

Data Science Bootcamp: Core

Mentor Course Handbook

CONFIDENTIAL

This document is for your eyes only. Please do not share with anyone else!

Table of Contents

Introduction	**2**
Course Overview	**3**
Course Projects	**4**
Video Call Sessions	**5**
Unit Overviews	**6**
Unit 1: Program Overview	**6**
Unit 2: What is Data Science	**7**
Unit 3: Problem Identification	**8**
Unit 4: The Python Data Science Stack	**8**
Unit 5: Creating Your Job Search Strategy	**9**
Unit 6: Applying the Data Science Method	**10**
Unit 7: Data Wrangling	**11**
Unit 8: SQL and Databases	**12**

Unit 9: Your Elevator Pitch and LinkedIn Profile 13

Unit 10: Statistics in EDA 14

Unit 11: Python Statistics in EDA 15

Unit 12: Effective Networking: Building Your Network 16

Unit 13: Machine Learning Overview 17

Unit 14: Supervised Learning 18

Unit 15: Unsupervised Learning 19

Unit 16: Feature Engineering 20

Unit 17: Informational Interviews 21

Unit 18: Machine Learning Applications 22

Unit 19: Finding the Right Job Titles and Companies 23

Unit 20: Data Storytelling 24

Specialization Tracks 25

Introduction

Thank you for mentoring students in the Data Science Bootcamp: Core. The insight you provide is central to each student's learning. This guide is designed to make it easier for you to access the information you need to create a smoother mentoring experience.

This guide provides:

- **A High-Level Overview of the Course**
- **Unit Overviews** — This is a description of each unit.
    - **Learning Objectives** — These are high-level objectives that indicate the goals of each unit.
    - **Coursework** — These are the activities that your student will complete while working on each unit.

- **Common Concerns** — This section describes common concerns students may have related to the topics covered in the unit.
- **Industry Expert Insight** — These are suggestions for ways you can share your industry experiences with your mentee to help them succeed.

Course Overview

Familiarizing Yourself with the Course
Before the course starts, please spend some time learning the ins and outs of the course in order to be prepared to address your mentee's concerns or questions. To prepare, please:

- Read through this handbook
- Read through all of the project instructions to familiarize yourself with what students will be asked to do
- Review each of the units of the course to better understand the topics and learning objectives addressed in each unit

The Course
Our Data Science Bootcamp: Core is designed to set students up to successfully land their first data science job and guides them through hands-on assignments that replicate the work that data scientists conduct on a daily basis. These projects are an invaluable addition to their portfolio and will help them to demonstrate their ability to work end-to-end through a data science project to potential employers. Students apply the skills and concepts they learned throughout the course to two capstone projects. By the end of this course, students will have:

- Learned the ins and outs of data science theory, tools, and skills
- Demonstrated their knowledge of the Data Science Method through course projects
- Created an interview-worthy portfolio to show off to potential employers

The course includes both technical units filled with videos, interactive resources, and projects, as well as an *optional* career curriculum designed to help each student land their dream job.

Prerequisites
Students in this course are expected to have:

- Six months experience using a general-purpose programming language, such as Python, R, Java and C/C++, including conditionals, loops, string manipulation, and writing and calling functions distributed over multiple source files
- Completed a university-level statistics course on probability and descriptive statistics, covering the principles of probability, basic probability distributions, and descriptive statistics, such as mean, median, and standard deviation
- Fluency in English (written and spoken)

Note: If you feel that your mentee doesn't meet the prerequisites, please contact us immediately at mentors@springboard.com.

The Course Learning Objectives
This course aims to:

- Equip students with the technical skills they'll need to build an impressive data science portfolio and launch a successful career as a data scientist.
- Empower students to secure a data science position through a proven job search strategy.

Course Projects

Portfolio Projects
Students will complete two projects that will be included in their portfolio.

Capstone 1 (Guided Capstone)	Your mentee's first capstone project comes up fairly early in the course. For this project, they'll be given a lightweight introduction to each step of the Data Science Method. They'll then be guided through each of those steps with helpful tips and instructions. This first capstone is designed to build their foundational understanding of each of these important steps, while also giving them an opportunity to practice each step before applying their knowledge to their second capstone.
Capstone 2	Your mentee's second capstone project follows the same Data Science Method steps as the first capstone, but this time with less guidance. Each of these project's steps will be their own submission and are interwoven throughout the core units of the course.

Case Studies and Mini-Projects
As your mentee works through the course, they'll encounter some stand-alone hands-on case studies that are designed to help them practice the crucial skills taught in this course. Students will add each of these completed case studies to their GitHub repository. Generally, your mentee will complete a case study or mini-project, get feedback from you — their mentor — and then apply the skills covered in the practice project to work they complete for their capstone projects.

Video Call Sessions (30 minutes)

Weekly calls with mentors are at the heart of our students' learning. The calls are also meant to help you **1)** better understand what your mentees are learning each week, **2)** make sure your mentees are on track, and **3)** provide feedback, encouragement, and inspiration for career advancement.

Before each call, please:

- Read through the **call agenda** your mentee has submitted
- Review any **submitted projects**. These will appear in your mentor dashboard.
- Review the **common concerns students may have about the course projects**
    - *These are listed under the description of the project in each unit overview*
- Review (*if needed*) the **unit overview** of the unit that your student is working on

During each call, be sure to:

- **Greet** your mentee and break the ice by asking about their day or finding common interests or hobbies you both share (1 minute)
- **Discuss the items listed in the call agenda,** including the work your mentee has done that week and any questions or concerns they may have
- **Tie up** the call when it's coming to an end by:
    - Reiterating key information and next steps
    - Ending on a positive and encouraging tone
    - Confirming the next 1-1 video call session

Unit Overviews

Unit 1: Program Overview

This unit introduces the course curriculum, gives an overview of course projects, explains how to navigate the course site, and discusses how to participate in the 1-1 video sessions.

Unit Learning Objectives
In this unit, students will:

- Become familiar with the course site
- Learn about the projects they will complete while working on this course
- Understand the course aims and outcomes
- Prepare for mentor 1-1 video call
- Determine how to pace the course

Coursework

- Schedule their first career coach call

Common Concerns

- Your mentee may need assistance setting up the initial 1-1 mentor video call. Please provide them support by emailing them instructions and offering other options for the initial contact.

Industry Expert Insights

- For the initial 1-1 session, use icebreakers, such as talking about your hobbies, interests, or locations, to make your mentee feel comfortable and to build a rapport.

Unit 2: What is Data Science?

This brief unit is intended to welcome your mentee to the world of data science. They'll learn the ins and outs of the data science world and will become familiar with the ways that data science impacts the world around them. They'll also be introduced to the Data Science Method (DSM), the structured approach they'll use to solve data science problems throughout this course. Let's dive in!

Unit Learning Objectives

- Become familiar with the field of data science, as well as the skills that data scientists rely on for success
- Learn about the Data Science Method, the series of steps that your mentee will follow to complete any data science project

Coursework
In this unit, your mentee will:

- Work through a series of resources that will introduce them to the world of data science

Common Concerns

- Students are still adjusting to the course, so gently remind them to refer back to Unit 1 to review the resources and support network available to them.
- There are many new terms that students are learning, so encourage your mentee to create a list or "cheat sheet" that they can refer back to for easy access.

Industry Expert Insights

- Share your experiences with data science and provide additional resources to help your mentee learn more about the field.
- Share real-world examples of data science in action and explore the ways that a data scientist would have been involved.
- Share your job hunting experience and any insights you can offer about the data science job market. Offer insights on career paths and any experiences of your own or your colleagues that are particularly interesting.

Unit 3: Problem Identification

This unit is all about problem identification, which is the process of defining a problem you want to solve and identifying what you want to accomplish by solving that problem. It also happens to be the first step of the Data Science Method. In this unit, your mentee will learn how to create a problem statement to define a problem and form a hypothesis.

Unit Learning Objectives

- Learn how to identify a problem, the first step in the Data Science Method
- Develop the ability to craft nuanced problem statements

Coursework
In this unit, your mentee will:

- Create problem statements for the Monalco Mining case study
- Complete the Nordic Sensing Co. case study

Common Concerns

- Students feel that defining a problem to solve is not important. Gently remind them that identifying the right problem to solve will save them time and frustration as they near the end of any project.

Industry Expert Insights

- Share your experience creating problem statements of your own and provide any tips or guidance related to problem identification

---

Unit 4: The Python Data Science Stack

The Python data science stack refers to the ecosystem within which Python exists. This unit focuses on two of the major players in this ecosystem: Python and pandas. This unit will equip your mentee with an understanding of all the most essential Python concepts as they relate to data science, including how to use functions and libraries to solve data science problems.

Unit Learning Objectives

- Learn Python data types, foundations, and standard libraries

- Learn pandas techniques and practice manipulating and merging dataframes with pandas
- Be able to write code with minimal hints in Jupyter Notebook

Coursework
In this unit, your mentee will:

- Work through several sets of Python exercises
- Apply their Python skills to the London Housing case study

Common Concerns

- Your mentee should already have a basic understanding of the foundations of Python — if they want to refresh their understanding of this important language, they can refer to the prework at the end of the Program Overview unit.

Industry Expert Insights

- Share any tips regarding the ways you have streamlined your work with Python or pandas

---

Unit 5: Creating Your Job Search Strategy (Optional)

Please note that this unit is optional but we encourage students who are hoping to get a data science job to complete this work.

Unit Overview

One of the main aims of this course is to make your mentee ready for a career as a data scientist. This unit introduces a job search strategy that we recommend and discusses the common mistakes made by job seekers.

⭐ **Please note that the career coach is responsible for assessing your mentee's LinkedIn profile and resume**. While we encourage you to review your mentee's resume and LinkedIn page if they ask, you are not responsible for grading these submissions.

Unit Learning Objectives

- Develop your career strategy to help you successfully land a job when you complete the course
- Appreciate the greater importance of network building and referrals over resumes, and start doing the upfront work to build a network
- Recognize some common misconceptions, like searching for job titles (rather than descriptions), or that good data science jobs are only limited to well-known tech firms
- Learn about how career services work and build rapport with your career coach

Coursework
In this unit, your mentee will:

- Watch videos and presentations related to job search strategies and common mistakes job seekers make

Common Concerns

- Mentees may feel that they don't want to focus on the career units right away. Please remind them that the career units are designed to help them kick off their job search early, which was proven to relate directly to past student success
- Mentees may be skeptical about the idea of applying for jobs via networking and referrals. Some mentees believe that if they're not sending out resumes right away, they won't get a job. Reinforce that networking is how a job search often works in technical fields.
- Mentees may experience "Imposter Syndrome," meaning that mentees often feel that their experience is of no real value. Give them a pep talk and help them see the real-world applications of their work.

Industry Expert Insights

- Provide tips on network building.
- Share examples of how you or your colleagues found data science jobs.
- Discuss the hiring process you went through.
- Provide examples of skills and qualities employers are looking for in a data scientist.

Unit 6: Applying the Data Science Method

In this unit, your mentee will get to see data science in action (and be part of that action)! The data science method is the process they'll rely on to solve any data science problem they come up against. They already learned about problem identification, which is the first step in this

process. Now, they'll work through each of the data science method's steps and get their first taste of an end-to-end data science project.

Unit Learning Objectives

- Gain first-hand experience with the end-to-end data science process you'll use to complete work for this course and throughout your career as a data scientist

Coursework
In this unit, your mentee will:

- Work through each step of the data science method by completing a guided capstone project

Common Concerns

- The guided capstone is the first major project of the course. Students may struggle with the work they need to do — please offer support and guidance.

Industry Expert Insights

- Share real-world examples of times that the steps of the Data Science Method helped you with your work.

Unit 7: Data Wrangling

Data wrangling is a crucial step in the process of preparing data so that your mentee can identify insights, perform EDA, and build models. By wrangling data at the beginning of any data science project, they'll be set up to explore their data efficiently. Your mentee was introduced to data wrangling concepts during the guided capstone. This unit takes a deeper dive into each of these concepts and introduces new scenarios and constraints that are important for them to understand.

Unit Learning Objectives

- Learn how to acquire and compile data through data collection
- Develop the skills to keep projects organized by building local file structures and keeping GitHub repo versions updated
- Learn how to review data types, build data profiles, and begin to develop an understanding of the feature dimensions of your data

- Be able to handle problems in your data such as null values, duplicates, and formatting issues

Coursework
In this unit, your mentee will:

- Work through three mini-projects:
    - An API mini-project
    - A NASA exercise
    - A text data handling exercise
- Complete the first three steps of their second capstone project

Common Concerns

- Students will reach out to you for feedback about their project ideas and proposal. Please review their submissions before your calls with your mentee.
- Boredom: This unit dives into the weeds of data wrangling, which might begin to seem tedious for some mentees. You can help your mentee by reminding them that data wrangling is a critical part of the data science process, and is essential to ensure that all further analysis is on a solid foundation.

Industry Expert Insights

- Share the importance of data wrangling and how it's a critical skill for a data scientist.
- Emphasize the amount of time and effort that data wrangling and cleaning can take in real-world projects.
- Give examples of common pitfalls during data wrangling and how to address them.

Unit 8: SQL and Databases

In this unit, your mentee will learn how to leverage Structured Query Language (SQL) to query relational database management systems. In other words, they'll use queries to understand the data contained in databases.

Unit Learning Objectives

- Explore the landscape of SQL and databases and gain an understanding of how you will use them in your career as a data scientist
- Write best practice queries in SQL and be able to work with relational databases in Python

Coursework
In this unit, your mentee will:

- Practice writing SQL to solve business questions related to a country club case study

Common Concerns

- Your mentees may have questions about databases, as they are a rather complicated topic.

Industry Expert Insights

- Share your own experiences with SQL and offer any tips or advice for how to use this tool efficiently.

Unit 9: Your Elevator Pitch and LinkedIn Profile (Optional)

Please note that this unit is optional but we encourage students who are hoping to get a data science job to complete this work.

Now that your mentee has identified the first steps of their career and job search strategy, it's important that they update their LinkedIn profile and begin crafting their elevator pitch. Both of these items will come in handy when they begin building relationships with professionals who can help them get hired.

⭐ Please note that the career coach is responsible for assessing your mentee's LinkedIn profile and resume. While we encourage you to review your mentee's resume and LinkedIn page if they ask, you are not responsible for grading these submissions.

Unit Learning Objectives

- Create a personal pitch and story to share with potential employers
- Repurpose your LinkedIn profile to reflect your journey and transition into data science and learn how to highlight your transferable skills

Coursework

In this unit, your mentee will:

- Create a personal pitch
- Revise their LinkedIn profile
- Book and have your coaching call to review your LinkedIn profile

Common Concerns

- Your mentee may not be sure what they should include on their LinkedIn page. Offer advice based on your own experiences with LinkedIn and job searching.
- Your mentee may feel uncomfortable “pitching” themselves. Please remind them that an elevator pitch is a great way to convey what they do and what they’ve achieved.

Industry Expert Insights

- If your mentee is nervous about sharing their elevator pitch, share the information you generally tell people when they ask what you do (if you are comfortable doing so.)

Unit 10: Statistics in EDA

Statistics lie at the heart of data science. Having a solid understanding of statistical concepts will allow your mentee to draw useful, nuanced conclusions from their data and will help them stand out from other candidates as they apply for data science jobs. In this unit, they’ll learn how to use statistics to draw insights from datasets. This unit was built around David Spiegelhalter's book, *The Art of Statistics* (aka *AoS*). In each subunit, they’ll read one or two chapters, test their understanding with a quiz, and review take-away notes.

Unit Learning Objectives

- Learn the fundamentals of statistics and develop an understanding of statistical concepts
- Become equipped with essential conceptual knowledge before diving into application statistics
- Understand that statistics is both a science and an art

Coursework

In this unit, your mentee will:

- Complete a frequentist inference exercise
- Work on an exercise involving visualizing basic probability and compound
- Work on a probability distribution exercise

- Complete several quizzes

Common Concerns

- Some mentees have a hard time seeing the importance of inferential statistics and may ask if it's useful and when it's used in practice. Explain that it's often used in the real world as a quick way to draw conclusions about a dataset and make decisions around it.

Industry Expert Insights

- Provide practical applications of inferential statistics from your professional experiences.
- Give tips and tricks for using different kinds of statistical methods.

Unit 11: Python Statistics in EDA

In this unit, your mentee will learn about how to apply Python to the sound statistical knowledge they acquired in the previous unit. In the first subunit, they'll look at statistical inference in Python. The second subunit focuses on data visualization in Python. Subunit three is dedicated to learning more about how to implement statistical hypothesis testing in Python. Finally, they'll look at some slightly more complex applications of statistics in Python, such as the statistical modeling methods of linear regression and classification.

Unit Learning Objectives

- Transfer statistical concepts into practical skills and learn how to apply Python to statistics in EDA
- Take a deep dive into statistical inference, hypothesis testing, and statistical modeling in Python through hands-on practices
- Incorporate learning from Data Visualization in Python and apply the techniques on your capstone

Coursework

In this unit, your mentee will:

- Complete a case study about frequentist inference
- Implement a hypothesis test via a permutation test in the Integrating Apps case study
- Work on a statistical modeling case study (and have the opportunity to get more practice via a second, optional case study)
- Complete the EDA step of their second capstone project

Common Concerns

- It's important to keep reminding mentees about the limitations of various inferential statistics techniques, the assumptions they make about the data, and how it's important to have a large enough dataset.
- Many mentees struggle with understanding the concept of statistical significance and p-values.
- It's worth discussing the concepts of "causation" and "correlation," and how to properly communicate statistical results to decision-makers without trying to suggest causation, and without trying to suggest decisions based on statistical results when the data scientist has no experience in the application field.

Industry Expert Insights

- Tell them about the ways that you use statistics to perform exploratory data analysis.

Unit 12: Effective Networking (Optional)

Please note that this unit is optional but we encourage students who are hoping to get a data science job to complete this work.

Effective networking is a cornerstone of this Data Science Bootcamp: Core course, as building a solid network in the data science field upfront often comes from obtaining referrals from target companies. The unit leads your mentee through their first attempt at building a data science network.

Unit Learning Objectives

- Learn about best practices for email and in-person communications
- Build your design network by using your first meetup as a way to build connections with people who share similar interests
- Learn how "cold emailing" can be used to reach out to people at your dream companies
- Learn about imposter syndrome, a common feeling that many professionals face, and share ways to overcome it to grow your confidence in networking

Coursework

In this unit, your mentee will:

- Attend a data science meetup or event and submit a brief summary of the meetup that they attended
- Book and have their coaching call to gain more insight into how to develop their networking strategy

Common Concerns

- Introverted and shy mentees: Many mentees are not used to the concept of networking and are nervous about going out and meeting strangers. This is particularly true of those who are naturally introverted or shy.
- Severe “imposter syndrome”: Some mentees are convinced that because they are new to the field of data science, they have nothing of value to offer experts. “Why would any experienced designer talk to me?” Help them understand that relationship building starts with listening and being genuinely curious about other people. Most people, including experienced data scientists, are excited to talk about themselves and their work, which can provide a lot of information.
- First time networking jitters: Some mentees have never been to or even heard of a meetup before and are nervous about going to one for the first time. Explain to them that meetups are very casual events, and there’s absolutely no pressure there. At this point, they’re still new and may not understand all the technical content people are talking about. That’s completely fine. It’s still useful to be curious, ask people about what they do, and listen well.

Industry Expert Insights

- Share how networking has helped you (or your colleagues/friends) with finding new opportunities.
- Share your experience going to or speaking at meetups and similar events.
- Give tips and tricks for building relationships with data science professionals.

Unit 13: Machine Learning Overview

Machine learning (ML) algorithms enable computers to learn from data, and even improve themselves, without being explicitly programmed. This unit takes a thorough look at the different facets of this practice and lays the foundation that your mentee will build on as you work through subsequent units about supervised learning, unsupervised learning, and advanced ML practices. In addition to learning about the fundamentals of machine learning, they’ll also become familiar with different ML use-cases and explore the Python libraries and packages most relevant to machine learning work. Let's get started!

Unit Learning Objectives

- Explore the fundamentals of machine learning and gain an understanding of the taxonomy of different types of ML algorithms
- Develop an understanding of best practices and common challenges that data scientists deal with when working on machine learning applications

Coursework

In this unit, your mentee will:

- Work through a series of introductory resources. This unit is intended to introduce your mentee to machine learning fundamentals.
- Complete a quiz to test their understanding of the topics introduced in the unit.

Common Concerns

- Mentees often wonder about the pros and cons of different machine learning techniques, when to apply which one, and how to pick one. Many mentees imagine that there's a straightforward method to decide which algorithm to apply and when.

Industry Expert Insights

- Provide practical examples of machine learning applications. Explain the problem that was solved and which algorithm was chosen and why.

Unit 14: Supervised Learning

Supervised Learning is the bread and butter of machine learning. In this unit, your mentee will learn to use supervised learning when they give their machine labeled training data and encode procedures for the machine to learn to assign those labels itself.

Unit Learning Objectives

- Develop an understanding of supervised learning and its common applications
- Be able to perform regression and classification techniques to solve real-world problems

Coursework

In this unit, your mentee will:

- Complete a logistic regression case study

- Complete a Decision Trees case study
- Complete a Random Forest case study
- Complete a Gradient Boosting case study
- Complete a Time Series case study

Common Concerns

- The topics covered in this and the next unit are pretty complicated. Offer guidance and advice if your mentee seems to be struggling with a particular section of this unit.

Industry Expert Insights

- Share the pros and cons of different algorithms with practical examples.
- Give tips and tricks for improving the performance of machine learning algorithms.

Unit 15: Unsupervised Learning

Unsupervised learning, as the name implies, requires a minimum amount of human supervision. It is a type of machine learning where the machine looks for patterns in a dataset with no pre-existing labels or classifications — the machine acts on this information without guidance. Typified by the clustering method, unsupervised learning is an efficient way to spot patterns in your data when you don't know in advance what you're looking for — a situation that is increasingly common as the quantity of our data increases. Clustering is a technique that involves grouping together different data points. Unsupervised learning is a great way to do exploratory data analysis and dimensionality reduction.

In this unit, your mentee will learn about clustering types, distance calculations and similarity measures, and principal component analysis and SVD.

Unit Learning Objectives

- Understand what unsupervised learning and how it is used
- Develop knowledge of common clustering types
- Be able to perform clustering techniques to solve real-world problems

Coursework

In this unit, your mentee will:

- Complete a distance metrics exercise
- Complete a Cosine Similarity exercise

- Complete a case study about customer segmentation using K-means clustering

Common Concerns

- The topics covered in this unit are pretty complicated. Offer guidance and advice if your mentee seems to be struggling with a particular section of this unit.

Industry Expert Insights

- Share the pros and cons of different algorithms with practical examples.
- Give tips and tricks for improving the performance of machine learning algorithms.

Unit 16: Feature Engineering

Not all data comes in a neat, tidy, numerical format. In fact, most data doesn't. To do proper machine learning, you almost always need to do some feature engineering; that is, you need to convert the data you've compiled to solve your problem into a feature matrix. This unit will be divided into two subunits. In the first, your mentee will learn how to deal with categorical features, image features, and text features. In the second, your mentee will dive into derived or secondary features, handling missing data, and automated feature engineering.

Unit Learning Objectives

- Be able to perform data transformation for categorical features, image features, and text features
- Learn best practice for deriving features, handling missing data, and automated feature engineering

Coursework

In this unit, your mentee will:

- Complete a series of interactive exercises
- Apply feature engineering techniques to step four of their second capstone: pre-processing and training data development

Common Concerns

- Your mentee may struggle to perform the pre-processing step or create a training dataset. Remind them to review the resources found in the Applying the Data Science Method unit and offer your own tips and advice.

Industry Expert Insights

- Share tales of your own feature engineering experiences.

Unit 17: Informational Interviews (Optional)

Please note that this unit is optional but we encourage students who are hoping to get a data science job to complete this work.

Informational interviews are one of the best ways to quickly gather in-depth information about a company (beyond what's available through internet searches.) This unit delves into improving social skills, participating in an informational interview, and developing long-lasting professional relationships.

Unit Learning Objectives

- Build your social skills to be able to confidently participate in an informational interview
- Develop the ability to listen carefully and ask the right questions during an informational interview
- Appreciate the importance of developing long-lasting professional relationships

Coursework

In this unit, your mentee will:

- Conduct informational interviews with data science professionals and submit a brief summary of their experience
- Book and have their coaching call to support them in developing their networking strategy

Common Concerns

- Some mentees have a hard time setting up informational interviews, so it's important to reassure them that it's normal and to continue reaching out to more people or help connect them with data science professionals.

- Some mentees are scared of rejection and are nervous to reach out to people for fear of appearing too forward or "weird." Reassure them that there's nothing wrong with a professional request for information.

Industry Expert Insights

- Share a story about an informational interview you've had. Talk about the takeaways from that story.
- Discuss how an informational interview helped you in your career, or how you ended up helping someone because they asked you for an informational interview.

Unit 18: Machine Learning Applications

This unit is all about modeling metrics and hyperparameter tuning. It's crucial that your mentee develop an ability to determine which modeling metric to use based on the type of response variable they have and the business problem they're trying to solve. Likewise, it's also essential to be able to communicate their model's performance in terms that a stakeholder can understand. The second important topic that is covered in this unit is model hyperparameter tuning. A hyperparameter is a parameter that is set before the machine learning process starts; it is a value that controls the learning process.

Learning Objectives

- Take a deep dive into the types of evaluation metrics for regression and classification
- Be able to choose the best evaluation metric for your machine learning project
- Learn best practice for model optimization and be able to apply the techniques to the capstone modeling step

Coursework

In this unit, your mentee will:

- Work through several interactive exercises focused on:
    - Model evaluation metrics
    - Grid search in KNN
    - Bayesian optimization
- Complete the modeling step of their second capstone project

Common Concerns

- Your mentee may need help or guidance to complete the modeling work required for their second capstone project. Be prepared to discuss their work during your calls.

Industry Expert Insights

- Share your experiences building models and relate your experiences to the work that your mentee is doing.

Unit 19: Finding the Right Jobs and Companies (Optional)

Please note that this unit is optional but we encourage students who are hoping to get a data science job to complete this work.

A dream job is the intersection of two phenomena: the right role at the right company. There are many job titles and a diverse array of companies in the data science domain. In this unit, students will research roles within the data science field that might be a good fit and identify industries and companies that offer those roles.

Unit Learning Objectives

- Investigate the different data science roles and titles
- Assess which role is the right fit for them
- Identify the companies and industries that offer their preferred data science role
- Recognize that there are numerous amazing companies doing innovative user experience work, not just a few well-known companies.

Coursework

In this unit, your mentee will:

- A brief summary describing at least 2-3 job titles that are right for them
- A list of 40-50 companies they'd like to join
- Book and have a coaching call to review the company list and job titles list

Mentee Challenges

- Mentees might struggle with how to know if a company would be a good fit for them.

Industry Expert Insights

- Share your own experiences looking for companies that you considered applying to

---

Unit 20: Data Storytelling

Being able to tell a compelling story about your data is an integral skill every data scientist must master. Data scientists need to tell stories with data when building data products and when communicating with employers or clients about those products. Crafting an effective story depends on a few crucial variables:

- Who your audience is
- What questions you're trying to answer
- Why your audience should care
- What your major insights are
- What impact you want to have by sharing your findings

Unit Learning Objectives

- Become familiar with different presentation techniques
- Learn how to communicate your insights in a compelling manner
- Learn how to apply presentation techniques for executive (C-suite), technical, and non-technical audiences

Coursework

In this unit, your mentee will:

- Prepare a presentation about a dataset of their choosing
- Finalize the documentation of their second capstone project

Common Concerns

- Communicating to a general audience: Many mentees, especially those with technical backgrounds, struggle to communicate at a level that’s appropriate for a more general audience. As a mentor, you may need to guide them and help them explain their ideas to a non-technical client.

Industry Expert Insights

- Emphasize the importance of data storytelling and communication skills for a data scientist. Many mentees come in with the intention of mastering the technical skills, but they underestimate the need for learning effective communication skills.
- Share examples of how you've developed and used communication and storytelling skills at work.

Specialization Tracks

The Generalist Track

The Business Insider Track

The Advanced Machine Learning Track