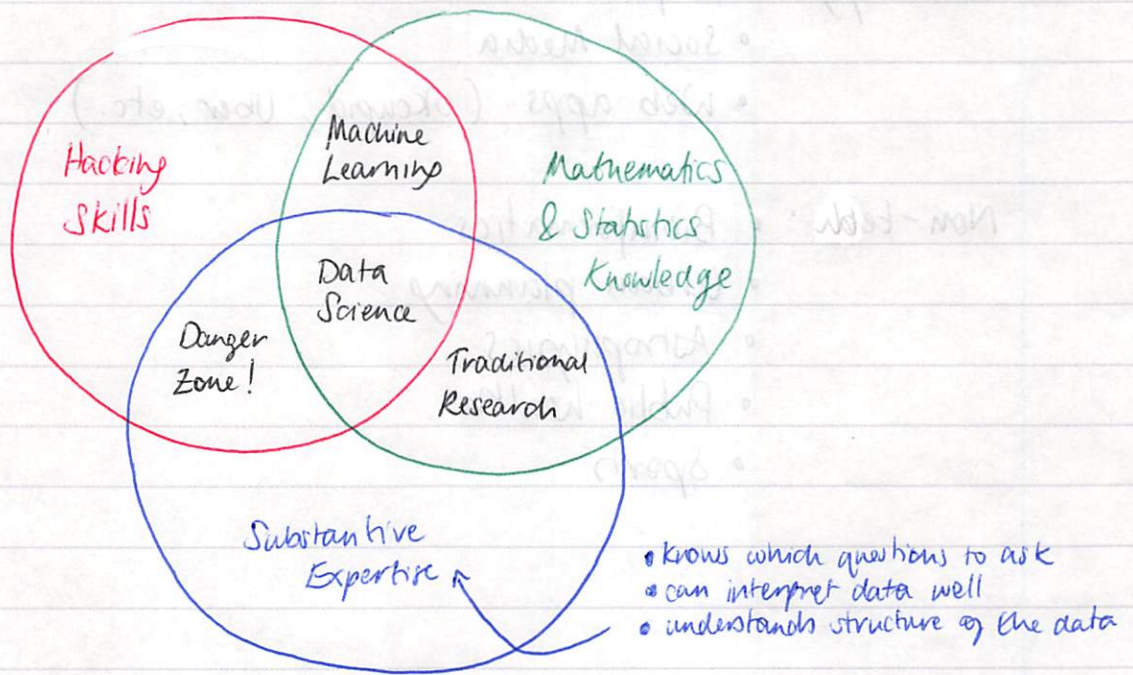


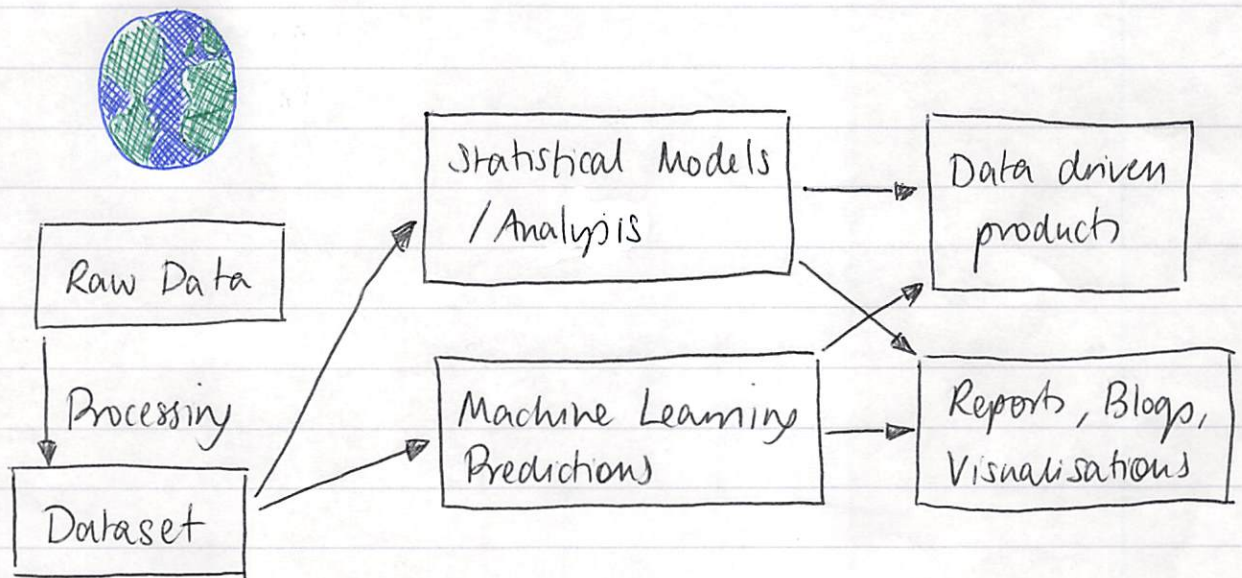
Introductory Material: Data Science

* What is a Data Scientist?

Person who is better at statistics than a software engineer and better at software engineering than a statistician



* What does a Data Scientist do?



* Simpson's Paradox

- See Berkeley admissions data
- vudlab.com/simpson

• Problems solved by Data Science

Technology :

- Netflix
- Social Media
- Web apps (OkCupid, Uber, etc.)

Non-tech :

- Bioinformatics
- Urban planning
- Astrophysics
- Public health
- Sports