3 data science projects that got me 12 interviews

3 Data Science Projects That Got Me 12 Interviews

By Anya Sharma, MS in Data Science, University of California, Berkeley

Published by Data Science Digest, a leading online publication for aspiring and experienced data scientists.

Edited by Dr. Emily Carter, PhD in Statistics, Harvard University.

Introduction:

Landing a data science job is competitive. But what if I told you that three meticulously crafted projects landed me twelve interviews? That's exactly what happened. This article details the journey behind "3 data science projects that got me 12 interviews," providing a behind-the-scenes look at the project selection, development, and presentation that ultimately led to my success. This isn't about luck; it's about strategic project selection and compelling storytelling. This article will help you understand how to create projects that stand out, demonstrating your skills to potential employers and showcasing your passion for data science.

1. Project 1: Predicting Customer Churn for a Telecom Company (The Classic)

My first project focused on a classic data science problem: customer churn prediction. I chose this because it's a universally understood challenge, relevant across many industries, and allows for the demonstration of a wide range of skills. I didn't just download a readily available dataset; instead, I scraped publicly available data from a fictional telecom company's website (using ethical and legal methods, of course!). This added a layer of complexity and showcased my data acquisition skills, a crucial aspect often overlooked in simpler projects.

The data cleaning process was far from straightforward. I encountered missing values, inconsistent data formats, and outliers that needed careful handling. My meticulous approach to data cleaning, documented thoroughly in my Jupyter Notebook, became a key selling point. I used various imputation techniques and explained my rationale behind each choice. This demonstrated not only my technical skills but also my analytical thinking and attention to detail.

For modeling, I employed several algorithms, including Logistic Regression, Random Forest, and Gradient Boosting. I didn't just compare their accuracy; I analyzed their performance metrics like precision, recall, and F1-score, considering the business implications of false positives and false negatives. This showcased my understanding of model evaluation and my ability to choose the most appropriate model for the specific problem. Finally, I created a compelling dashboard using Tableau, visualizing the key findings and presenting actionable insights. This visually appealing presentation of complex data helped make my project stand out. This project, a prime example of the power behind "3 data science projects that got me 12 interviews", played a crucial role in securing several interviews.

2. Project 2: Analyzing Social Media Sentiment Towards a Specific Brand (The Current Event)

My second project was more contemporary and relevant to current events. I chose to analyze social media sentiment towards a specific brand, leveraging the power of Natural Language Processing (NLP). This project allowed me to demonstrate my ability to work with unstructured data, a skill highly valued in the industry. I

scraped tweets related to a particular brand, cleaned the data, and used NLP techniques like sentiment analysis to gauge public opinion.

This project wasn't just about running pre-built models. I experimented with different NLP libraries, compared their performance, and fine-tuned the models to improve accuracy. This demonstrated my proactive approach to problem-solving and my desire to push beyond the basics. I also incorporated topic modeling to identify the key themes driving positive and negative sentiment, providing a deeper understanding of the brand's public perception. Again, I presented my findings in an interactive dashboard, making the insights easily digestible and visually appealing. The innovative approach to this classic problem within "3 data science projects that got me 12 interviews" further strengthened my portfolio.

3. Project 3: Developing a Recommendation System for an E-commerce Platform (The Innovative)

My third project involved developing a recommendation system for a fictional e-commerce platform. This project allowed me to showcase my skills in collaborative filtering and content-based filtering, demonstrating my proficiency in a specific area of machine learning with significant industry application. I created a dataset simulating e-commerce transactions, adding a layer of creativity and showcasing my understanding of data generation and manipulation.

This project challenged me to optimize the recommendation system for accuracy and efficiency. I experimented with different algorithms, evaluated their performance using metrics like precision@k and recall@k, and optimized the system for scalability. I even went a step further and considered the ethical implications of recommendation systems, discussing potential biases and ways to mitigate them in my project documentation. This demonstrated my awareness of the broader societal impact of data science, a crucial aspect often overlooked but highly appreciated by employers. The innovative nature of this project in the context of "3 data science projects that got me 12 interviews" proved particularly effective.

Why These Projects Worked:

The success of these "3 data science projects that got me 12 interviews" wasn't accidental. They were carefully chosen and executed, emphasizing the following key elements:

Relevance: Each project addressed real-world problems with clear business implications.
Technical Proficiency: I demonstrated proficiency in various data science techniques, including data cleaning, modeling, visualization, and NLP.
Innovation: I didn't just replicate existing projects; I added my own creative twist and explored different approaches.
Storytelling: I presented my projects not as mere technical exercises, but as compelling narratives, clearly articulating the problem, my approach, my findings, and the insights gained.
Presentation: My Jupyter Notebooks and dashboards were meticulously organized, well-documented, and visually appealing.

Conclusion:

The journey behind "3 data science projects that got me 12 interviews" demonstrates that building a strong portfolio is key to landing a data science job. It's not about quantity, but quality. Focus on selecting relevant projects that allow you to showcase your skills, demonstrate your creativity, and tell a compelling story. Invest time in refining your presentation skills and meticulously documenting your work. Remember, your projects are your best ambassadors, showcasing your capabilities and passion to potential employers. With the right approach, you too can significantly increase your interview opportunities.

FAQs:

1. How long did it take to complete each project? Each project took approximately 2-3 weeks of dedicated work, depending on complexity.

1. What tools and technologies did you use? I primarily used Python, Pandas, Scikit-learn, NLTK, Tableau, and Jupyter Notebooks.

1. Where did you find the datasets? I used a combination of publicly available datasets and created my own synthetic datasets for some projects.

1. Did you use any specific templates for your projects? No, I developed my own project structures, focusing on clarity and ease of understanding.

1. How did you tailor your projects to specific job applications? I adapted my project descriptions and highlighted the relevant skills based on the specific job requirements.

1. What was the most challenging aspect of creating these projects? Data cleaning and feature engineering proved to be the most challenging but rewarding aspects.

1. How important was the visualization of your findings? Visualization was crucial in making my projects more engaging and easier to understand.

1. What advice would you give to someone just starting out with data science projects? Start small, focus on one problem at a time, and gradually increase the complexity.

1. What are the next steps after securing interviews? Practice your technical skills, work on your communication skills, and thoroughly research the companies you're interviewing with.

Related Articles:

1. Building a Killer Data Science Portfolio: Project Ideas and Best Practices: A comprehensive guide on creating impressive data science projects.
2. Top 10 Data Science Projects to Boost Your Resume: A curated list of popular and effective project ideas.

3. Mastering Data Visualization for Data Science Projects: Techniques for creating compelling data visualizations.
4. The Ultimate Guide to Data Cleaning for Data Science Projects: A detailed explanation of data cleaning techniques.
5. From Zero to Hero: A Step-by-Step Guide to your First Data Science Project: A beginner-friendly tutorial on building your first data science project.
6. Data Science Project Showcase: 5 Real-World Examples: A collection of inspiring data science projects from successful professionals.
7. Landing your Dream Data Science Job: The Complete Guide: Advice on the job application process, resume writing, and interview preparation.
8. Data Science Project Ideas for Beginners: 3 Easy Projects to Start With: Three straightforward data science projects for beginners.
9. Advanced Data Science Projects: Challenging Yourself to Master Complex Techniques: Ideas for experienced data scientists looking for challenging projects.

3 data science projects that got me 12 interviews: Data Science in Production Ben Weber, 2020 Putting predictive models into production is one of the most direct ways that data scientists can add value to an organization. By learning how to build and deploy scalable model pipelines, data scientists can own more of the model production process and more rapidly deliver data products. This book provides a hands-on approach to scaling up Python code to work in distributed environments in order to build robust pipelines. Readers will learn how to set up machine learning models as web endpoints, serverless functions, and streaming pipelines using multiple cloud environments. It is intended for analytics practitioners with hands-on experience with Python libraries such as Pandas and scikit-learn, and will focus on scaling up prototype models to production. From startups to trillion dollar companies, data science is playing an important role in helping organizations maximize the value of their data. This book helps data scientists to level up their careers by taking ownership of data products with applied examples that demonstrate how to: Translate models developed on a laptop to scalable deployments in the cloud Develop end-to-end systems that automate data science workflows Own a data product from conception to production The accompanying Jupyter notebooks provide examples of scalable pipelines across multiple cloud environments, tools, and libraries (github.com/bgweber/DS_Production). Book Contents Here are the topics covered by Data Science in Production: Chapter 1: Introduction - This chapter will motivate the use of Python and discuss the discipline of applied data science, present the data sets, models, and cloud environments used throughout the book, and provide an overview of automated feature engineering. Chapter 2: Models as Web Endpoints - This chapter shows how to use web endpoints for consuming data and hosting machine learning models as endpoints using the Flask and Gunicorn libraries. We'll start with scikit-learn models and also set up a deep learning endpoint with Keras. Chapter 3: Models as Serverless Functions - This chapter will build upon the previous chapter and show how to set up model endpoints as serverless functions using AWS Lambda and GCP Cloud Functions. Chapter 4: Containers for Reproducible Models - This chapter will show how to use containers for deploying models with Docker. We'll also explore scaling up with ECS and Kubernetes, and building web applications with Plotly Dash. Chapter 5: Workflow Tools for Model Pipelines - This chapter focuses on scheduling automated workflows using Apache Airflow. We'll set up a model that pulls data from BigQuery, applies a model, and saves the results. Chapter 6: PySpark for Batch Modeling - This chapter will introduce readers to PySpark using the community

edition of Databricks. We'll build a batch model pipeline that pulls data from a data lake, generates features, applies a model, and stores the results to a No SQL database. Chapter 7: Cloud Dataflow for Batch Modeling - This chapter will introduce the core components of Cloud Dataflow and implement a batch model pipeline for reading data from BigQuery, applying an ML model, and saving the results to Cloud Datastore. Chapter 8: Streaming Model Workflows - This chapter will introduce readers to Kafka and PubSub for streaming messages in a cloud environment. After working through this material, readers will learn how to use these message brokers to create streaming model pipelines with PySpark and Dataflow that provide near real-time predictions. Excerpts of these chapters are available on Medium (@bgweber), and a book sample is available on Leanpub.

3 data science projects that got me 12 interviews: *Data Science For Dummies* Lillian Pierson, 2021-08-20 Monetize your company’s data and data science expertise without spending a fortune on hiring independent strategy consultants to help What if there was one simple, clear process for ensuring that all your company’s data science projects achieve a high a return on investment? What if you could validate your ideas for future data science projects, and select the one idea that’s most prime for achieving profitability while also moving your company closer to its business vision? There is. Industry-acclaimed data science consultant, Lillian Pierson, shares her proprietary STAR Framework – A simple, proven process for leading profit-forming data science projects. Not sure what data science is yet? Don’t worry! Parts 1 and 2 of Data Science For Dummies will get all the bases covered for you. And if you’re already a data science expert? Then you really won’t want to miss the data science strategy and data monetization gems that are shared in Part 3 onward throughout this book. Data Science For Dummies demonstrates: The only process you’ll ever need to lead profitable data science projects Secret, reverse-engineered data monetization tactics that no one’s talking about The shocking truth about how simple natural language processing can be How to beat the crowd of data professionals by cultivating your own unique blend of data science expertise Whether you’re new to the data science field or already a decade in, you’re sure to learn something new and incredibly valuable from Data Science For Dummies. Discover how to generate massive business wins from your company’s data by picking up your copy today.

3 data science projects that got me 12 interviews: Build a Career in Data Science Emily Robinson, Jacqueline Nolis, 2020-03-24 Summary You are going to need more than technical knowledge to succeed as a data scientist. Build a Career in Data Science teaches you what school leaves out, from how to land your first job to the lifecycle of a data science project, and even how to become a manager. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the technology What are the keys to a data scientist’s long-term success? Blending your technical know-how with the right “soft skills” turns out to be a central ingredient of a rewarding career. About the book Build a Career in Data Science is your guide to landing your first data science job and developing into a valued senior employee. By following clear and simple instructions, you’ll learn to craft an amazing resume and ace your interviews. In this demanding, rapidly changing field, it can be challenging to keep projects on track, adapt to company needs, and manage tricky stakeholders. You’ll love the insights on how to handle expectations, deal with failures, and plan your career path in the stories from seasoned data scientists included in the book. What's inside Creating a portfolio of data science projects Assessing and negotiating an offer Leaving gracefully and moving up the ladder Interviews with professional data scientists About the reader For readers who want to begin or advance a data science career. About the author Emily Robinson is a data scientist at Warby Parker. Jacqueline Nolis is a data science consultant and mentor. Table of Contents: PART 1 - GETTING STARTED WITH DATA SCIENCE 1. What is data science? 2. Data science companies 3. Getting the skills 4. Building a portfolio PART 2 - FINDING YOUR DATA SCIENCE JOB 5. The search: Identifying the right job for you 6. The application: Résumés and cover letters 7. The interview: What to expect and how to handle it 8. The offer: Knowing what to accept PART 3 - SETTLING INTO DATA SCIENCE 9. The first months on the job 10. Making an effective analysis 11. Deploying a model into production 12.

Working with stakeholders PART 4 - GROWING IN YOUR DATA SCIENCE ROLE 13. When your data science project fails 14. Joining the data science community 15. Leaving your job gracefully 16. Moving up the ladder

3 data science projects that got me 12 interviews: R for Data Science Hadley Wickham, Garrett Grolemund, 2016-12-12 Learn how to use R to turn raw data into insight, knowledge, and understanding. This book introduces you to R, RStudio, and the tidyverse, a collection of R packages designed to work together to make data science fast, fluent, and fun. Suitable for readers with no previous programming experience, R for Data Science is designed to get you doing data science as quickly as possible. Authors Hadley Wickham and Garrett Grolemund guide you through the steps of importing, wrangling, exploring, and modeling your data and communicating the results. You'll get a complete, big-picture understanding of the data science cycle, along with basic tools you need to manage the details. Each section of the book is paired with exercises to help you practice what you've learned along the way. You'll learn how to: Wrangle—transform your datasets into a form convenient for analysis Program—learn powerful R tools for solving data problems with greater clarity and ease Explore—examine your data, generate hypotheses, and quickly test them Model—provide a low-dimensional summary that captures true signals in your dataset Communicate—learn R Markdown for integrating prose, code, and results

3 data science projects that got me 12 interviews: Data Science from Scratch Joel Grus, 2015-04-14 Data science libraries, frameworks, modules, and toolkits are great for doing data science, but they’re also a good way to dive into the discipline without actually understanding data science. In this book, you’ll learn how many of the most fundamental data science tools and algorithms work by implementing them from scratch. If you have an aptitude for mathematics and some programming skills, author Joel Grus will help you get comfortable with the math and statistics at the core of data science, and with hacking skills you need to get started as a data scientist. Today’s messy glut of data holds answers to questions no one’s even thought to ask. This book provides you with the know-how to dig those answers out. Get a crash course in Python Learn the basics of linear algebra, statistics, and probability—and understand how and when they're used in data science Collect, explore, clean, munge, and manipulate data Dive into the fundamentals of machine learning Implement models such as k-nearest Neighbors, Naive Bayes, linear and logistic regression, decision trees, neural networks, and clustering Explore recommender systems, natural language processing, network analysis, MapReduce, and databases

3 data science projects that got me 12 interviews: Dear Data Giorgia Lupi, Stefanie Posavec, 2016-09-13 Equal parts mail art, data visualization, and affectionate correspondence, Dear Data celebrates the infinitesimal, incomplete, imperfect, yet exquisitely human details of life, in the words of Maria Popova (Brain Pickings), who introduces this charming and graphically powerful book. For one year, Giorgia Lupi, an Italian living in New York, and Stefanie Posavec, an American in London, mapped the particulars of their daily lives as a series of hand-drawn postcards they exchanged via mail weekly—small portraits as full of emotion as they are data, both mundane and magical. Dear Data reproduces in pinpoint detail the full year's set of cards, front and back, providing a remarkable portrait of two artists connected by their attention to the details of their lives—including complaints, distractions, phone addictions, physical contact, and desires. These details illuminate the lives of two remarkable young women and also inspire us to map our own lives, including specific suggestions on what data to draw and how. A captivating and unique book for designers, artists, correspondents, friends, and lovers everywhere.

3 data science projects that got me 12 interviews: Interactive Data Visualization for the Web Scott Murray, 2013-03-15 Create and publish your own interactive data visualization projects on the Web, even if you have no experience with either web development or data visualization. It’s easy with this hands-on guide. You’ll start with an overview of data visualization concepts and simple web technologies, and then learn how to use D3, a JavaScript library that lets you express data as visual elements in a web page. Interactive Data Visualization for the Web makes these skills available at an introductory level for designers and visual artists without programming experience,

journalists interested in the emerging data journalism processes, and others keenly interested in visualization and publicly available data sources. Get a practical introduction to data visualization, accessible for beginners Focus on web-based tools that help you publish your creations quickly to a wide audience Learn about interactivity so you can engage users in exploring your data

3 data science projects that got me 12 interviews: Recommendation Engines Michael Schrage, 2020-09-01 How companies like Amazon, Netflix, and Spotify know what you might also like: the history, technology, business, and societal impact of online recommendation engines. Increasingly, our technologies are giving us better, faster, smarter, and more personal advice than our own families and best friends. Amazon already knows what kind of books and household goods you like and is more than eager to recommend more; YouTube and TikTok always have another video lined up to show you; Netflix has crunched the numbers of your viewing habits to suggest whole genres that you would enjoy. In this volume in the MIT Press's Essential Knowledge series, innovation expert Michael Schrage explains the origins, technologies, business applications, and increasing societal impact of recommendation engines, the systems that allow companies worldwide to know what products, services, and experiences you might also like.

3 data science projects that got me 12 interviews: Data Science Projects with Python Stephen Klosterman, 2019-04-30 Gain hands-on experience with industry-standard data analysis and machine learning tools in Python Key FeaturesTackle data science problems by identifying the problem to be solvedIllustrate patterns in data using appropriate visualizationsImplement suitable machine learning algorithms to gain insights from dataBook Description Data Science Projects with Python is designed to give you practical guidance on industry-standard data analysis and machine learning tools, by applying them to realistic data problems. You will learn how to use pandas and Matplotlib to critically examine datasets with summary statistics and graphs, and extract the insights you seek to derive. You will build your knowledge as you prepare data using the scikit-learn package and feed it to machine learning algorithms such as regularized logistic regression and random forest. You'll discover how to tune algorithms to provide the most accurate predictions on new and unseen data. As you progress, you'll gain insights into the working and output of these algorithms, building your understanding of both the predictive capabilities of the models and why they make these predictions. By then end of this book, you will have the necessary skills to confidently use machine learning algorithms to perform detailed data analysis and extract meaningful insights from unstructured data. What you will learnInstall the required packages to set up a data science coding environmentLoad data into a Jupyter notebook running PythonUse Matplotlib to create data visualizationsFit machine learning models using scikit-learnUse lasso and ridge regression to regularize your modelsCompare performance between models to find the best outcomesUse k-fold cross-validation to select model hyperparametersWho this book is for If you are a data analyst, data scientist, or business analyst who wants to get started using Python and machine learning techniques to analyze data and predict outcomes, this book is for you. Basic knowledge of Python and data analytics will help you get the most from this book. Familiarity with mathematical concepts such as algebra and basic statistics will also be useful.

3 data science projects that got me 12 interviews: *The Data Science Design Manual* Steven S. Skiena, 2017-07-01 This engaging and clearly written textbook/reference provides a must-have introduction to the rapidly emerging interdisciplinary field of data science. It focuses on the principles fundamental to becoming a good data scientist and the key skills needed to build systems for collecting, analyzing, and interpreting data. The Data Science Design Manual is a source of practical insights that highlights what really matters in analyzing data, and provides an intuitive understanding of how these core concepts can be used. The book does not emphasize any particular programming language or suite of data-analysis tools, focusing instead on high-level discussion of important design principles. This easy-to-read text ideally serves the needs of undergraduate and early graduate students embarking on an "Introduction to Data Science" course. It reveals how this discipline sits at the intersection of statistics, computer science, and machine learning, with a distinct heft and character of its own. Practitioners in these and related fields will find this book

perfect for self-study as well. Additional learning tools: Contains “War Stories,” offering perspectives on how data science applies in the real world Includes “Homework Problems,” providing a wide range of exercises and projects for self-study Provides a complete set of lecture slides and online video lectures at www.data-manual.com Provides “Take-Home Lessons,” emphasizing the big-picture concepts to learn from each chapter Recommends exciting “Kaggle Challenges” from the online platform Kaggle Highlights “False Starts,” revealing the subtle reasons why certain approaches fail Offers examples taken from the data science television show “The Quant Shop” (www.quant-shop.com)

3 data science projects that got me 12 interviews: Cracking the Data Science Interview Maverick Lin, 2019-12-17 Cracking the Data Science Interview is the first book that attempts to capture the essence of data science in a concise, compact, and clean manner. In a Cracking the Coding Interview style, Cracking the Data Science Interview first introduces the relevant concepts, then presents a series of interview questions to help you solidify your understanding and prepare you for your next interview. Topics include: - Necessary Prerequisites (statistics, probability, linear algebra, and computer science) - 18 Big Ideas in Data Science (such as Occam's Razor, Overfitting, Bias/Variance Tradeoff, Cloud Computing, and Curse of Dimensionality) - Data Wrangling (exploratory data analysis, feature engineering, data cleaning and visualization) - Machine Learning Models (such as k-NN, random forests, boosting, neural networks, k-means clustering, PCA, and more) - Reinforcement Learning (Q-Learning and Deep Q-Learning) - Non-Machine Learning Tools (graph theory, ARIMA, linear programming) - Case Studies (a look at what data science means at companies like Amazon and Uber) Maverick holds a bachelor's degree from the College of Engineering at Cornell University in operations research and information engineering (ORIE) and a minor in computer science. He is the author of the popular Data Science Cheatsheet and Data Engineering Cheatsheet on GCP and has previous experience in data science consulting for a Fortune 500 company focusing on fraud analytics.

3 data science projects that got me 12 interviews: *Practical Data Science Cookbook* Prabhanjan Tattar, Tony Ojeda, Sean Patrick Murphy, Benjamin Bengfort, Abhijit Dasgupta, 2017-06-29 Over 85 recipes to help you complete real-world data science projects in R and Python About This Book Tackle every step in the data science pipeline and use it to acquire, clean, analyze, and visualize your data Get beyond the theory and implement real-world projects in data science using R and Python Easy-to-follow recipes will help you understand and implement the numerical computing concepts Who This Book Is For If you are an aspiring data scientist who wants to learn data science and numerical programming concepts through hands-on, real-world project examples, this is the book for you. Whether you are brand new to data science or you are a seasoned expert, you will benefit from learning about the structure of real-world data science projects and the programming examples in R and Python. What You Will Learn Learn and understand the installation procedure and environment required for R and Python on various platforms Prepare data for analysis by implement various data science concepts such as acquisition, cleaning and munging through R and Python Build a predictive model and an exploratory model Analyze the results of your model and create reports on the acquired data Build various tree-based methods and Build random forest In Detail As increasing amounts of data are generated each year, the need to analyze and create value out of it is more important than ever. Companies that know what to do with their data and how to do it well will have a competitive advantage over companies that don't. Because of this, there will be an increasing demand for people that possess both the analytical and technical abilities to extract valuable insights from data and create valuable solutions that put those insights to use. Starting with the basics, this book covers how to set up your numerical programming environment, introduces you to the data science pipeline, and guides you through several data projects in a step-by-step format. By sequentially working through the steps in each chapter, you will quickly familiarize yourself with the process and learn how to apply it to a variety of situations with examples using the two most popular programming languages for data analysis—R and Python. Style and approach This step-by-step guide to data science is full of hands-on examples of real-world data

science tasks. Each recipe focuses on a particular task involved in the data science pipeline, ranging from readying the dataset to analytics and visualization

3 data science projects that got me 12 interviews: Data Science for Business Foster Provost, Tom Fawcett, 2013-07-27 Written by renowned data science experts Foster Provost and Tom Fawcett, Data Science for Business introduces the fundamental principles of data science, and walks you through the data-analytic thinking necessary for extracting useful knowledge and business value from the data you collect. This guide also helps you understand the many data-mining techniques in use today. Based on an MBA course Provost has taught at New York University over the past ten years, Data Science for Business provides examples of real-world business problems to illustrate these principles. You'll not only learn how to improve communication between business stakeholders and data scientists, but also how participate intelligently in your company's data science projects. You'll also discover how to think data-analytically, and fully appreciate how data science methods can support business decision-making. Understand how data science fits in your organization—and how you can use it for competitive advantage Treat data as a business asset that requires careful investment if you're to gain real value Approach business problems data-analytically, using the data-mining process to gather good data in the most appropriate way Learn general concepts for actually extracting knowledge from data Apply data science principles when interviewing data science job candidates

3 data science projects that got me 12 interviews: Data Science for Undergraduates National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Board on Science Education, Division on Engineering and Physical Sciences, Committee on Applied and Theoretical Statistics, Board on Mathematical Sciences and Analytics, Computer Science and Telecommunications Board, Committee on Envisioning the Data Science Discipline: The Undergraduate Perspective, 2018-11-11 Data science is emerging as a field that is revolutionizing science and industries alike. Work across nearly all domains is becoming more data driven, affecting both the jobs that are available and the skills that are required. As more data and ways of analyzing them become available, more aspects of the economy, society, and daily life will become dependent on data. It is imperative that educators, administrators, and students begin today to consider how to best prepare for and keep pace with this data-driven era of tomorrow. Undergraduate teaching, in particular, offers a critical link in offering more data science exposure to students and expanding the supply of data science talent. Data Science for Undergraduates: Opportunities and Options offers a vision for the emerging discipline of data science at the undergraduate level. This report outlines some considerations and approaches for academic institutions and others in the broader data science communities to help guide the ongoing transformation of this field.

3 data science projects that got me 12 interviews: Data Science Projects with Python Stephen Klosterman, 2021-07-29 Gain hands-on experience of Python programming with industry-standard machine learning techniques using pandas, scikit-learn, and XGBoost Key FeaturesThink critically about data and use it to form and test a hypothesisChoose an appropriate machine learning model and train it on your dataCommunicate data-driven insights with confidence and clarityBook Description If data is the new oil, then machine learning is the drill. As companies gain access to ever-increasing quantities of raw data, the ability to deliver state-of-the-art predictive models that support business decision-making becomes more and more valuable. In this book, you'll work on an end-to-end project based around a realistic data set and split up into bite-sized practical exercises. This creates a case-study approach that simulates the working conditions you'll experience in real-world data science projects. You'll learn how to use key Python packages, including pandas, Matplotlib, and scikit-learn, and master the process of data exploration and data processing, before moving on to fitting, evaluating, and tuning algorithms such as regularized logistic regression and random forest. Now in its second edition, this book will take you through the end-to-end process of exploring data and delivering machine learning models. Updated for 2021, this edition includes brand new content on XGBoost, SHAP values, algorithmic fairness, and the ethical

concerns of deploying a model in the real world. By the end of this data science book, you'll have the skills, understanding, and confidence to build your own machine learning models and gain insights from real data. What you will learnLoad, explore, and process data using the pandas Python packageUse Matplotlib to create compelling data visualizationsImplement predictive machine learning models with scikit-learnUse lasso and ridge regression to reduce model overfittingEvaluate random forest and logistic regression model performanceDeliver business insights by presenting clear, convincing conclusionsWho this book is for Data Science Projects with Python – Second Edition is for anyone who wants to get started with data science and machine learning. If you're keen to advance your career by using data analysis and predictive modeling to generate business insights, then this book is the perfect place to begin. To quickly grasp the concepts covered, it is recommended that you have basic experience of programming with Python or another similar language, and a general interest in statistics.

3 data science projects that got me 12 interviews: *Ask a Manager* Alison Green, 2018-05-01 From the creator of the popular website Ask a Manager and New York's work-advice columnist comes a witty, practical guide to 200 difficult professional conversations—featuring all-new advice! There's a reason Alison Green has been called "the Dear Abby of the work world." Ten years as a workplace-advice columnist have taught her that people avoid awkward conversations in the office because they simply don't know what to say. Thankfully, Green does—and in this incredibly helpful book, she tackles the tough discussions you may need to have during your career. You'll learn what to say when • coworkers push their work on you—then take credit for it • you accidentally trash-talk someone in an email then hit "reply all" • you're being micromanaged—or not being managed at all • you catch a colleague in a lie • your boss seems unhappy with your work • your cubemate's loud speakerphone is making you homicidal • you got drunk at the holiday party Praise for Ask a Manager "A must-read for anyone who works . . . [Alison Green's] advice boils down to the idea that you should be professional (even when others are not) and that communicating in a straightforward manner with candor and kindness will get you far, no matter where you work."—Booklist (starred review) "The author's friendly, warm, no-nonsense writing is a pleasure to read, and her advice can be widely applied to relationships in all areas of readers' lives. Ideal for anyone new to the job market or new to management, or anyone hoping to improve their work experience."—Library Journal (starred review) "I am a huge fan of Alison Green's Ask a Manager column. This book is even better. It teaches us how to deal with many of the most vexing big and little problems in our workplaces—and to do so with grace, confidence, and a sense of humor."—Robert Sutton, Stanford professor and author of The No Asshole Rule and The Asshole Survival Guide "Ask a Manager is the ultimate playbook for navigating the traditional workforce in a diplomatic but firm way."—Erin Lowry, author of Broke Millennial: Stop Scraping By and Get Your Financial Life Together

3 data science projects that got me 12 interviews: The Art of Statistics David Spiegelhalter, 2019-09-03 In this important and comprehensive guide to statistical thinking (New Yorker), discover how data literacy is changing the world and gives you a better understanding of life's biggest problems. Statistics are everywhere, as integral to science as they are to business, and in the popular media hundreds of times a day. In this age of big data, a basic grasp of statistical literacy is more important than ever if we want to separate the fact from the fiction, the ostentatious embellishments from the raw evidence -- and even more so if we hope to participate in the future, rather than being simple bystanders. In The Art of Statistics, world-renowned statistician David Spiegelhalter shows readers how to derive knowledge from raw data by focusing on the concepts and connections behind the math. Drawing on real world examples to introduce complex issues, he shows us how statistics can help us determine the luckiest passenger on the Titanic, whether a notorious serial killer could have been caught earlier, and if screening for ovarian cancer is beneficial. The Art of Statistics not only shows us how mathematicians have used statistical science to solve these problems -- it teaches us how we too can think like statisticians. We learn how to clarify our questions, assumptions, and expectations when approaching a problem, and -- perhaps even more importantly -- we learn how to responsibly interpret the answers we receive. Combining

the incomparable insight of an expert with the playful enthusiasm of an aficionado, The Art of Statistics is the definitive guide to stats that every modern person needs.

3 data science projects that got me 12 interviews: *Ace the Data Science Interview* Kevin Huo, Nick Singh, 2021

3 data science projects that got me 12 interviews: Deep Learning for Coders with fastai and PyTorch Jeremy Howard, Sylvain Gugger, 2020-06-29 Deep learning is often viewed as the exclusive domain of math PhDs and big tech companies. But as this hands-on guide demonstrates, programmers comfortable with Python can achieve impressive results in deep learning with little math background, small amounts of data, and minimal code. How? With fastai, the first library to provide a consistent interface to the most frequently used deep learning applications. Authors Jeremy Howard and Sylvain Gugger, the creators of fastai, show you how to train a model on a wide range of tasks using fastai and PyTorch. You'll also dive progressively further into deep learning theory to gain a complete understanding of the algorithms behind the scenes. Train models in computer vision, natural language processing, tabular data, and collaborative filtering Learn the latest deep learning techniques that matter most in practice Improve accuracy, speed, and reliability by understanding how deep learning models work Discover how to turn your models into web applications Implement deep learning algorithms from scratch Consider the ethical implications of your work Gain insight from the foreword by PyTorch cofounder, Soumith Chintala

3 data science projects that got me 12 interviews: Python Data Science Essentials Alberto Boschetti, Luca Massaron, 2016-10-28 Become an efficient data science practitioner by understanding Python's key concepts About This Book Quickly get familiar with data science using Python 3.5 Save time (and effort) with all the essential tools explained Create effective data science projects and avoid common pitfalls with the help of examples and hints dictated by experience Who This Book Is For If you are an aspiring data scientist and you have at least a working knowledge of data analysis and Python, this book will get you started in data science. Data analysts with experience of R or MATLAB will also find the book to be a comprehensive reference to enhance their data manipulation and machine learning skills. What You Will Learn Set up your data science toolbox using a Python scientific environment on Windows, Mac, and Linux Get data ready for your data science project Manipulate, fix, and explore data in order to solve data science problems Set up an experimental pipeline to test your data science hypotheses Choose the most effective and scalable learning algorithm for your data science tasks Optimize your machine learning models to get the best performance Explore and cluster graphs, taking advantage of interconnections and links in your data In Detail Fully expanded and upgraded, the second edition of Python Data Science Essentials takes you through all you need to know to suceed in data science using Python. Get modern insight into the core of Python data, including the latest versions of Jupyter notebooks, NumPy, pandas and scikit-learn. Look beyond the fundamentals with beautiful data visualizations with Seaborn and ggplot, web development with Bottle, and even the new frontiers of deep learning with Theano and TensorFlow. Dive into building your essential Python 3.5 data science toolbox, using a single-source approach that will allow to to work with Python 2.7 as well. Get to grips fast with data munging and preprocessing, and all the techniques you need to load, analyse, and process your data. Finally, get a complete overview of principal machine learning algorithms, graph analysis techniques, and all the visualization and deployment instruments that make it easier to present your results to an audience of both data science experts and business users. Style and approach The book is structured as a data science project. You will always benefit from clear code and simplified examples to help you understand the underlying mechanics and real-world datasets.

3 data science projects that got me 12 interviews: Janice VanCleave's A+ Science Fair Projects Janice VanCleave, 2003-08-08 A fabulous collection of science projects, explorations,techniques, and ideas! Looking to wow the judges at the science fair this year? Everyone'sfavorite science teacher is here to help. Janice VanCleave's A+Science Fair Projects has everything you need to put together awinning entry, with detailed advice on properly planning yourproject, from choosing a topic and collecting your facts todesigning experiments and presenting

your findings. Featuring all-new experiments as well as time-tested projectscollected from Janice VanCleave's A+ series, this easy-to-followguide gives you an informative introduction to the science fairprocess. You get thirty-five complete starter projects on varioustopics in astronomy, biology, chemistry, earth science, andphysics, including explorations of: * The angular distance between celestial bodies * The breathing rate of goldfish * Interactions in an ecosystem * Nutrient differences in soils * Heat transfer in the atmosphere * Magnetism from electricity * And much more! You'll also find lots of helpful tips on how to develop your ownideas into unique projects. Janice VanCleave's A+ Science FairProjects is the ideal guide for any middle or high school studentwho wants to develop a stellar science fair entry.

3 data science projects that got me 12 interviews: *Python for Data Analysis* Wes McKinney, 2017-09-25 Get complete instructions for manipulating, processing, cleaning, and crunching datasets in Python. Updated for Python 3.6, the second edition of this hands-on guide is packed with practical case studies that show you how to solve a broad set of data analysis problems effectively. You'll learn the latest versions of pandas, NumPy, IPython, and Jupyter in the process. Written by Wes McKinney, the creator of the Python pandas project, this book is a practical, modern introduction to data science tools in Python. It's ideal for analysts new to Python and for Python programmers new to data science and scientific computing. Data files and related material are available on GitHub. Use the IPython shell and Jupyter notebook for exploratory computing Learn basic and advanced features in NumPy (Numerical Python) Get started with data analysis tools in the pandas library Use flexible tools to load, clean, transform, merge, and reshape data Create informative visualizations with matplotlib Apply the pandas groupby facility to slice, dice, and summarize datasets Analyze and manipulate regular and irregular time series data Learn how to solve real-world data analysis problems with thorough, detailed examples

3 data science projects that got me 12 interviews: Smarter Data Science Neal Fishman, Cole Stryker, 2020-04-14 Organizations can make data science a repeatable, predictable tool, which business professionals use to get more value from their data Enterprise data and AI projects are often scattershot, underbaked, siloed, and not adaptable to predictable business changes. As a result, the vast majority fail. These expensive quagmires can be avoided, and this book explains precisely how. Data science is emerging as a hands-on tool for not just data scientists, but business professionals as well. Managers, directors, IT leaders, and analysts must expand their use of data science capabilities for the organization to stay competitive. Smarter Data Science helps them achieve their enterprise-grade data projects and AI goals. It serves as a guide to building a robust and comprehensive information architecture program that enables sustainable and scalable AI deployments. When an organization manages its data effectively, its data science program becomes a fully scalable function that's both prescriptive and repeatable. With an understanding of data science principles, practitioners are also empowered to lead their organizations in establishing and deploying viable AI. They employ the tools of machine learning, deep learning, and AI to extract greater value from data for the benefit of the enterprise. By following a ladder framework that promotes prescriptive capabilities, organizations can make data science accessible to a range of team members, democratizing data science throughout the organization. Companies that collect, organize, and analyze data can move forward to additional data science achievements: Improving time-to-value with infused AI models for common use cases Optimizing knowledge work and business processes Utilizing AI-based business intelligence and data visualization Establishing a data topology to support general or highly specialized needs Successfully completing AI projects in a predictable manner Coordinating the use of AI from any compute node. From inner edges to outer edges: cloud, fog, and mist computing When they climb the ladder presented in this book, businesspeople and data scientists alike will be able to improve and foster repeatable capabilities. They will have the knowledge to maximize their AI and data assets for the benefit of their organizations.

3 data science projects that got me 12 interviews: Doing Data Science Cathy O'Neil, Rachel Schutt, 2013-10-09 Now that people are aware that data can make the difference in an

election or a business model, data science as an occupation is gaining ground. But how can you get started working in a wide-ranging, interdisciplinary field that's so clouded in hype? This insightful book, based on Columbia University's Introduction to Data Science class, tells you what you need to know. In many of these chapter-long lectures, data scientists from companies such as Google, Microsoft, and eBay share new algorithms, methods, and models by presenting case studies and the code they use. If you're familiar with linear algebra, probability, and statistics, and have programming experience, this book is an ideal introduction to data science. Topics include: Statistical inference, exploratory data analysis, and the data science process Algorithms Spam filters, Naive Bayes, and data wrangling Logistic regression Financial modeling Recommendation engines and causality Data visualization Social networks and data journalism Data engineering, MapReduce, Pregel, and Hadoop Doing Data Science is collaboration between course instructor Rachel Schutt, Senior VP of Data Science at News Corp, and data science consultant Cathy O'Neil, a senior data scientist at Johnson Research Labs, who attended and blogged about the course.

3 data science projects that got me 12 interviews: Data Science Job: How to become a Data Scientist Przemek Chojecki, 2020-01-31 We're living in a digital world. Most of our global economy is digital and the sheer volume of data is stupendous. It's 2020 and we're living in the future. Data Scientist is one of the hottest job on the market right now. Demand for data science is huge and will only grow, and it seems like it will grow much faster than the actual number of data scientists. So if you want to make a career change and become a data scientist, now is the time. This book will guide you through the process. From my experience of working with multiple companies as a project manager, a data science consultant or a CTO, I was able to see the process of hiring data scientists and building data science teams. I know what's important to land your first job as a data scientist, what skills you should acquire, what you should show during a job interview.

3 data science projects that got me 12 interviews: *Storytelling with Data* Cole Nussbaumer Knaflic, 2015-10-09 Don't simply show your data—tell a story with it! Storytelling with Data teaches you the fundamentals of data visualization and how to communicate effectively with data. You'll discover the power of storytelling and the way to make data a pivotal point in your story. The lessons in this illuminative text are grounded in theory, but made accessible through numerous real-world examples—ready for immediate application to your next graph or presentation. Storytelling is not an inherent skill, especially when it comes to data visualization, and the tools at our disposal don't make it any easier. This book demonstrates how to go beyond conventional tools to reach the root of your data, and how to use your data to create an engaging, informative, compelling story. Specifically, you'll learn how to: Understand the importance of context and audience Determine the appropriate type of graph for your situation Recognize and eliminate the clutter clouding your information Direct your audience's attention to the most important parts of your data Think like a designer and utilize concepts of design in data visualization Leverage the power of storytelling to help your message resonate with your audience Together, the lessons in this book will help you turn your data into high impact visual stories that stick with your audience. Rid your world of ineffective graphs, one exploding 3D pie chart at a time. There is a story in your data—Storytelling with Data will give you the skills and power to tell it!

3 data science projects that got me 12 interviews: Visualize This Nathan Yau, 2011-06-13 Practical data design tips from a data visualization expert of the modern age Data doesn't decrease; it is ever-increasing and can be overwhelming to organize in a way that makes sense to its intended audience. Wouldn't it be wonderful if we could actually visualize data in such a way that we could maximize its potential and tell a story in a clear, concise manner? Thanks to the creative genius of Nathan Yau, we can. With this full-color book, data visualization guru and author Nathan Yau uses step-by-step tutorials to show you how to visualize and tell stories with data. He explains how to gather, parse, and format data and then design high quality graphics that help you explore and present patterns, outliers, and relationships. Presents a unique approach to visualizing and telling stories with data, from a data visualization expert and the creator of flowingdata.com, Nathan Yau Offers step-by-step tutorials and practical design tips for creating statistical graphics, geographical

maps, and information design to find meaning in the numbers Details tools that can be used to visualize data-native graphics for the Web, such as ActionScript, Flash libraries, PHP, and JavaScript and tools to design graphics for print, such as R and Illustrator Contains numerous examples and descriptions of patterns and outliers and explains how to show them Visualize This demonstrates how to explain data visually so that you can present your information in a way that is easy to understand and appealing.

3 data science projects that got me 12 interviews: *Multivariable Calculus* James Stewart, 2011-09-27 Success in your calculus course starts here! James Stewart's CALCULUS, 7e, International Metric texts are world-wide best-sellers for a reason: they are clear, accurate, and filled with relevant, real-world examples. With MULTIVARIABLE CALCULUS, 7e, International Metric Edition Stewart conveys not only the utility of calculus to help you develop technical competence, but also gives you an appreciation for the intrinsic beauty of the subject. His patient examples and built-in learning aids will help you build your mathematical confidence and achieve your goals in the course!

3 data science projects that got me 12 interviews: Drawdown Paul Hawken, 2017-04-18 • New York Times bestseller • The 100 most substantive solutions to reverse global warming, based on meticulous research by leading scientists and policymakers around the world “At this point in time, the Drawdown book is exactly what is needed; a credible, conservative solution-by-solution narrative that we can do it. Reading it is an effective inoculation against the widespread perception of doom that humanity cannot and will not solve the climate crisis. Reported by-effects include increased determination and a sense of grounded hope.” —Per Espen Stoknes, Author, What We Think About When We Try Not To Think About Global Warming “There’s been no real way for ordinary people to get an understanding of what they can do and what impact it can have. There remains no single, comprehensive, reliable compendium of carbon-reduction solutions across sectors. At least until now. . . . The public is hungry for this kind of practical wisdom.” —David Roberts, Vox “This is the ideal environmental sciences textbook—only it is too interesting and inspiring to be called a textbook.” —Peter Kareiva, Director of the Institute of the Environment and Sustainability, UCLA In the face of widespread fear and apathy, an international coalition of researchers, professionals, and scientists have come together to offer a set of realistic and bold solutions to climate change. One hundred techniques and practices are described here—some are well known; some you may have never heard of. They range from clean energy to educating girls in lower-income countries to land use practices that pull carbon out of the air. The solutions exist, are economically viable, and communities throughout the world are currently enacting them with skill and determination. If deployed collectively on a global scale over the next thirty years, they represent a credible path forward, not just to slow the earth’s warming but to reach drawdown, that point in time when greenhouse gases in the atmosphere peak and begin to decline. These measures promise cascading benefits to human health, security, prosperity, and well-being—giving us every reason to see this planetary crisis as an opportunity to create a just and livable world.

3 data science projects that got me 12 interviews: 25 Totally Terrific Science Projects Michael Gravois, 2010 Engaging, easy-to-make projects provide a hands-on way for students to share what they have learned about key science topics.

3 data science projects that got me 12 interviews: Human-in-the-Loop Machine Learning Robert Munro, Robert Monarch, 2021-07-20 Machine learning applications perform better with human feedback. Keeping the right people in the loop improves the accuracy of models, reduces errors in data, lowers costs, and helps you ship models faster. Human-in-the-loop machine learning lays out methods for humans and machines to work together effectively. You'll find best practices on selecting sample data for human feedback, quality control for human annotations, and designing annotation interfaces. You'll learn to dreate training data for labeling, object detection, and semantic segmentation, sequence labeling, and more. The book starts with the basics and progresses to advanced techniques like transfer learning and self-supervision within annotation workflows.

3 data science projects that got me 12 interviews: *Don't Trust Your Gut* Seth Stephens-Davidowitz, 2022-05-10 Seth Stephens-Davidowitz is more than a data scientist. He is a prophet for how to use the data revolution to reimagine your life. Don’t Trust Your Gut is a tour de force—an intoxicating blend of analysis, humor, and humanity.” — Daniel H. Pink, #1 New York Times bestselling author of When, Drive, and To Sell Is Human Big decisions are hard. We consult friends and family, make sense of confusing “expert” advice online, maybe we read a self-help book to guide us. In the end, we usually just do what feels right, pursuing high stakes self-improvement—such as who we marry, how to date, where to live, what makes us happy—based solely on what our gut instinct tells us. But what if our gut is wrong? Biased, unpredictable, and misinformed, our gut, it turns out, is not all that reliable. And data can prove this. In Don’t Trust Your Gut, economist, former Google data scientist, and New York Times bestselling author Seth Stephens-Davidowitz reveals just how wrong we really are when it comes to improving our own lives. In the past decade, scholars have mined enormous datasets to find remarkable new approaches to life’s biggest self-help puzzles. Data from hundreds of thousands of dating profiles have revealed surprising successful strategies to get a date; data from hundreds of millions of tax records have uncovered the best places to raise children; data from millions of career trajectories have found previously unknown reasons why some rise to the top. Telling fascinating, unexpected stories with these numbers and the latest big data research, Stephens-Davidowitz exposes that, while we often think we know how to better ourselves, the numbers disagree. Hard facts and figures consistently contradict our instincts and demonstrate self-help that actually works—whether it involves the best time in life to start a business or how happy it actually makes us to skip a friend’s birthday party for a night of Netflix on the couch. From the boring careers that produce the most wealth, to the old-school, data-backed relationship advice so well-worn it’s become a literal joke, he unearths the startling conclusions that the right data can teach us about who we are and what will make our lives better. Lively, engrossing, and provocative, the end result opens up a new world of self-improvement made possible with massive troves of data. Packed with fresh, entertaining insights, Don’t Trust Your Gut redefines how to tackle our most consequential choices, one that hacks the market inefficiencies of life and leads us to make smarter decisions about how to improve our lives. Because in the end, the numbers don’t lie.

3 data science projects that got me 12 interviews: *The Data Science Handbook* Carl Shan, Henry Wang, William Chen, Max Song, 2015-05-03 The Data Science Handbook is a curated collection of 25 candid, honest and insightful interviews conducted with some of the world's top data scientists.In this book, you'll hear how the co-creator of the term 'data scientist' thinks about career and personal success. You'll hear from a young woman who created her own data scientist curriculum, subsequently landing her a role in the field. Readers of this book will be left with war stories, wisdom and

3 data science projects that got me 12 interviews: *Getting Started with Data Science* Murtaza Haider, 2015-12-14 Master Data Analytics Hands-On by Solving Fascinating Problems You’ll Actually Enjoy! Harvard Business Review recently called data science “The Sexiest Job of the 21st Century.” It’s not just sexy: For millions of managers, analysts, and students who need to solve real business problems, it’s indispensable. Unfortunately, there’s been nothing easy about learning data science–until now. Getting Started with Data Science takes its inspiration from worldwide best-sellers like Freakonomics and Malcolm Gladwell’s Outliers: It teaches through a powerful narrative packed with unforgettable stories. Murtaza Haider offers informative, jargon-free coverage of basic theory and technique, backed with plenty of vivid examples and hands-on practice opportunities. Everything’s software and platform agnostic, so you can learn data science whether you work with R, Stata, SPSS, or SAS. Best of all, Haider teaches a crucial skillset most data science books ignore: how to tell powerful stories using graphics and tables. Every chapter is built around real research challenges, so you’ll always know why you’re doing what you’re doing. You’ll master data science by answering fascinating questions, such as: • Are religious individuals more or less likely to have extramarital affairs? • Do attractive professors get better teaching evaluations? • Does

the higher price of cigarettes deter smoking? • What determines housing prices more: lot size or the number of bedrooms? • How do teenagers and older people differ in the way they use social media? • Who is more likely to use online dating services? • Why do some purchase iPhones and others Blackberry devices? • Does the presence of children influence a family's spending on alcohol? For each problem, you'll walk through defining your question and the answers you'll need; exploring how others have approached similar challenges; selecting your data and methods; generating your statistics; organizing your report; and telling your story. Throughout, the focus is squarely on what matters most: transforming data into insights that are clear, accurate, and can be acted upon.

3 data science projects that got me 12 interviews: Introduction to Probability Joseph K. Blitzstein, Jessica Hwang, 2014-07-24 Developed from celebrated Harvard statistics lectures, Introduction to Probability provides essential language and tools for understanding statistics, randomness, and uncertainty. The book explores a wide variety of applications and examples, ranging from coincidences and paradoxes to Google PageRank and Markov chain Monte Carlo (MCMC). Additional application areas explored include genetics, medicine, computer science, and information theory. The print book version includes a code that provides free access to an eBook version. The authors present the material in an accessible style and motivate concepts using real-world examples. Throughout, they use stories to uncover connections between the fundamental distributions in statistics and conditioning to reduce complicated problems to manageable pieces. The book includes many intuitive explanations, diagrams, and practice problems. Each chapter ends with a section showing how to perform relevant simulations and calculations in R, a free statistical software environment.

3 data science projects that got me 12 interviews: Data Science Strategy For Dummies Ulrika Jägare, 2019-06-12 All the answers to your data science questions Over half of all businesses are using data science to generate insights and value from big data. How are they doing it? Data Science Strategy For Dummies answers all your questions about how to build a data science capability from scratch, starting with the "what" and the "why" of data science and covering what it takes to lead and nurture a top-notch team of data scientists. With this book, you'll learn how to incorporate data science as a strategic function into any business, large or small. Find solutions to your real-life challenges as you uncover the stories and value hidden within data. Learn exactly what data science is and why it's important Adopt a data-driven mindset as the foundation to success Understand the processes and common roadblocks behind data science Keep your data science program focused on generating business value Nurture a top-quality data science team In non-technical language, Data Science Strategy For Dummies outlines new perspectives and strategies to effectively lead analytics and data science functions to create real value.

3 data science projects that got me 12 interviews: *Solving Data Science Case Studies with Python* Aman Kharwal, 2021-06-28 This book is specially written for those who know the basics of the Python programming language as well as the necessary Python libraries you need for data science like NumPy, Pandas, Matplotlib, Seaborn, Plotly, and Scikit-learn. This book aims to teach you how to think while solving a business problem with your data science skills. To achieve the goal of this book, I started by giving you all the knowledge you need to have before you apply for your first data science job. The technical skills and soft skills you need to become a Data Scientist are also discussed in this book. Next, you'll find some of the best data science case studies that will help you understand what your approach should be while solving a business problem. Ultimately, you will also find some of the most important data science interview questions with their solutions at the end. I hope this book will add a lot of value to your data science skills and that you will feel confident in your entire journey to become Data Scientist.

3 data science projects that got me 12 interviews: *Social Science Research* Anol Bhattacherjee, 2012-04-01 This book is designed to introduce doctoral and graduate students to the process of conducting scientific research in the social sciences, business, education, public health, and related disciplines. It is a one-stop, comprehensive, and compact source for foundational concepts in behavioral research, and can serve as a stand-alone text or as a supplement to research

readings in any doctoral seminar or research methods class. This book is currently used as a research text at universities on six continents and will shortly be available in nine different languages.

3 data science projects that got me 12 interviews: *Deploying Machine Learning* Robbie Allen, 2019-05 Increasingly, business leaders and managers recognize that machine learning offers their companies immense opportunities for competitive advantage. But most discussions of machine learning are intensely technical or academic, and don't offer practical information leaders can use to identify, evaluate, plan, or manage projects. Deploying Machine Learning fills that gap, helping them clarify exactly how machine learning can help them, and collaborate with technologists to actually apply it successfully. You'll learn: What machine learning is, how it compares to big data and artificial intelligence, and why it's suddenly so important What machine learning can do for you: solutions for computer vision, natural language processing, prediction, and more How to use machine learning to solve real business problems -- from reducing costs through improving decision-making and introducing new products Separating hype from reality: identifying pitfalls, limitations, and misconceptions upfront Knowing enough about the technology to work effectively with your technical team Getting the data right: sourcing, collection, governance, security, and culture Solving harder problems: exploring deep learning and other advanced techniques Understanding today's machine learning software and hardware ecosystem Evaluating potential projects, and addressing workforce concerns Staffing your project, acquiring the right tools, and building a workable project plan Interpreting results -- and building an organization that can increasingly learn from data Using machine learning responsibly and ethically Preparing for tomorrow's advances The authors conclude with five chapter-length case studies: image, text, and video analysis, chatbots, and prediction applications. For each, they don't just present results: they also illuminate the process the company undertook, and the pitfalls it overcame along the way.

3 data science projects that got me 12 interviews: Cracking the Coding Interview Gayle Laakmann McDowell, 2011 Now in the 5th edition, Cracking the Coding Interview gives you the interview preparation you need to get the top software developer jobs. This book provides: 150 Programming Interview Questions and Solutions: From binary trees to binary search, this list of 150 questions includes the most common and most useful questions in data structures, algorithms, and knowledge based questions. 5 Algorithm Approaches: Stop being blind-sided by tough algorithm questions, and learn these five approaches to tackle the trickiest problems. Behind the Scenes of the interview processes at Google, Amazon, Microsoft, Facebook, Yahoo, and Apple: Learn what really goes on during your interview day and how decisions get made. Ten Mistakes Candidates Make -- And How to Avoid Them: Don't lose your dream job by making these common mistakes. Learn what many candidates do wrong, and how to avoid these issues. Steps to Prepare for Behavioral and Technical Questions: Stop meandering through an endless set of questions, while missing some of the most important preparation techniques. Follow these steps to more thoroughly prepare in less time.

Additional Resources

The Complete Collection of Data Science Cheat Sheets

Tidyverse. The Tidyverse contains a complete data science solution from importing data to creating visually simulating data reports. • Tidyverse for Beginners • Data visualization with ggplot2 • Data ...

Solved End-to-End big Data and Data Science Projects

We've got you covered with top 25 solved end-to-end big data and data science projects that you should explore and get your hands dirty to master the skills and concepts required to land your ...

Data Science with Python Career Program - Capstone Project

• Data Exploration is very useful step in the process of analytics and model building as it helps us to remove

any redundant data & also construct new features to enhance the analysis and model ...

FINAL YEAR PROJECT REPORT - Sathyabama Institute of ...

Tools for reading and writing data between in-memory data structures and different file formats. Data alignment and integrated handling of missing data. Reshaping and pivoting of data sets.

30+ Data Science Mini Project Ideas For College Students

In this blog, we'll delve into why Data Science Mini Projects matter and guide you on how to choose the right project. We've even curated a list of 30+ mini project ideas for college students. Stay ...

Applying Scrum in Data Science Projects - Computer

In this study, we selected the Scrum approach and integrated it into the CRISP-DM methodology for data science projects using a Design Science Research approach. This new methodology was ...

Playbook for Project Management in Data Science and Artificial ...

Two established thought leaders, the Project Management Institute and NASSCOM's Center of Excellence for Data Science and AI have come together to bring to you a unique perspective on ...

Data Collection methods (Questionnaire & Interview)

Accurate and systematic data collection is critical to conducting scientific research. Data collection allows us to collect information that we want to collect about our study participants. Depending ...

Chapter 02 Process of Data Science Projects - GitHub Pages

•Generic process for data science projects with six phases •Discovery, data preparation, model planning, model building, communication of results, and operationalization

The Risk Management Process for Data Science: Gaps in ...

To enhance the field's understanding of risk management for data science projects via a qualitative study involving 16 organizations, this paper explores how organizations identify and mitigate ...

NPSAS:12 Computer-Aided Data Extraction Usability Interviews

Section 3 contains a description of each of these four CADE modes. interview with the respondent. This interview protocol was developed by NCES with feedback from the Westat research team. A ...

Cracking Data Science Interview - www2.internationalinsurance

1. Understanding the Data Science Interview Landscape Before we dive into specific strategies, let's understand the beast we're facing. Data science interviews are notoriously rigorous, often ...

Data Cleaning and Preprocessing for Data Science Beginners

Data Cleaning: Clean the collected data by identifying and correcting errors, removing duplicates and irrelevant observations, and handling missing values. Data Integration: Integrate data from ...

Why Data Science Projects Fail - arXiv.org

We prepared a list of questions and interviewed experienced Data Scientists, Senior Leader who actively manage data science and data analytic team, and stakeholders part of data science ...

M.Tech. in Data Science (MDS) Information Brochure (Year ...

Data Sciences (MDS) program in Academic Year 2018. the MDS is a self-paced program of 48 credits that can be taken over 3-4 years. Students will do 24 credits of coursework in the first ...

Beyond the Hype: Why Do Data-Driven Projects Fail?

identify the reasons that make data-driven projects fall short of expectations, multiple rounds of qualitative semi-structured interviews with domain experts with different roles in data-driven ...

Masters Thesis for Data Science - ichatz.me

TechnipFMC acts in three distinct segments: Subsea, Offshore/Onshore and surface projects. These projects include offshore oil and gas exploration and extraction platforms, rigs, crude oil ...

Best Practices in Structuring Data Science Projects - Springer
In particular, we surveyed three sources of information on how to structure projects: common manage-ment methodologies, community best practices, and data sharing plat-forms. The ...

Comparative Analysis of Process Models for Data Science ...
In this work we perform a comparative analysis of existing process models for data science projects. Emerging data science project models are analysed in a structured-literature review , taking into ...

The Complete Collection of Data Science Cheat Sheets
Tidyverse. The Tidyverse contains a complete data science solution from importing data to creating visually simulating data reports. • Tidyverse for Beginners • Data visualization with ...

Solved End-to-End big Data and Data Science Projects
We've got you covered with top 25 solved end-to-end big data and data science projects that you should explore and get your hands dirty to master the skills and concepts required to land your ...

Data Science with Python Career Program - Capstone Project
• Data Exploration is very useful step in the process of analytics and model building as it helps us to remove any redundant data & also construct new features to enhance the analysis and ...

FINAL YEAR PROJECT REPORT - Sathyabama Institute of ...
Tools for reading and writing data between in-memory data structures and different file formats. Data alignment and integrated handling of missing data. Reshaping and pivoting of data sets.

30+ Data Science Mini Project Ideas For College Students
In this blog, we'll delve into why Data Science Mini Projects matter and guide you on how to choose the right project. We've even curated a list of 30+ mini project ideas for college ...

Applying Scrum in Data Science Projects - Computer
In this study, we selected the Scrum approach and integrated it into the CRISP-DM methodology for data science projects using a Design Science Research approach. This new methodology ...

Playbook for Project Management in Data Science and ...
Two established thought leaders, the Project Management Institute and NASSCOM's Center of Excellence for Data Science and AI have come together to bring to you a unique perspective ...

Data Collection methods (Questionnaire & Interview)
Accurate and systematic data collection is critical to conducting scientific research. Data collection allows us to collect information that we want to collect about our study participants. Depending ...

Chapter 02 Process of Data Science Projects - GitHub Pages
•Generic process for data science projects with six phases •Discovery, data preparation, model planning, model building, communication of results, and operationalization

The Risk Management Process for Data Science: Gaps in ...
To enhance the field's understanding of risk management for data science projects via a qualitative study involving 16 organizations, this paper explores how organizations identify and ...

NPSAS:12 Computer-Aided Data Extraction Usability ...
Section 3 contains a description of each of these four CADE modes. interview with the respondent. This interview protocol was developed by NCES with feedback from the Westat ...

Cracking Data Science Interview - www2.internationalinsurance

1. Understanding the Data Science Interview Landscape Before we dive into specific strategies, let's understand the beast we're facing. Data science interviews are notoriously rigorous, often ...

Data Cleaning and Preprocessing for Data Science Beginners

Data Cleaning: Clean the collected data by identifying and correcting errors, removing duplicates and irrelevant observations, and handling missing values. Data Integration: Integrate data from ...

Why Data Science Projects Fail - arXiv.org
We prepared a list of questions and interviewed experienced Data Scientists, Senior Leader who actively manage data science and data analytic team, and stakeholders part of data science ...

M.Tech. in Data Science (MDS) Information Brochure (Year ...
Data Sciences (MDS) program in Academic Year 2018. the MDS is a self-paced program of 48 credits that can be taken over 3-4 years. Students will do 24 credits of coursework in the first ...

Beyond the Hype: Why Do Data-Driven Projects Fail?
identify the reasons that make data-driven projects fall short of expectations, multiple rounds of qualitative semi-structured interviews with domain experts with different roles in data-driven ...

Masters Thesis for Data Science - ichatz.me
TechnipFMC acts in three distinct segments: Subsea, Offshore/Onshore and surface projects. These projects include offshore oil and gas exploration and extraction platforms, rigs, crude oil ...

Best Practices in Structuring Data Science Projects - Springer
In particular, we surveyed three sources of information on how to structure projects: common manage-ment methodologies, community best practices, and data sharing plat-forms. The ...

Comparative Analysis of Process Models for Data Science ...
In this work we perform a comparative analysis of existing process models for data science projects. Emerging data science project models are analysed in a structured-literature review , ...

3 Data Science Projects That Got Me 12 Interviews

3 Data Science Projects That Got Me 12 Interviews

Related Articles

- 3 wire 220v welder plug wiring diagram
- 3 examples of mechanical waves
- 3 5 3 defense diagram

Back to Home