CodeAcademy Report

ML Team 6 - Sion Kim, Ethan Lee, Kaylyn Phan, Justin Ngo

**Metrics**

* Number of total posts: ~25.8 k
* Number of Categories: 4 categories. Get Help, Community, Projects, and FAQ
* Number of Subcategories: 29 in Get Help and FAQ, 12 in Community, 4 in Projects

**Navigation and Display**

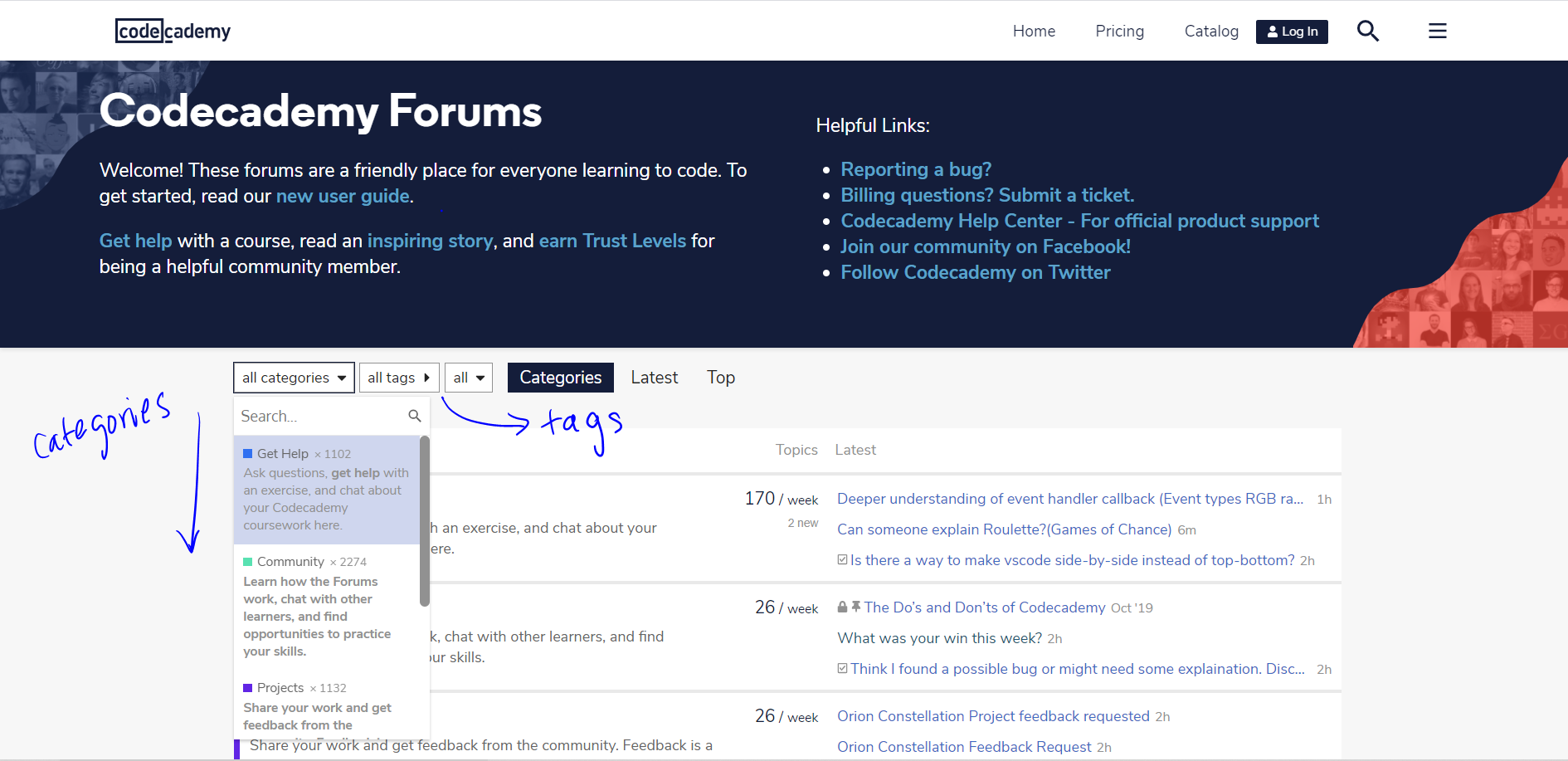
* Users can filter through topics by:
  + Categories
  + Tags
  + Solved/unsolved/all
  + Latest and top posts
* For each topic:
  + The category, sub-category, and tags are shown at the top of the post and at the sticky top bar.
  + Categories and subcategories are represented by colored squares
    - Categories are colored and Subcategories are split-colored
  + Tags are represented by gray squares
* Visually, it has a similar look to the STEM-Away forum.
  + We suspect that Codecademy forum is built off Discourse, like STEM-Away

**Detailed metrics**

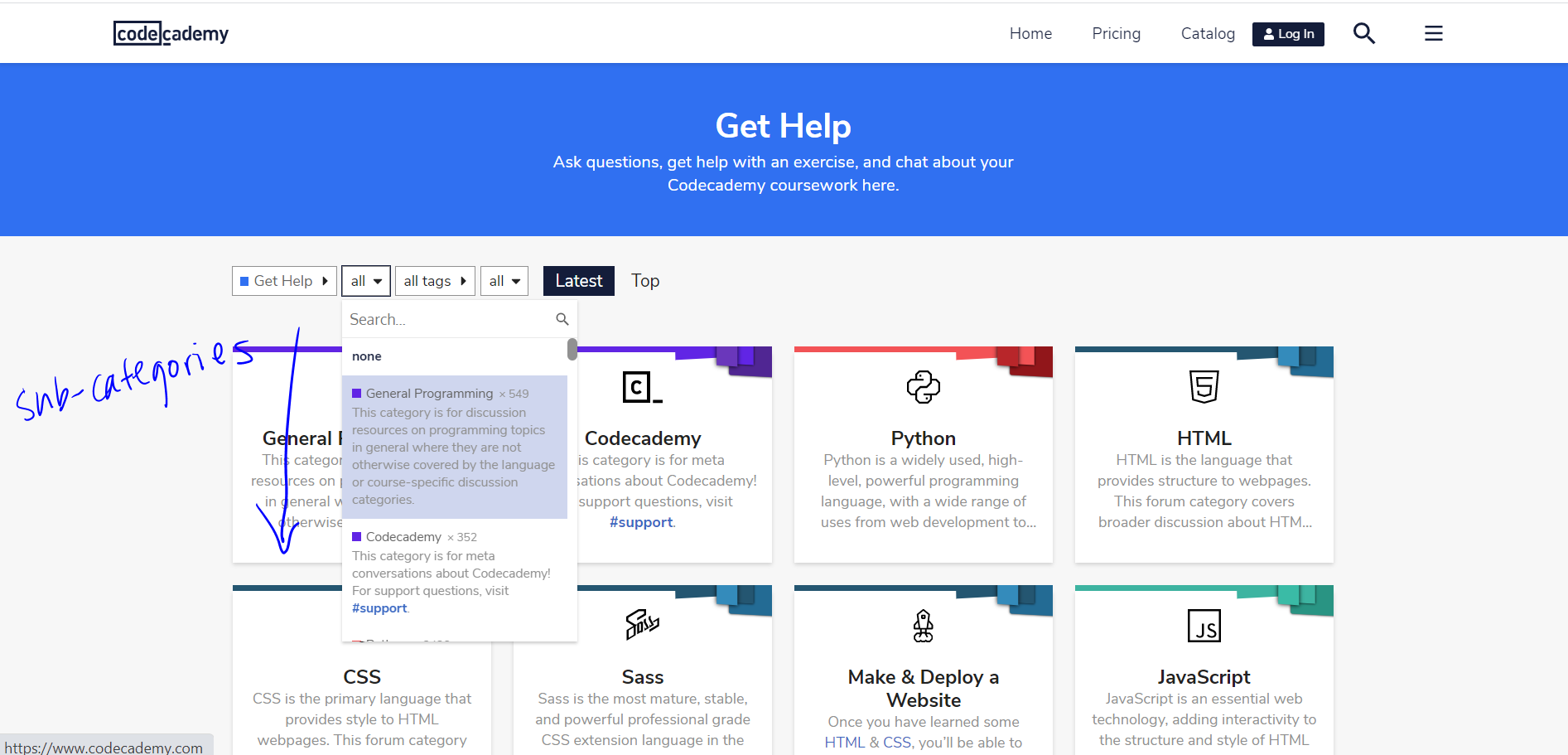
* **Tags System**
  + Popular Tags: advanced-topics, blocks-sorting, classes, conditionals-control-flow-py, functions, lists-dictionaries, loops-py, python-syntax
  + Are Posts Properly Tagged: Yes, most of the time posts are tagged appropriately
  + Are Tags Related to Keywords: Yes, tags are related to keywords (see popular tags for examples)
* **Categories**
  + Are Most of the Posts In a Reasonable Category: Yes, most of the posts are in an appropriate category
* **Titles**
  + Are Most Titles Informative: Yes, most titles are informative: questions contain the phrase “question” or the question mark character, projects are titled “project”, etc.
* **Content**
  + Popular Topics: Python: How to count the number of letters in a string? ; Ls on windows command prompt ; SyntaxError: Invalid shorthand property initializer ;
  + Keywords: What, Why, How, Python, Characters
* **Comments**
  + What People Comment About: Programming, providing coding help, answering coding questions, asking specific questions relating to the original post’s topic such as programming language, frameworks, help, and projects.
  + Are Comments Relevant: Yes, the comments are relevant to the forum and relevant for the purpose of training a topic recommendation and forum classification for STEM-Away because the forum classified posts based on the programming languages and frameworks related to the post.

**Useful places to scrape**

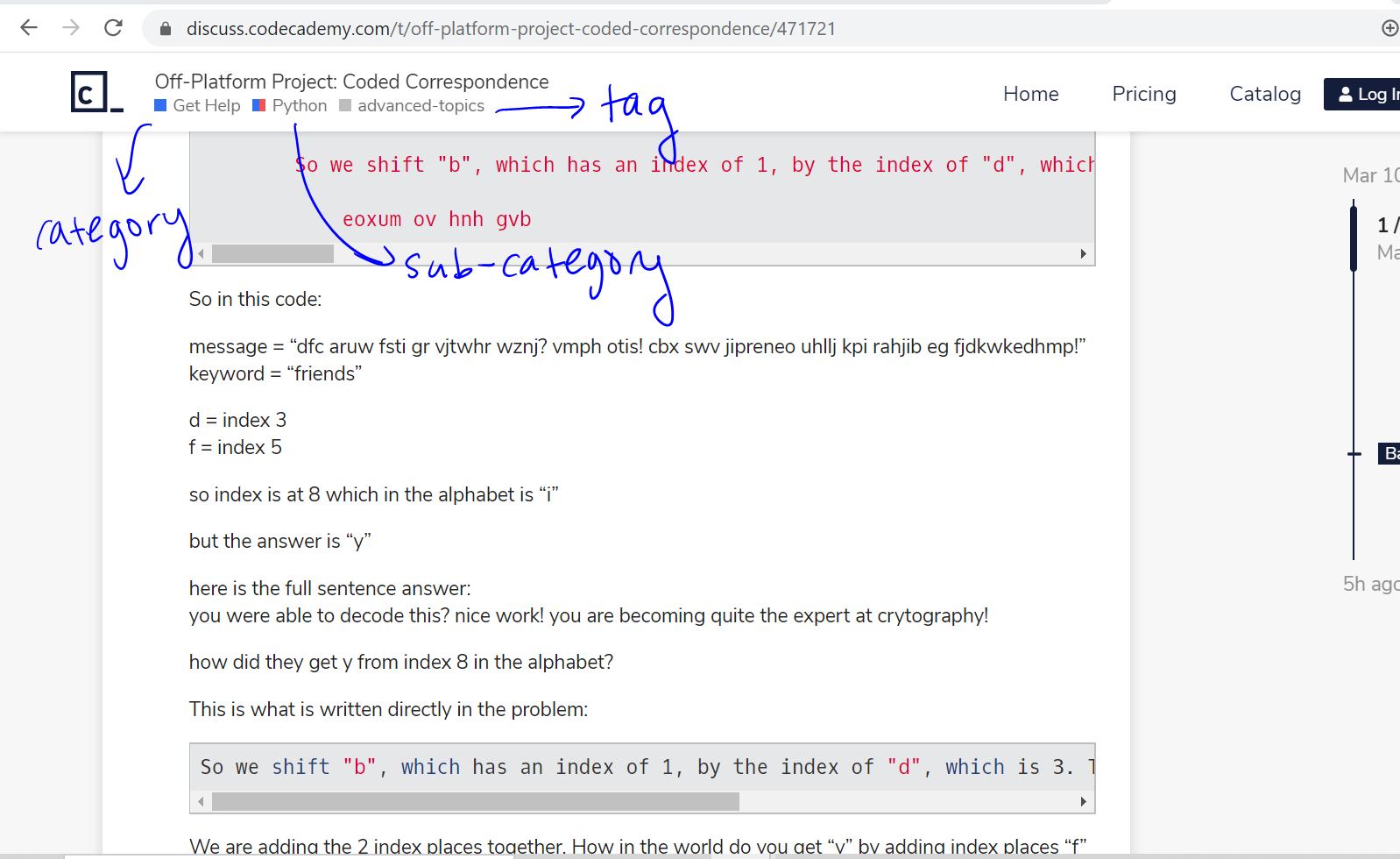
|  |  |  |
| --- | --- | --- |
| **Category/Subcategory** | **Number of topics** | **Range of replies in topic** |
| Community | 7000 | 0-3200 |
| Python | 3436 | 0-136 |
| Python FAQ | 2394 | 0-132 |
| HTML | 2233 | 0-58 |
| CSS | 2092 | 0-59 |
| JavaScript | 3237 | 0-72 |
| JavaScript FAQ | 925 | 0-102 |
| Careers, Journeys, and Stories | 344 | 0-51 |
| Personal Projects | 184 | 0-38 |
| Skill Path Projects | 264 | 0-19 |

**Figures**

***Home Page***



***Category Page***



***Topic Page***

**Conclusion**

After looking at the metrics, **we conclude that the forum is suitable for our project** because there is a sufficient, large amount of technical data as well as general data related to STEM topics that can be collected from the forum.