APPM2720 Homework 7

This is a homework to explore least squares fitting.

The Bolder Boulder race results is in BB200.txt. The file contains results on the top 200 finishers from the 2014 Bolder Boulder 10K (6.2 miles) footrace. The columns are: place, names, sex/age division (M or F and age cateogry), place in the division, times for each mile, average mile time and the total time. Read in these data using

```
results <- read.table("BB200.txt",
        header=TRUE, skip=2 )
```

Source the handy function `convertTime` from the HW7.R file.
and convert the MILE1 and TIME times to numeric

- Fit a line by least squares using the mile 1 split (MILE1) to predict the final time (TIME) in the race. Interpret the slope as: "If a runner decreases his/her split by 10 seconds then they tend to decrease their finishing time by ??? seconds. "
- Make a diagonostic plot of the residuals (yaxis) against the predicted values (xaxis) (the components **residuals** and **fitted.values** from the **lm** output.) Do you see any patterns?
- Assuming that the statistical assumptions of the least fit are satisfied find a 95% interval for the estimated slope. Use the standard error in the second column in the coefficients table from the summary of the lm fit. Specifically if **fit** is your **lm** fit object

```
look <- summary(fit)
look$coefficients
```

will give you the table. `look$coefficients[2,2]` ` is the relevant standard error in this case.

- Repeat this analysis where the response is the place in the race (PLACE), not the finishing time. Based on the residuals and any other plots does this model fit better or worse?