

ПОХОДЫ ПО СОБЕСЕДОВАНИЯМ – КАКОЙ РАБОТОДАТЕЛЬ ВАМ НЕ НУЖЕН

Крючкова Анна

mail.to.me.ann@gmail.com

Поздравляю! Быть data scientist-ом – это круто!

- Но есть обратная сторона медали:

После больших технологических компаний и небольших лабораторий бум на потребности в DS докатился до массового рынка, где data science стал священным Граалем будущих успехов компании

«Наш бизнес спасет data scientist! Google только на них и держится»

Ваш руководитель не DS

- И о DS он слышал только по радио.

Признаки: некорректное использование терминов (нужны знания machine learning, python и crowdfunding), общие слова без углубления в специфику

+	-
Свобода творчества	Очень много времени придется потратить на ликбез руководителя
	Руководитель точно будет разочарован (и, возможно, даже вами, а не DS)
	До DS вы дойдете не так скоро - данных мало и / или они не готовы к DS

У вас нет конкретных задач

- Разберемся по мере работы!

Признаки: «отвечать за ds направление», «участвовать в проектах, максимизирующих выручку»

+	-
Свобода творчества	Придумывать задачи самому – прекрасная идея, если вы умеете их продавать и продавливать.
	Много сил уйдет на презентации и уговоры
	Скорее всего, год-полтора вам придется разворачивать инфраструктуру под придуманные вами задачи

Никаких вводных тестов*

- (что, с одной стороны, не может не радовать).

Признаки: вас спросили только об успехах на предыдущем месте работы

+	-
Меньше стресса	А кто составит вашу великолепную команду? Не окажется ли там знатный выдумщик (конечно, речь не о вас)

* да, это самый спорный пункт

У вас нет команды

- Вы DS Generalist, немного DE , разработчик и PM

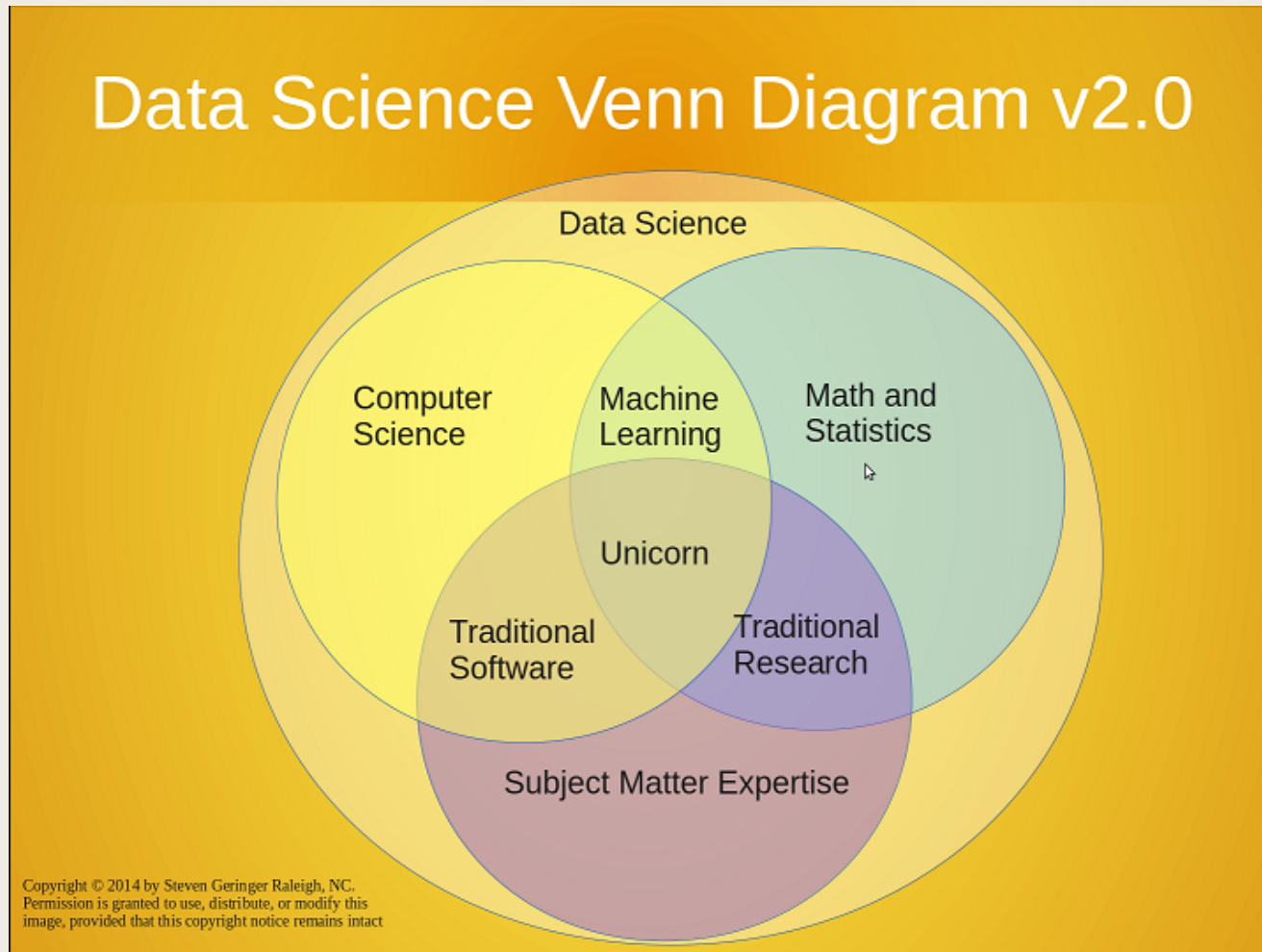
Признаки: спектр ваших задач покрывает все от выбора инфраструктуры, до построения моделей и вывода их в прод.

+	-
Неплохо быть универсальным бойцом	Не специализируясь на чем-то конкретном, можно быстро потерять более глубокие навыки
	Главное – не перегореть от той части задач, которая вам не по душе

А чему бы вы хотели посвящать 80% своего времени?

- Определитесь, какая сторона DS вам нравится больше:
 - *Вы любите простые модели на БОЛЬШИХ данных?*
 - *Любите составлять сложные модели на простых датасетах?*
 - *Работать с картинками? Текстом? Временными рядами? Бизнес-фичами?*
 - *А может, вы любите настраивать инфраструктуру для хранения данных?*
 - *Или выводить модели в продакшн?*

Самое время для диаграммы



Выбираем специализацию

кто	что	технологии	пример
Data engineer	Управление pipeline данных	Apache Spark, Hadoop, Hive, Kafka, SQL.	Очистка датасета без загрузки в память
Data analyst	Получение инсайтов из данных на границе с бизнесом	Python, SQL, Tableau и даже Excel.	Как объяснить падение оборотов? Что заставляет пользователей возвращаться?
Data scientist	Исследование датасетов, предсказание бизнес-величин	Python (scikit-learn, Pandas, PyTorch), SQL, Spark	Как построить модель, recommending пользователю релевантный товар
Machine learning engineer	Построение, оптимизация и деплой моделей в продакшн	Python, Javascript	Как имплементировать модель в java script app Как уменьшить время выполнения модели
Machine learning researcher	Поиск новых путей решения сложных проблем	Python (TensorFlow, PyTorch)	Как я могу использовать тут state-of-the-art методы?