

Programming Languages

Arnab Panja

2023-05-06

Tidy Tuesday - Programming Languages

This paper analyzes the sample data collected and published in Tidy Tuesday Week 12 dataset.

1. Load Dataset

Let us first load the dataset.

```
languages <- readr::read_csv(file = 'https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/programming_languages/programming_languages.csv')
```

2. Check Dataset Summary

Let us get a glimpse of the dataset

```
glimpse(languages)
```

```
## Rows: 4,303
## Columns: 49
## $ pldb_id      <chr> "java", "javascript", "c", "python", ~
## $ title        <chr> "Java", "JavaScript", "C", "Python", ~
## $ description  <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ type         <chr> "pl", "pl", "pl", "pl", "queryLangua~
## $ appeared     <dbl> 1995, 1995, 1972, 1991, 1974, 1985, ~
## $ creators     <chr> "James Gosling", "Brendan Eich", "De~
## $ website      <chr> "https://oracle.com/java/", NA, NA, ~
## $ domain_name  <chr> NA, NA, NA, "python.org", NA, "isocp~
## $ domain_name_registered <dbl> NA, NA, NA, 1995, NA, 2012, NA, NA, ~
## $ reference    <chr> NA, "https://www.w3schools.com/js/js~
## $ isbndb       <dbl> 400, 349, 78, 339, 177, 128, 111, 15~
## $ book_count   <dbl> 401, 351, 78, 342, 182, 128, 116, 15~
## $ semantic_scholar <dbl> 37, 48, 19, 52, 37, 6, 7, 37, 26, 9, ~
## $ language_rank <dbl> 0, 1, 2, 3, 4, 6, 5, 7, 8, 9, 10, 12~
## $ github_repo  <chr> NA, NA, NA, NA, NA, NA, NA, NA, "htt~
## $ github_repo_stars <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 3399~
## $ github_repo_forks <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 7234~
## $ github_repo_updated <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 2022~
## $ github_repo_subscribers <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 1441~
## $ github_repo_created <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 2011~
## $ github_repo_description <chr> NA, NA, NA, NA, NA, NA, NA, NA, "The~
## $ github_repo_issues <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 392, ~
## $ github_repo_first_commit <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 1999~
## $ github_language <chr> "Java", "JavaScript", "C", "Python", ~
## $ github_language_tm_scope <chr> "source.java", "source.js", "source.~
## $ github_language_type <chr> "programming", "programming", "progr~
```

| | |
|---|---|
| ## \$ github_language_ace_mode | <chr> "java", "javascript", "c_cpp", "pyth~ |
| ## \$ github_language_file_extensions | <chr> "java jav", "js _js bones cjs es es6~ |
| ## \$ github_language_repos | <dbl> 11529980, 16046489, 2160271, 9300725~ |
| ## \$ wikipedia | <chr> "https://en.wikipedia.org/wiki/Java_~ |
| ## \$ wikipedia_daily_page_views | <dbl> 5242, 4264, 6268, 7204, 3084, 4307, ~ |
| ## \$ wikipedia_backlinks_count | <dbl> 11543, 8982, 10585, 6849, 4159, 1094~ |
| ## \$ wikipedia_summary | <chr> "Java is a general-purpose computer ~ |
| ## \$ wikipedia_page_id | <dbl> 15881, 9845, 6021, 23862, 29004, 720~ |
| ## \$ wikipedia_appeared | <dbl> 1995, 1995, 2011, 1991, 1986, 1998, ~ |
| ## \$ wikipedia_created | <dbl> 2001, 2001, 2001, 2001, 2001, 2001, ~ |
| ## \$ wikipedia_revision_count | <dbl> 7818, 6131, 7316, 6342, 4153, 1487, ~ |
| ## \$ wikipedia_related | <chr> "javascript pizza ada csharp eiffel ~ |
| ## \$ features_has_comments | <lgl> TRUE, TRUE, TRUE, TRUE, TRUE, NA, TR~ |
| ## \$ features_has_semantic_indentation | <lgl> FALSE, FALSE, FALSE, TRUE, FALSE, NA~ |
| ## \$ features_has_line_comments | <lgl> TRUE, TRUE, TRUE, TRUE, TRUE, NA, FA~ |
| ## \$ line_comment_token | <chr> "//", "//", "//", "#", "--", "//", N~ |
| ## \$ last_activity | <dbl> 2022, 2022, 2022, 2022, 2022, 2022, ~ |
| ## \$ number_of_users | <dbl> 5550123, 5962666, 3793768, 2818037, ~ |
| ## \$ number_of_jobs | <dbl> 85206, 63993, 59919, 46976, 219617, ~ |
| ## \$ origin_community | <chr> "Sun Microsystems", "Netscape", "Bel~ |
| ## \$ central_package_repository_count | <dbl> NA, NA, 0, NA, 0, 0, 0, 0, NA, NA, N~ |
| ## \$ file_type | <chr> "text", "text", "text", "text", "tex~ |
| ## \$ is_open_source | <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, ~ |