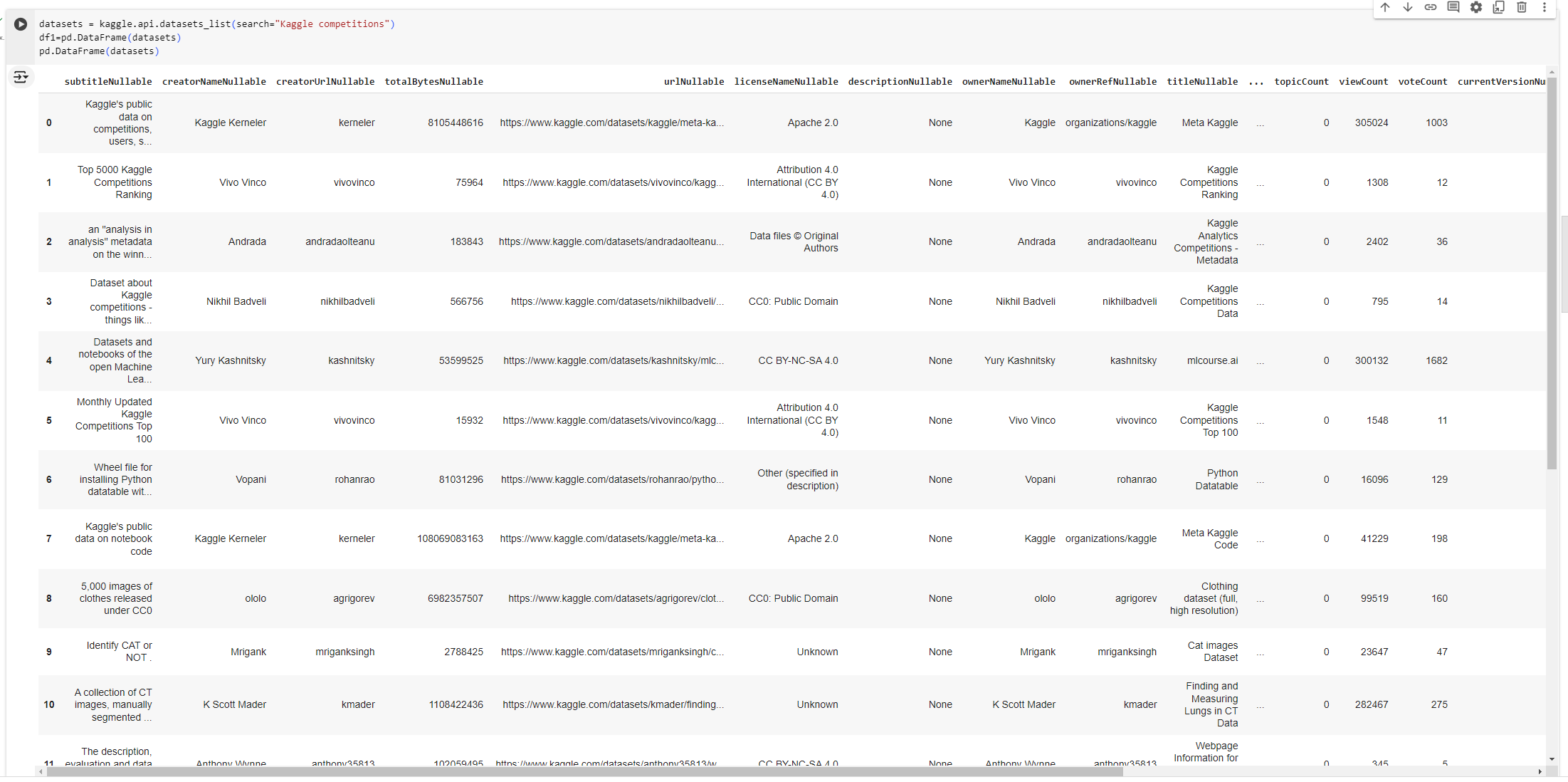
**Практическая работа 1. Начало работы с API**

**Выполнил: Григорьев П.К., БД-231**

**Вариант 3:** анализ участников конкурсов: Географический анализ участников конкурсов по большим данным.

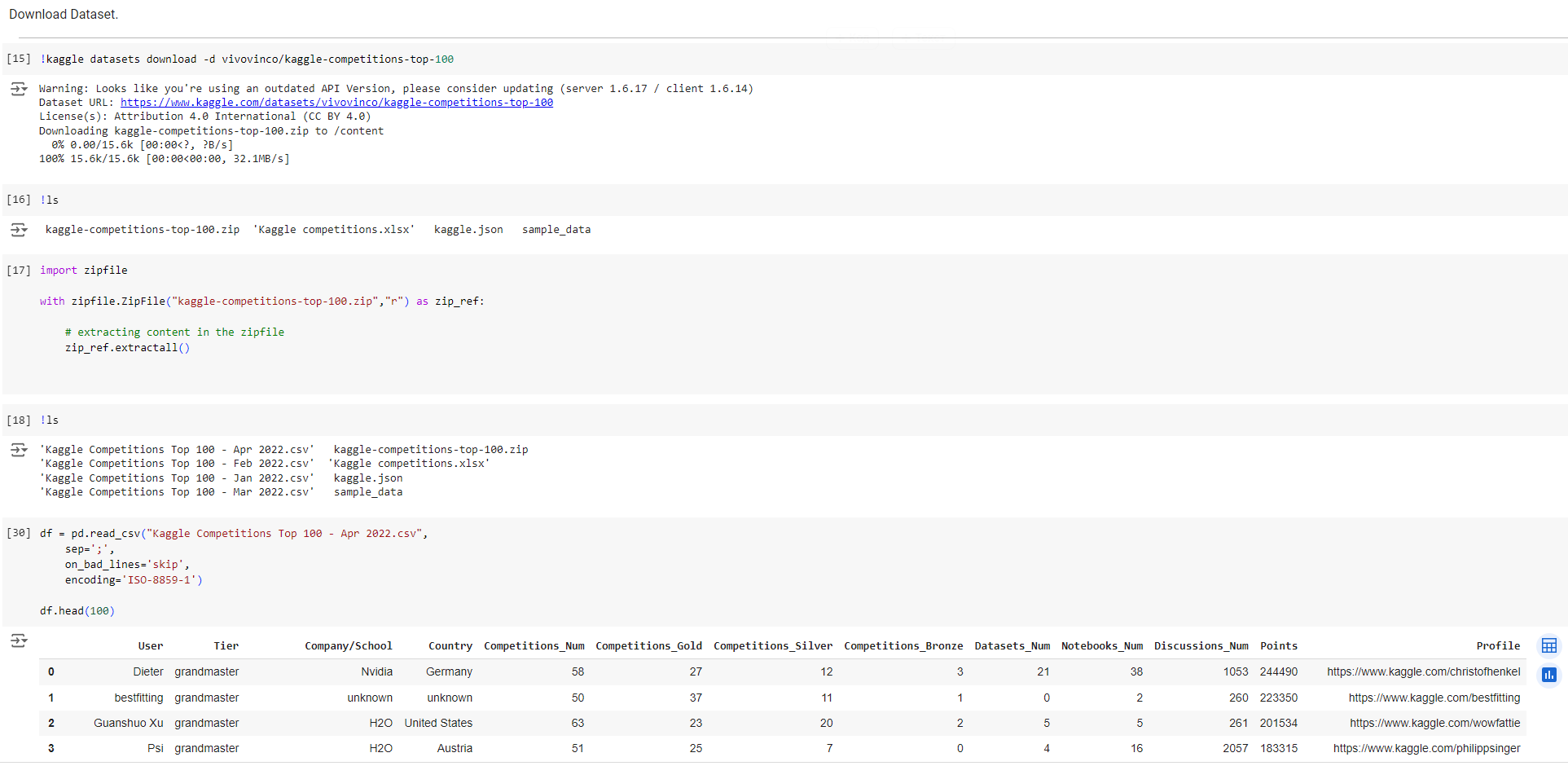
Краткий отчет:

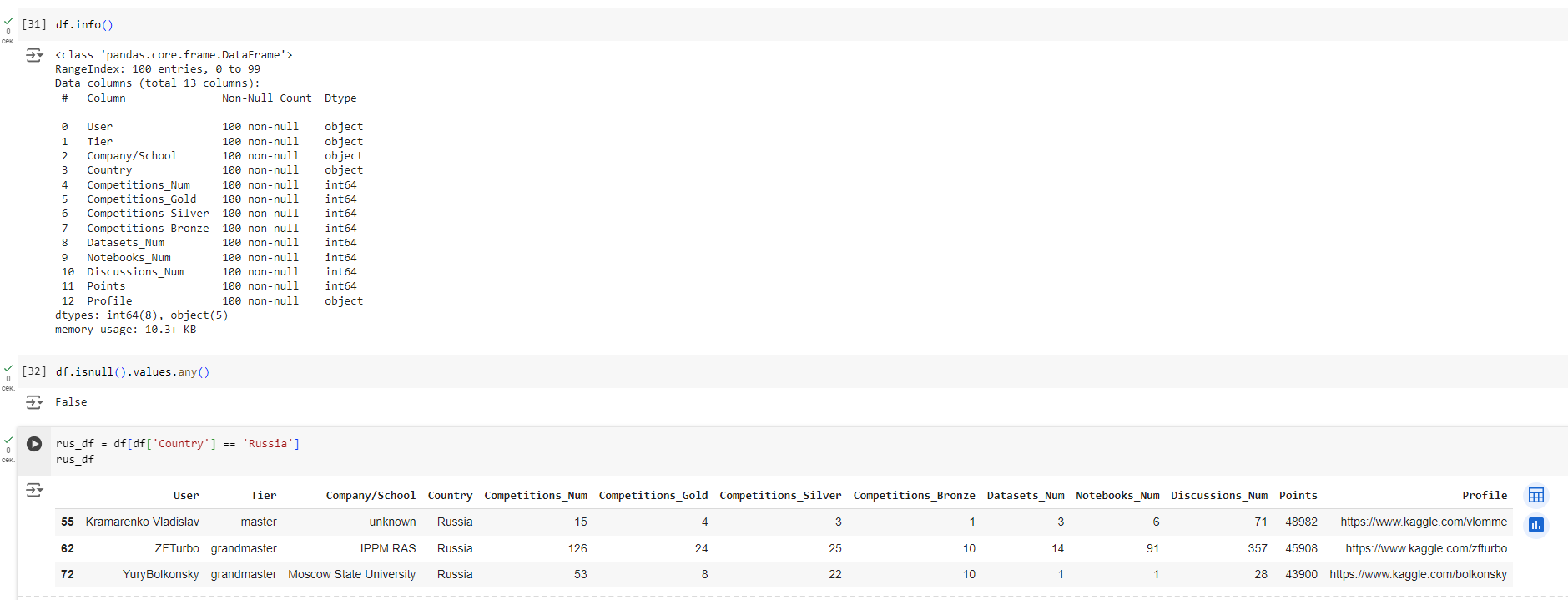
С помощью API Kaggle по ключевому слову "Kaggle competitions" был найден датасет “Kaggle Competitions Top 100”, содержащий информацию о 100 лучших участниках конкурсов от Kaggle по работе с данными на момент апреля 2022 года:



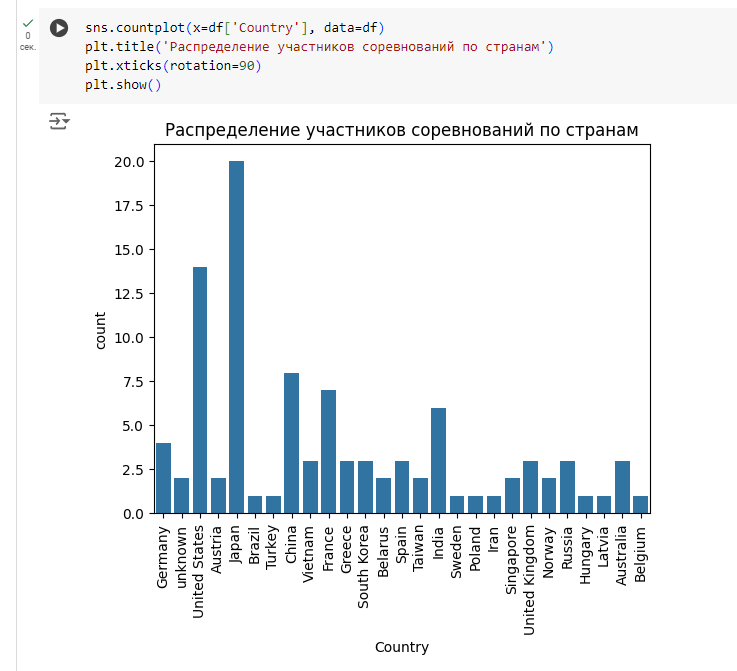
Был выбран именно указанный датасет, так как он содержит информацию о странах участников, без чего провести географический анализ будет невозможно.

Датасет был загружен в Colab и преобразован в датафрейм pandas:

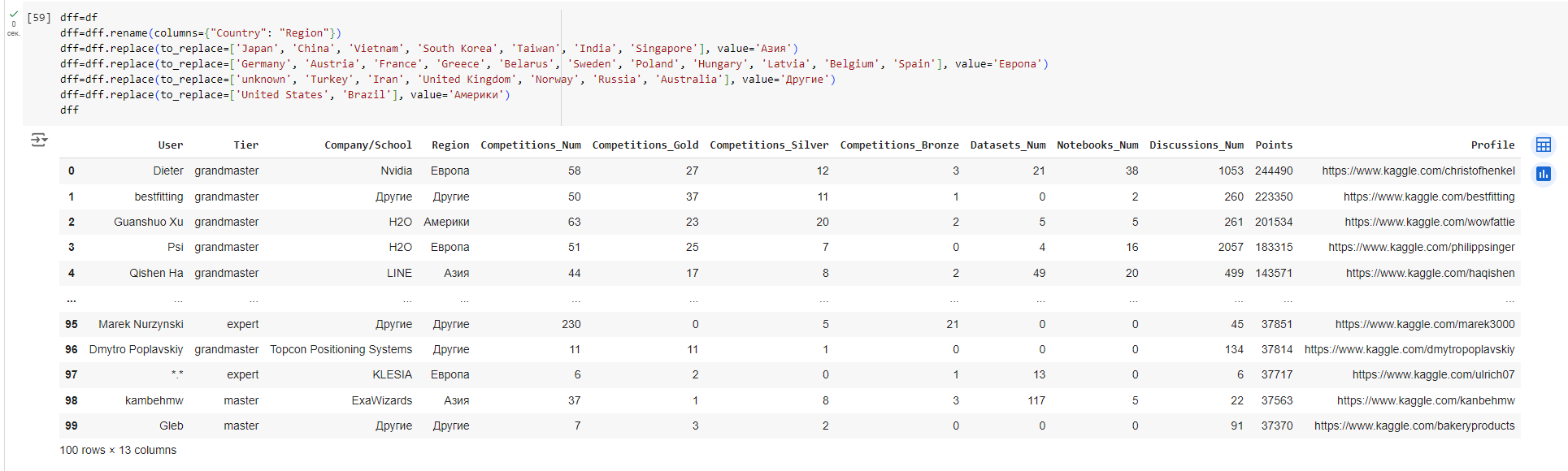




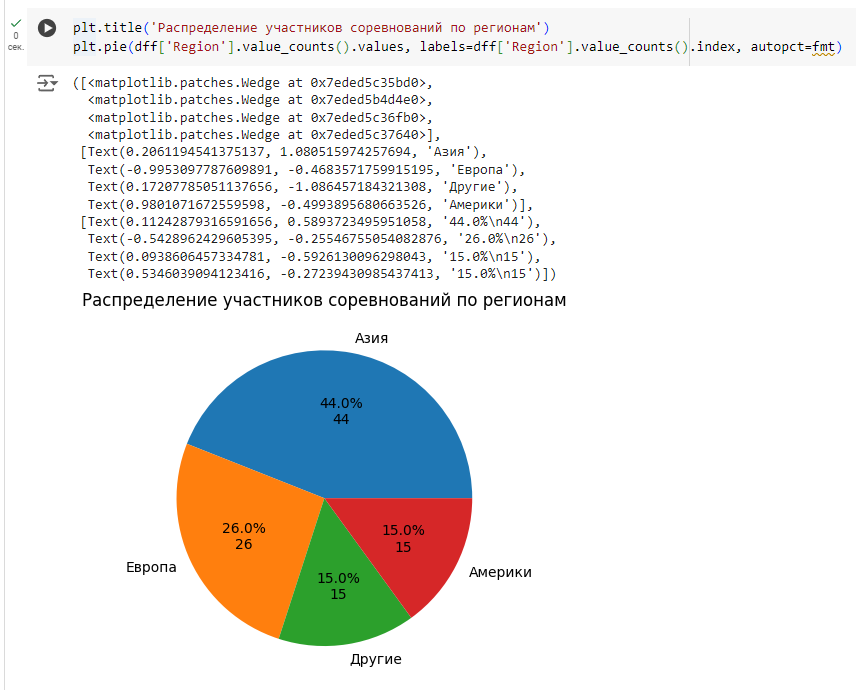
Был построен график для изучения распределения участников конкурсов по странам:



Для анализа распределения участников по регионам датафрейм был преобразован:



Была построена круговая диаграмма для изучения распределения участников по регионам:



В рамках географического анализа можно сделать следующие выводы и наблюдения:

1. В топе 100 участников соревнований (на момент апреля 2022), проводимых Kaggle, больше всего представителей азиатского региона - 44% (в особенности участников из Японии), что может говорить о активном развитии интереса к данной тематике в азиатских странах;
2. Вторая страна по количеству участников соревнований в топ 100 – США, количество участников в топ 100 из которой превышает количество участников из таких стран как Китай и Индия;
3. В топ 100 представлено 3 участника соревнований из России.