# A3 Redux Report

*Pankaj Tripathi, Kartik Mahaley*

*February 11, 2016*

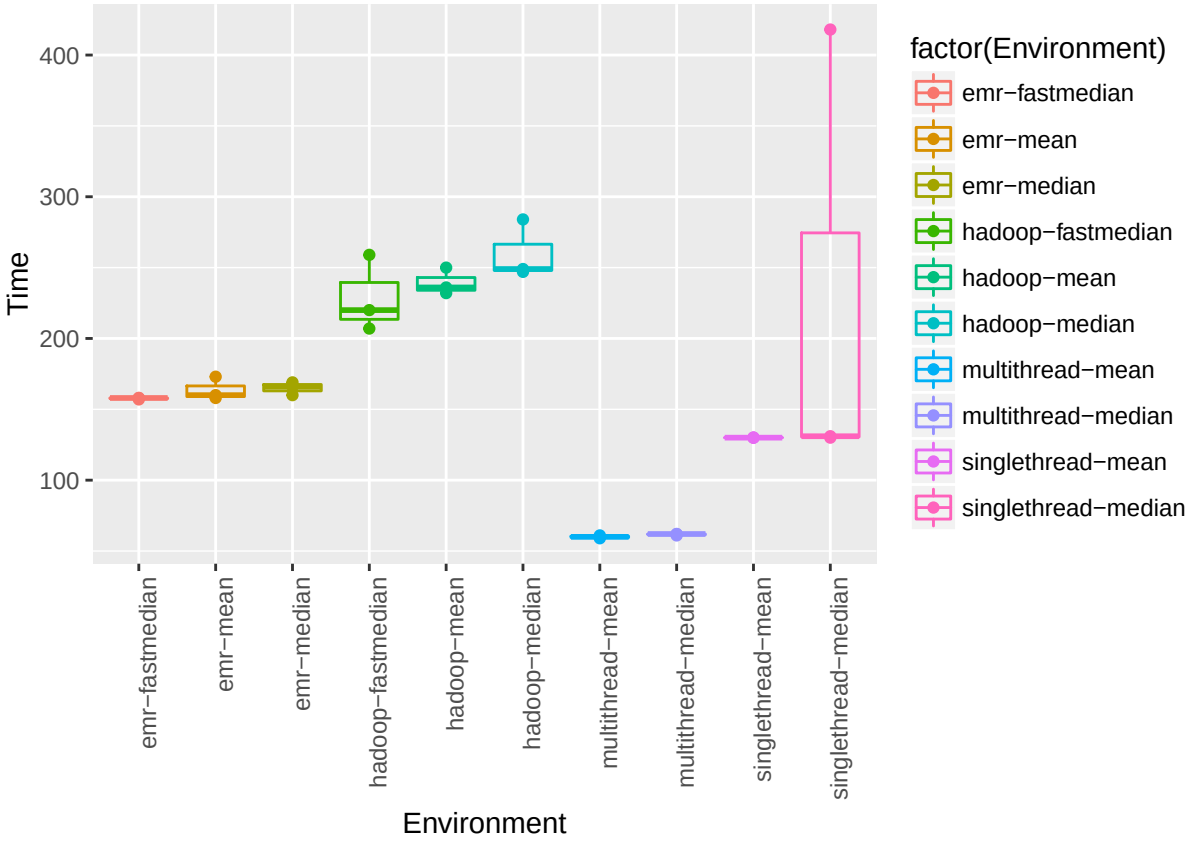
This is a report generated from R and it has the required results of our assignment A3.

For A3, we are resusing our code from A0, A1 and A2 with some changes to suffice the requirement of A3. We have made few modifications in the code given by prof so that it suits our program. For instance we removed input and output property from the perf.txt instead we are taking those parameter from the user while running Make file. There are other so subtle changes.

We are taking mean/median/fastmedian for each airline irrespective of its year.

Here are our analysis: As per our analysis for 653 MB gzipped data on i5-8gb Mac machine and 653MB on i5-8gb Linux machine we found that performance in much better in multi thread mode followed by single thread mode followed by AWS EMR cloud and worst when run with Hadoop in local psuedo distributed mode.calculating mean was faster than calculating median or fast median. We had an outlier in single thread mode while calculating mean for the third time. The outlier can be seen in the plot below.

```
## Loading required package: ggplot2
```




Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.