

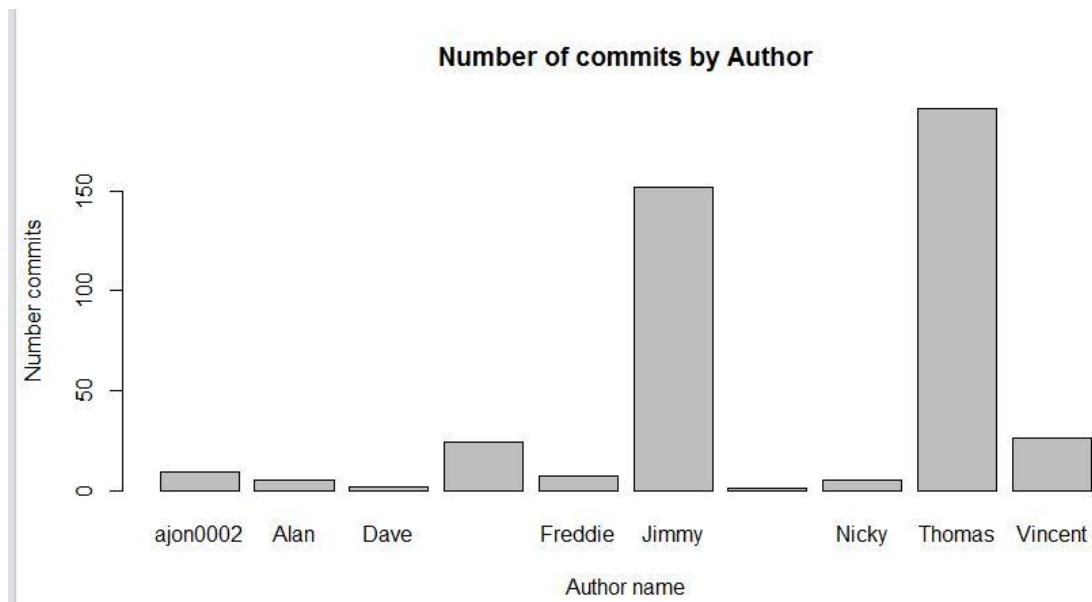
CA4 - Perform analysis on 5000 line dataset

The following describes my four interesting statistics based on analysing the log file

1) Most versus least productive coders

The first statistic relates to who were the most productive coders in terms of the number of commits over the period reviewed. Figure 1 shows a chart generated in R of the number of commits made by each individual. As is evident from the chart, Thomas and Jimmy were by far the most productive with 191 commits and 152 commits respectively. Between the two of them they were responsible for 81% of the total number of commits. All other individuals made far fewer commits apart from Vincent who made 26 commits. Dave and Murari were the least productive with just two commits and a single commit respectively.

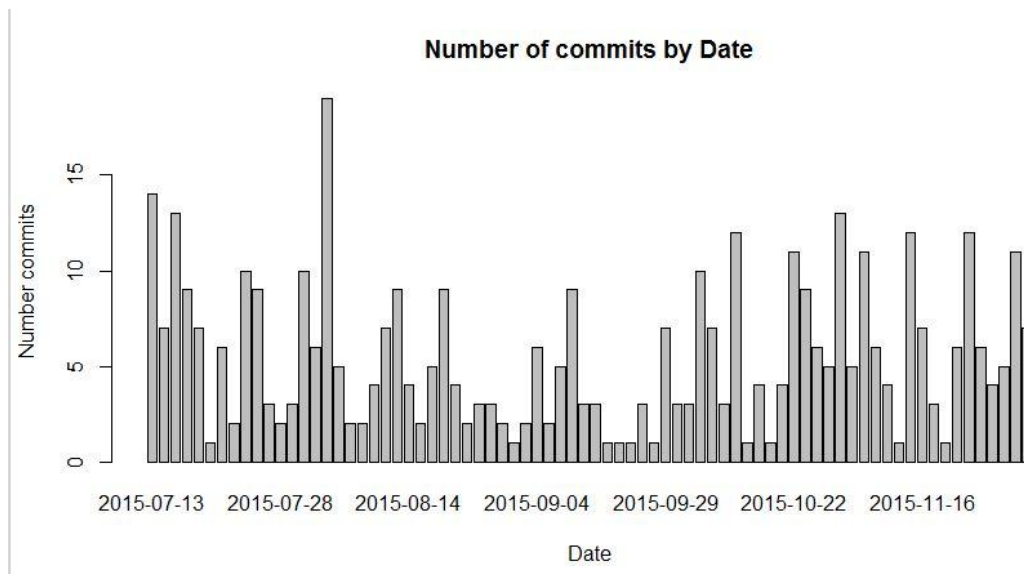
Figure 1 - Number of commits by author



2) Date with most commits

The second statistic relates to the date during which there was most commits. Figure 2 overleaf shows a chart generated in R of the number of commits made on each date between 13th July and 27th November 2015. As can be seen from the chart, the single most productive date in terms of numbers of commits made was the 4th August when 19 commits were made. The least productive date was 20th July on which just a single commit was made. July generally was the most productive month and September was the least productive. This may be due to deadlines for a product update or new release set for July.

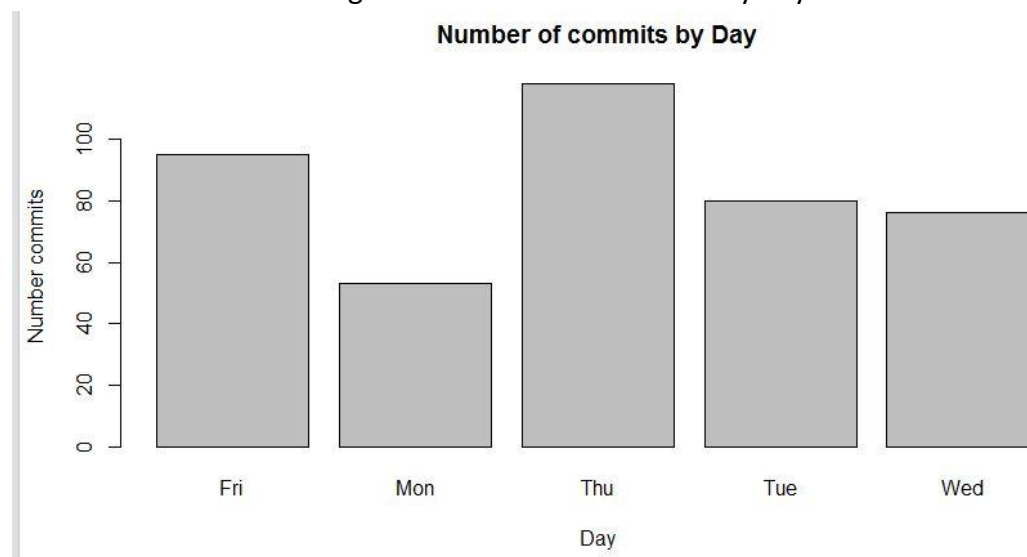
Figure 2 - Number of commits by date



3) Day of the week with most commits

The third statistic relates to the most productive day of the week in terms of commits. Figure 3 shows a chart of the total number of commits made on each day over the period. As can be seen from the chart, Thursday is the most productive day with a total of 118 out of the 422 commits made on Thursdays. Monday is the least productive day with a total of just 53 commits over the period. This can best be explained by the fact that many people will take Mondays off as annual leave coupled with bank holidays falling on Mondays. Tuesday and Wednesday had similar numbers of commits (80 and 76) and Friday was the second most productive day with 95 commits over the period.

Figure 3 - Number of commits by day



4) Number of commits with more than 30 updated files

One final statistic relates to the commits where large numbers of files were updated as part of the commit. In particular there was a focus on those commits where greater than 30 files were updated as part of the commit. There were 12 commits where the number of files updated was greater than 30. The most number of updated files within a single commit was 298 by Thomas on 23 July 2015. Figure 4 shows a chart showing the date versus the number of updated files where the number was greater than 30.