

Odisee  
DE CO-HOGESCHOOL

# Data Science lifecycle



Jens Baetens



# Data science – Wat is het?



# Hoe verloopt een data science project?



## Data Science voorbeeld – verwerken van de verzamelde data

- ▣ Spoor fouten op: Verkeerde ingave, labelling, ruis, ...
- ▣ Zorg dat de privacy in de data gegarandeerd is
- ▣ Reduceer het volume: verwijder onnodige data/duplicaten/ ...
- ▣ Zoek naar reeds bestaande patronen
  - ▢ ML is een zeer krachtige tool om patronen te zoeken



## Data Science voorbeeld – model bouwen op basis van de beschikbare data

- ▣ Kies 1 of meerdere ML-technieken
- ▣ Train met de verzamelde data
- ▣ Valideer de resultaten en vergelijk verscheidene technieken

## Data Science voorbeeld – Communiceer over je resultaten

- ▣ Hoe goed werkt het model?
- ▣ Wat kunnen we eruit leren?
- ▣ Doe dit op maat van je doelpubliek!

Landen	Odds*
Engeland	7.50
Belgie	8.00
Frankrijk	4.50
Nederland	15.00
Spanje	8.50
Duitsland	11.00
Portugal	8.50
Italie	7.50



## Wat is Data Science?

- ▣ Data Science is het vermogen om de juiste data
  - te selecteren
  - te begrijpen
  - te verwerken
  - te visualiseren

met als doel een bepaalde waarde uit de data te halen



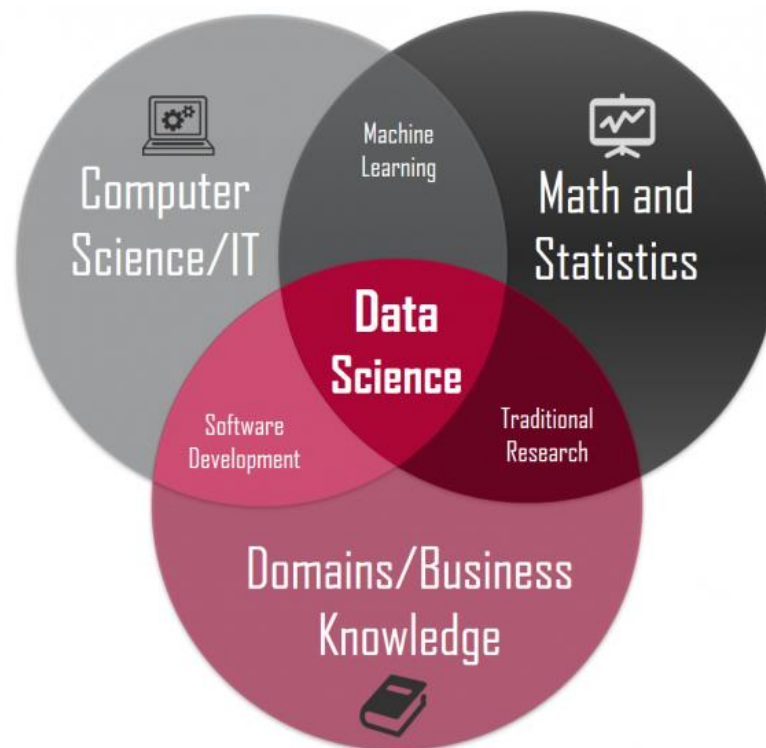
**2** MILLION+  
VIEWS

# WHAT IS DATA SCIENCE?

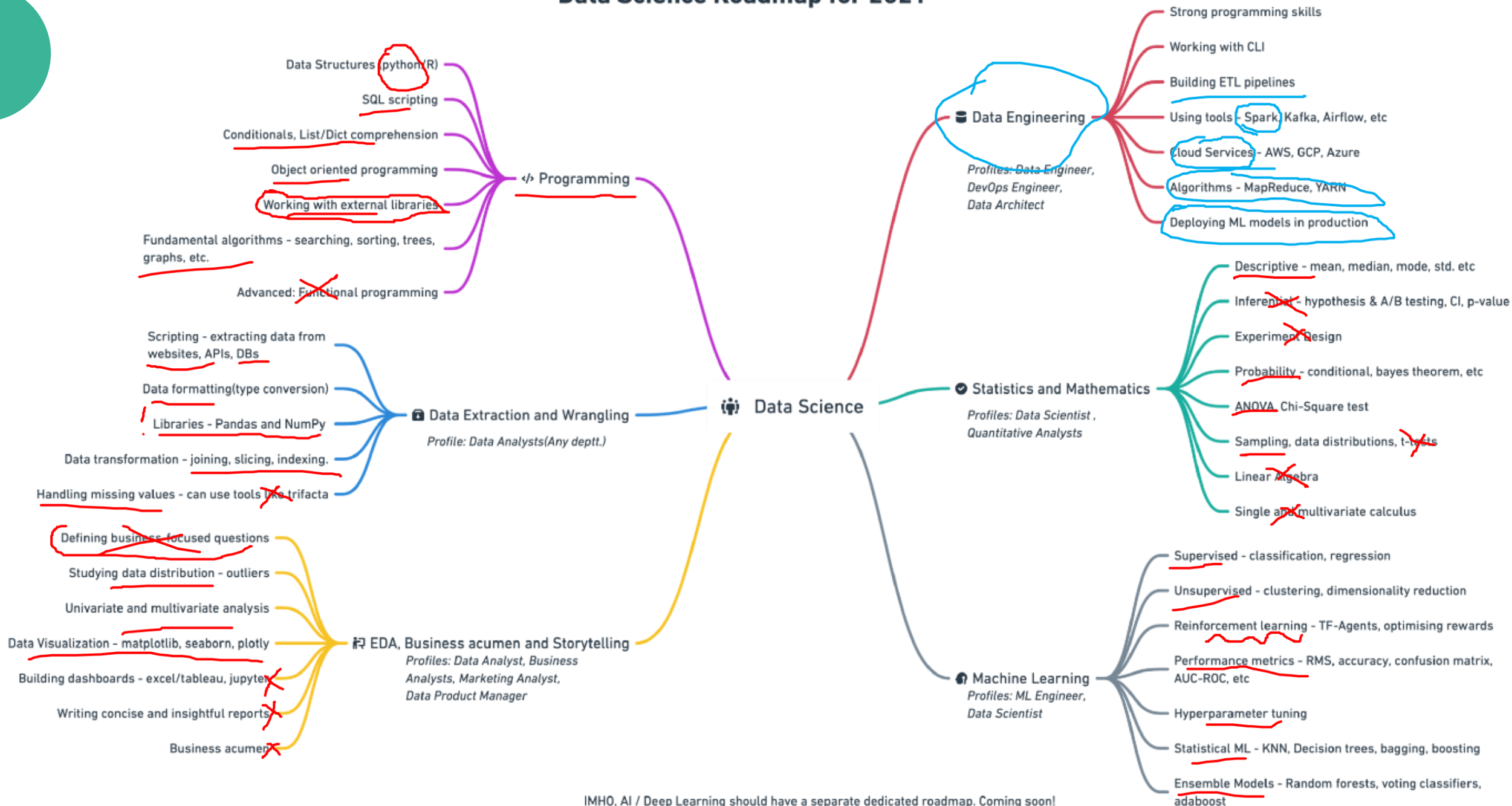


## Wat zijn de benodigde vaardigheden?

- ▣ Combinatie van IT, statistiek en domein kennis



# Data Science Roadmap for 2021



IMHO, AI / Deep Learning should have a separate dedicated roadmap. Coming soon!