SEBASTIAN FITZEK

Introduction to DATA SCIENCE

A PYTHON Path for a Non-computer Scientist

comunicare.ro

Redactor: Anca Milu-Vaidesegan
Tehnoredactor: Cristian Lupeanu

Editura COMUNICARE.RO este departament în cadrul
Școlii Naționale de Studii Politice și Administrative,
Facultatea de Comunicare și Relații Publice.

Editura COMUNICARE.RO
SNSPA, Facultatea de Comunicare și Relații Publice
Str. Povernei, nr. 6, sector 1
010643, București
România
Tel.: 0372.177.150
www.edituracomunicare.ro
e-mail: editura@comunicare.ro

ISBN 978-973-711-642-0 (ediție electronică)

Contents

Foreword

Asking yourself the most important questions is a good way to introduce a book dedicated to Data Science because it helps to set the stage for the material that will be covered in the book. By asking questions, you can get a sense of what the book will be about and what you can expect to learn from it. Additionally, asking questions can help to engage the reader and get them thinking about the topic at hand, which can make the material more interesting and relevant to their own experiences. So, the first legitimate question is why is Data Science today a discipline of great importance for the present and future of master students? For sure, Data Science is a discipline of great importance today because it allows organizations to make better decisions by leveraging the vast amounts of data that are generated in today's world. With the help of Data Science, organizations can gain insights into their operations and customers, and use that information to improve their products, services, and overall business strategies. Additionally, Data Science is a rapidly growing field, with many job opportunities for individuals with the right skills and training. As a result, pursuing a master's degree in communication and Data Science can open a wide range of career possibilities for students.

Another key question is why should communication students be the ones to start and deepen Data Science or what is the connection between communication and public relations students and Data Science? Communication and public relations students should be interested in studying Data Science because it can help them better understand and analyze the vast amounts of data that are generated in today's world. Data Science can provide communication and public relations students with the skills to glean meaningful insights from data, which can be utilized to elevate and refine their practices in these areas. Learning Data Science, students can also gain the ability to extract valuable insights from data that can be applied to enhance their work in communication and public relations. Additionally, Data Sci-

ence can help communication and public relations students to better understand the needs and preferences of their target audience, and to create more effective communication strategies and campaigns. Finally, the skills and knowledge gained from studying Data Science can be highly valuable in the job market and can help communication and public relations students stand out from their peers and advance their careers.

A thirty-concrete key question would be how can it open a career and what are the shortcuts to success for any student wishing to specialize in this field? Studying Data Science can open a wide range of career possibilities for students. Data Science professionals may find employment as data scientists, data analysts, machine learning engineers, or business intelligence analysts. Alternatively, those with a strong background in Data Science may choose to pursue careers as data scientists, data analysts, machine learning engineers, or business intelligence analysts. In these roles, individuals can work in a variety of industries, including technology, finance, healthcare, and government, to help organizations make better decisions using data. To succeed in this field, students should be prepared to learn a wide range of technical skills, such as programming, statistics, and machine learning, as well as soft skills, such as problem-solving and communication. Additionally, students can gain a competitive advantage by participating in internships or other hands-on learning experiences, as well as by staying up to date with the latest developments in the field.

However, the field is very broad and then we should ask ourselves which parts or structures of Data Science are worth learning in the early stages. In this sense, what is worth studying in a master's program of only 2 years, having this limited time, and what is worth didactically deepened in the early stages of initiation? In a master's program with a limited time frame, students need to prioritize the key concepts and skills that are most essential for success in Data Science. Some of the most important areas to focus on in the early stages of a Data Science program include:

- In **Part I** of this book, we will learn **Fundamental Programming in Python**: **Data Science** involves working with large and complex

datasets, and students will need to be proficient in **Python**, to manipulate and analyze that data.

- In **Part II** of this book, we will learn about **Data Visualization** and other key procedures for a communication and data science specialist**.** Data visualization is an important tool for communicating the results of data analyses, and students will need to learn how to create clear and effective visualizations to effectively communicate their findings. **Web scraping** is an important tool for **data scientists** because it allows them **to extract data from websites** and turn it into structured, usable data that can be **analyzed** and **visualized**.
- In **Part III**: **Introduction to business statistics in Data Science**. This part aims to provide a fundamental understanding of statistics and its role in data science and business, as well as equip students with the tools and skills to apply statistical analysis to real-world situations.
- **Part IV** of this book is optional and is intended for readers who wish to further their knowledge in communication and data science. It covers the topic of **Machine learning** and is geared towards those who want to rapidly progress in this field and delve into advanced areas of data science. For the basics of this discipline, I recommend delving into just the first two parts, and optionally part three.

By focusing on these key areas, students can gain a solid foundation in Data Science, which they can then build upon in more advanced coursework and real-world experiences. The four stages of study broadly make up the overall structure of this book. By learning programming, statistics, machine learning, and data visualization, students can gain a solid foundation in Data Science, which they can then build upon in more advanced coursework and real-world experiences. These four areas are essential for success in Data Science because they provide the tools and techniques needed to manipulate, analyze, and interpret data, as well as to communicate the results of those analyses to others.

As the author of this book, I agree that organizing information into theoretical, example and practical components is an effective way to help students assimilate key information quickly and efficiently. By providing a mix of conceptual and practical material, students can

gain a deep understanding of the subject matter and apply this knowledge to real-world problems and scenarios. In addition, by providing examples and exercises, students can see how the concepts they are learning apply in practice, which can help to reinforce their understanding and facilitate the transfer of knowledge to new situations. Overall, this approach can help to engage and motivate learners and help them develop the skills and knowledge they need to succeed as communication and data scientists.

I hereby wish my students and all my readers to use this book to the fullest and learn to love Data Science professionally and then teach others as I have lovingly taught them. The best way to learn about any subject is to approach it with an open mind and a willingness to try new things. To maximize the benefit of this book, I would encourage my students to engage with the material actively, asking questions, trying out the examples and exercises, and seeking out additional resources and opportunities to learn more. Additionally, I encourage them to connect with other Data Science enthusiasts and professionals, either through online communities or in-person events, to learn from each other and stay up to date with the latest developments in the field. By adopting this approach, my students can develop a deep understanding of Data Science and they can become skilled and knowledgeable professionals who are well-equipped to teach others.