAS A RANKCORRELATION PERFORMANCE OF 78.98. AT A MAILVOLUME OF 20%, ON THIS VALIDATION SET, THE CARAVAN-POLICY-INTEREST IS 20.3%. THIS IS A HIGHLY ROBUST MODEL.

As a last step, to get an understanding of those customers that have a high caravan-policy-interest, a chi-squared test was performed. By setting a cutoff in the model score at a mailvolume of 20%, two classes are generated. Those customers within the class with a score above this cutoff are compared with the total dataset via a chi-squared test. This results in a summary of these customers (in class 1) as shown below.

Summary for class 1, projected select rate : 20.00%, penetration rate : 19.37%
 Projected probability of caravan policy holder behaviour 19.35%

```

Most likely characteristics
-----
mhkoop      8, 9
mskd        1, 0
pinboed     0
mgodrk      0, 2, 1
mberboer    0
moplhoog    2, 1, 3
agezong     0
pwapart     2
pfiets      0
pbrand      5, 2, 3, 4, 1, 6, 7, 8
mfgekind    2, 0, 5, 3, 4, 1
pwerkt      0
pplezier    0

Sub Group
Attributes:
  aplezier: 2, 1
  pinboed: 2, 1, 3, 6, 5, 4
  pplezier: 0

abystand    0
ppersong    0
awaoreg     0
aplezier    0
mgodov      0, 2
pbrom       0
atractor    0
mshoofd 2, 8, 3, 9, 1
minkgem 4
ppersaut 6
  
```

The summary above speaks for itself. In this summary, the most important variables are underlined.

SUMMARY

It is important to notice that for the mailing we are aiming at a mailvolume of 20%. As a result, small interesting subgroups (like *pplezier* value unequal 0, people having boat policies), are of less importance. The most decisive characteristic of caravan policy holders is that *ppersaut* has a value of 6.