Mary Hennessy 10345121

Analysis of Dataset for CA 3:

PROJECT STATISTICS

**1. Examine data by time** looking at the time of day commits were carried out and how busy each day of the week was commit wise.

* The hours in table below are divided so as ‘05’ represents 05:00 to 05:59 etc
* In this first table we see commit rate peaked around 2pm with 22.75% of all commits made between 2-3pm – it might be reasonable to assume that work was carried out throughout the morning and early afternoon to prepare for committing but without other figures it is difficult to make any assertions re time spent on project.
* It is always of interest to me to see that most work was carried out during regular working day with less than 5% of commits outside the 8a.m – 5pm time frame. This suggests that working here is a regular 8-5 job
* From the days of the week it is interesting to see that from an eighth of commits happening on Mondays work builds steadily throughout the week before falling off slightly – but I think not worryingly so, on the Fridays – 50% of all commits happen on Thursday or Friday.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

COMMITS by HOUR of DAY

Hour of Day: 05 :commits: 1 ( 0.24 %)

Hour of Day: 06 :commits: 3 ( 0.71 %)

Hour of Day: 07 :commits: 2 ( 0.47 %)

Hour of Day: 08 :commits: 12 ( 2.84 %)

Hour of Day: 09 :commits: 33 ( 7.82 %)

Hour of Day: 10 :commits: 30 ( 7.11 %)

Hour of Day: 11 :commits: 57 ( 13.51 %)

Hour of Day: 12 :commits: 31 ( 7.35 %)

Hour of Day: 13 :commits: 56 ( 13.27 %)

Hour of Day: 14 :commits: 96 ( 22.75 %)

Hour of Day: 15 :commits: 52 ( 12.32 %)

Hour of Day: 16 :commits: 39 ( 9.24 %)

Hour of Day: 17 :commits: 4 ( 0.95 %)

Hour of Day: 18 :commits: 2 ( 0.47 %)

Hour of Day: 19 :commits: 1 ( 0.24 %)

Hour of Day: 20 :commits: 3 ( 0.71 %)

**busiest time of day: hour commencing 14:00 with 22.75% of all commits**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

COMMITS by DAY of WEEK

Thu 118 ( 27.96 %)

Fri 95 ( 22.51 %)

Tue 80 ( 18.96 %)

Wed 76 ( 18.01 %)

Mon 53 ( 12.56 %)

**busiest day of the week was: Thu : in total 118 commits happened on Thursday's**

**2. Examine data by person:**

* Over 80% of the work on this project was carried out by two people as seen in table below. This is not unusual and others who made fewer commits could have been managers or specialists amending details or even interns
* I’m not sure about the 24 commits by /OU=Domain Control Validated/CN=svn.synchronoss.net and think these may warrant further investigation i.e. maybe if it’s a generic code it is better to have a name attached for reference (or could be an automated checking system but unlikely as the comments made were ‘human’ sounding (e.g. 'Moved app recommendations icons into client' )and similar to other comments made by the other people)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

COMMITS per PERSON

(45.26)% viacheslav.vdovenko 191

(36.02)% robert.hanaway 152

(6.16)% vnai0001 26

(5.69)% /OU=Domain Control Validated/CN=svn.synchronoss.net 24

(2.13)% ajon0002 9

(1.66)% fled0001 7

(1.18)% nsab0001 5

(1.18)% apat0004 5

(0.47)% dred0001 2

(0.24)% murari.krishnan 1

Person with most commits on this project is viacheslav.vdovenko with a total of 191 commits

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**3. Examine overview of commits on the data**

* In total 422 commits over 76 days which was an average of 5.55 commits per day for the duration of the project.
* Initially I wasn’t going to look at the per month figures but think it gives a nice overview of the project lifecycle. Here nearly a quarter of the commits were in beginning i.e. the latter half of July and except for a deep trough in Sept commits were very evenly spaced in October and November.
* I looked at the breakdown of types of jobs done per commit and the breakdown from the table shows most of the work (almost three quarters) was additions and modifications with 25% deletions. While it is good to see fewer deletions and without knowing more re expected rates on a project like this 25% of previous commits having to be deleted might warrant investigation whereas if the deletes were of an old system being replaced that would be quite acceptable.
* Finally I was interested to see that this was 5000+ lines about the project rather than the actual project itself so very easy to see from this how data/metadata can be generated and needs to be managed correctly

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Total number of commits: 422.0

Total no. of project days = 76.0

Average number of commits per day = 5.55

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

JOBS by proportion/type carried out in COMMITs

Total additions(A) for this project = 1056 (35.0 %)

Total deletions(D) for this project = 767 (25.0 %)

Total modifications(M) for this project = 1186 (39.0 %)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Total number of lines in this document: 5255

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_