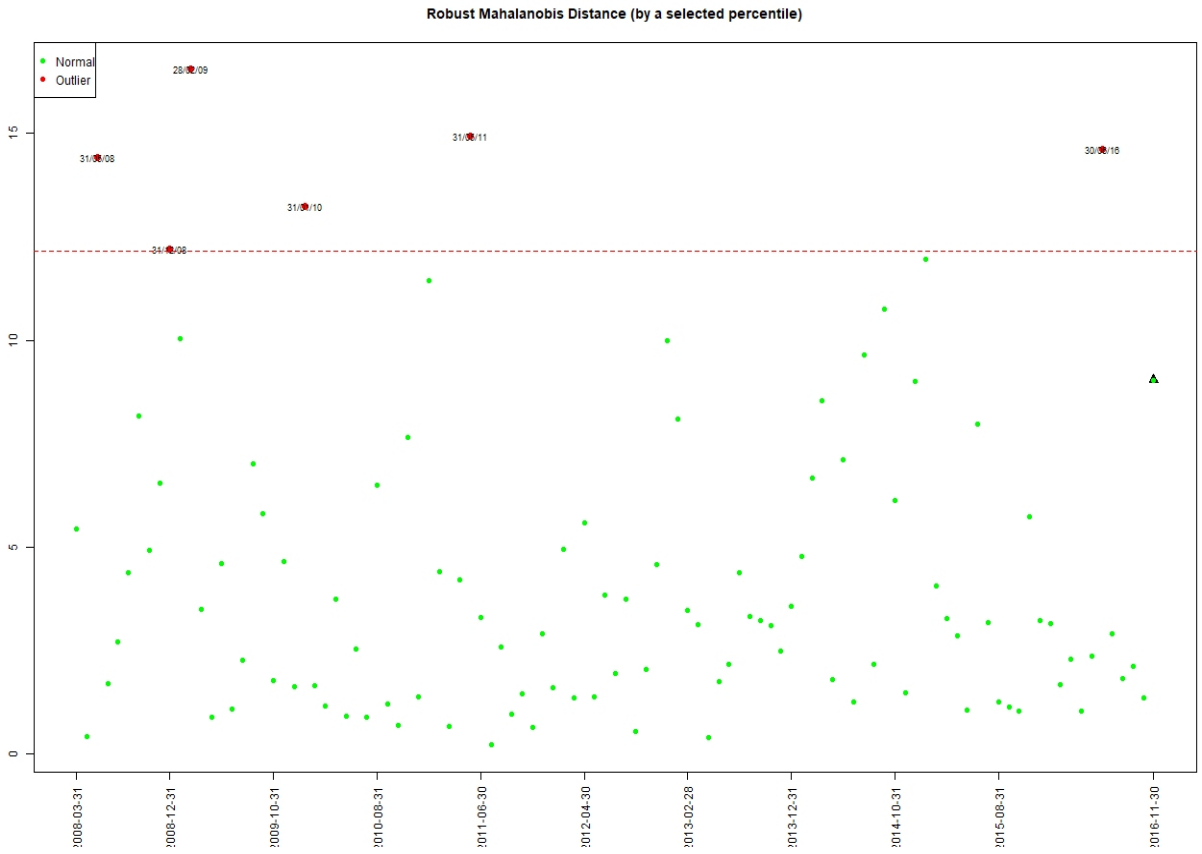
Mahalanobis Distance method - MD

The Mahalanobis distance measures distance relative to the centroid - a base or central point which can be thought of as an overall mean for multivariate data. The centroid is a point in multivariate space where all means from all variables intersect. The larger the MD, the further away from the centroid a data point is. Let x_t be the vector of p observations at day t then, a Mahalanobis distance is defined as:

$$MHD_t = (x_t - M)' S^{-1} (x_t - M)$$

Where, M is the vector of the p variables' means and S is the Var-Cov matrix. MHD in MODS is robust to outliers as it is measured using M as vector of medians rather than means and using MAD (Median Absolute Deviation) with a multiplication factor of 1.4286 (see the help file on the Hampel's method) rather than the common Var-Cov matrix. Screen example is depicted below:



References

Ben-Gal, I. (2005). Outlier detection. In Data mining and knowledge discovery handbook (131-146). Springer, Boston, MA.