## Crush Your DS/ML Interviews: A Student's Roadmap

This handout is your 10-step guide to preparing for Data Science and Machine Learning interviews, with top resources hand-picked to help you learn, review, and practice efficiently. Start at the left, and work your way right or pick based on what you need most!

ML Pipeline (scikit-learn) If you're new to interviews, begin with Python Basics and SQL. Then follow the sections in order. DataCamp Scikit-learn Cheat Sheet START HERE **Pandas Wrangling** 6 a) DataCamp Pandas Cheat Sheet ML System Design b) Kaggle Pandas Cheatsheet Notebook 1/2 ML System Design Interview Cheat Sheet **Python** a) DataCamp Python Cheat Sheet b) UTC "Python for Data Science" PDF SQL a) SQL Interview Cheat Sheet b) Mode SQL Tutorial Interactive Lessons Statistics & Probability a) MIT Statistics Cheat Sheet **Evaluation Metrics** b) KDnuggets' 21 Data-Science Cheat Sheets (Stats section) Stanford CS229 ML Metrics Cheat Sheet (accuracy, F1, AUC, R<sup>2</sup>, etc.) **EDA & Visualization** a) Datasans EDA Cheatsheet b) DataCamp Seaborn Cheat Sheet





Behavioural (STAR)

uConnect STAR Method Interview Cheat Sheet

## **Mock Interview Reps**

Pramp / Exponent Data-Science Interview platform (live peer practice)

## **Quick Tips & Common Pitfalls**

- Practice solving problems out loud to simulate real interview conditions.
- Don't just read solutions but write and test your code from scratch.
- Focus on high-frequency SQL topics: JOIN, GROUP BY, WINDOW FUNCTIONS.
- Go through Leetcode's SQL 50
- Know how to compare models using metrics (Accuracy vs. F1,  $R^2$  vs. Adjusted  $R^2$ ).
- Prepare 2-3 STAR stories that show ownership, problemsolving, and teamwork.
- Be ready to explain trade-offs in model choice, feature selection, and data cleaning.
- Review a few end-to-end ML projects (Your own or others) as this helps answer system design questions.