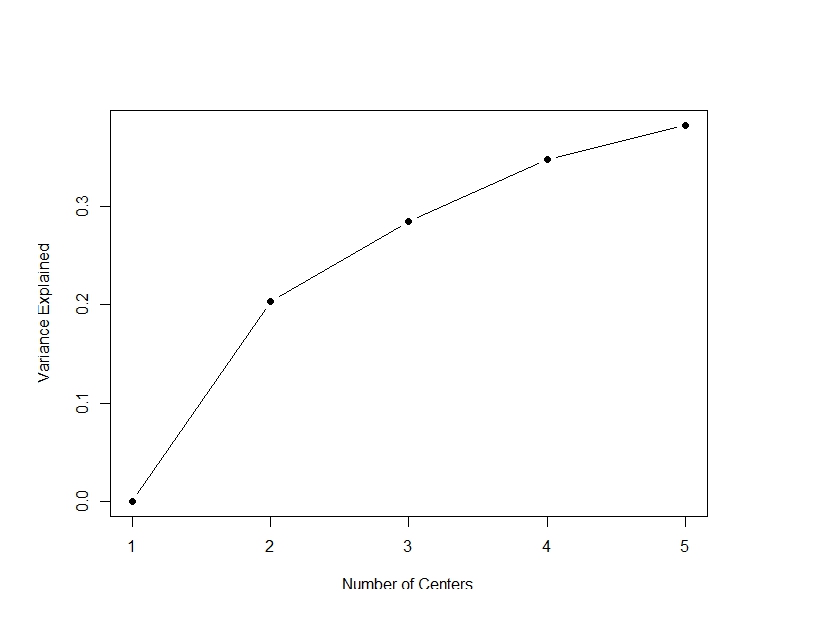
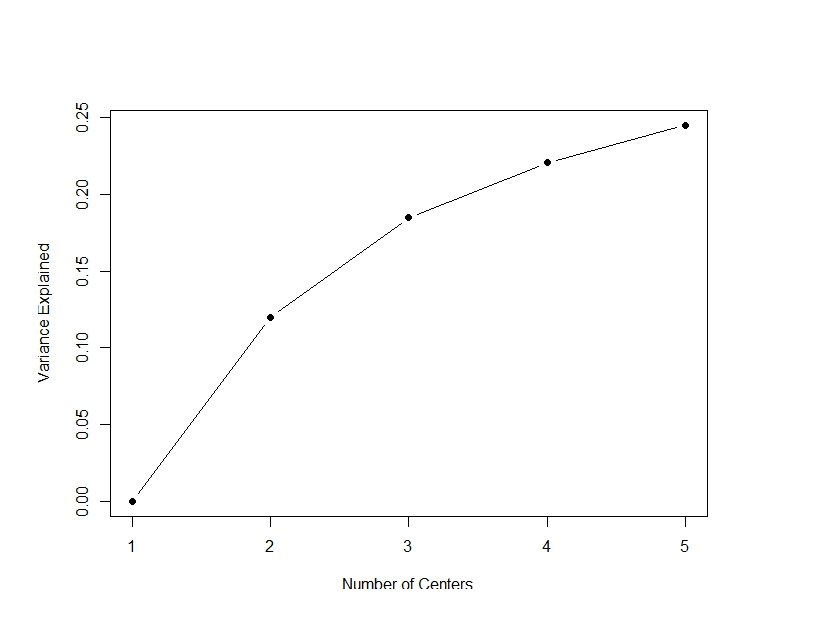
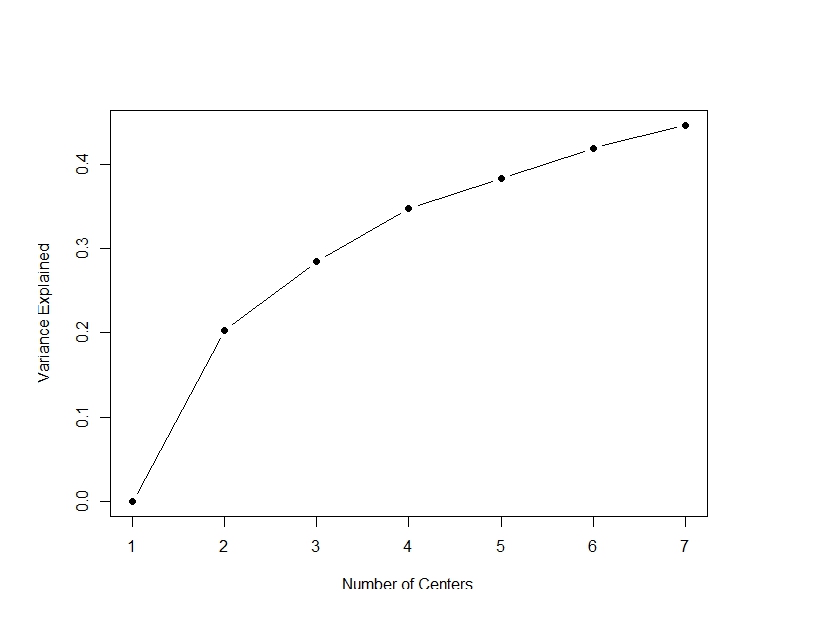
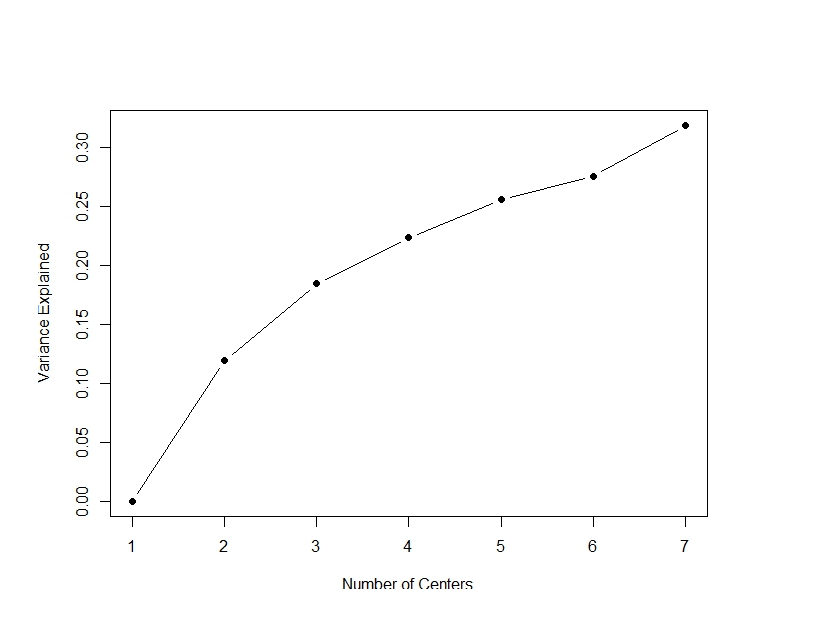
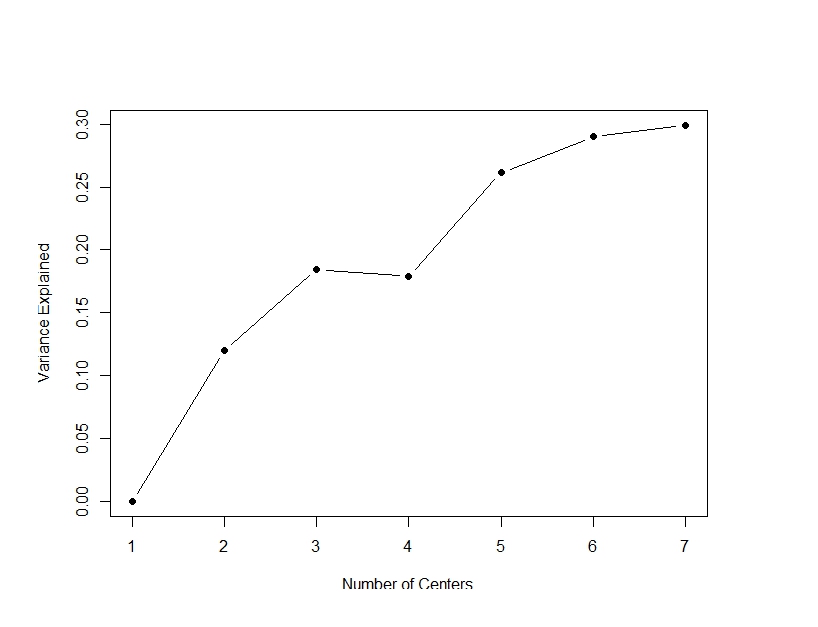
Dataset 1, 5 centers: 

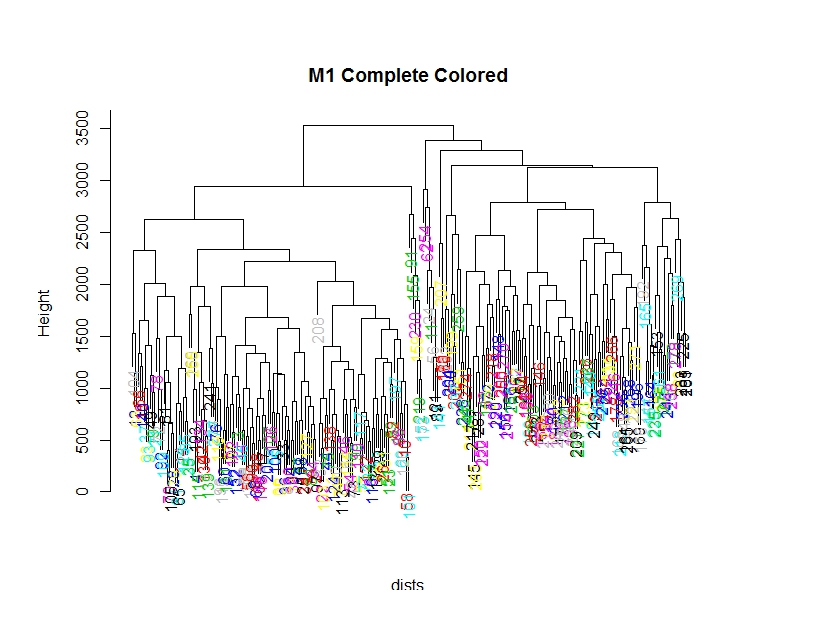
Dataset 2, 5 centers: 

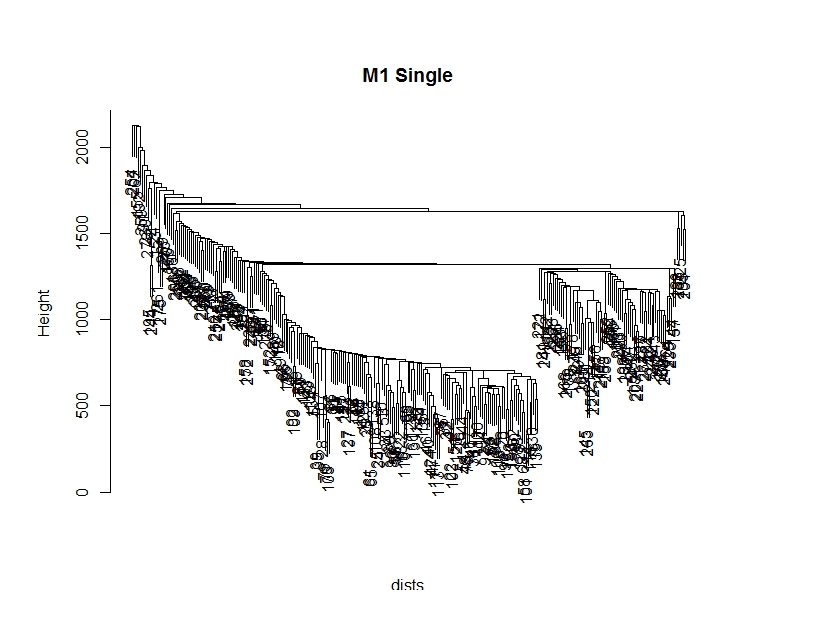
Dataset 1, 7 centers: 

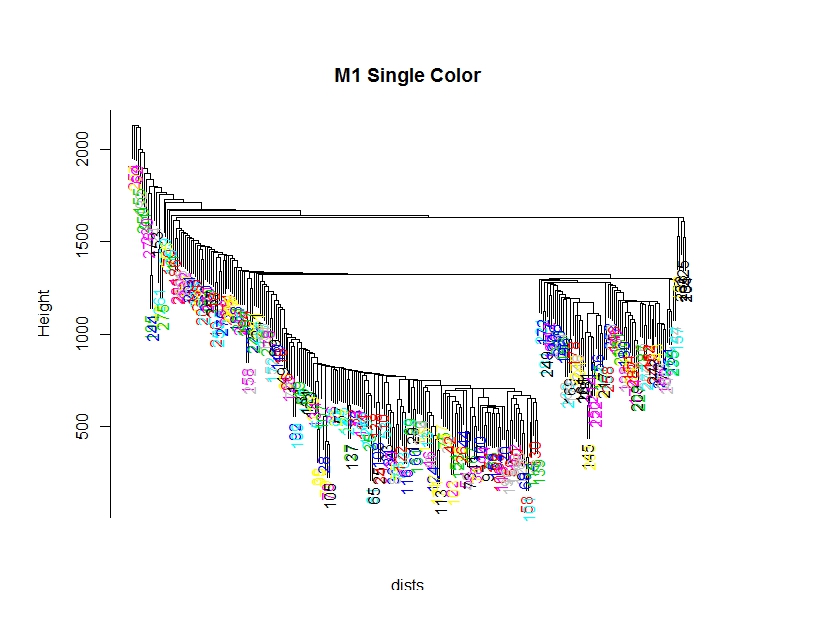
Dataset 2, 7 centers:

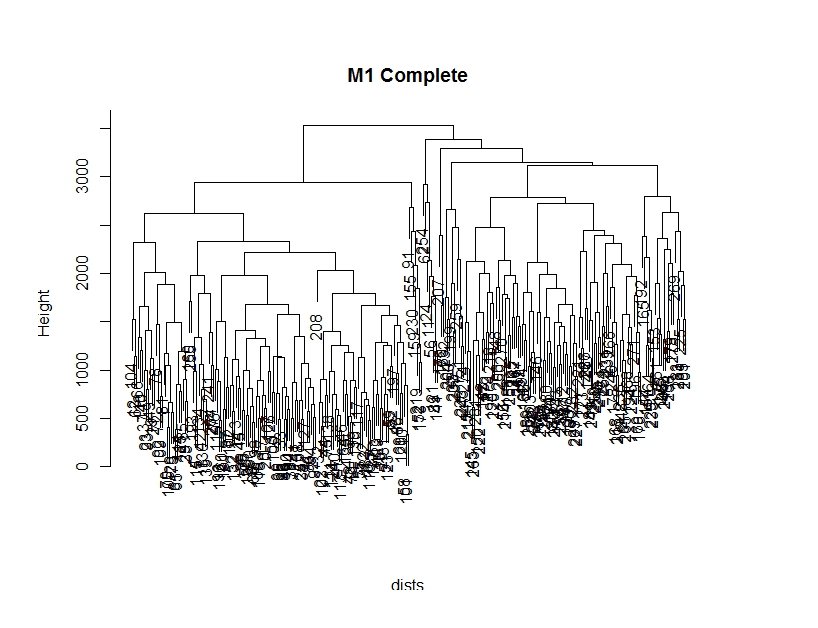
Dataset 2, 7 centers: 

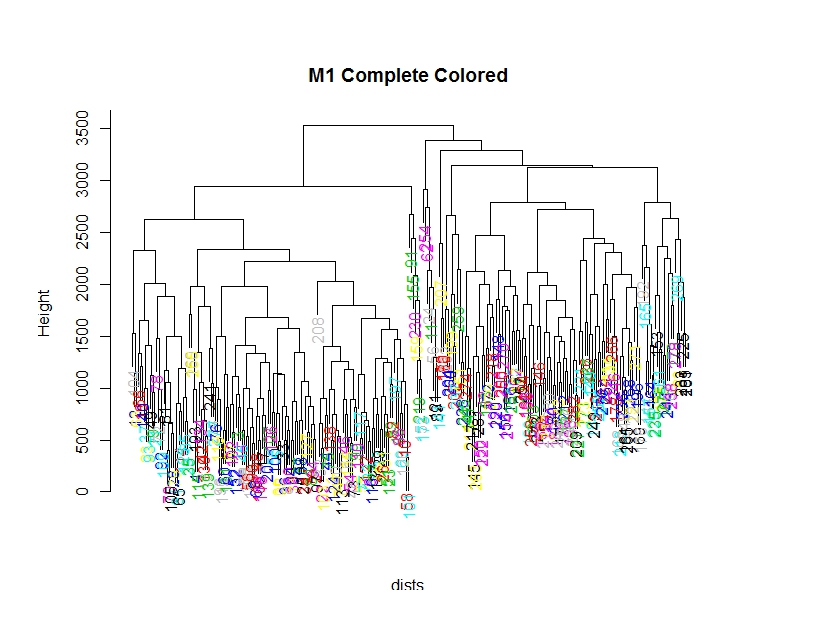
From one particular instance of running the k means. Odd results.

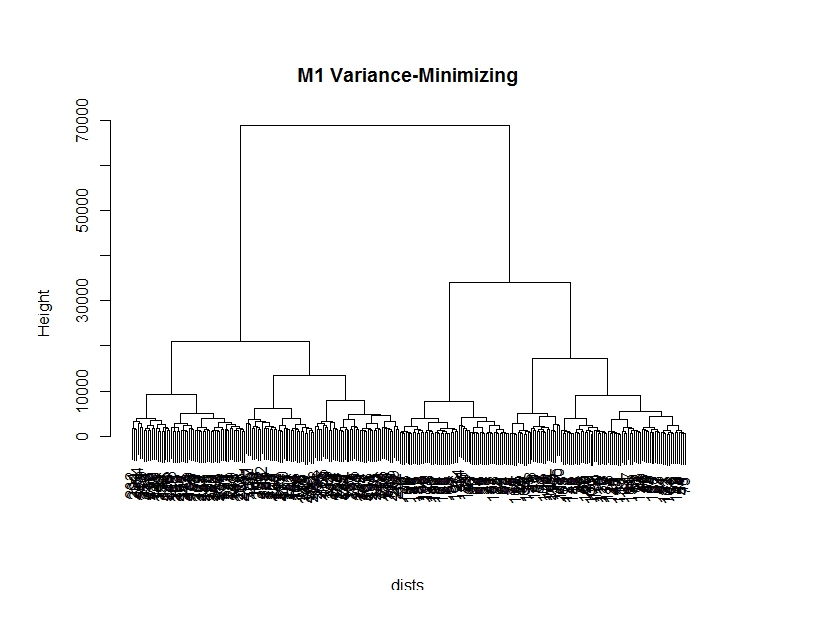


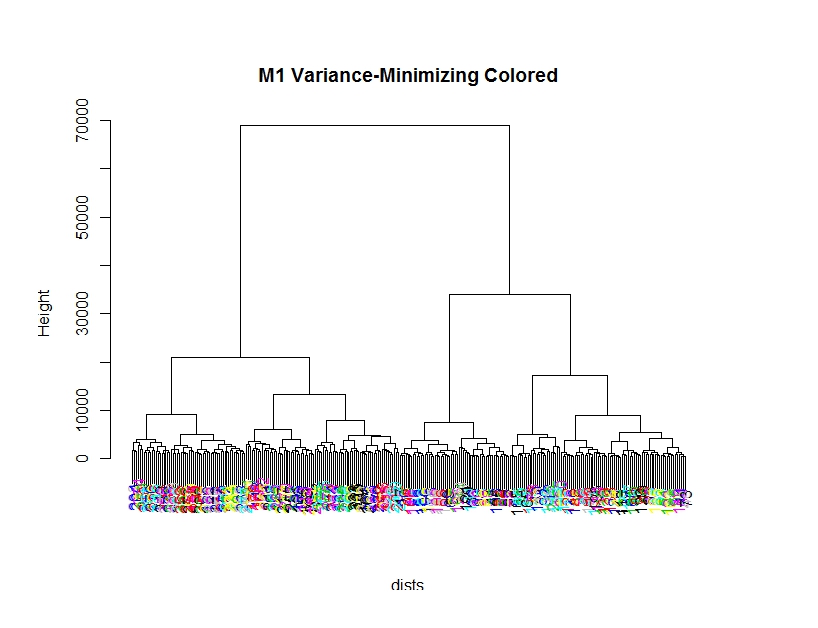


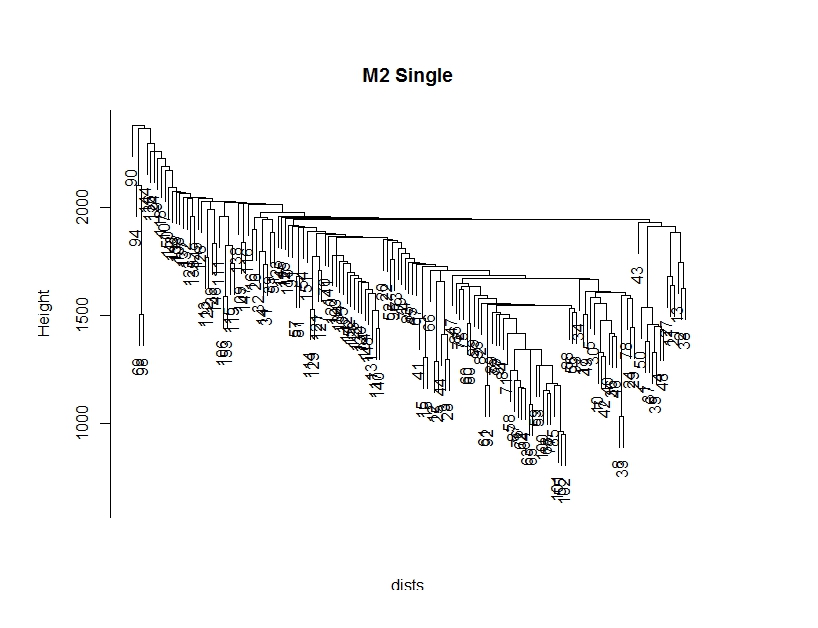


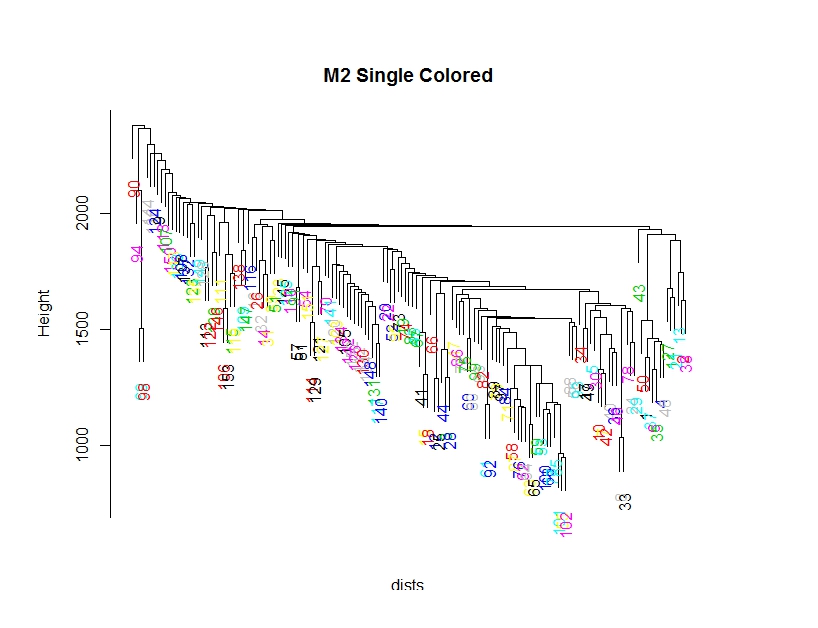


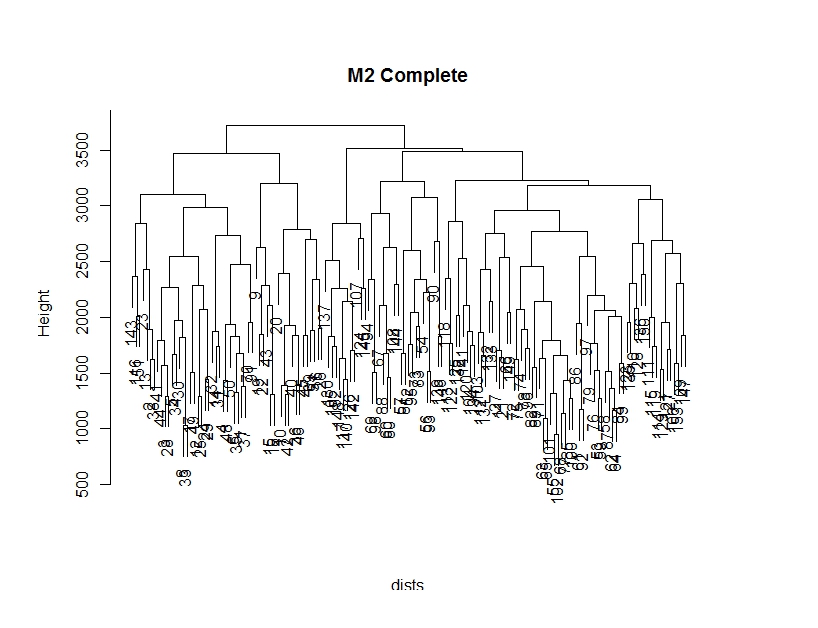


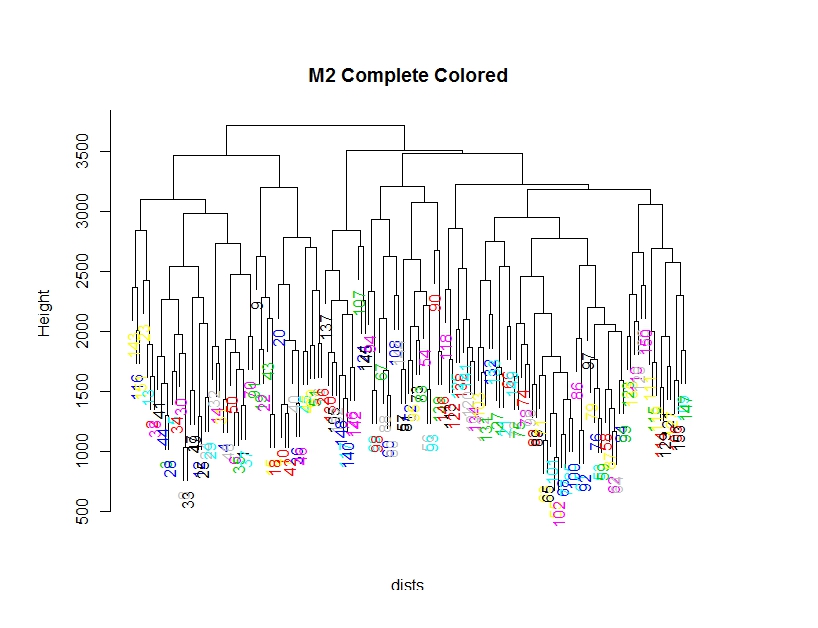


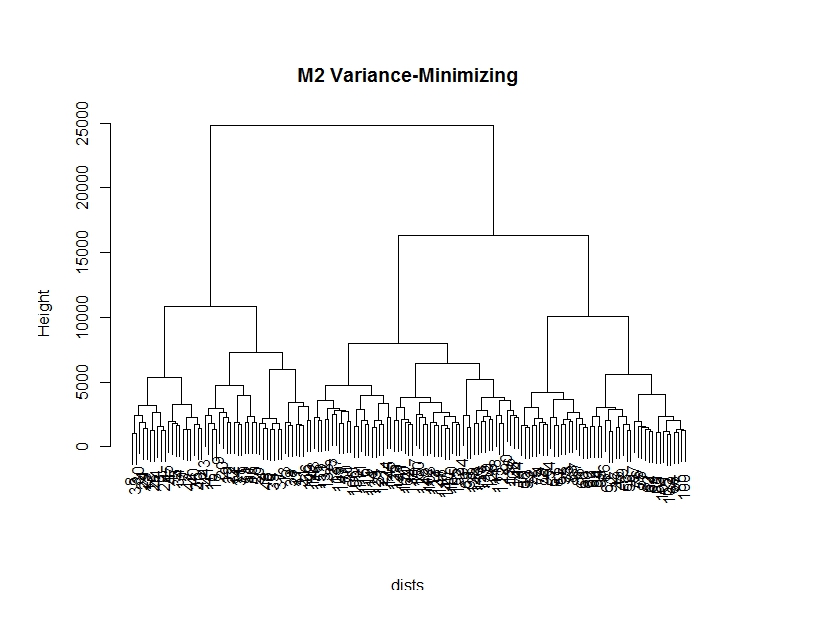


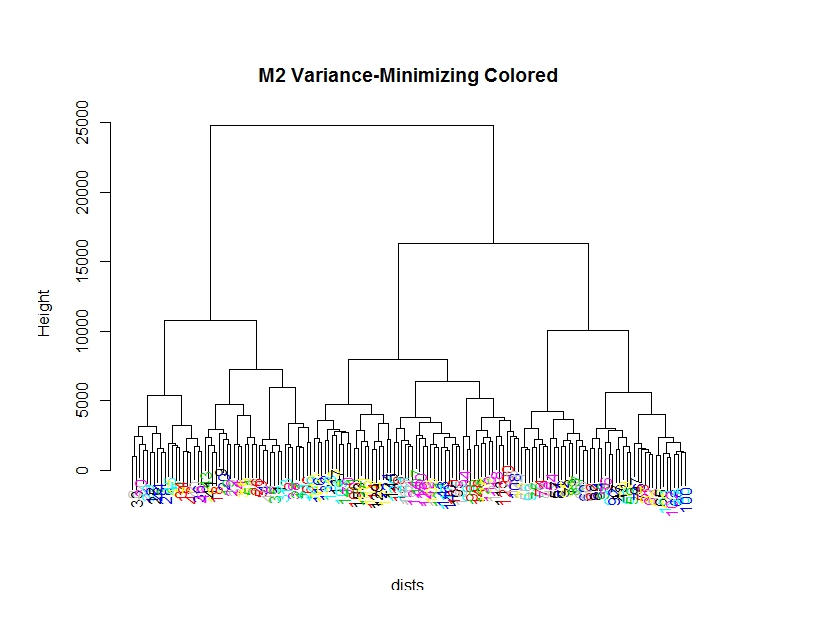






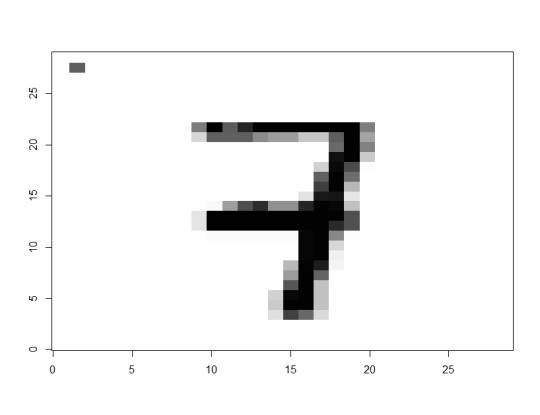
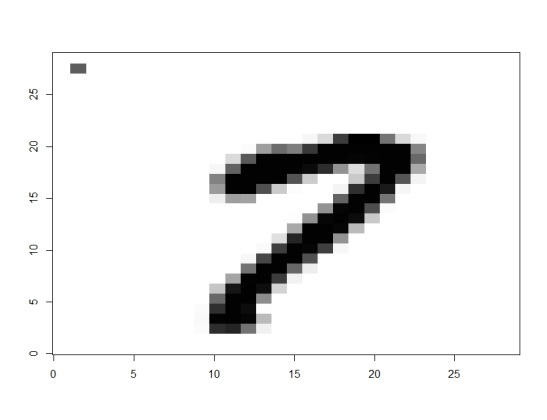






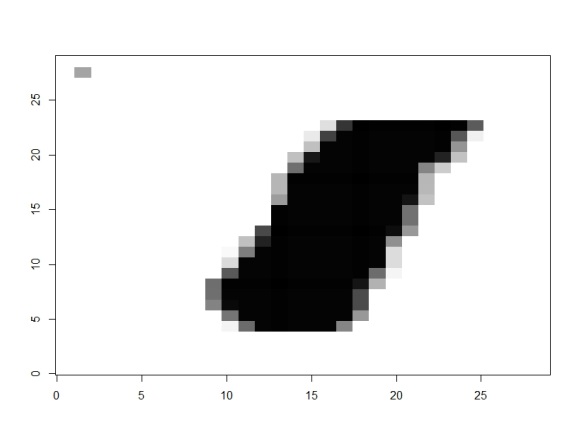
Variance-minimizing is the most useful linkage-type for these data sets. In fact, it’s the only useful linkage-type of the three; single present unintelligible data, and complete is rather unappealing – but it does at least indicate the “outliers”, the oddly or poorly written digits from the dataset. In contrast, variance-minimizing linkage presents the data in a neatly organized fashion, and seems to have succeeded in grouping it together in a meaningful manner.

Variance-minimizing clustering sorted each dataset into two main categories for the first dataset, and three main categories. This is sensible, as dataset 1 had two different digits (1 and 7), and dataset 2 had three different digits (0, 2, and 6). Beyond that, further branches appear due to the differing accepted styles of numbers. For example:

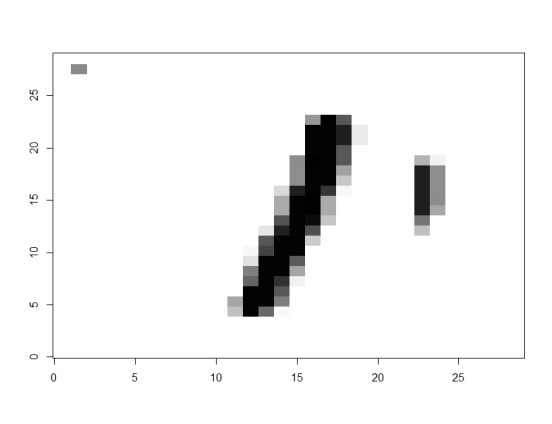


Pictured here are two different styles of 7’s from the first dataset, explaining the secondary branch in the right cluster of the variance-minimizing linkage. Other branches follow in a similar manner, but beyond the secondary branch, this is typically due to carelessness.

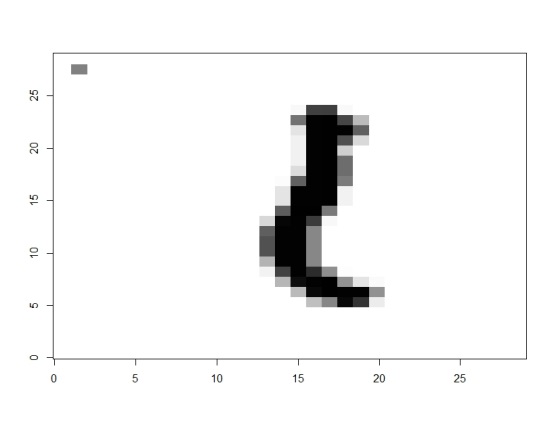
I can’t really explain the number of clusters in the complete and single linkage-types, and I don’t think counting them would be very useful. I have noticed, however, that rather than simply recording which pixels are darkened, the sorting seems to consider the proportionality of the digit itself (which makes sense, considering the mathematics involving distances behind this process). I’d wager that many of the misclassifications are due to skewed proportions. Below I’ve gathered some examples of the “outliers”, as indicated by the complete linkage-type.

 m1 91

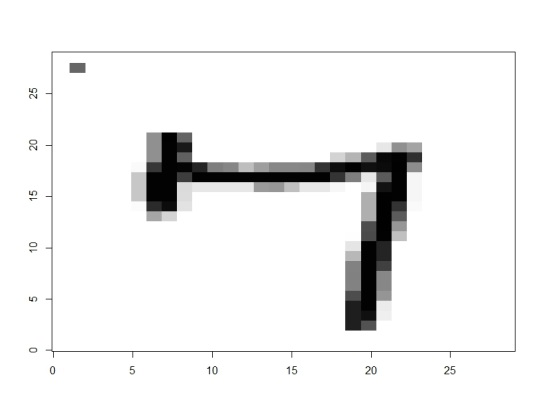
It’s just so much thicker than all the other 1’s.

 m1 117

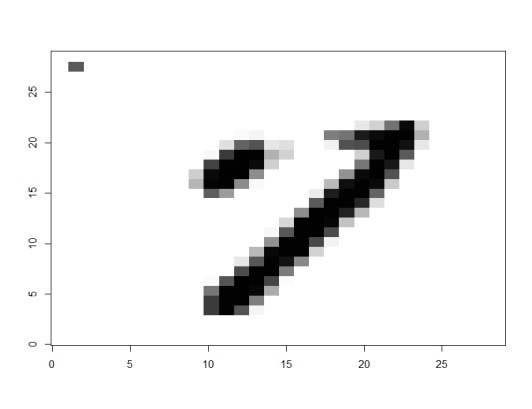
This one may have appeared to be a different digit entirely – ones are typically a continuous shape.

 m1 126

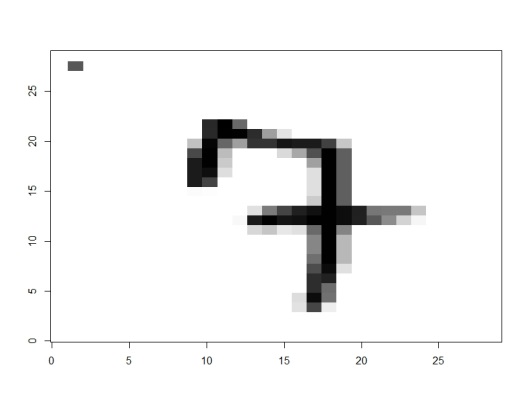
I believe this may be an upside-down question mark without a period.  
 That or a shadowy figure in a hat, hunched in the corner of a train car on a dark, stormy night.

 m1 153

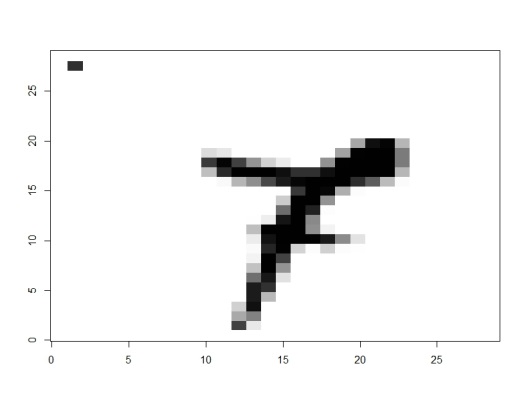
I wouldn’t classify this as a seven, either. Good job, R. I say this one is correctly classified.

 m1 166

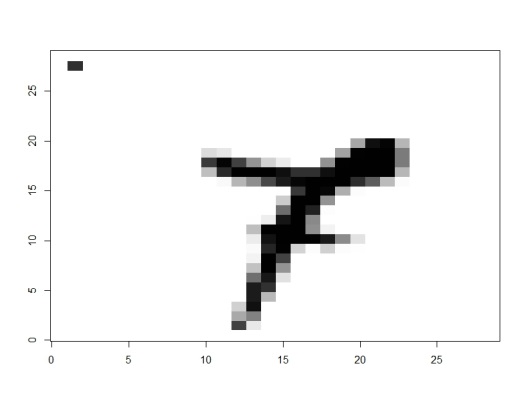
This seven is a metaphor for my broken faith in humanity.

 m1 167

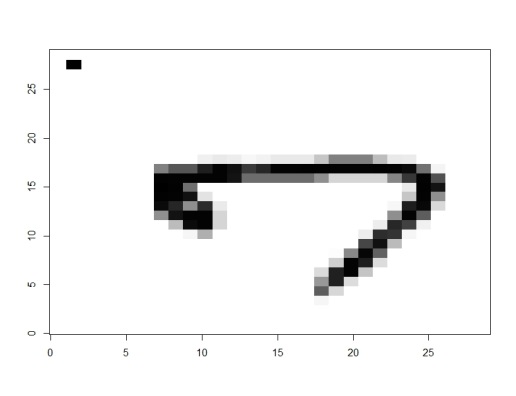
This was probably misclassified because of the weird little hook at the front.  
 Sevens don’t have noses, people. Come on.

 m1 208

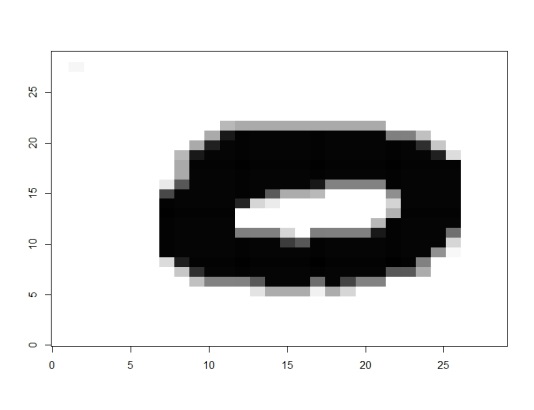
This was misclassified because it was upside-down.  
Rotating the digit, we see that it is not a digit at all:



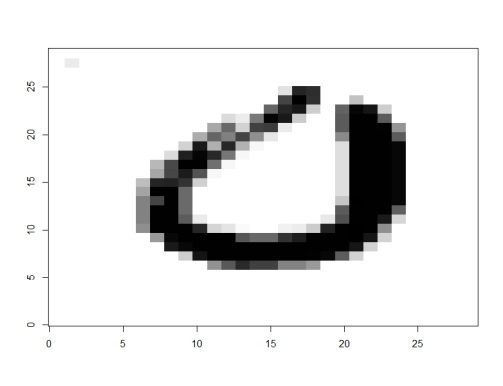
…but instead a wizard’s hat, and therefore shouldn’t be in the dataset.

 m1 254

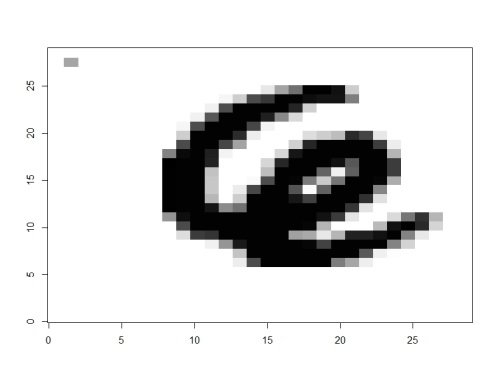
Most sevens are longer in their diagonal line than their horizontal top edge; not so with this one!

 m2 9

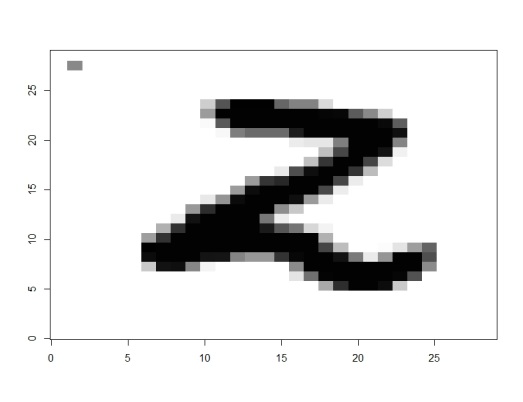
Likely misclassified due to being much thicker than all the other zero’s in the dataset.

 m2 20

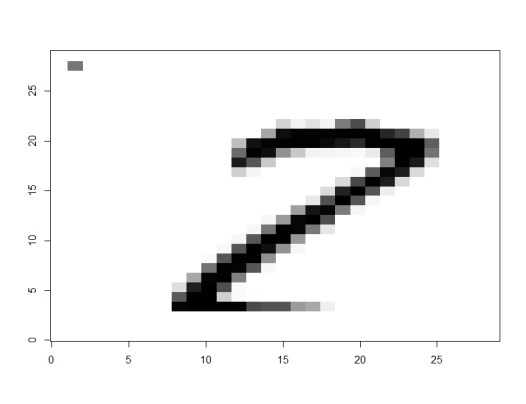
This zero, like a few others, had a gap in it. This was handled all right,  
 but produced a different category from the continuous zeroes.

 m2 90

This was misclassified because it looks more like an flailing octopus than a six.   
 (The right tentacle made for a much different set of calculations than the other sixes,  
 because generally the “loop” also forms the bottom of the six.)

 m2 118

This was much straighter at the top and middle sections than most of the other two’s.  
 I think it looks more like a “z”.

 m2 137

And this one looks more like a seven because of its proportions.