

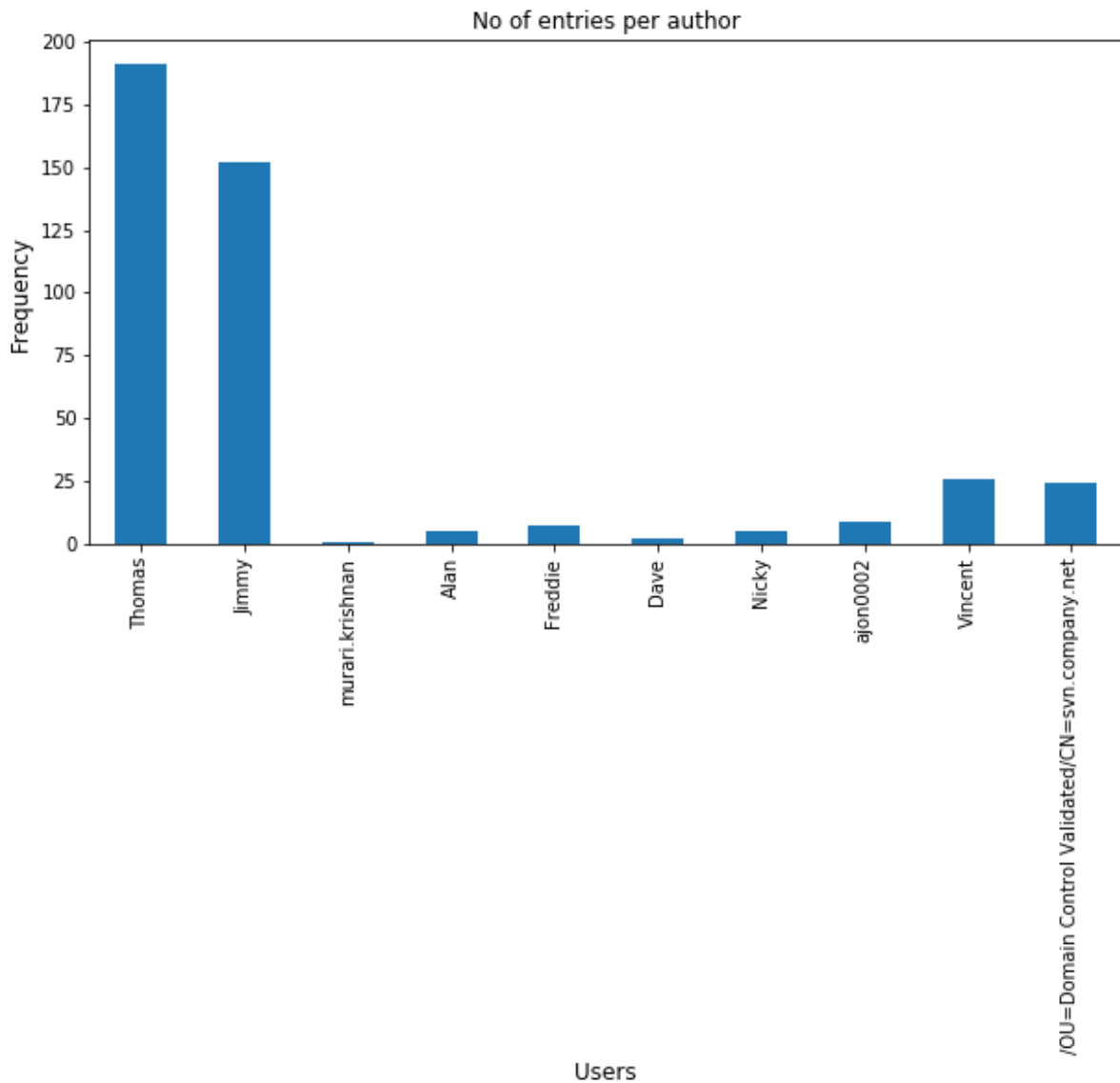
CA 4

Programming for Big Data

Andrew Doran-Sherlock

10361540

## Analysis of Commits Dataset



It is clearly evident that Thomas and Jimmy are very frequent users and are submitting updates more than the rest of the users combined in the timeframe examined.

Time	Commits
05:00	1
06:00	3
07:00	2
08:00	12
09:00	33
10:00	30
11:00	57
12:00	31
13:00	56
14:00	96
15:00	52
16:00	39
17:00	4
18:00	2
19:00	1
20:00	3

The times of the commits were rounded to the nearest hour. This shows that 14:00 was by far the most popular time for the commits and that the vast majority were performed between 11:00 and 16:00 inclusive. If these users were meant to be operating in the traditional 09:00 – 17:00 working times, it would mean that for approximately half the working day, they are not sending any commits.