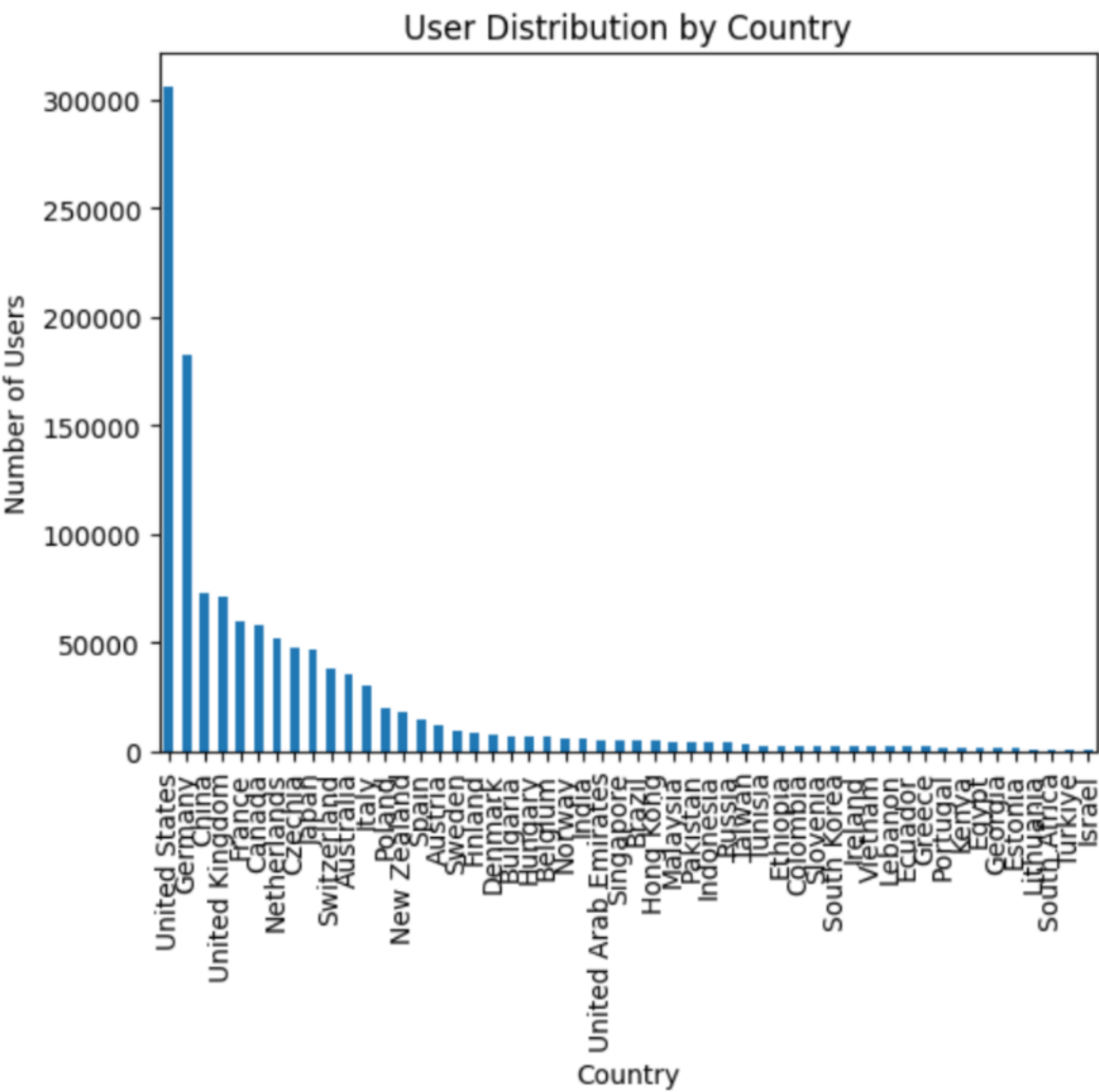


数据洞察报告

1. 人口统计分析

1) 国家和地区分布



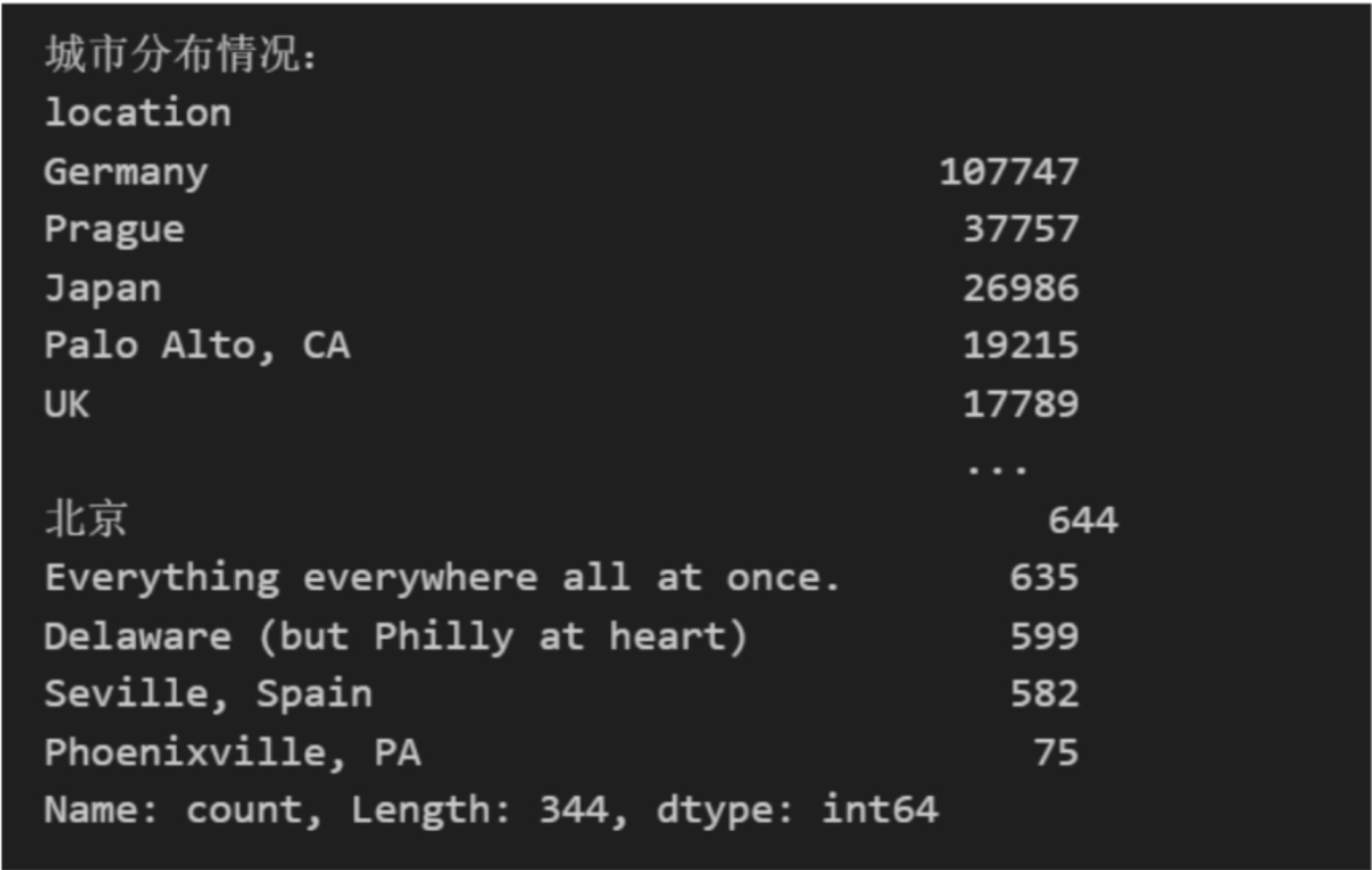
美国用户数量远远超过其他国家，达到了接近 30000 人的规模。这可能是因为美国是全球科技和互联网行业最为发达的国家之一，拥有众多的科技企业和创新中心，吸引了大量相关领域的人才，这些人才有较大可能参与到相关平台的使用中。

英国、加拿大、德国的用户数量也较多这些国家同样拥有发达的科技产业和较高的互联网普及率，并且在国际交流和技术合

作方面比较活跃，其用户有较多机会参与到相关平台的使用中。

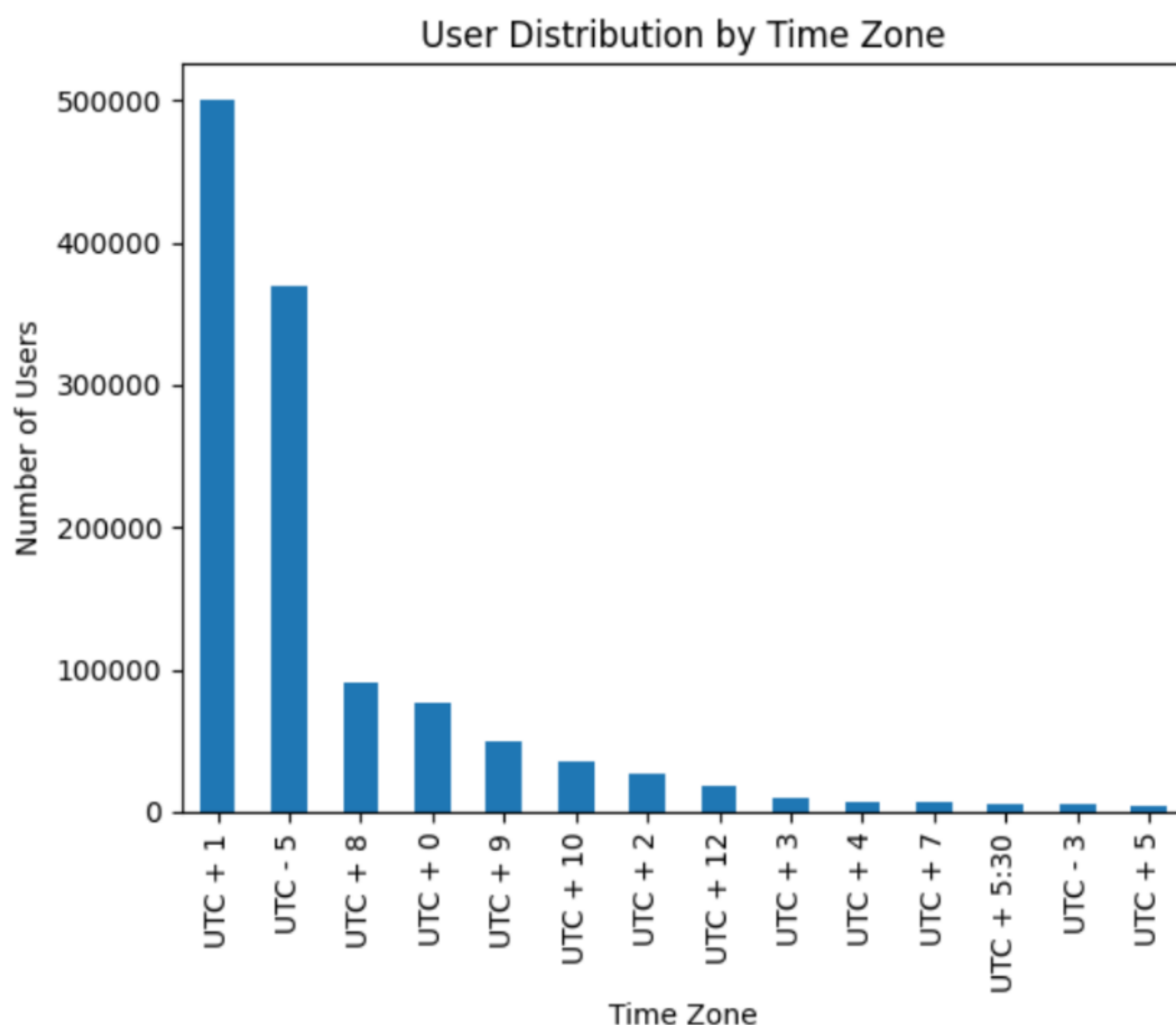
而随着国家经济和科技发展水平的下降，用户数量也逐渐减少。发展中国家的科技基础设施可能相对薄弱，互联网的普及程度和相关技术人才的数量相对较少，导致这些国家的用户数量较少。

2) 城市级别分布



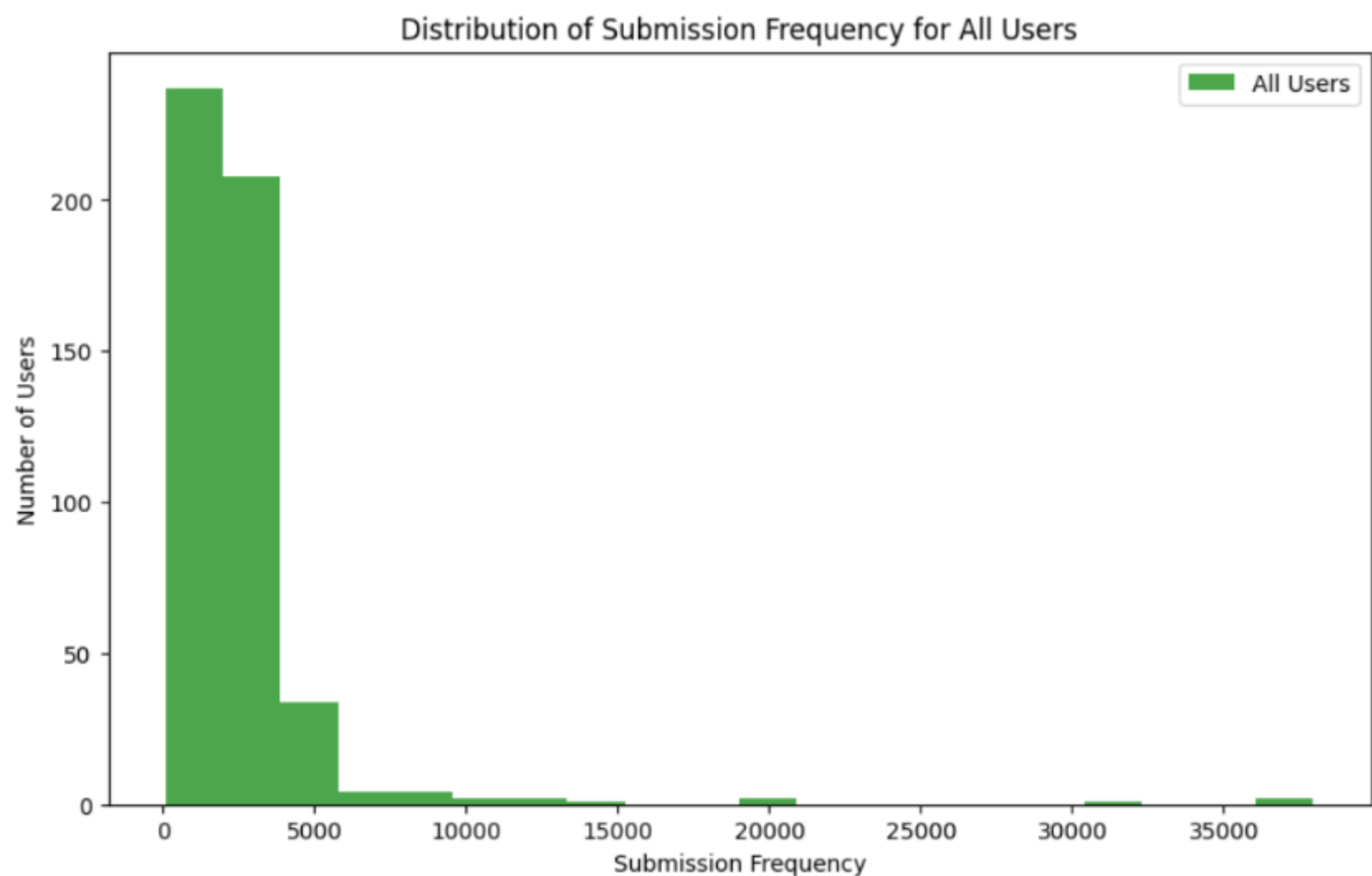
根据图标，我们发现德国的用户数量最多，达到 107747 人。这表明德国在这个数据集中是人口最为集中的地区。布拉格、日本、加利福尼亚州帕洛阿尔托和英国也有较多的人口，分别为 37757 人、26986 人、19215 人和 17789 人。这些城市或国家在数据集中显示出较高的人口聚集度。

3) 时区分布



从图中可以看出：UTC+1时区的用户数量最多，接近 50000 人，远远超过其他时区的用户数量。这是欧洲的部分发达国家所处时区；UTC-5 时区用户也相对较多，这是美国所在时区约有 40000 人。其他时区的用户数量相对较少，大部分时区的用户数量在 10000 人以下，甚至有一些时区的用户数量非常少，接近 0。这种用户分布可能与这些时区所覆盖的地区的人口密度、互联网使用习惯、业务需求等因素有关。

2. 协作行为分析



通过对提交次数的分析，我们区分出了高活跃用户与低活跃度用户，其中提交次数在 3000 次以上的用户有 124 名，提交次数最多的用户提交了 4509 次，而低活跃用户提交次数都在 1500 次以下，也有 124 名。

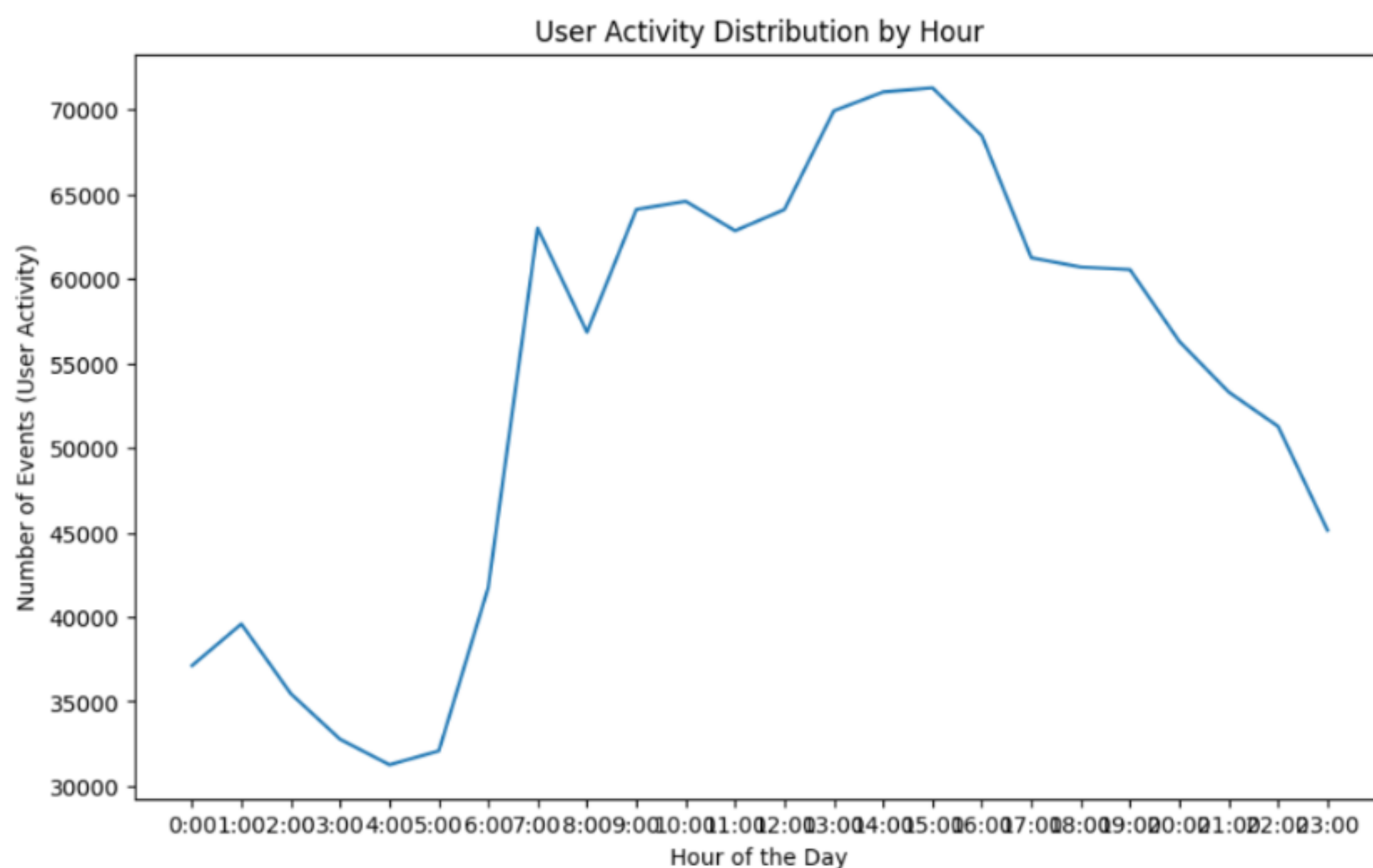
高活跃用户：

user_id	
13564	3140
26967	3214
27350	4509
32321	3284
47313	3356
...	
66076021	3735
74887181	2997
88161975	3394
91018726	4098
100913391	3177

低活跃用户:

user_id	
2621	796
23304	1278
47792	1337
57224	1225
94272	1116
...	
62513924	1068
62625502	621
71636191	1158
83238715	923
89584709	862

3. 探究用户活跃时间分布



从图中可以看出：

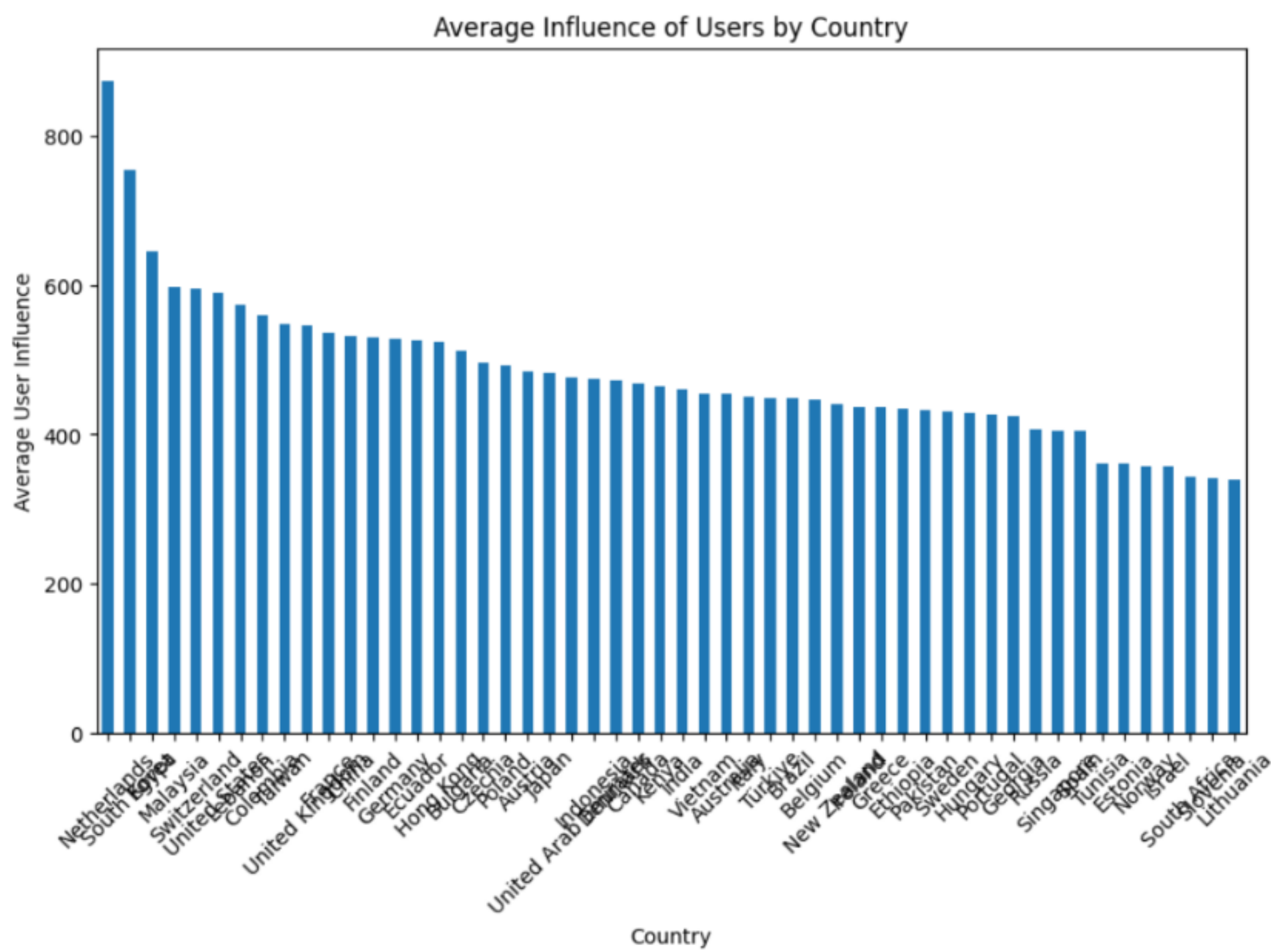
用户活动在一天中的高峰期大约在 11:00、12:00 和 13:00 左右，此时的用户活动数量（Number of Events）达到了接近

70000 的峰值。这可能是因为这些时间段是大多数人工作或学习的午休时间，人们有更多的空闲时间来进行相关活动。

用户活动在凌晨 00:00 到 04:00 左右处于低谷，活动数量在 30000 到 35000 左右，这符合大多数人在夜间休息的生活规律。

从 04:00 开始，用户活动数量逐渐上升，在 11:00 - 13:00 达到顶峰，然后逐渐下降。从 17:00 开始，用户活动数量下降得较为明显，进入夜间后保持在较低水平。

4. 探究不同国家用户的平均影响力



从图中可以看出，荷兰的用户平均影响力最高，接近 800。这表明荷兰用户在相关领域可能具有较高的影响力。新加坡、瑞士、卢森堡、英国、德国等国家的用户平均影响力也

较高，都在 600 以上。这些国家通常具有发达的科技和经济，可能是其用户具有较高影响力的原因之一。

越南、新西兰、希腊等国家的用户平均影响力相对较低，在 300 - 400 之间。苏丹、津巴布韦等国家的用户平均影响力最低，在 200 - 300 之间。

总体而言，不同国家用户的平均影响力与该国的科技、经济、教育水平和行业集中程度等因素密切相关。