1. Find the project’s github url from this list:

<https://docs.google.com/spreadsheets/d/1tYh-ilK5x9UFTJJep2VR2iuxDFPYsl-D9glJkTpTXsQ/edit#gid=547704908>

Based on the github url, derive the github api url. The github api url usually follow this pattern:

e.g. <https://api.github.com/repos/apache/carbondata>

A json file will can be retrieved from the url.

Fill in the github api url back to the datasheet.

1. Run R script “data\GitAndroidAnalysis\repo\_analysis\firstcommit-date.R” to get the first commit of the project.

Run R script “data\GitAndroidAnalysis\repo\_analysis\ get\_project\_duration.R” to get the duration of the project.

Run R script “data\GitAndroidAnalysis\repo\_analysis\email-collection.R” to get a list of emails for the developers.

Run R script “data\GitAndroidAnalysis\repo\_analysis\ active\_contributors\_every\_30.R” to get a list of emails for the developers.

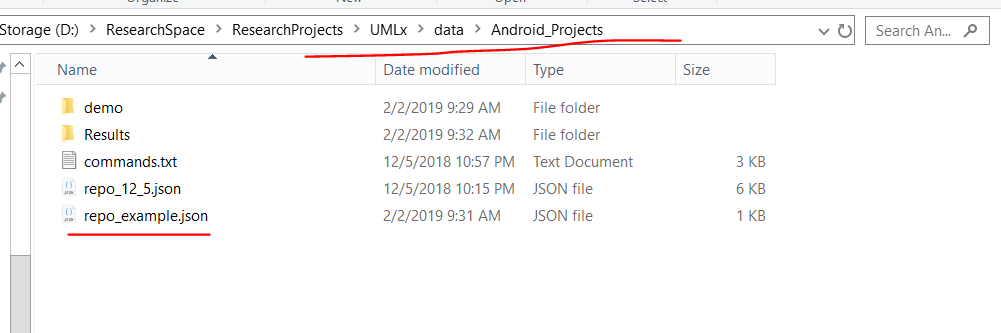
Fill in the information derived back to the datasheet:

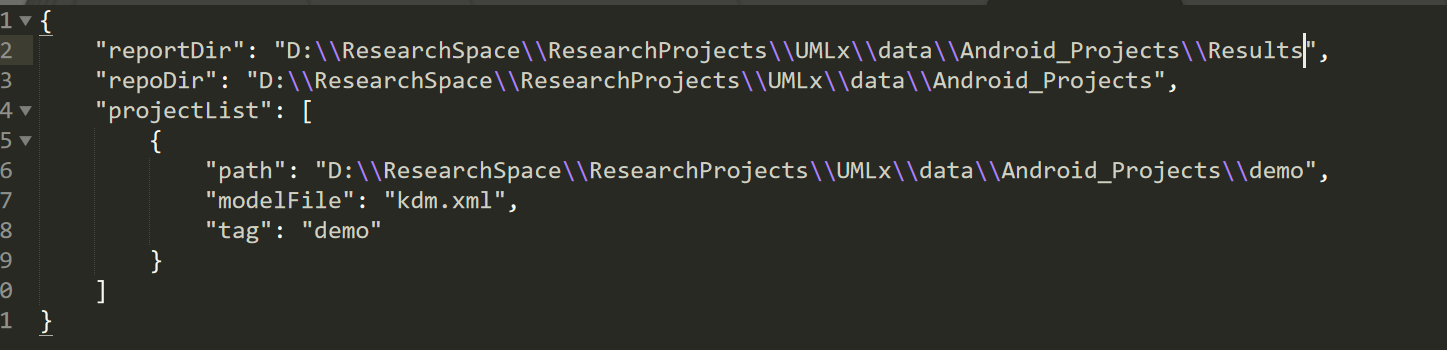
<https://docs.google.com/spreadsheets/d/1tYh-ilK5x9UFTJJep2VR2iuxDFPYsl-D9glJkTpTXsQ/edit#gid=547704908>

1. Run the command below to derive the data for lines of source code.

Command to source code of code.

1. Create a config file to reference the project source.





* You can add more projects to this configuration file to further do the analysis.

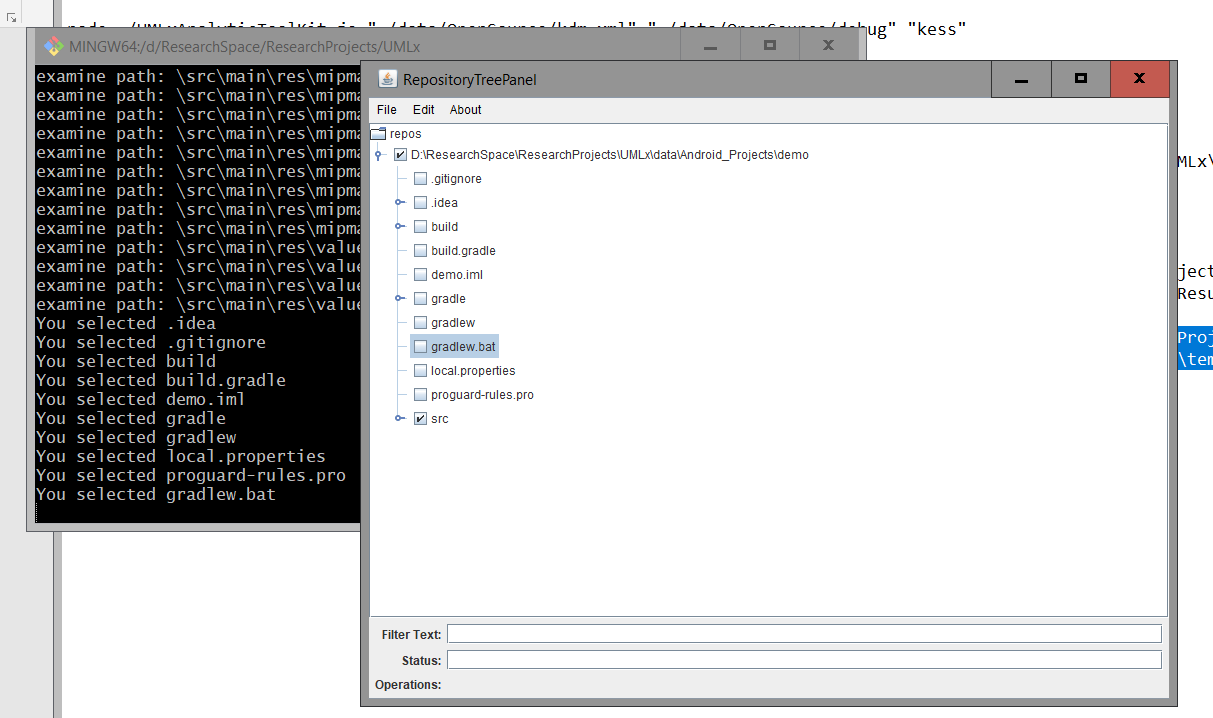
1. Scan the repo

node --max\_old\_space\_size=10240 ".\utils\OpenSourceProjectAnalysis.js" --scan-repo "D:\\ResearchSpace\\ResearchProjects\\UMLx\\data\\Android\_Projects\\repo\_example.json"

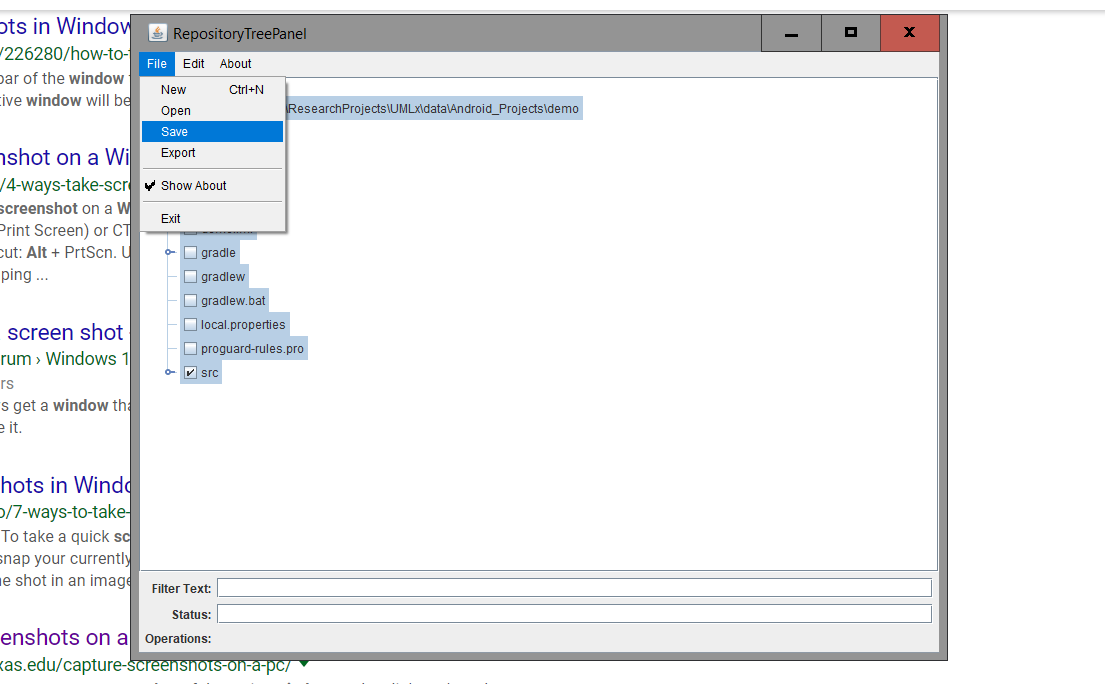
[skip this step]

1. Select the source code files using the following command:

node --max\_old\_space\_size=10240 ".\utils\OpenSourceProjectAnalysis.js" --select-files "D:\\ResearchSpace\\ResearchProjects\\UMLx\\data\\Android\_Projects\\repo\_example.json"

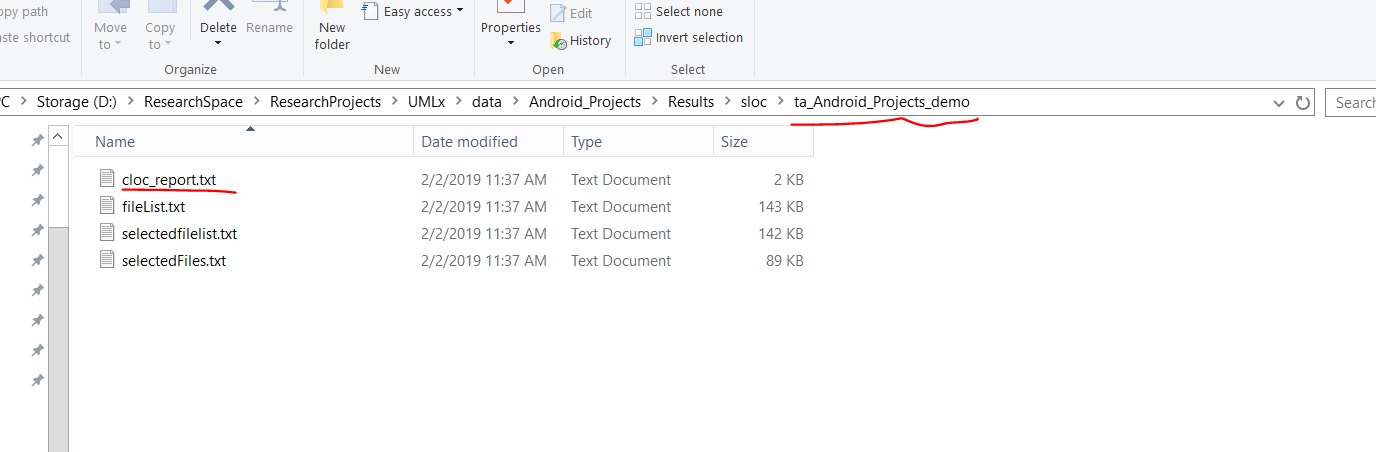


Save the selected files



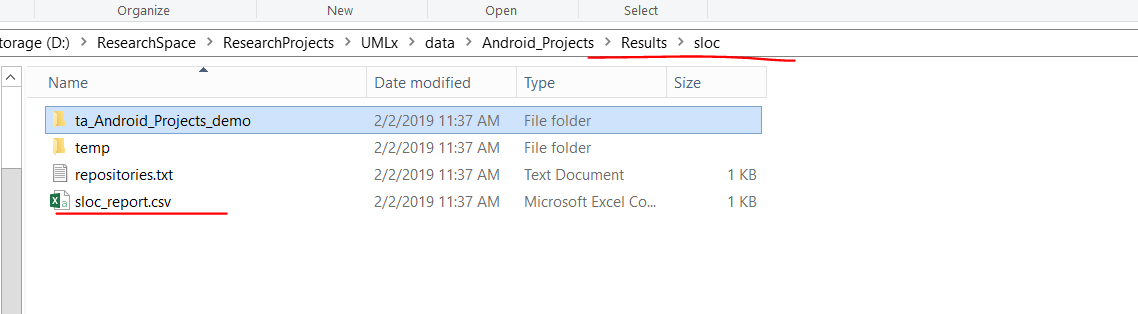
1. Analyse the lines of source code:

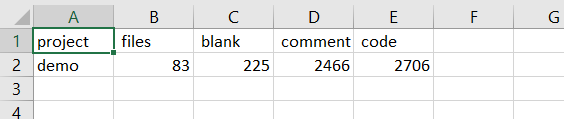
node --max\_old\_space\_size=10240 ".\utils\OpenSourceProjectAnalysis.js" --analyse-sloc "D:\\ResearchSpace\\ResearchProjects\\UMLx\\data\\Android\_Projects\\repo\_example.json"



1. Generate the report

node --max\_old\_space\_size=10240 ".\utils\OpenSourceProjectAnalysis.js" --generate-sloc-report "D:\\ResearchSpace\\ResearchProjects\\UMLx\\data\\Android\_Projects\\repo\_example.json"





1. Run the R script to understand the correlations and the stepwise model

UMLx\data\FeatureSelection\gzh\_report.Rmd