**Programming for Big data**

**CA 4**

**Student Name:**

**Jassem Alhaj Tamer**

**Student ID:**

**10168910**

**DATA ANALYSIS**

According to data set given for analysis, more than one tool have been used to analyse and infer statistical piece of information for this dataset with supporting evidence of interestingness.

When applying Python, R studio and Tableau, we could in fear:

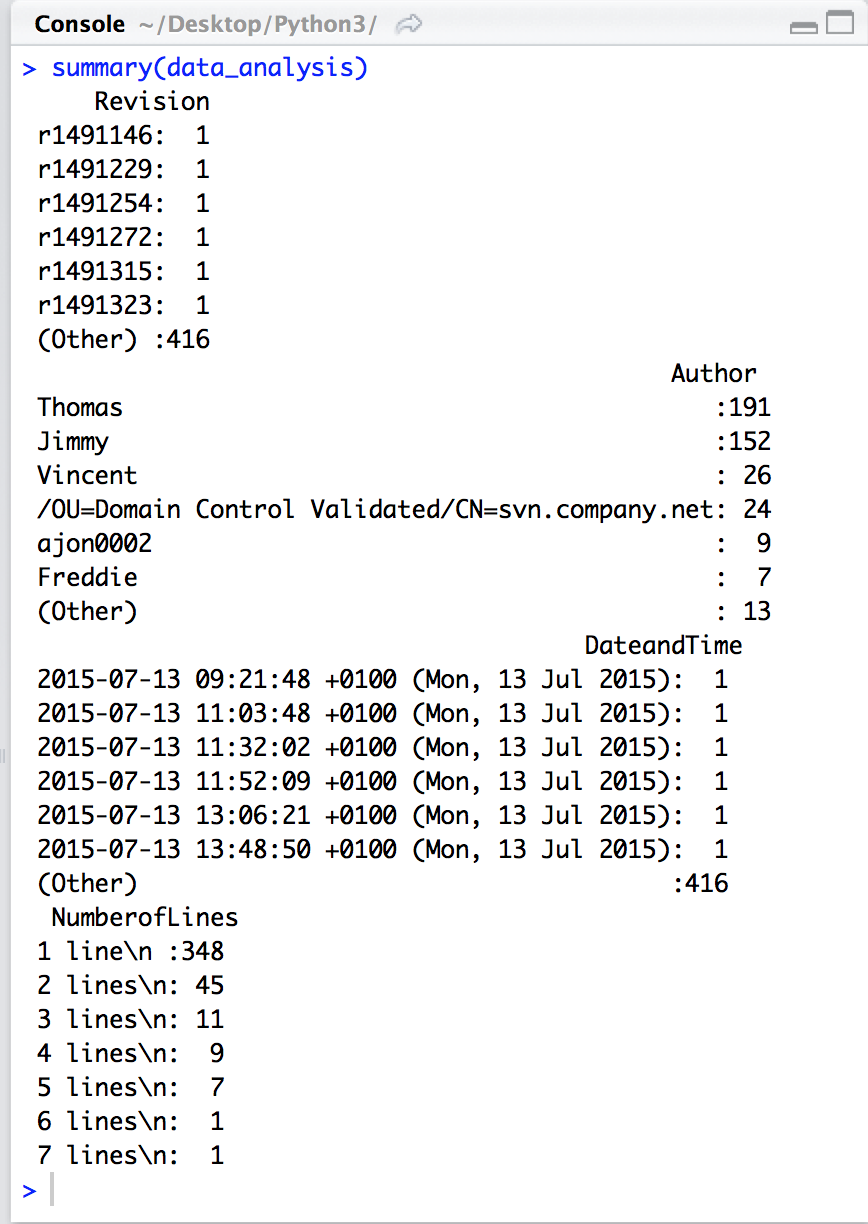
1. the file has 5255 Lines and 422 commits.
2. The first and last author who commit was “Thomas” . The first commit was at evening and last was at morning.
3. The total number of people who commit the file is 10 :

['Thomas', 'Vincent', '/OU=Domain Control Validated/CN=svn.company.net', 'Jimmy', 'Freddie', 'Dave', 'ajon0002', 'murari.krishnan', 'Nicky', 'Alan']

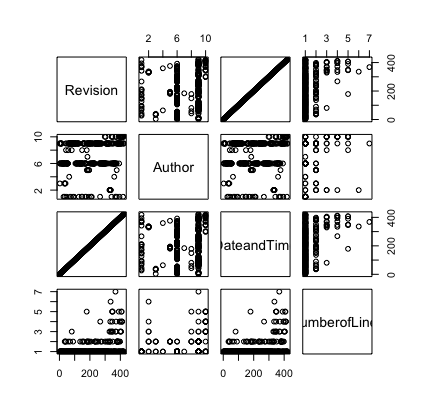
1. “Thomas” has the biggest portion of commits with 190 commits.
2. “Jimmy” has 152 commits
3. “Vincent” has 26 commits
4. The Author “Thomas” has the highest number of lines with 348 lines modified in the file.
5. Merge a revision with its committer and changes related to this revision:

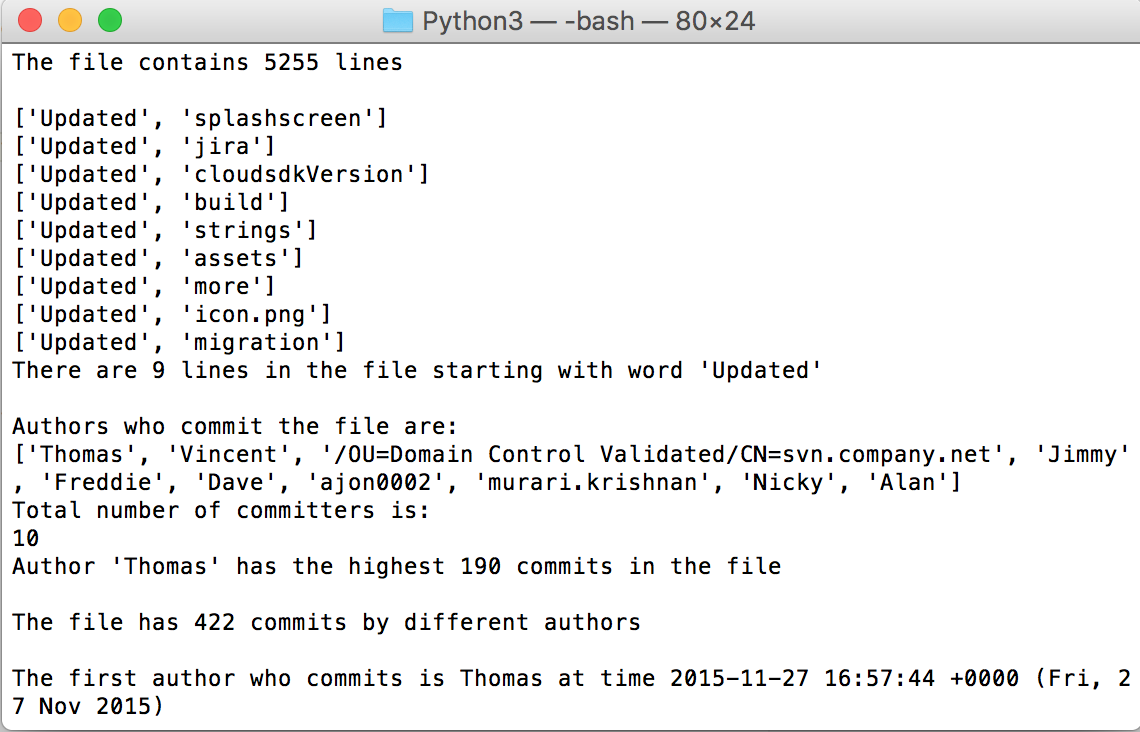
svn merge -**r1491322**:1491323 by **Jimmy** with the comment **Added application identifier to auth call and the changes Changed paths**:,M /cloud/personal/client-international/android/branches/android-15.2-solutions/clients/client-att/build-config/prod.properties,M /cloud/personal/client-international/android/branches/android-15.2-solutions/libs/ui/src/com/biscay/client/android/ui/atp/AtpAuthActivity.java

The following figure illustrate analysis in RStudio.



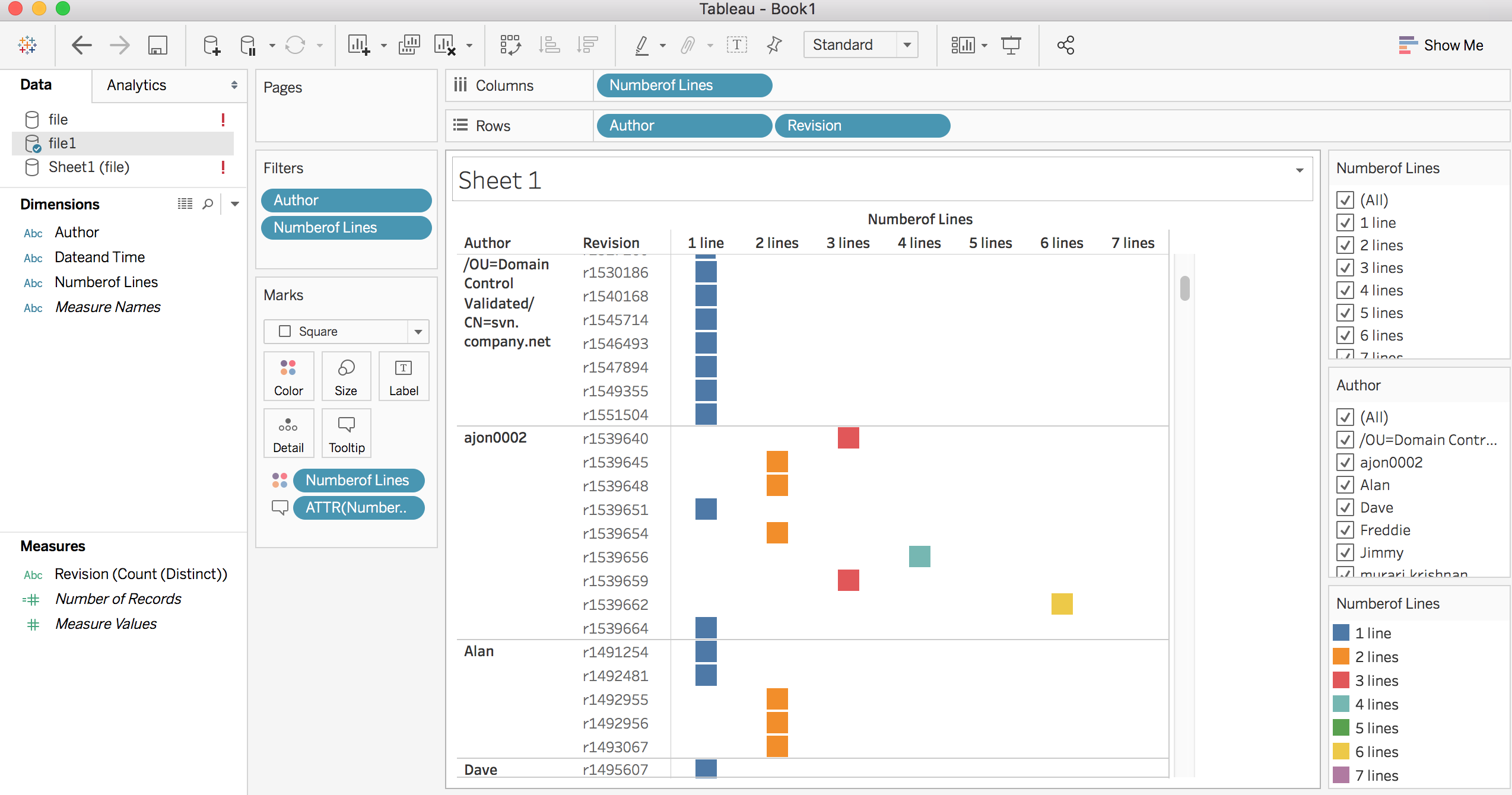
As each commit has a revision No., R data summary shows the files has 422 revisions which gives another indication that the file has 422 commits.





We could also see that there are 9 lines starts with the word “Updated”.

Bu using Tableau regarding this analysis, the following figure has been captured:



Author “Alan” did 5 commits in the file. Revision number r1491254 shows the author modified 1 line whereas “Ajoon0002” could modify 3 lines by his revision r1539640

More importantly, commits have been done in year 2015 between July and November. Most of those commits done at evening. For example: “ajoon0002” did all commits in November 02-11-2015 at evening. “Alan” did commits in July from 13-07-2015 to 16-07-2015 at evening.

