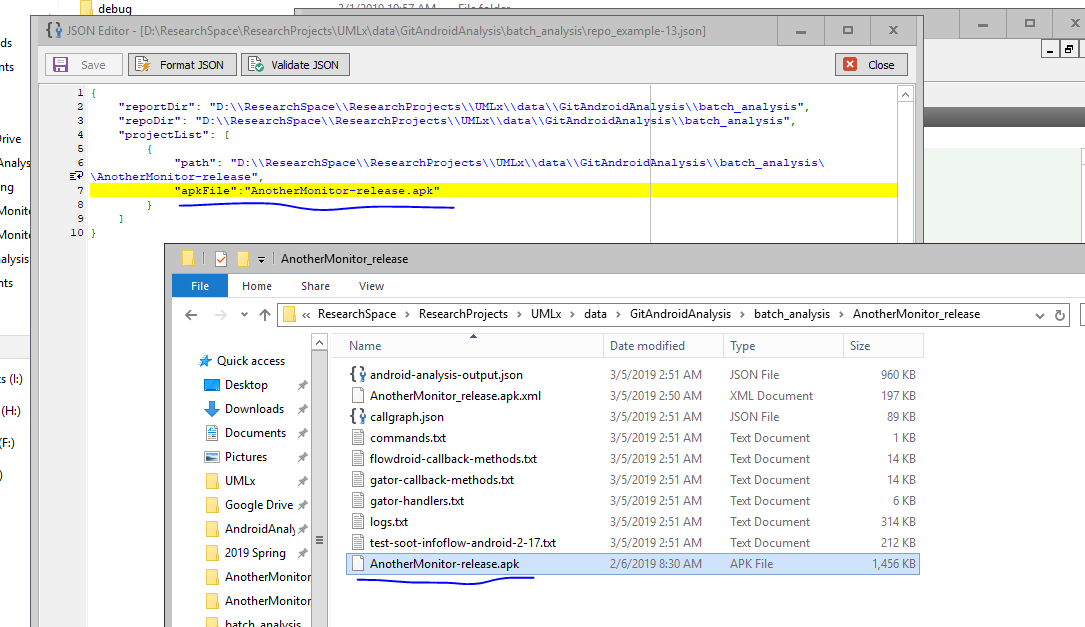
1. ***Android Data Collection:***
2. ***Access the google drive android apps repo to get the list of android applications:***

***https://docs.google.com/spreadsheets/d/1gozRUgml6nCy\_evCyQWzLrabhuKbRgMBJLD3Itg-NOQ/edit#gid=547704908***

***find the apks from the folder:***

[***https://drive.google.com/drive/u/1/folders/1Dl9wFfMH-nhlMe4aAmW6HMP3023rMhid***](https://drive.google.com/drive/u/1/folders/1Dl9wFfMH-nhlMe4aAmW6HMP3023rMhid)

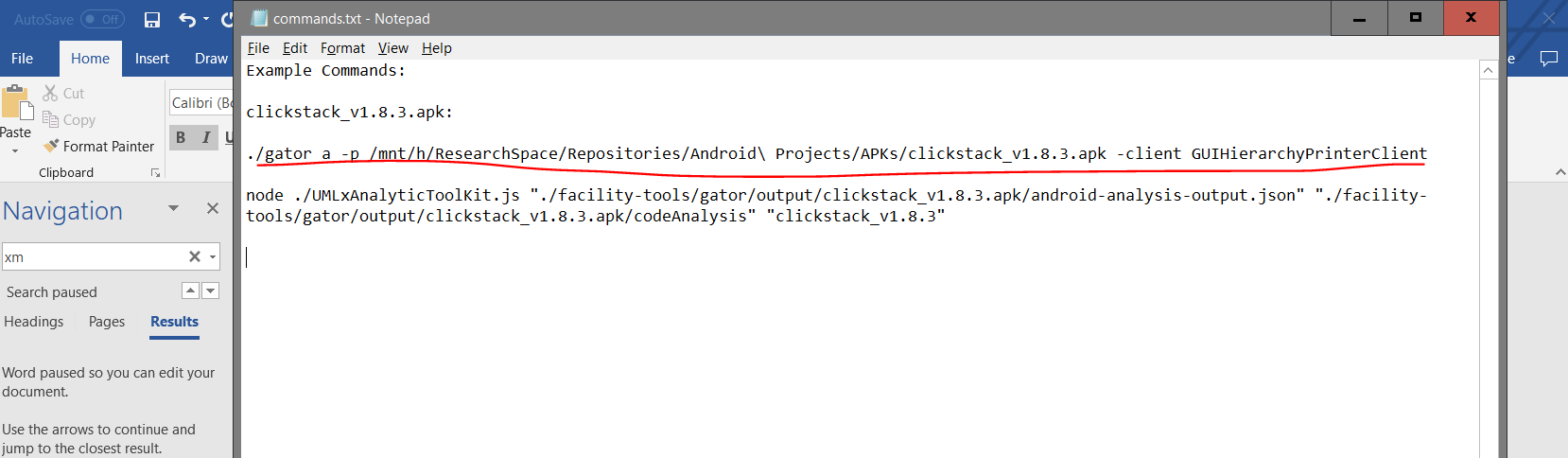
1. Crate a configuration file to point to the apk file:



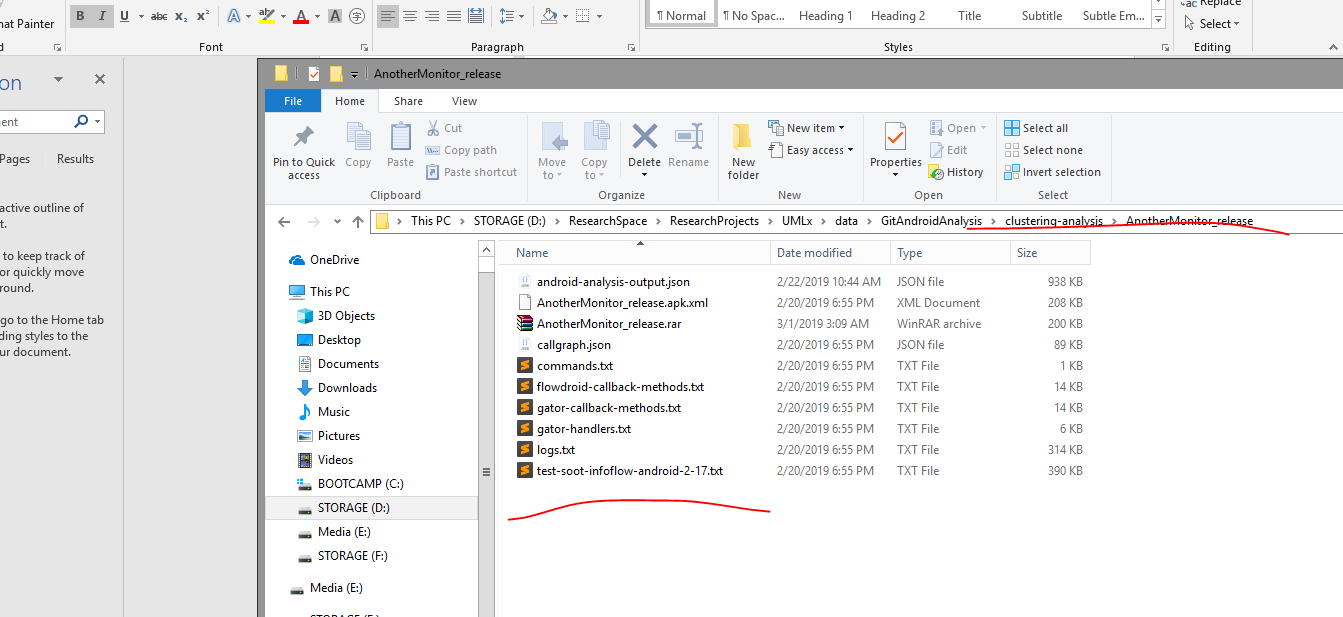
1. ***Run the command to drive the callgraph and output data***

node --max\_old\_space\_size=10240 ".\utils\AndroidProjectAnalysis.js" --analyse-android-apks "D:\\ResearchSpace\\ResearchProjects\\UMLx\\data\\GitAndroidAnalysis\\batch\_analysis\\repo\_example-13.json”

***Replace the path in red with the path to the actual apk and run the android analysis.***



***The results can be found at the folder:***



***Update the gator generated files into this folder:***

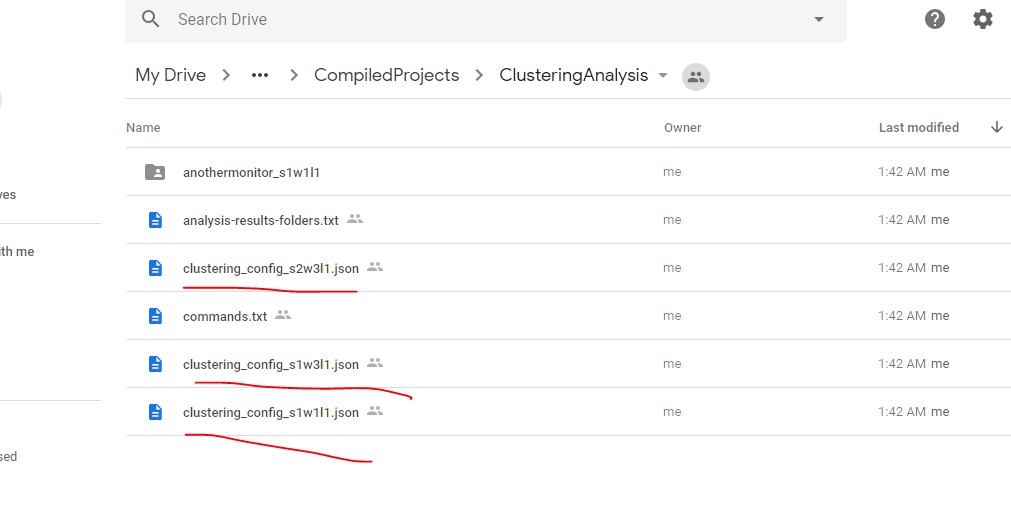
[***https://drive.google.com/drive/u/1/folders/1Y1-eWs4lL92z4RmCZkHWSPEnvpsNM8N2***](https://drive.google.com/drive/u/1/folders/1Y1-eWs4lL92z4RmCZkHWSPEnvpsNM8N2)

***and update the datasheet:***

[***https://docs.google.com/spreadsheets/d/1gozRUgml6nCy\_evCyQWzLrabhuKbRgMBJLD3Itg-NOQ/edit#gid=547704908***](https://docs.google.com/spreadsheets/d/1gozRUgml6nCy_evCyQWzLrabhuKbRgMBJLD3Itg-NOQ/edit#gid=547704908)

1. ***Run the clustering analysis.***

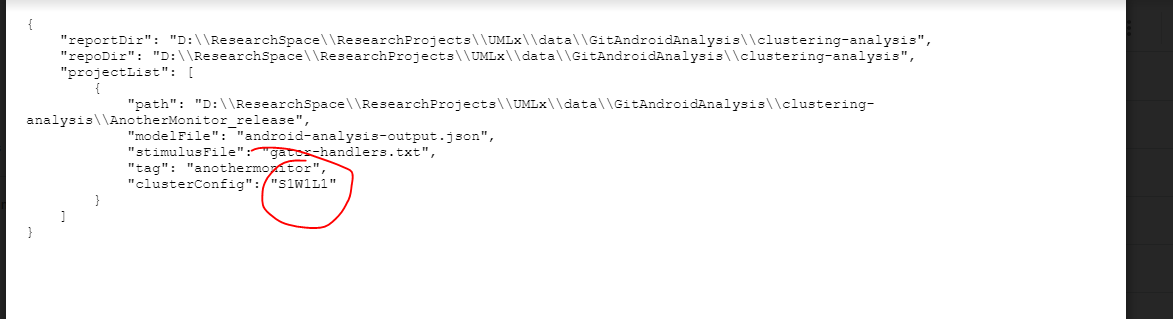
***For each of the set of gator analysis result files, create three configuration files to point to the generated files. Those json files can be created based on the json file used in the first command – only the “clusteringCofig” field needs to be updated.***



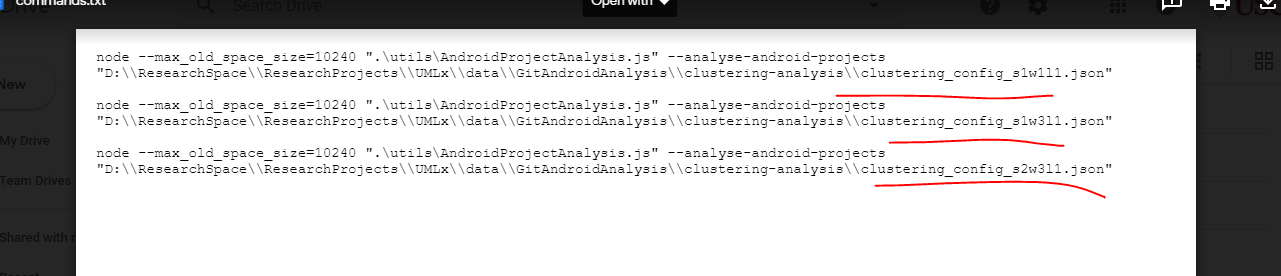
***Some of examples of the configuration files are here:***

[***https://drive.google.com/drive/u/1/folders/1cgxBTZwCC1FpkkE\_QTh0XeMzp2FQ1R-Y***](https://drive.google.com/drive/u/1/folders/1cgxBTZwCC1FpkkE_QTh0XeMzp2FQ1R-Y)

***for example, for clustering\_config\_s1w1l1.json, the following field has to be consistent with the file name:***

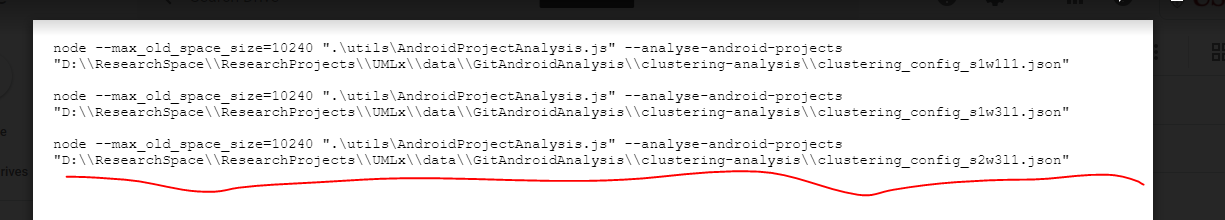


***For each of the three configurations files, run the corresponding command.***

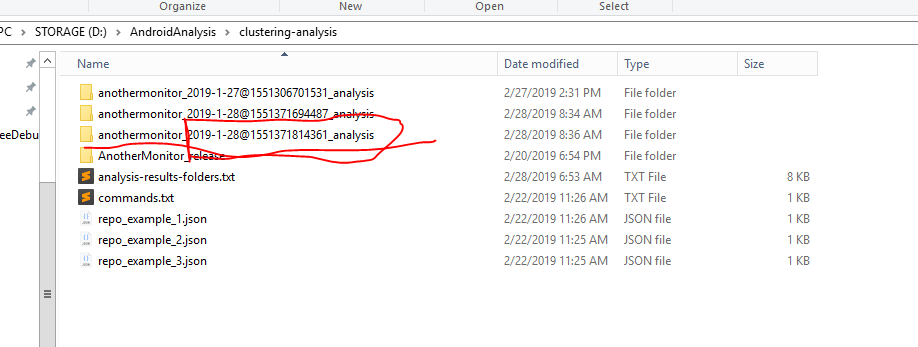


* ***The path should be replaced to your local environment.***

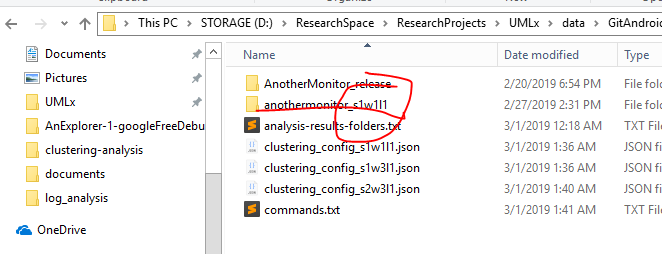
***The example commands are provided at this link:***



***Three folders are generated for each of the configuration file. An example of the generated folder is provided here:***



***Update the part of name of the folder to the same name of the config file.***



***Update the generated analysis files into this google drive folder:***

[***https://drive.google.com/drive/u/1/folders/1cgxBTZwCC1FpkkE\_QTh0XeMzp2FQ1R-Y***](https://drive.google.com/drive/u/1/folders/1cgxBTZwCC1FpkkE_QTh0XeMzp2FQ1R-Y)

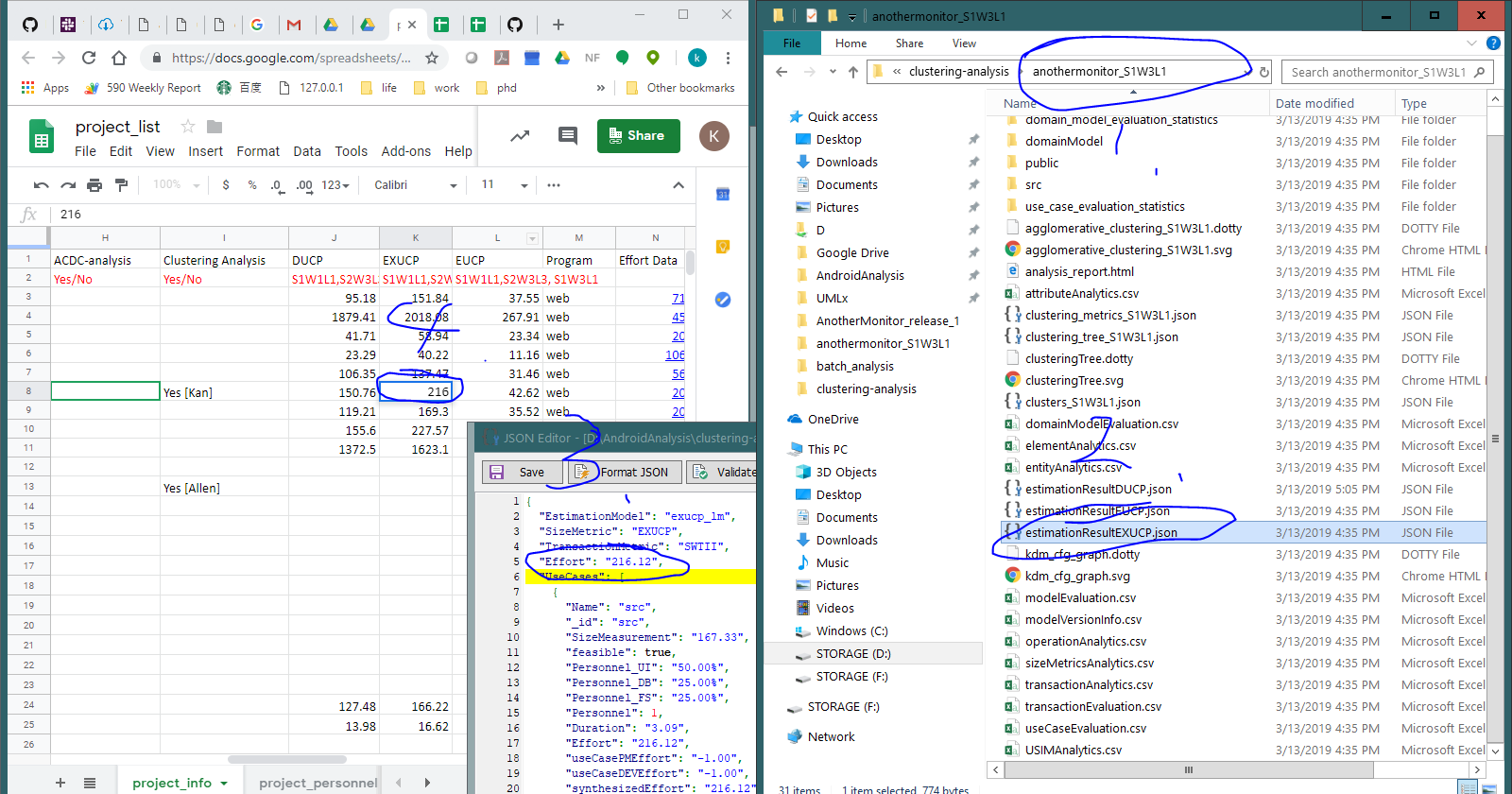
***and update the data sheets:***

[***https://docs.google.com/spreadsheets/d/1gozRUgml6nCy\_evCyQWzLrabhuKbRgMBJLD3Itg-NOQ/edit***](https://docs.google.com/spreadsheets/d/1gozRUgml6nCy_evCyQWzLrabhuKbRgMBJLD3Itg-NOQ/edit)

***and :***

[***https://docs.google.com/spreadsheets/d/17sTqNy4dq91VEoVFAE065I24IlNV6qfBH1mWRysSUK8/edit#gid=0***](https://docs.google.com/spreadsheets/d/17sTqNy4dq91VEoVFAE065I24IlNV6qfBH1mWRysSUK8/edit#gid=0)

1. ***Update the estimates with the analysis results.***



***For each configuration file, in the results folder, find the json files “estimationResultEUCP.json”, “estimationResultEXUCP.json”, “estimationResultDUCP.json”, and find the effort estimate as shown in the circle “3” in the screenshot. Update the number into the datasheet: https://docs.google.com/spreadsheets/d/1gozRUgml6nCy\_evCyQWzLrabhuKbRgMBJLD3Itg-NOQ/edit#gid=547704908***