

https://github.com/SnellJ2/cs472project/tree/mine_repository/repo_mining

[Jonas-Snell_CollectFiles.py](#) - Uses github api to collect commit data, which is parsed to get only the source files that are also programming files.

```
MatthewRollings,2021-05-24T07:35:11Z,rootbeerlib/src/main/java/com/scottiyab/rootbeer/RootBeerNative.java
ScottAlexander-Bown,2021-05-08T14:18:37Z,rootbeerlib/src/main/java/com/scottiyab/rootbeer/RootBeerNative.java
```

Data Interpretation:

Scatter plot showing the number of months since start (Y-axis, 0 to 80) versus the file number (X-axis, 0 to 16). The plot displays data points for 15 different individuals, each represented by a unique color and shape. The data points are scattered across the plot, with some individuals showing a clear upward trend in months since start as the file number increases, while others remain relatively flat or show a downward trend.

Legend (Individuals):

- MatthewRollings
- ScottAlexander-Bown
- DanielKutik
- altvnn
- AndyBarber
- MohammedEzzat
- FriederBluemle
- Fi5t
- mat
- IvanVlasov
- matthew
- SlimNamouchi
- AliiWaseem
- LeonKarabchevsky
- BIOCATCH\Leon.Karabchevsky
- vyas
- leocadiotine

On the other hand **Matthew Rollings** committed around **20 changes**.

Though it's a bit cluttered due to the amount of contributors, it's possible to match color to name and get a rough idea who had a high and low amount of commits. Of course this can't be taken as the full story on who is productive to the project, but it is a good place to start.