

Введение в статистику

Мини-кейс #0А

В кресле руководителя

Описание кейса



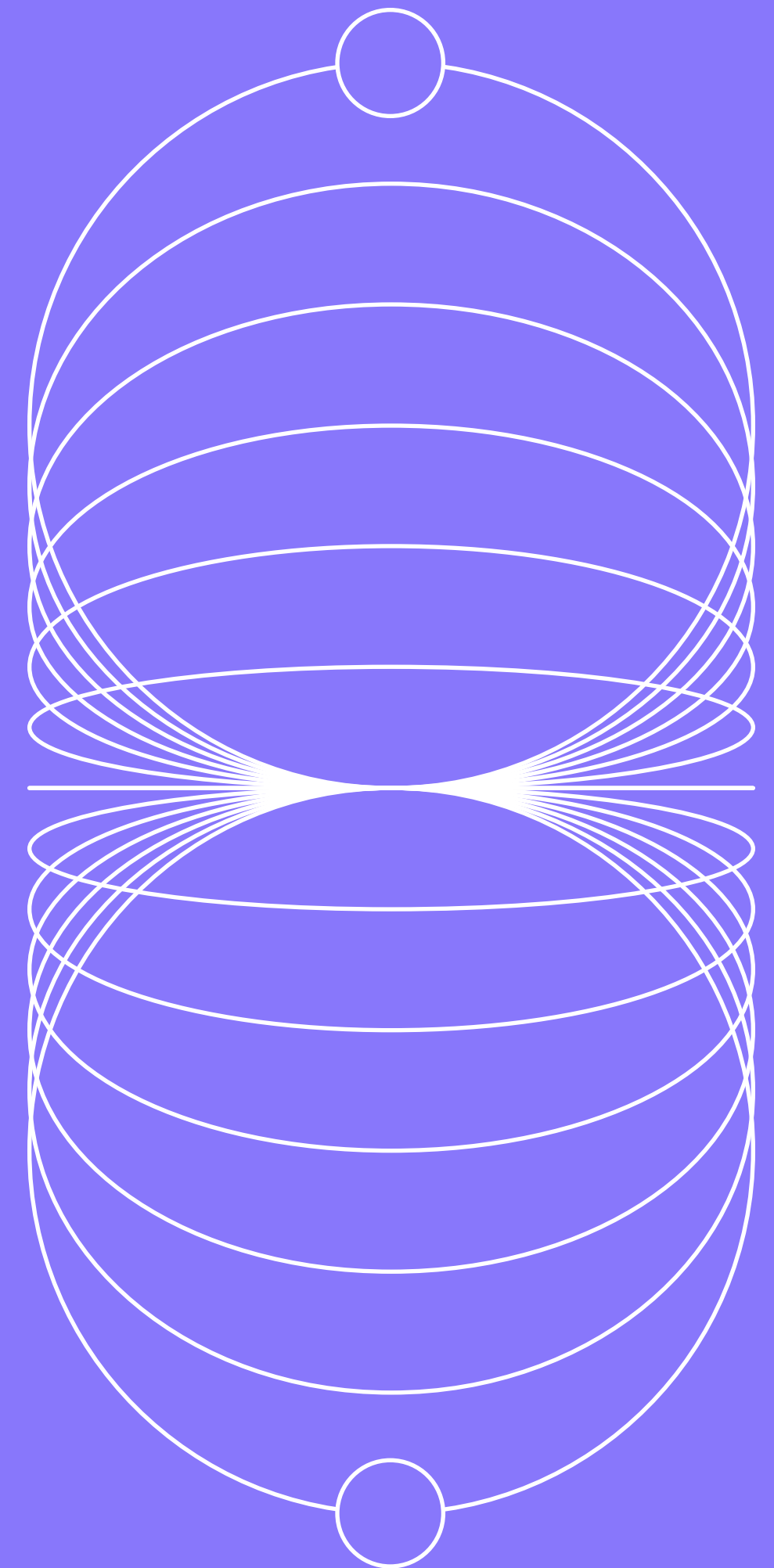
Кейс

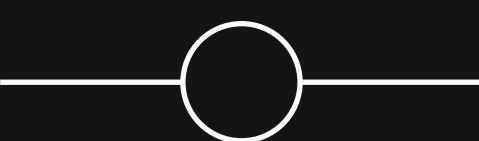
- Ты руководитель отдела маркетинга в компании, которая обучает и подбирает специалистов в сфере Data Science
- Ты готовишь презентацию для руководства
- Андрей и Борис — стажёры у тебя в подчинении
- Ты поставил им задачу сделать два слайда для твоей презентации
- Только что ты получил слайды от Андрея и Бориса




Задача

Оценить результат работы стажёров, чтобы потом обсудить с руководством их будущее в компании





Как оценивать работу стажёров?



Критерии оценки — в конце этой презентации



Скорее к слайдам!



A decorative graphic consisting of two white, overlapping, wave-like shapes that form a central lens-like area. The background is a solid green color.

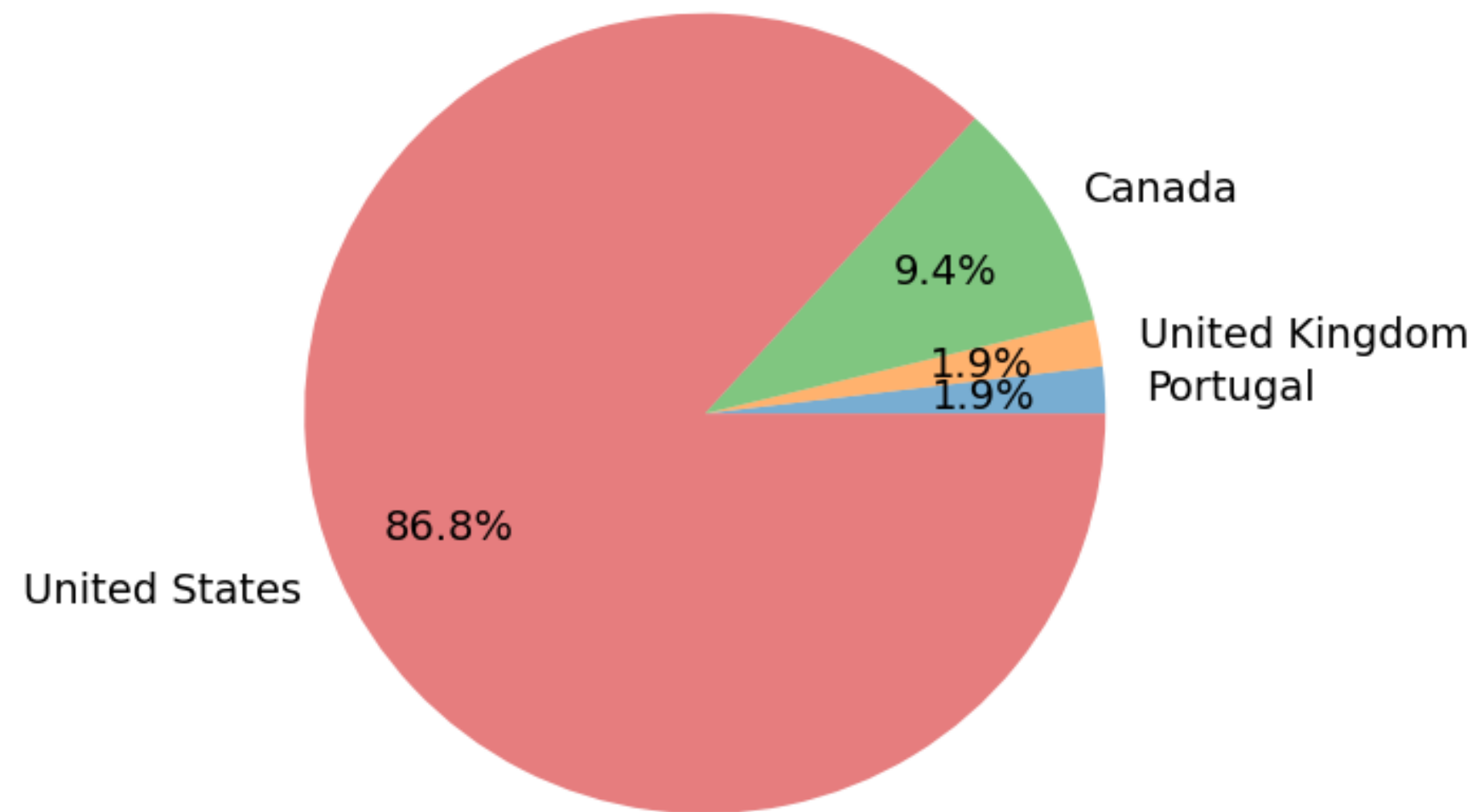
Слайды Андрея

Как выявить текущий спрос на вакансию “data-scientist” в разных странах?

- Датасет международных вакансий в сфере Data Science и IT
- 87 стран
- 6599 вакансий
- 106 вакансий “Data Scientist” за январь 2024
- Найдём страны, откуда большинство вакансий

Почти все вакансии из США

Доля вакансий "Data Scientist"
в январе 2024 года в разных странах



- В США наибольший спрос на специалистов "Data Scientist"
- Спрос в США превосходит суммарный спрос в остальном мире
- За январь 2024 в некоторых странах Европы нет вакансий

Ссылка на расчёты по датасету

https://colab.research.google.com/drive/1SaT6TU6X3RH_2G_8_TQkVjDC7hz0h8jd?usp=sharing



Слайды Бориса

Проделанная работа

1.1

- Проверили датасет на предмет пропущенных значений:

```
Number of missing values: 0
```



- Проанализировали данные по 87 странам:

```
Number of countries in datatable: 87
```



- Выявили количество текущий вакансий "Data Scientist" :

```
Number of 'Data Scientist' vacancies: 106
```

https://colab.research.google.com/drive/1Emgn9BCjg4XlvkJQq_vlX2kwftQsIEpU?usp=sharing

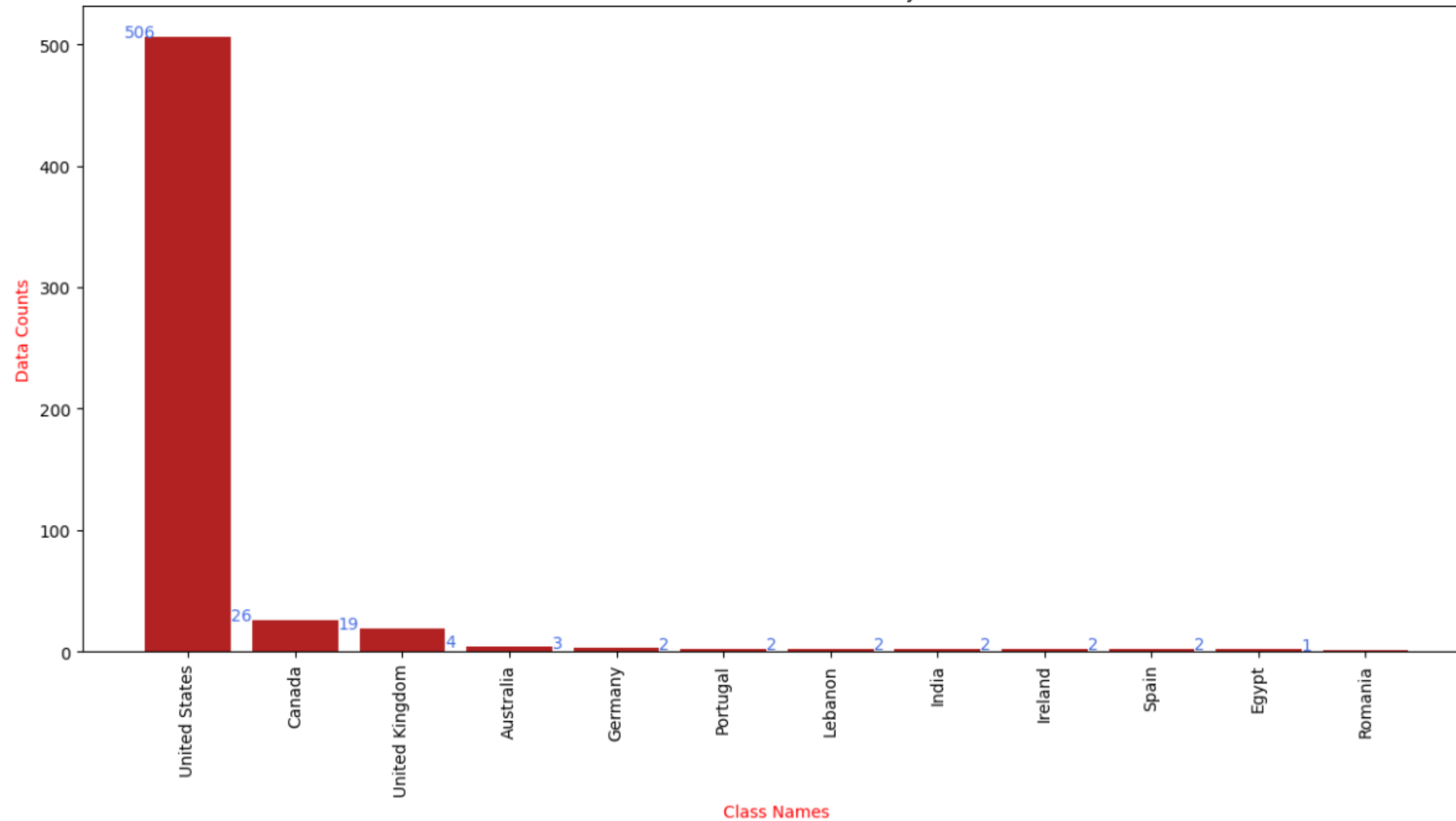
Получили количества вакансий "Data Scientist" по каждой стране

employee_residence	
United States	92
Canada	10
Portugal	2
United Kingdom	2

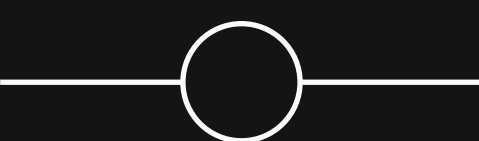


Распределение вакансий "Data Scientist"



Dataset Class Imbalance Analysis



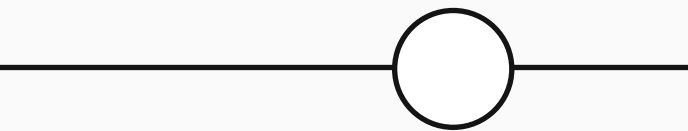
- Больше всего вакансий в США, это видно на графике, потому что столбец самый высокий



**Пока ты в кресле
руководителя — оцени
работу Андрея и Бориса!**



**Критерии оценивания
— на следующем слайде**



Требования к слайдам

Если слайды или Jupyter Notebook не приложен, решение кейса **оценивается в 0 баллов**

Понятность/внешний вид

- Внешний вид презентации
- не мешает воспринимать информацию
- Понятно на какие вопросы отвечает каждый слайд
- Содержимое таблиц, графиков понятно из слайда без необходимости открывать исходный датасет
- Выводы явно сформулированы

МАКСИМУМ 4 БАЛЛА

Обоснованность

- Выводы основаны на таблицах, графиках, показателях, полученных из данных
- Таблицы и графики получены скриншотом или картинкой из Jupyter Notebook, поэтому их можно перепроверить
- Выводы явно сформулированы

МАКСИМУМ 4 БАЛЛА

Реакция заказчика

- **0 баллов**
Не принимает, ищет другого исполнителя
- **1 балл**
Частично принимает, считает необходимым отдать на доработку текущему исполнителю
- **2 балла**
Принимает, готов пересылать слайды от своего имени, под свою ответственность

МАКСИМУМ 2 БАЛЛА

Максимальная оценка — **10 баллов**

Бонусные баллы могут поднять основную оценку, но не выше **10 баллов**, даже если в сумме получится 11 или 12 баллов