Out of 4001 comments 69% were requests for changes

Key Observations for comments which were merged :

* The sample data is mostly from 2016
* The merges are maximum in the year 2016
* These merges are mostly done in 4th quarter and least in 3rd quarter
* The least merges were June and highest in December
* Maximum merges were done in 18th of every month and least on 5th of every month
* Tuesday has maximum merges with Friday being the least

Key observations about comments :

* Most of the comments are from 2016 and 2017
* 1st and 2nd quarter has highest comments with 3rd quarter having significantly low comments
* The comments are least in July and highest in Feb, March , November, December
* The comments are highest on Sundays and lest on Fridays and Saturdays
* 16th of every Month seems to have highest comments

Key observations about comments which are requests for changes:

* Highest’s requests were in the 1st and 4th quarter
* March month has highest requests for change and July has least
* 16th of every month has highest requests changes
* Sundays Monday are highest requests days and Friday, Saturday are least

Out of 1220 requests for changes only 284 were merged

Key observations for these records

* They are highest in 4th quarter and least in 3rd quarter
* They are highest in December
* The merges were done mostly on 18th and 21st of every month
* The merge were mostly done on Tuesday

The average resolution time is 3.7days . Hence, most of the requests to change were done on 16th of every month and they were resolved by 21st

The data is very small to confidently claim the above observations

We notice that most comments were on 16th of every month this can be because code review could be done every fortnight which needs to be verified with the teams

By using NLP-Natural Language Processing

-the comments can be classified as requests for changes , requests for test , issues and general comments

We can try to cluster this by applying word2vec and KNN