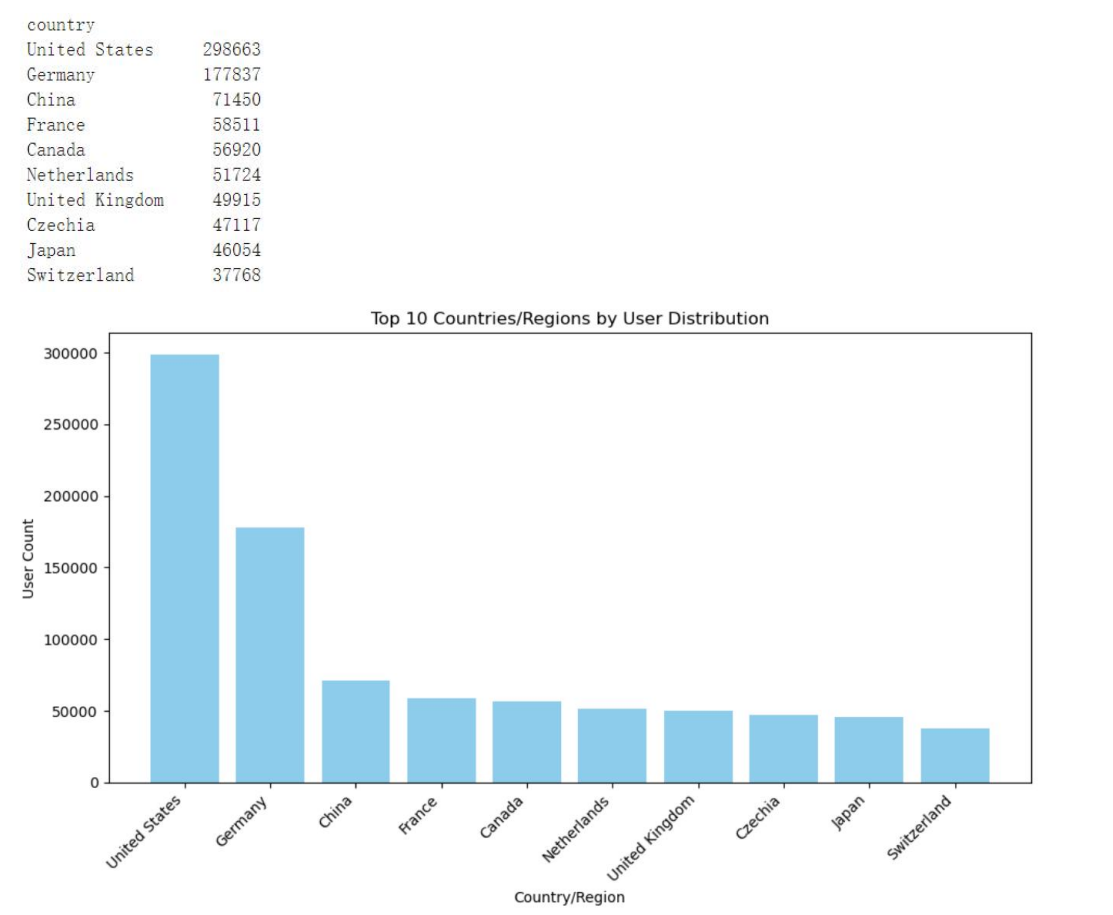


数据洞察报告

一、国家和地区分布



从条形图中可以看出，美国拥有最多的开发者用户，远超其他国家地区。其次是德国，用户数量也达到了一个较高的水平。中国和法国紧随其后，分别位列第三和第四。其他国家地区的用户数量虽然也不少，但与美国和德国相比仍有较大差距。

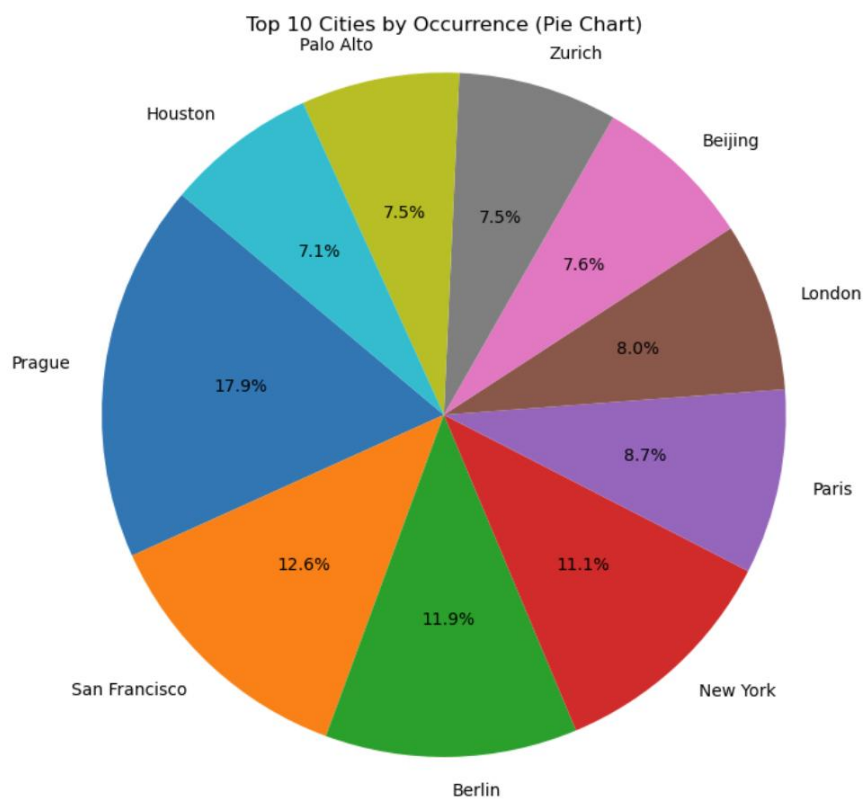
由此可知，美国是全球最大的开发者集中地，用户数量远超其他国家地区，这与其发达的科技产业、丰富的创新资源和完善的创业生态密切相关。德国作为欧洲的经济和科技强国，也吸引了大量开发者，其严谨的教育体系、强大的工业实力和良好的创新环境为开发者提供了广阔的发展空间。中国作为全球最大的互联网市场之一，也拥有庞大的开发者群体。近年来，随着国内科技产业的快速发展和创新环境的不断改善，中国的开发者数量持续增长，展现出强大的发展潜力和活力。法国、加拿大、荷兰、英国等国家地区也拥有一定数量的开发者用户，这些地区在科技、教育、经济等方面都具有较强的实力，为开发者提供了良好的发展环境。值得注意的是，捷克和瑞士等小国也表现不俗，这可能与它们在教育、科研和创新方面的持续投入有关。这些国家虽然规模不大，但在科技和创新领域却取得了显著成就，吸引了众多开发者的关注和参与。

随着全球科技产业的不断发展和创新环境的持续改善，预计未来会有更多国家和地区成为开发者的聚集地。特别是在亚洲和非洲等新兴市场，随着互联网和移动技术的普及，开发者数量有望快速增长。这些地区的开发者将成为全球科技产业发展的重要力量，推动科技创新和产业升级。

数据洞察报告

二、城市级别分布

city	
Prague	38582
San Francisco	27321
Berlin	25747
New York	24050
Paris	18883
London	17249
Beijing	16373
Zurich	16255
Palo Alto	16149
Houston	15409



通过饼图，我们发现布拉格的开发者数量占比最高，紧随其后的是旧金山和柏林，。纽约、巴黎、伦敦等城市的开发者数量也占据了相当大的比例。

饼图直观地展示了各个城市开发者数量的相对大小，使我们能够迅速识别出技术热点区域。布拉格、旧金山、柏林和纽约可以被视为全球技术热点区域。这些城市不仅拥有庞大的开发者群体，还吸引了众多科技公司和初创企业入驻，形成了良好的技术创新和创业生态。布拉格作为东欧的领头羊，其技术实力不容小觑。随着全球数字化转型的加速推进，布拉格有望在未来几年内继续保持其技术领先地位，并吸引更多国际企业和人才前来投资和发展。旧金山和柏林作为美国和欧洲的技术中心，其创新能力和技术实力一直备受瞩目。这两个城市不仅拥有众多知名科技公司和初创企业，还孕育了众多颠覆性的技术和产品，对全球科技发展产生了深远影响。纽约作为全球金融中心之一，其技术实力同样不容忽视。随着金融科技和人工智能等领域的快速发展，纽约正在逐步成为全球技术创新的重要一环。

未来，随着全球数字化转型的加速推进，这些城市有望继续保持其技术领先地位，并引领全球科技发展的潮流。

数据洞察报告

三、时区分布

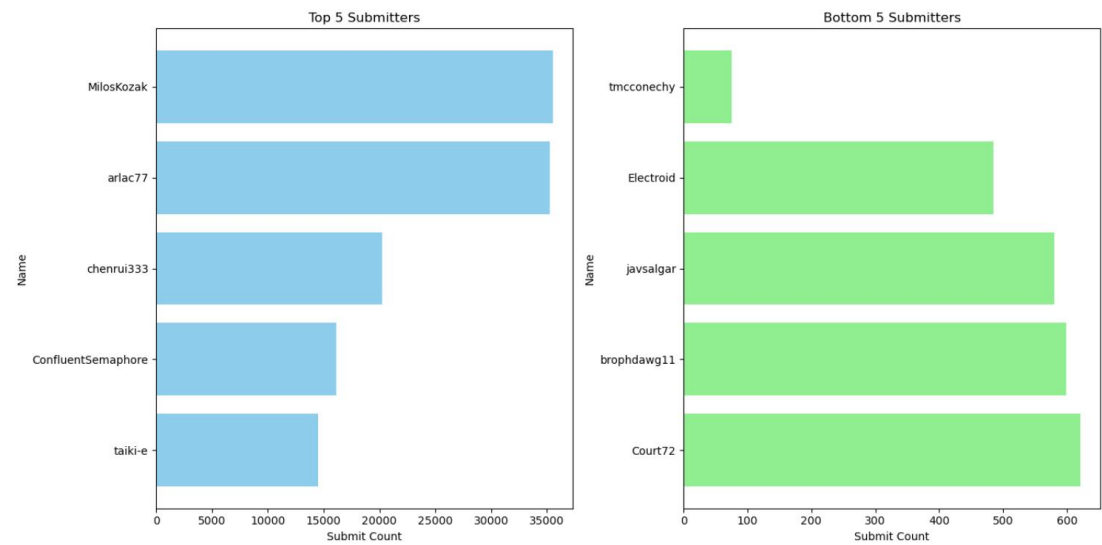
```
timezone_offset
+08:00      1245524
```

从时区分布来看，所有用户均位于+08:00 时区，这主要涵盖了中国的标准时间以及部分东南亚和澳大利亚地区的夏令时。这意味着我们的用户群体主要集中在这一时区范围内，他们可能共享相似的作息习惯和工作时间安排。

四、提交频率

```
Top 5 Submitters:
user_id      name      submit_count
345  11146458  MilosKozak  35532
52   158862   arlac77     35237
179  1580956   chenrui333  20214
448  40306929  ConfluentSemaphore  16149
454  43724913  taiki-e     14492

Bottom 5 Submitters:
user_id      name      submit_count
130  814283   tmcconechy  75
232  3238291  Electroid   485
150  1133777  javsalgar   580
181  1609022  brophdawg11  599
475  62625502  Court72     621
```



在高活跃用户中，用户 ID 为 11146458（MilosKozak）和用户 ID 为 158862（arlac77）的提交次数遥遥领先，分别达到了 35532 次和 35237 次。这两位用户可能是社区中的核心成员或高度活跃的用户，对 Github 平台的贡献度极高。接下来是用户 ID 为 1580956（chenrui333），提交次数为 20214 次，虽然与前两名用户相比有所下降，但仍然保持了较高的活跃度。用户 ID 为 40306929（ConfluentSemaphore）和用户 ID 为 43724913（taiki-e）的提交次数分别为 16149 次和 14492 次，虽然不及前三名，但在所有用户中仍然属于高活跃群体。

在低活跃用户中，用户 ID 为 814283（tmcconechy）的提交次数最少，仅为 75 次。这可能是由于该用户是新加入的用户，或者对平台的参与度不高。用户 ID 为 3238291（Electroid）、

数据洞察报告

1133777 (javsalgar)、1609022 (brophdawg11) 和 62625502 (Court72) 的提交次数分别为 485 次、580 次、599 次和 621 次。这些用户的活跃度相对较低，但仍有一定的参与。

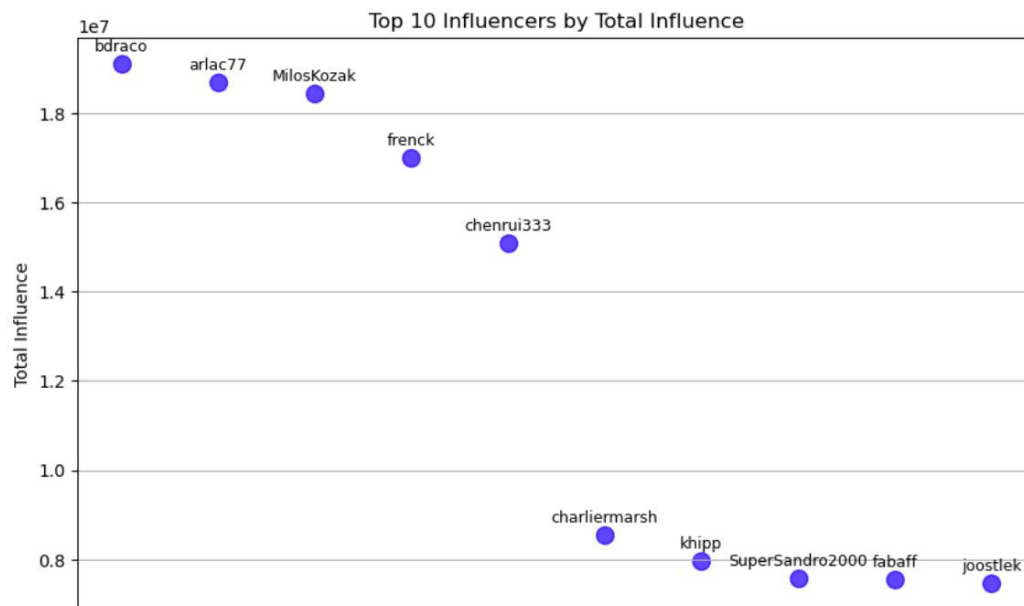
通过分析高活跃用户和低活跃用户的提交次数，我们可以发现用户之间的活跃度存在显著差异。高活跃用户对平台的贡献度远高于低活跃用户。

对于高活跃用户，平台可以给予更多的关注和奖励，如提供专属的优惠、特权或荣誉标识，以进一步激发其参与度和忠诚度。对于低活跃用户，平台可以通过推送个性化内容、优惠活动或邀请其参与特定活动等方式，提高其参与度和活跃度。同时，也可以分析这些用户的行为特征，找出导致其活跃度低的原因，并采取相应的改进措施。

五、总影响力

Top 10 Influencers:

	name	total_influence
119	bdraco	1.911306e+07
52	arlac77	1.870270e+07
345	MilosKozak	1.844511e+07
63	frenck	1.700514e+07
179	chenrui333	1.508810e+07
162	charliermarsh	8.540971e+06
340	khipp	7.954755e+06
313	SuperSandro2000	7.572737e+06
40	fabaff	7.538915e+06
310	joostlek	7.465266e+06



根据总影响力排序，前 10 名用户分别是 bdraco、arlac77、MilosKozak、frenck、chenrui333、charliermarsh、khipp、SuperSandro2000、fabaff 和 joostlek。这些用户的总影响力介于 1911 万至 746 万之间，显示出较高的社会或网络影响力。同时，前五名的总影响力显著高于第六到十名。

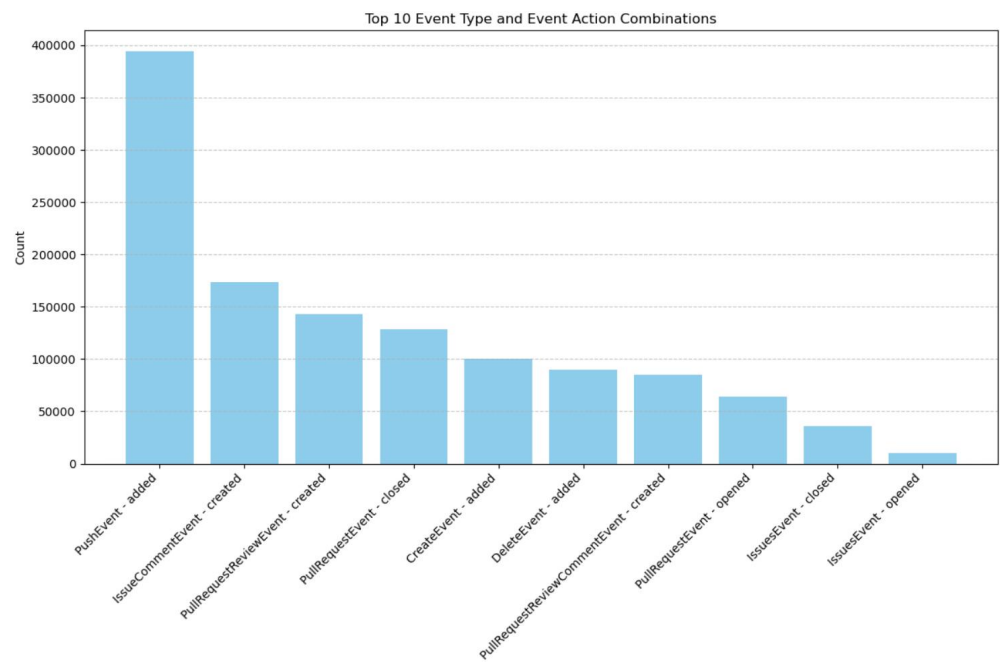
对于 Github 平台来说，识别并关注这些精英用户是至关重要的，因为他们可能具有更大的传播力和影响力。了解精英用户的内容创作和互动模式，可以为其他用户提供有益的参考和借鉴。为了平衡影响力分布，平台也可以考虑采取措施来鼓励和支持更多用户参与内容创作和社区互动，从而形成一个更加健康和多元的用户生态。

数据洞察报告

六、事件类型与事件动作的关联

事件类型与事件动作的关联分析（前10个常见组合）：

	event_type	event_action	count
17	PushEvent	added	394469
5	IssueCommentEvent	created	173514
16	PullRequestReviewEvent	created	143107
11	PullRequestEvent	closed	128783
1	CreateEvent	added	100026
2	DeleteEvent	added	89586
15	PullRequestReviewCommentEvent	created	85084
12	PullRequestEvent	opened	64375
6	IssuesEvent	closed	35817
7	IssuesEvent	opened	10371



在所有事件类型与事件动作的组合中，PushEvent 与 added 动作的组合最为常见，出现了 394,469 次。这表明在 GitHub 上，代码推送（Push）并添加新内容的行为非常频繁。IssueCommentEvent 与 created 动作的组合出现了 173,514 次，显示了在问题（Issue）评论区的活跃交流。PullRequestReviewEvent 与 created 动作的组合出现了 143,107 次，反映了在拉取请求（Pull Request）审查过程中的高参与度。PullRequestEvent 与 closed 动作的组合出现了 128,783 次，而与其 opened 动作的组合出现了 64,375 次。这表明在拉取请求的生命周期中，关闭操作相对更常见，但打开操作也不容忽视。CreateEvent 与 added 动作的组合出现了 100,026 次，而 DeleteEvent 与 added 动作的组合也达到了 89,586 次。这显示了创建和删除操作在 GitHub 上的重要性。IssuesEvent 与 closed 动作的组合出现了 35,817 次，而与其 opened 动作的组合出现了 10,371 次。这表明在问题（Issue）的生命周期中，关闭操作远多于打开操作。

综上所述，PushEvent - added、IssueCommentEvent - created 以及 PullRequest 相关的事件类型与事件动作组合在监测的期间内最为常见。这些发现对于理解代码仓库中的活动模式、优化工作流程以及提高团队协作效率具有重要意义。

鉴于 PushEvent - added 组合的流行程度，可以进一步优化代码推送的流程，确保代码质量的同时提高推送效率。对于 IssueCommentEvent - created 组合的高频出现，建议加强问题

数据洞察报告

讨论的管理，确保问题得到及时解决。针对 **PullRequest** 相关事件的高频发生，可以优化代码审查的流程，提高审查效率和质量。通过实施这些举措，可以进一步提高代码仓库的管理效率，促进团队协作和项目的顺利进行。