

CA2 - Process an Unstructured Data File - 30%

Module Title: Programming Essentials

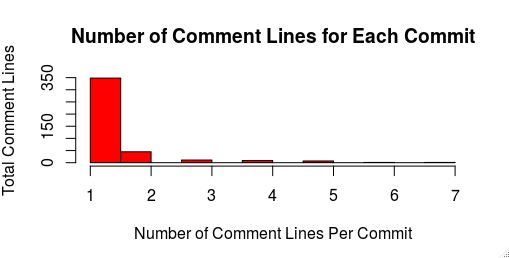
Module Code: B8IT102

Module Leader: Darren Redmond

Student Name: Karina Jonina

Student Code: 10543032

**Interesting Fact #1: The Number of Comment Lines for Each Commit**



*Graph*: The day of the week that where the code was most amended was Thursday, followed by Friday.

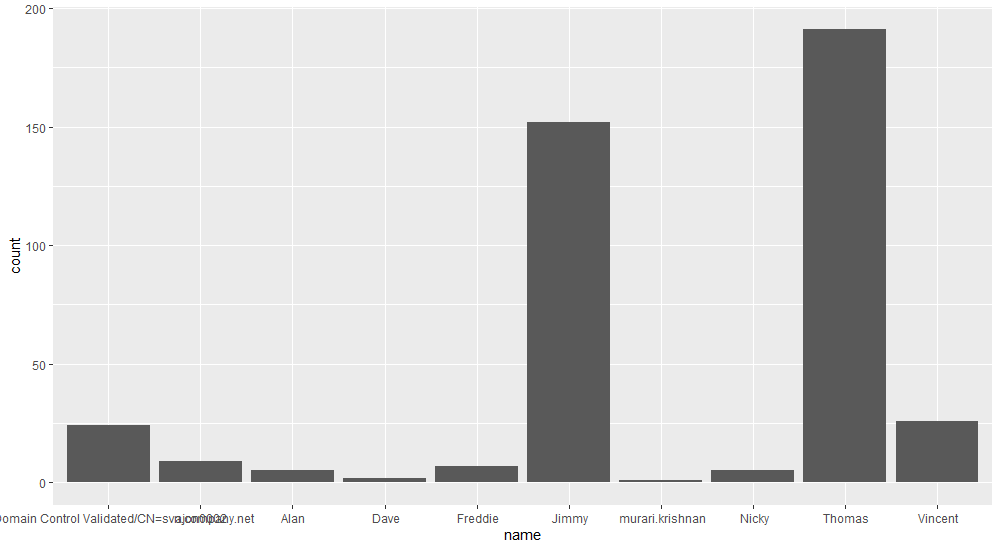
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Lines | *1 Line* | *2 Lines* | *3 Lines* | *4 Lines* | *5 Lines* | *6 Lines* | *7 Lines* |
| Count | 348 | 45 | 11 | 9 | 7 | 1 | 1 |

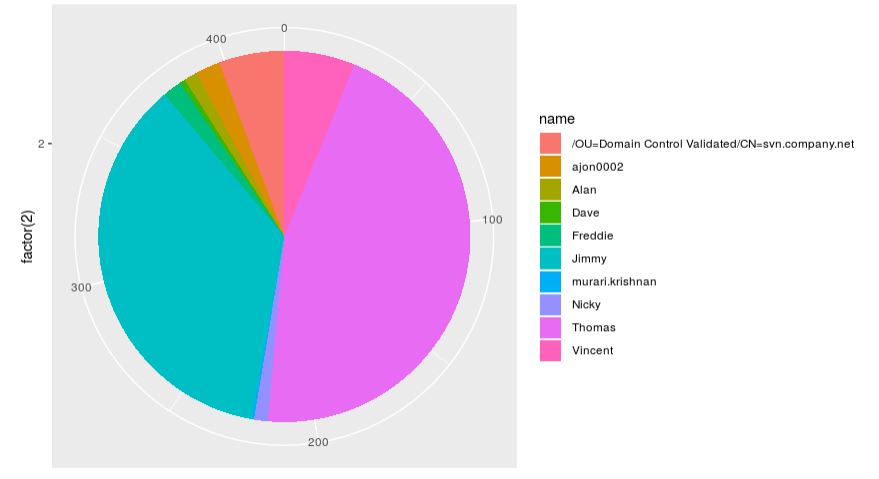
*Table* : The number of lines that was added to the program.

*Why is this interesting?*

This is interesting to see because this would suggest that developers make small amendments to the code that require only one or two lines of comment. Otherwise, this suggests that developers are under-commenting on code and if one of the developers were to leave the project, it would be very difficult for other developers to pick up where the developers had left of.

**Interesting Fact #2: The Most Dedicated Team-Player**

The Number of Commits Made by Each Developer*Graph:* It is clearly visible that Thomas and Jimmy contribute to the code.



*Pie chart:* Examining the work distribution with the project.

|  |  |
| --- | --- |
| *Name of Developer* | *Number of Amendments* |
| Thomas | 191 |
| Jimmy | 152 |
| Vincent | 26 |
| Domain Control | 24 |
| ajon0002 | 9 |
| Freddie | 7 |
| Nicky | 5 |
| Alan | 5 |
| Dave | 2 |
| Murari Krishnan | 1 |

*Table*: The number of times each developer committed code.

|  |  |  |  |
| --- | --- | --- | --- |
|  | *N* | *M* | *SD* |
| Domain Control | 24 | 1 | 0 |
| ajon0002 | 9 | 2.68 | 1.58 |
| Alan | 5 | 1.6 | 0.55 |
| Dave | 2 | 1 | 0 |
| Freddie | 7 | 2 | 0 |
| Jimmy | 152 | 1.01 | 0.114 |
| Murari | 1 | 1 | n/a |
| Nicky | 5 | 2.80 | 1.30 |
| Thomas | 191 | 1.23 | 0.67 |
| Vincent | 26 | 3.08 | 1.49 |

*Table:* Reporting average lines that were amended by each developer.

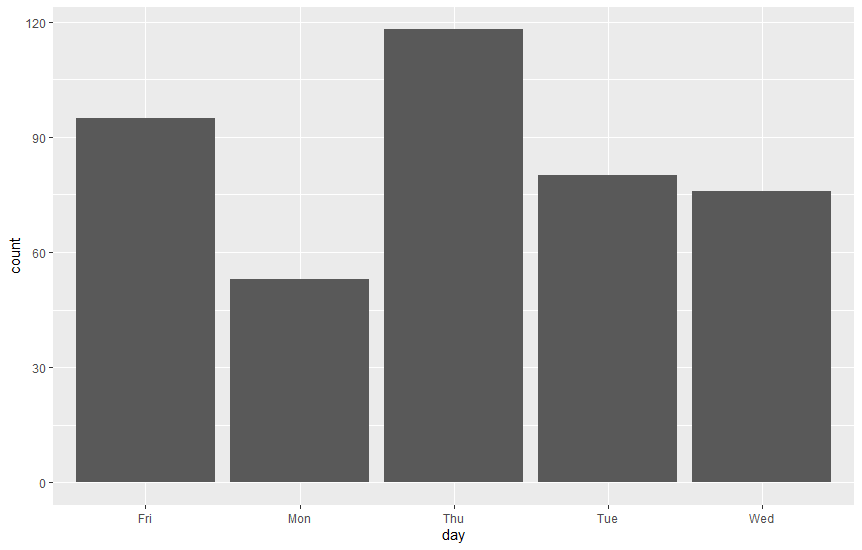
A One-Way ANOVA was run to examine whether there is a significant difference between the number of lines by each developer. There is a significant relationship between the two variables,  *F* (9, 412) = 34.76, *p >* .05. TurkeyHSD showed significant between Thomas and the rest of the developers, between Jimmy and the rest of the developers. There was no significant difference between Jimmy and Thomas (*p* = 0.07)

*Why is it interesting?*

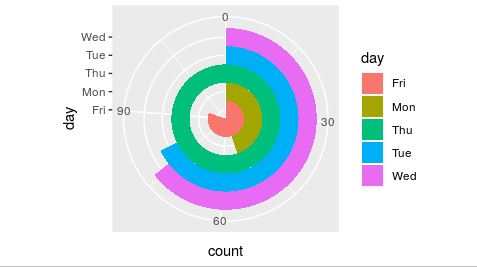
This is important information to show that Thomas and Jimmy are the main developers on the project. They comment most on the code and do most of the amendments on the code. This is also important if there is dispute on who gets the most credit for completing a project, raises or promotions within the company.

**Interesting Fact 3: The Most Productive Day of the Week**

Number of Commits Per Weekday



*Graph*: The day of the week that where the code was most amended was Thursday, followed by Friday.



*Graph* : The day of the week that where the code was most amended was Thursday, followed by Friday.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Day | *Thursday* | *Friday* | *Tuesday* | *Wednesday* | *Monday* |
| Count | 118 | 95 | 80 | 76 | 53 |

*Table* : The number of times code was committed to the program.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *1 Line* | *2 Lines* | *3 Lines* | *4 Lines* | *5 Lines* | *6 Lines* | *7 Lines* |
| Friday | 81 | 14 | 0 | 0 | 0 | 0 | 0 |
| Monday | 41 | 6 | 3 | 1 | 1 | 1 | 0 |
| Thursday | 95 | 13 | 2 | 4 | 4 | 0 | 0 |
| Tuesday | 70 | 3 | 2 | 3 | 1 | 0 | 1 |
| Wednesday | 61 | 9 | 4 | 1 | 1 | 0 | 0 |

*Table:* The number of comment lines that were submitted on each day of the week.

*Relationship between weekday and number of lines?*

A One-Way ANOVA was run to examine whether there is a significant difference between the number of lines on each day of the week. There is a no significant relationship between the two variables,  *F* (4, 417) = 1.476, *p <* .05. Therefore, Turkey HSD was not run.

|  |  |  |  |
| --- | --- | --- | --- |
|  | *N* | *M* | *SD* |
| Fri | 95 | 1.15 | 0.35 |
| Mon | 53 | 1.45 | 1.05 |
| Thu | 118 | 1.38 | 0.95 |
| Tue | 80 | 1.32 | 1.02 |
| Wed | 76 | 1.32 | 0.75 |

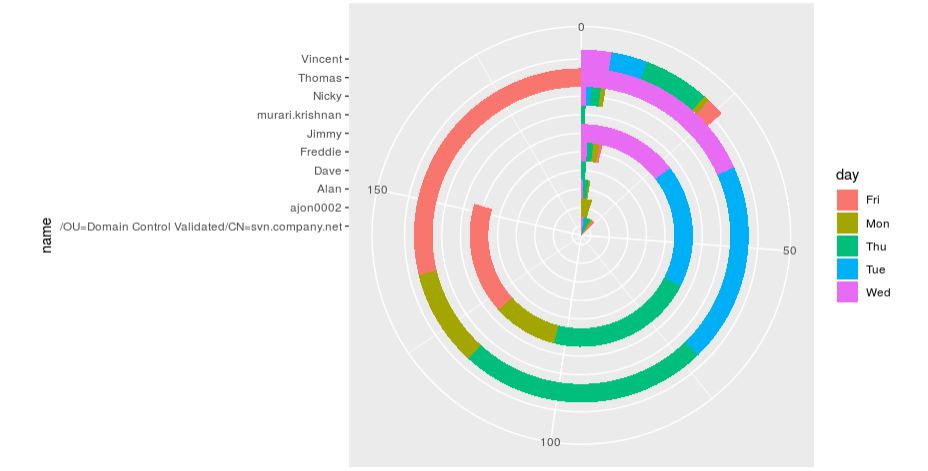
*Table:* Reporting average comment lines that were submitted on each day of the week.

*Why is it interesting?*

Thursday seemed to be the most productive day of the week, followed by Friday. However, Monday is seems to be the least productive day of the day. Perhaps, the developers are working harder at the weekend to meet weekly deadlines. However, there is no significant difference between comment lines, therefore the developers distribute their workload throughout the week.

This is also important to examine whether the developers have a good work/life balance and do not work on weekdays.

**Interesting Fact #4: The Number of Lines Amended by Each Developer**



*Graph* : Examining who works the most on what day of the week.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Friday | Monday | Tuesday | Wednesday | Thursday | Total |
| Thomas | 55 | 18 | 46 | 37 | 35 | 191 |
| Jimmy | 31 | 18 | 41 | 34 | 28 | 152 |
| Vicent | 3 | 1 | 11 | 6 | 5 | 26 |
| Domain Control | 5 | 3 | 10 | 2 | 4 | 24 |
| ajon0002 | 0 | 9 | 0 | 0 | 0 | 9 |
| Freddie | 1 | 2 | 2 | 0 | 2 | 7 |
| Nicky | 0 | 1 | 2 | 1 | 1 | 5 |
| Alan | 0 | 1 | 3 | 0 | 1 | 5 |
| Dave | 0 | 0 | 2 | 0 | 0 | 2 |
| murari.krishnan | 0 | 0 | 1 | 0 | 0 | 1 |

*Table 3*: The number of commits conducted by each developer on each day of the week.

*Relationship between weekday and developer?*

Chi-square was run to examine the relationship between these two ordinal categories (day, names). There is a significant relationship between the two variables,  *X*2 (36, *N* = 422) = 95.04, *p*> .05.

*Why is this important?*

It is clear from the table when each of the developers contributed to the comment and on what day of the week. It is important to know whether the developers were working on the project throughout the week or whether they cramp the corrections before the weekend.