

Data Scientist:

Simple ML pipeline
Feature Engineering
Feature Selection
Model metrics
Basic Algos

} Beginner

Hyperparameter tuning
Applied ML for Business
Algorithm Internals
Model Interpretability
Automating pipelines
Read research

} Intermediate

Research open source
Understands trade offs
production ML
ML experiment

} Advanced

Embedding ML into existing processes

Skills:-

Lit hub !!!

Base



• (R / python) + SQL

• Basic maths

↳ calculus

↳ matrix algebra

• Basic Statistics

↳ probability distributions

↳ hypothesis testing

} only need enough to understand papers

} experiment design

Applied -

Common packages
End-to-End ML project

Bonus?

• Docker

• Dash boards

• Blog / Tutorials

• Deep Learning

• Cloud Computing

↳ you should write

Different Types of Data Scientists:

A - Analyst - Insights, experiments, dashboards, exploration

B - Builder - software, Auto ML, cloud computing, packages

C - Consultant - project management, embedded solutions, sales

Uni-Corn
Data Scientist

