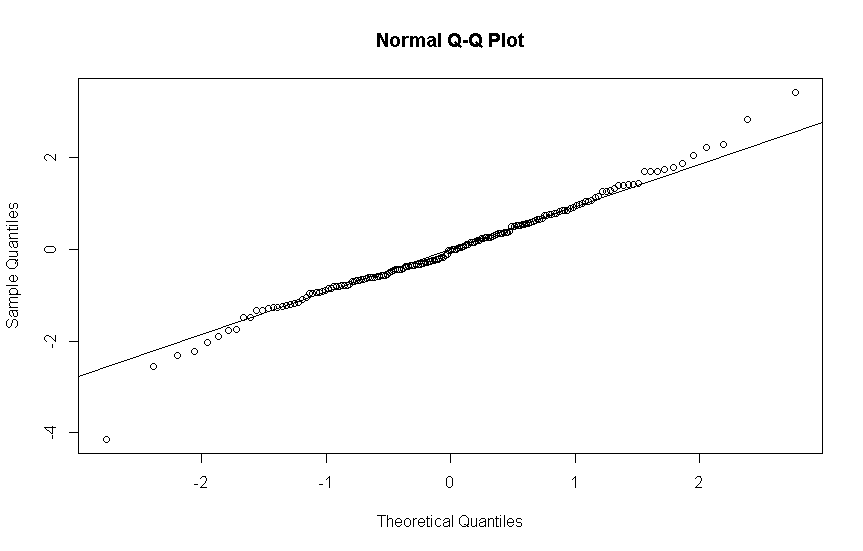
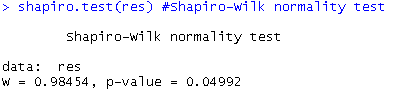
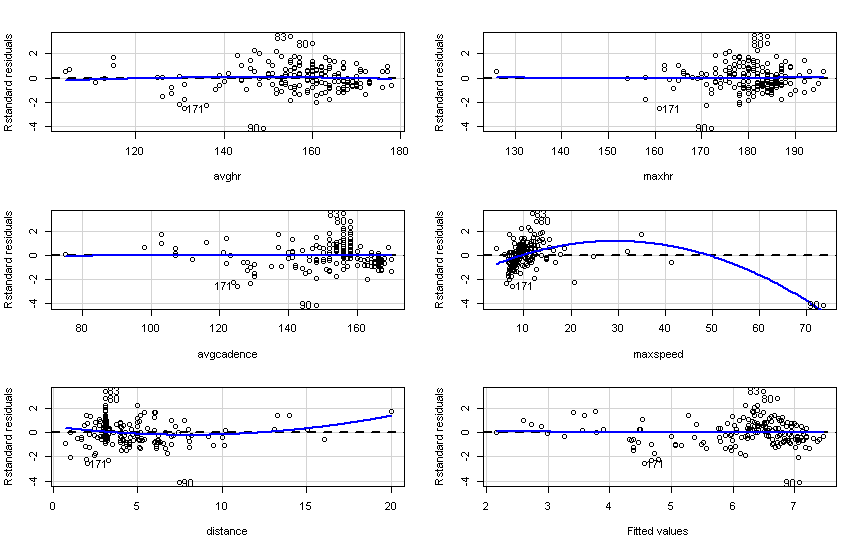


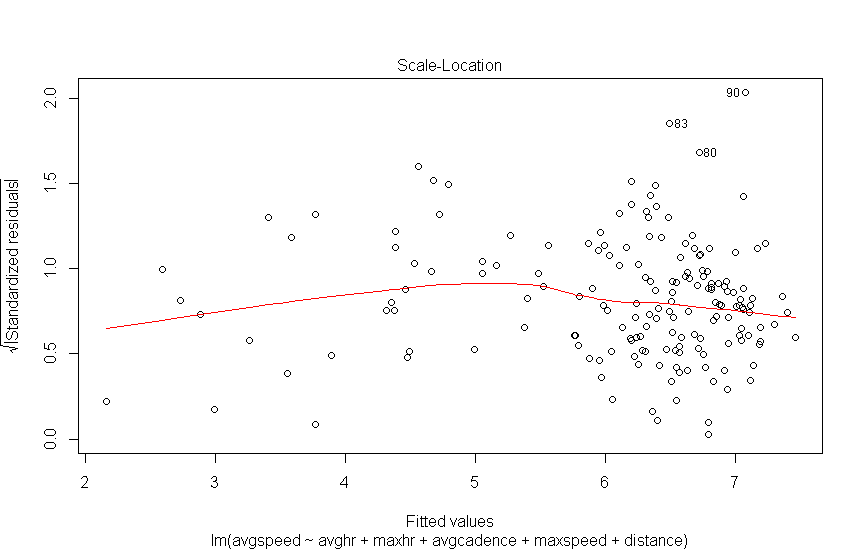
**The individual scatterplots of the response ‘avgspeed’ with these predictors suggests that there is a significant linear relationship that would be found in a simple linear regression model for all predictors except ‘maxhr’ and ‘maxspeed’.**

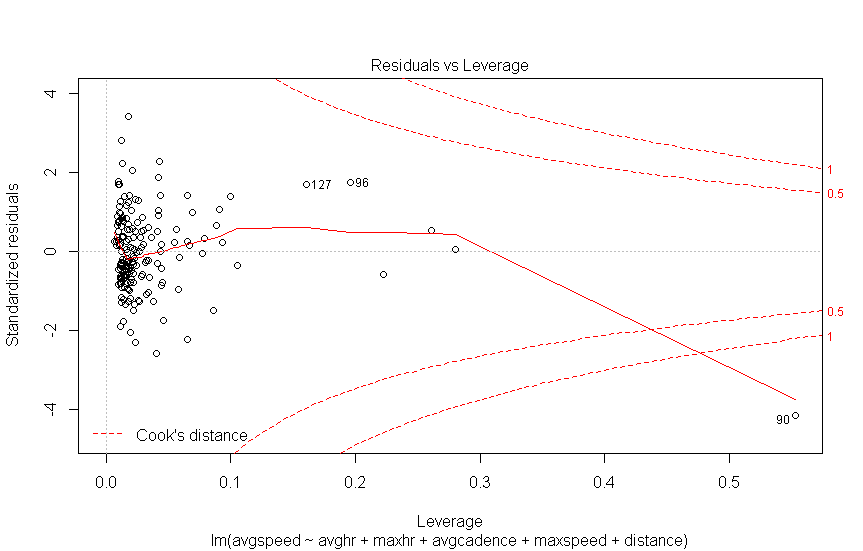
**From the Normal QQ Plot we see a very slight ‘heavy-tailed’ distribution. The p-value is 0.05 which is rather small, leading us to question normality.**

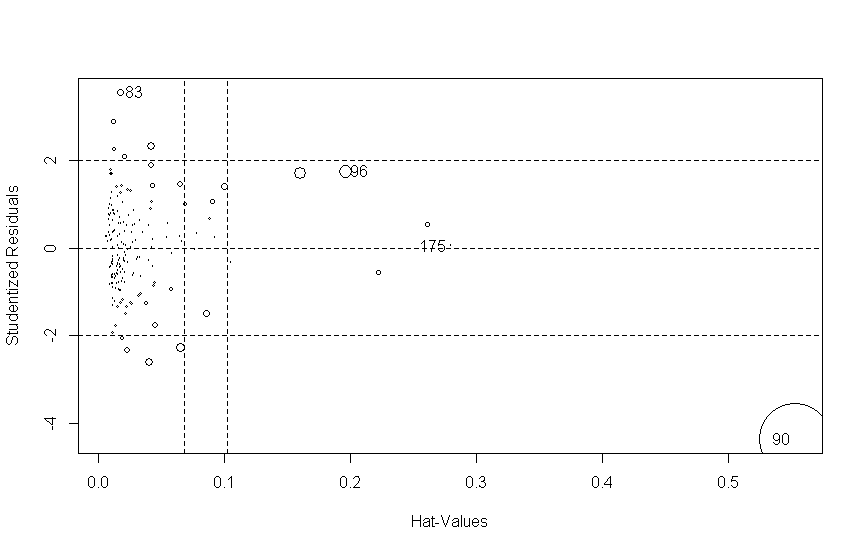


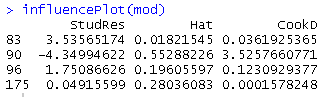
**The linearity assumption looks good for most of the plots above as there are no obvious patterns. We do see an interesting quadratic shape in maxspeed and distance plots against standardized residuals implying that linearity does not hold for those predictors.**

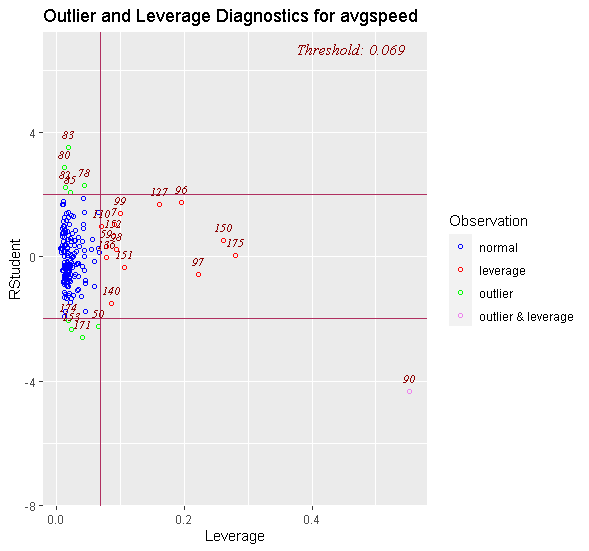


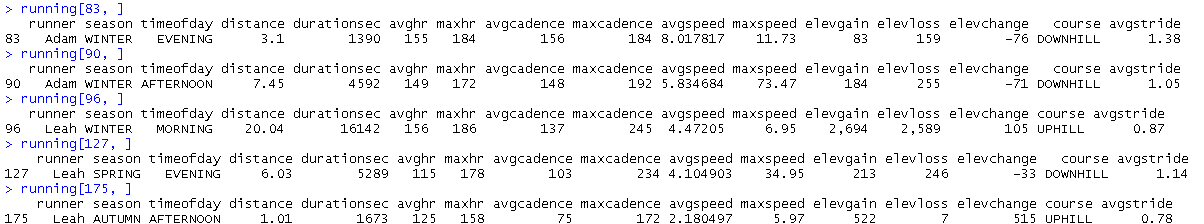
**The spread-location plot is used to access homoscedasticity. As there is no apparent pattern it feels safe to assume homoscedasticity isn’t violated.**



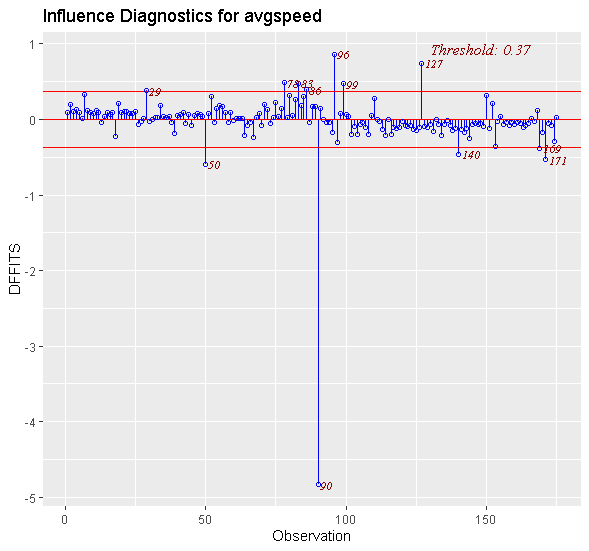




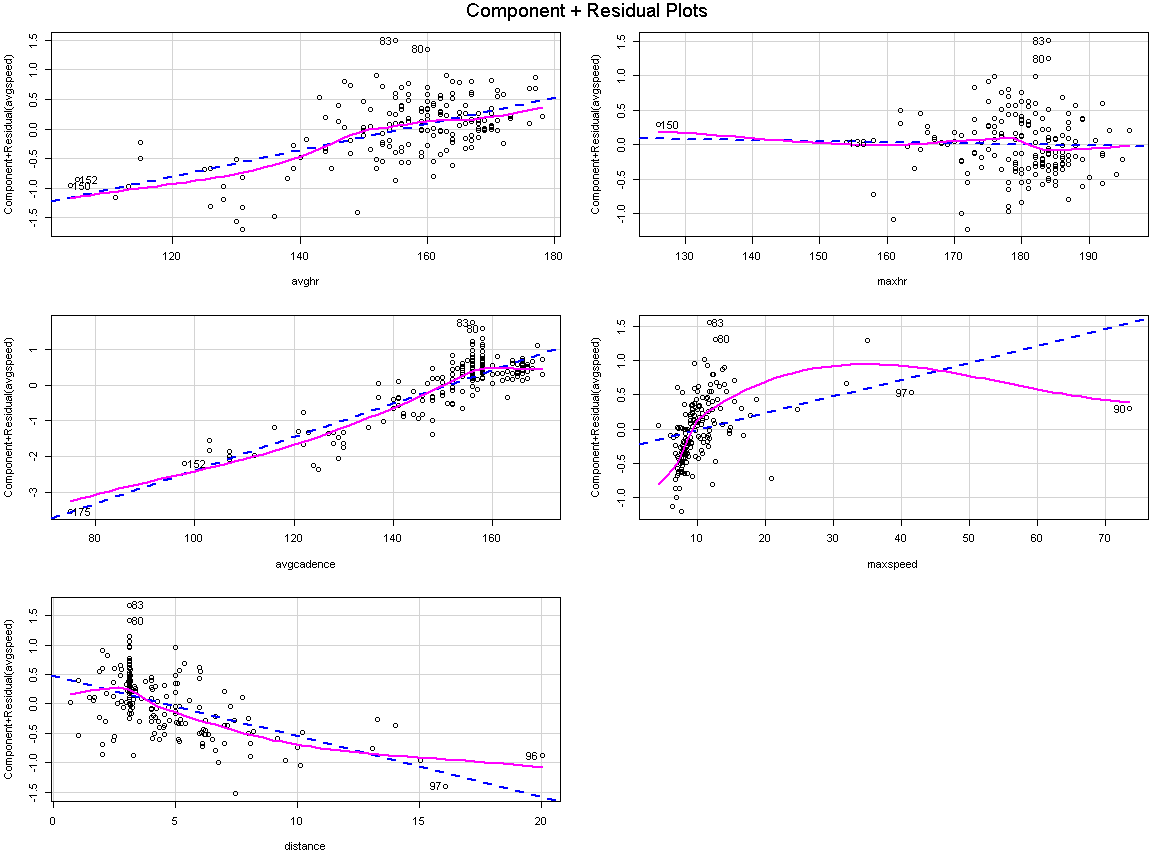




**Observations 90 and 127 are clearly outliers (high residuals) due to maxspeed. I think these were likely GPS measurement errors recorded by the smartwatch. I also believe observation 175(large hat value – high leverage) may have been a walk rather than a “run”, hence the very slow average speed of 2 mph and low average cadence. Observation 96 (outlier – high residual) was a run where Leah was training for a half marathon and did complete a 20 mile run/walk. Observation 83 was just a really good run for me where I achieved a very high average speed – unfortunately based on the other observations it appears to be marked as an outlier (high residual).**



**Observation 90 has a rather exciting DFFITS value of almost -5. Possibly due to the 70 mph top speed. None of the other observations are as extreme, but the two next closest are observation 127 which appears to be influential for a 35 mph top speed and observation 96 because of the extreme distance (20+ miles).**



**Average Heart Rate and Average Cadence both appear to clear upward trend, suggesting that they play a significant role in predicting Average Speed.**

**Max Heart Rate has no increasing or decreasing trend in the points, suggesting it does not belong as a predictor in the model.**

**The points for max speed have an upward trend that doesn’t appear linear. This can suggest that maxspeed plays in important nonlinear role in predicting Average Speed. (Or it could be caused by my impressive 70 mph max speed in observation 90)**

**Distance has a clear downward trend, suggesting that it plays a significant role in predicting Average Speed. I’m guessing the downward part suggests that it negatively affects Average Speed, which makes sense from a logical viewpoint.**