**Simple analysis of revisions for the period July 2015 to end November 2015.**

Introduction: As part of the CA students were asked to clean and process 5,000 objects and discuss some of the more interesting facts arising from the analysis.

Methodology:

First the raw data on the text file was cleaned by using python code. It was tested and saved as csv. File. The statistical package use for analysis was excel.

Strengths:

It only took a few second to submit the python code through the console. The output a single csv file could then be used to study the data and analysis the data.

Weakness:

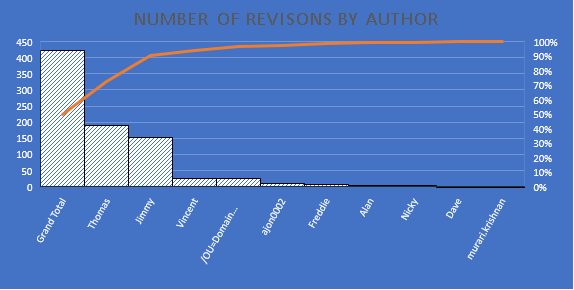
The data was analysis in a vacuum with no background as to why it was generated. Suggested for a deeper analysis would be to include data on the time it took each author to executive the revisions. Was the author the same person who wrote the original code. Why was this period of time looked at and who requested the initial data and why? Was it one product or more than one?

**Analysis:**

Number of revisions:

A total of 422 revisions were made to programme in the period July to November 2015. Of this amount 24 revisions were not attributed to an individual and were preparations for upgrades and not necessarily revision to clean bugs.

Revisions by author:



The greatness number of revisions were made by programmer Thomas. Thomas made 191 revision to programmes. Jimmy was second with 152.

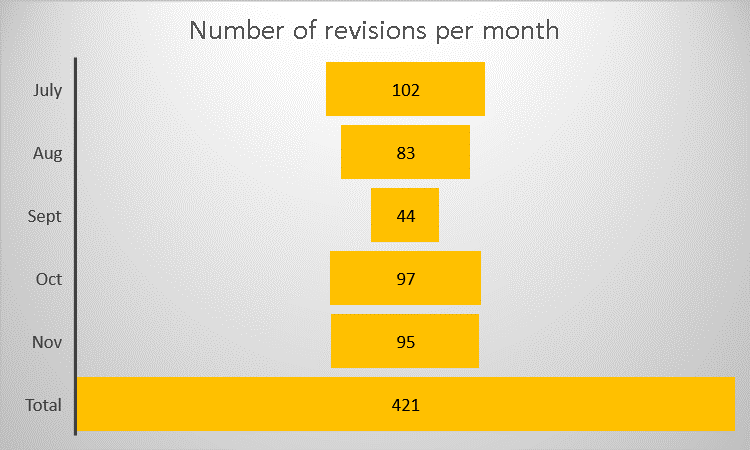
The lowest number of revision was attributed to Murari with 1.

|  |  |  |
| --- | --- | --- |
| Author | Total | % of Total |
|  | 422 | 100% |
| **Thomas** | **191** | **45.26%** |
| **Jimmy** | **152** | **36.02%** |
| **Total** |  | **81.28** |

Thomas and Jimmy account for 81.28 percent of the total of 422 revisions. This fact ensures that average would be skewed. Thomas completed on average 27.28571429 over a five-day working week. However, Jimmy daily average was greater at 30.4. which could mean that Jimmy more efficient with his time another possible reason is that Thomas’s may have had more difficult revisions to make.

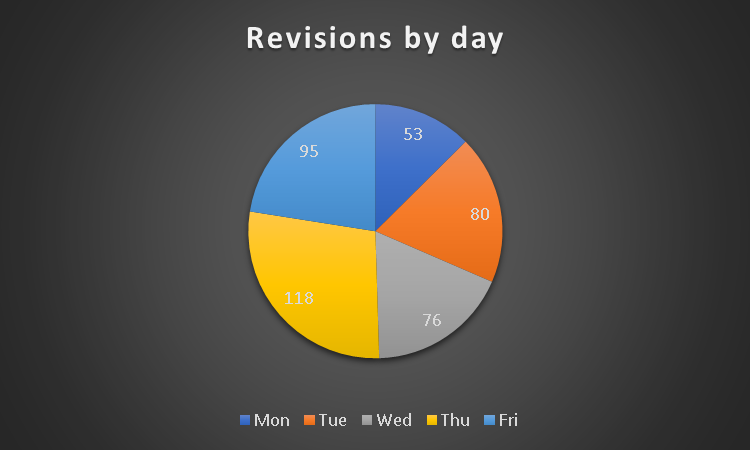
For future analysis, it would be interesting to see how much time the authors spent on each revision and how much time they spent on new products. Additionally, for non-programmers a data directory for the revisions would be a useful accompaniment to deepen the understanding of the analysis.

Revisions by month:



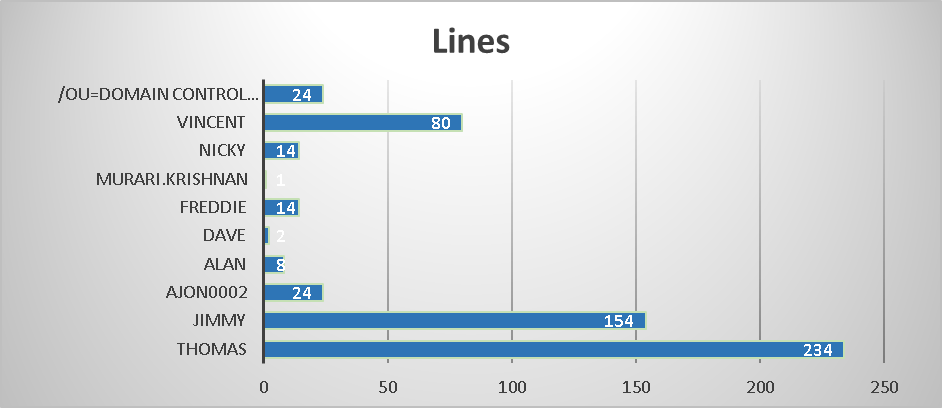
July was the month when most revisions were made and September was the month with the least revisions. This may be linked to the fact that there was more [gradle-release] prepare for next development iterations in this month. Here is would be interesting to see the number of tests made to the new upgrades.

However, October only had one gradle-release] prepare for next development iterations in this month yet had 97 Revisions and a 121 of scripts written.



Thursday was the day when most revisions were made with over 118 of these 46 revisions were accredited to Thomas. His most productive day was Fridays with over 55 revisions.

Number of lines per Author:



Thomas wrote more lines that any other author. In terms of the number of lines Jimmy came a close second. Murari with one line was the author who wrote the least number of lines and the least number of revisions.

Conclusions:

The data generated some interesting finding with tentative links to the number of upgrades and the number of revisions. There is further scope for analysis with some additional information such as how many of the authors were involved in the initial generations of the products. How many tests were made of the products before launch. How many bugs were found? Finally, the amount of time it took to write the revisions would be interesting to see.