|  |  |  |
| --- | --- | --- |
| **variable** | **description** | **type** |
| paper\_id | Unique identifier of the paper |  |
| indegrees\_of\_cited\_papers\_sum | Sum of indegrees of papers cited by the paper in focus |  |
| indegrees\_of\_cited\_papers\_avg | Average of indegrees of papers cited by the paper in focus |  |
| recency\_of\_cited\_papers\_avg | Average recency of papers cited by the paper in focus (days?) |  |
| max\_time\_difference\_bw\_cited\_papers | Maximum time difference between dates of publication of the cited papers |  |
| Outdegree | Outdegree of the paper in focus |  |
| outdegrees\_of\_cited\_papers\_sum | Sum of outdegrees of papers cited by the paper in focus |  |
| outdegrees\_of\_cited\_papers\_avg | Average of outdegrees of papers cited by the paper in focus |  |
| Submitter | The email id and Name of the submitter |  |
| submission\_date | The date of submission of the paper and revision dates (if the paper is revised) |  |
| Title | Title of the paper |  |
| Authors | Authors of the paper |  |
| Comments | General comments regarding the paper like number of pages, format, number of figures etc. |  |
| report\_no | Metadata |  |
| journal\_ref | Metadata |  |
| Abstract | A paragraph to summarise the paper |  |
| submitter\_email | Cleaned email id of submitter (if available) |  |
| submitter\_name | Cleaned Name of submitter (if available) |  |
| Submitter\_details | Cleaned unique identifier for each submitter (Mostly email, but if email is NA, then we take the name) |  |
| is\_revised | Whether the paper was revised or not |  |
| times\_revised | If revised, the number of times a paper was revised |  |
| first\_submission\_datetime | The first submission date with time |  |
| first\_submission\_date | The first submission date |  |
| num\_of\_pages | Number of pages of the paper (derived from comments) |  |
| Format | Format of the paper (derived from the comments) |  |
| journal\_counts | Showing the number of times a particular journal occurs in the data |  |
| first\_365\_days | Metadata |  |
| label | Metadata |  |
| label\_name | Metadata |  |
| submitter\_counts | Number of papers submitted by the submitter | Int |
| submitter\_active | Submitter activity marker. 1 if submitter has submitted 5 or more papers, 0 if 4 or less | Int |
| journal\_popularity | Journal popularity marker. 1 if there are 13 or more papers in the database published in this journal, 0 if 12 or less | Int |
| format\_latex | Marker showing whether the paper was submitted in latex format. 1 if yes, 0 if no | int |
| format\_revtex | Marker showing whether the paper was submitted in revtex format. 1 if yes, 0 if no | Int |
| format\_harvmac | Marker showing whether the paper was submitted in harvmac format. 1 if yes, 0 if no | Int |
| format\_plaintex | Marker showing whether the paper was submitted in plaintex format. 1 if yes, 0 if no | Int |
| datedelta | Number of days passed between the paper being added to graph and the date of first publication. Note: multiple outliers exist. | int |
| Citations\_till\_date | For each submitter, the citations s/he has received till the date s/he publishes a new paper. (based on date added graph) |  |