

# File\_Split v1.0

The **DE Quick Start tutorial** provides an introduction to basic DE functionality and navigation.

Please work through the documentation and add your comments on the bottom of this page, or email comments to [support@cyverse.org](mailto:support@cyverse.org).

## ***Rationale and background:***

Splitting a large file into a number of small lines based on line number is essential for many applications or analysis. This is especially useful if you have a long list of id's and then you want to use them parallelly using [HT Analysis Path List](#) file.

## **Pre-Requisites:**

A CyVerse account (Register for a CyVerse account at <https://user.cyverse.org>).

An up-to-date Java-enabled web browser. (Firefox recommended. If you wish to work with your own large datasets and upload them using iCommands, Chrome is not suitable due to its issues in utilizing 64-bit Java.)

## **Mandatory arguments**

1. Inputs
  - a. Input File: Path to the input file that contains the list of id's
  - b. Number of lines per file: How many lines that you want to split the input file into. The number of output files corresponds to number of lines that you want to split into
  - c. File\_prefix: Prefix of the file that gets created (Default is "x")
2. Outputs
  - a. Output folder: Name of the output folder (Default is "Output")

## **Test run**

The test data for this app is located at `/iplant/home/shared/iplantcollaborative/example_data/OSG-RMTA/sra_id_list.txt`

The 'sra\_id\_list.txt' file contains a list of 15 SRA id's, one SRA ID per line

```
SRR3403881
SRR5227388
SRR5227389
SRR5227390
SRR5227391
SRR5227392
SRR5227393
SRR5227394
SRR5227395
SRR5227396
SRR5227397
SRR5227398
SRR5227399
SRR5227400
SRR5227401
```

1. Inputs
  - a. Input File: `/iplant/home/shared/iplantcollaborative/example_data/OSG-RMTA/sra_id_list.txt`
  - b. Number of lines per file: 5
  - c. File\_prefix: `osg_rmta_`
2. Outputs
  - a. Output folder: `OSG-RMTA-File-Split`

After a successful run, you'll get 3 files each containing 5 SRA Id's with the prefix `osg_rmta_` in `OSG-RMTA-File-Split` folder