Python for Finance

Data analysis, financial modeling, and portfolio management

Dmytro Zherlitsyn

www.bpbonline.com

First Edition 2024

ISBN: 978-93-55516-893

LIMITS OF LIABILITY AND DISCLAIMER OF WARRANTY

The information contained in this book is true to correct and the best of author's and publisher's knowledge. The author has made every effort to ensure the accuracy of these publications, but publisher cannot be held responsible for any loss or damage arising from any information in this book.

All trademarks referred to in the book are acknowledged as properties of their respective owners but BPB Publications cannot guarantee the accuracy of this information.

Dedicated to

My beloved wife:

Darya

About the Author

Dmytro Zherlitsyn, a professor and doctor of science, has dedicated over 20 years to university teaching, business training, financial consulting, scientific research and data analysis. He has authored over 250 academic publications (e-learning courses, textbooks, scientific papers and monographs) in economics, finance, data science, system analysis and software engineering. Dmytro headed the Economic Cybernetics department and co-led several data science and business improvement projects. His current roles include: researcher at the University of National and World Economy in Bulgaria and professor at the technical university "Metinvest Polytechnic" in Ukraine. His teaching comprises pioneering courses in Python for data analysis and applied statistics, aligning with his professional focus on using Python to drive financial insights and innovations. His work encompasses the development of predictive models for business and market analysis, including advanced regression, simulation and ML methods for financial sectors and the cryptocurrency market.

About the Reviewers

- **Dheerendra Panwar** is a seasoned professional in the field of Edge, IoT, and Machine Learning with over ten years of experience. He earned his master's degree in embedded electrical and computer systems from San Francisco State University, further fortifying his expertise in the domain. Throughout his career, he has contributed significantly to various IoT projects, ranging from manufacturing and smart cities to the retail and energy sectors. Having worked in large organizations and startups, he has a comprehensive understanding of the intricacies of IoT/edge technologies and their practical applications.

- **Yash Yennam** is an experienced data scientist and ML engineer with a robust background in developing data-driven solutions and comprehensive ML models. He excels in transforming complex datasets into actionable insights, particularly within the domains of geographic information systems, credit risk and market risk. His expertise encompasses various ML frameworks such as TensorFlow, Scikit-Learn, and PyTorch, in addition to proficiency in big data tools like Spark and Hadoop.

  Yash has led projects involving data orchestration pipelines, model deployment, and real-time data processing. His contributions have been instrumental in driving business impact through innovative solutions and operational efficiencies. He possesses a strong foundation in computer science and maintains a passion for continuous learning. Yash is dedicated to advancing the fields of AI and ML, with a keen interest in the latest developments and their practical applications. Beyond his professional endeavors, he is an avid reader of science fiction and non-fiction, and he enjoys playing the guitar. A technology enthusiast, Yash continually explores the fields of astrophysics, engineering, and stock markets.

- **Suryakant** is a distinguished expert in the fields of ML, deep learning, and AI. With a wealth of experience and proven proficiency in these cutting-edge disciplines, Suryakant leads pioneering efforts in generative AI solutions. His extensive knowledge and expertise are instrumental in exploring the transformative potential of this burgeoning technology.

  Having authored several research papers and a seminal book on hyperautomation, Suryakant's contributions to the field are widely recognized. With a steadfast commitment to pushing the boundaries of AI innovation, he is dedicated to harnessing the power of generative AI to drive creativity and innovation. Positioned as a key influencer in the realm of generative AI, Suryakant's relentless pursuit of excellence and passion for leveraging data-driven insights exemplify his status as a prominent figure in this dynamic and rapidly evolving domain.

Acknowledgement

I want to express my deepest gratitude to my family and friends for their unwavering support and encouragement throughout the writing of this book, especially my wife Darya.

I am also profoundly grateful to BPB Publications for their unwavering guidance and expertise. Their role in bringing this book to fruition has been invaluable. The journey of revising this book was made possible by the valuable participation and collaboration of reviewers, technical experts, and editors.

I would also like to acknowledge the valuable contributions of my colleagues and students. They have taught me so much and provided valuable feedback on my results.

Finally, I would like to thank all the readers who have taken an interest in my book and for their support in making it a reality.

Preface

Financial markets are complex systems involving many participants, interconnected entities, as well as models, methods and technologies. In recent years, the financial sector has experienced an "explosion" due to the catastrophic growth in information volumes. Classical methods of financial data analysis need to gain competitive ground. Therefore, financial analysts increasingly use IT technologies, particularly Python programming tools, for quick decision-making and profit generation.

Python, a high-level programming language, has become a staple in various fields, including finance. Its straightforward syntax, vast user community, and extensive range of libraries and tools make it a practical and powerful tool for financial data analysis. This book utilizes key Python libraries such as Pandas, NumPy, SciPy, Statsmodels, Matplotlib, Seaborn, Scikit-learn, Prophet, and others, empowering you with the tools you need to excel in financial analysis.

This book introduces fundamental concepts for analyzing financial markets and supporting investment decisions. These concepts, including time-series analysis, graphical analysis, technical and fundamental analysis, asset pricing, portfolio theory, investment and trading strategies, risk assessment, and the basics of financial machine learning, are more than just theoretical. We bring them to life with real-world examples of analyzing financial market dynamics, forecasting future trends, optimizing investment portfolios, assessing strategies, and managing financial risks, making the content engaging and applicable to your work.

With this book, you will gain Python programming basics, its primary libraries for data analysis, and their integration with the core financial concepts.

Chapter 1: Getting Started with Python for Finance - explains foundational knowledge of Python's role in finance and its advantages over other programming languages. The installation and configuration of Python on local computers or using the Google Colab cloud platform are described. This chapter provides an overview of the top libraries for solving financial problems with Python. It also illustrates the fundamentals of the Python programming language, including syntax, operators, and basic data structures, including those related to financial data analysis.

Chapter 2: Python Tools for Data Analysis: Primer to Pandas and NumPy - presents an overview of the essential Python control and data structures operations, built-in functions, and primary libraries for financial data analysis (NumPy and Pandas). The chapter provides practical examples close to actual financial data to explore foundational tools and operations crucial for such manipulation. Learn to create and manage arrays with

NumPy and handle tabular data effortlessly with Pandas, gearing you to derive insightful outcomes from the financial data analyses.

Chapter 3: Financial Data Manipulation with Python - covers the foundational concepts of financial data, explores various open data sources, and investigates their role in finance. The practical skills with Python will be expanded while collecting real-world financial datasets by importing and structuring information. In the chapter, explain how to use the benefits of yfinance, pandas_datareader, quandl and other Python libraries, as well as CSV and Excel data files, APIs, and web-scraping tools. The practical results create datasets for analyzing financial data and making informed decisions using Python.

Chapter 4: Exploratory Data Analysis for Finance - allows the reader to learn essential exploratory data analysis skills for finance. The data-transforming processes and patterns used to inspect and clean financial data and related mathematical operations are described. The chapter provides hands-on experience in data visualization using Matplotlib and Seaborn, as well as understanding the descriptive statistics metrics and moving average data to determine financial trends. Investment returns and risk statistics, as well as explanatory, visual, and correlation analysis tools, are explored. This skill set helps to make informed, data-driven investment decisions and prepares the typical analytical stages.

Chapter 5: Investment and Trading Strategies - gives special attention to investment, analytical and trading strategies, with the ability to integrate technical, fundamental, and graphical analysis into your trading strategy. The chapter delves into core investment principles and metrics, offering practical insights and advanced candlestick chart techniques. Utilizing Python's visualization tools, financial market data is brought to life, highlighting significant patterns and interpreting market indicators. Essential graphical and technical analysis tools are covered, enabling the generation and testing of trading strategies. Real-world market data is used to solidify understanding, preparing readers to navigate the dynamic nature of trading with a thorough grasp of risk and return dynamics and equipping them to make informed decisions using sophisticated analytical tools.

Chapter 6: Asset Pricing and Portfolio Management - details Python tools for estimating investment portfolio parameters and regression model parameters. The basics of modern portfolio theory are covered to inform long-term investment strategies. Foundational portfolio theories, such as Markowitz's model and the Sharpe Ratio criteria, are examined. Statistical tools and regression models are used to quantify the risk-return ratio for making investment decisions. The power of Python statistical libraries, such as Statsmodels and SciPy, is highlighted for regression analysis and to find optimum solutions mathematically.

Chapter 7: Time Series Analysis and Financial Data Forecasting - applies traditional time series analysis in financial forecasting with Python, pointing out the core limitations of these models. Various forecasting techniques are explored, from exponential smoothing to advanced SARIMAX models, revealing the challenges faced in volatile financial

markets. The chapter describes why the Mean Absolute Percentage Error (MAPE) metric can sometimes yield better results with actual financial time series data variations. This insight is pivotal for applying more robust and adaptive forecasting techniques, including machine learning.

Chapter 8: Risk Assessment and Volatility Modelling – explains sophisticated principles of probability theory with executable Python code, leading to a deep understanding of financial risk and volatility principles. Proficiency is gained in applying Python's computational capabilities to financial risk assessment and volatility modelling. Understanding how to use Python tools with key probabilistic distributions is achieved through the computation of VaR and aVaR. The power of Monte Carlo simulations is used by applying randomizing or stochastic methods to real-world examples of option price prediction and VaR estimation. By comprehensively exploring ARCH and GARCH models, the ability to anticipate and model financial volatility is developed.

Chapter 9: Machine Learning and Deep Learning in Finance – explore the ML world as it applies to the financial sector. An understanding of fundamental theories, models, and steps for applying ML to analyze and predict financial data is provided. The chapter focuses on the practical utility of the scikit-learn library, demonstrating how to implement machine learning models such as clustering and regression and employ feature engineering to enhance model performance. Description of the scikit-learn, XGBoost, and lightGBM libraries and evaluation of the basic machine learning models using appropriate tools for financial applications are covered. The chapter lays the foundation for applying skills in using Python-based regression and clustering techniques, understanding the importance of cross-validation, and performing hyperparameter tuning to improve model accuracy.

Chapter 10: Time Series Analysis and Forecasting with FB Prophet Library – describes the FB Prophet library for advanced time series analysis and forecasting in finance. This chapter provides a detailed understanding of Prophet's functionalities, from executing basic operations to exploiting advanced features for more accurate forecasting. Techniques for applying Prophet to various financial datasets are covered, enabling the forecasting of market trends, evaluation of investment risks, and making well-informed financial decisions. The tools are provided to construct, assess, and refine complex forecasting models, employ cross-validation techniques, tune hyperparameters, and combine Prophet with machine learning methods for enhanced financial decision-making.

Appendix A: Python Code Examples for Finance – contains the main code examples from this book.

Appendix B: Glossary – outlines the meaning of keywords and definitions.

Appendix C: Valuable Resources – describes key resources for future development of new Python programming and self-development

Code Bundle and Coloured Images

Please follow the link to download the
Code Bundle and the ***Coloured Images*** of the book:

https://rebrand.ly/6b3c55

The code bundle for the book is also hosted on GitHub at
https://github.com/bpbpublications/Python-for-Finance.
In case there's an update to the code, it will be updated on the existing GitHub repository.

We have code bundles from our rich catalogue of books and videos available at **https://github.com/bpbpublications**. Check them out!

Errata

We take immense pride in our work at BPB Publications and follow best practices to ensure the accuracy of our content to provide with an indulging reading experience to our subscribers. Our readers are our mirrors, and we use their inputs to reflect and improve upon human errors, if any, that may have occurred during the publishing processes involved. To let us maintain the quality and help us reach out to any readers who might be having difficulties due to any unforeseen errors, please write to us at :

errata@bpbonline.com

Your support, suggestions and feedbacks are highly appreciated by the BPB Publications' Family.

Piracy

If you come across any illegal copies of our works in any form on the internet, we would be grateful if you would provide us with the location address or website name. Please contact us at **business@bpbonline.com** with a link to the material.

If you are interested in becoming an author

If there is a topic that you have expertise in, and you are interested in either writing or contributing to a book, please visit **www.bpbonline.com**. We have worked with thousands of developers and tech professionals, just like you, to help them share their insights with the global tech community. You can make a general application, apply for a specific hot topic that we are recruiting an author for, or submit your own idea.

Reviews

Please leave a review. Once you have read and used this book, why not leave a review on the site that you purchased it from? Potential readers can then see and use your unbiased opinion to make purchase decisions. We at BPB can understand what you think about our products, and our authors can see your feedback on their book. Thank you!

For more information about BPB, please visit **www.bpbonline.com**.

Join our book's Discord space

Join the book's Discord Workspace for Latest updates, Offers, Tech happenings around the world, New Release and Sessions with the Authors:

https://discord.bpbonline.com

Table of Contents

Chapter 1
Getting Started with Python for Finance

Introduction

In the constantly evolving field of finance, professionals and enthusiasts must stay up to date with the latest tools and methodologies. One such tool that has gained substantial traction in finance is the Python programming language. This chapter will introduce the dynamic world of Python and its potential in finance and financial data analysis. Initially conceived in the late 1980s for students' software engineering skills training, Python has become an impressive part of the world of finance and FinTech as one of the most popular analytical tools. With its simplicity and expansive ecosystem of libraries, the chapter has become an indispensable asset for financial analysts, quantitative researchers, and investment bankers. This might prompt the following questions:

- What makes Python different from other programming languages for finance and data analysis aims?
- Why have major financial institutions and individual investors applied Python tools as a core of data analytics and financial modelling?

We will learn more about the Python programming language's core principles for data analysis and the finance sector decision-making to answer these questions. We will also compare and show Python's advantages as an analysis tool with other major programming languages, like Java, Julia, or R. Furthermore, to ensure that we are well-equipped to use Python's power, we will guide the essentials of Python installation and the intricacies of

setting up **integrated programming environments** (**IDEs**) for Python coding. We will also introduce the pivotal packages and libraries for finance and data analysis. Python is well-known for its friendly syntaxes. The chapter will describe essential principles of syntax, basic operations, control flow, and data types.

So, whether you are a seasoned finance professional looking to enhance your analytical toolkit or a programming enthusiast keen on financial data analysis, this chapter promises to be an enlightening starting point for a further journey in the world of Python for Finance. This combination revolutionized the financial world.

Structure

This chapter covers the following topics:

- Finance principles and contemporary trends in data analysis
- Comparison of analytical tools for various programming languages
- Installing Python and using IDEs for financial data analysis
- Overview of the mainly used Python libraries for finance
- Python essentials: Syntax, basic operations, control flow and data types

Objectives

By the end of this chapter, you will have a foundational knowledge of Python's role in finance and its advantages over other programming languages. You will investigate installing and configuring Python on local computers or using the Google Colab cloud platform. This chapter provides an overview of the top libraries for solving financial problems with Python. It also illustrates the fundamentals of the Python programming language, including syntax, operators, and basic data structures, including those related to financial data analysis.

Finance principles and contemporary trends in data analysis

Typically, the emergence of the notion of finance is intertwined with the inception of the earliest states and the nascent trading and market relations. However, finance relates to managing financial assets in its contemporary practical meaning. Hence, the paramount objective in financial management is maximizing potential returns and profits from utilizing these assets.

Therefore, the term *Finance* often refers to operations in the financial markets, particularly transactions involving the profitability bay or sale of financial assets (stocks, bonds, derivatives, cryptocurrencies, etc.). This book will explore the primary analytical algorithms using the Python programming language in this domain. Nonetheless, most of

the analytical or managerial tools discussed can be used for other financial tasks, such as individual or corporate budget planning, risk assessment and forecasting, and formulating analytical reports to make financial decisions. For those unacquainted with the finance details, this book will describe the essential principles of managerial decisions on financial markets and core terms of financial assets. It delves into the world of financial information and computation.

Inherently quantitative entities, finance invariably involves the numerical representation of outcomes: stock price, profits, turnovers, or past losses. This intrinsic quantitativeness positions finance near the modern world of information technology. In some respects, the valuation of financial assets is dictated by informational factors. For instance, the global financial crisis of 2008-2009 was instigated by discernible fundamental factors. However, this crisis's precursors, or the weak signals, manifested well before its peak. Individual investors and financial institutions that astutely identified these informational cues either minimized their losses or even capitalized on the repercussions of the crisis. The cinematic depiction in *The Big Short* (2015) is recommended for an insightful exploration of this theme. The film artistically demonstrates examples of the informational aspects of finance and might even inspire risky financial instincts.

The simultaneous development of innovations in the field of finance and IT technologies has led to a sharp increase in the volume of data, including financial data, which requires the adoption of advanced data processing methodologies. According to data from Statista.com, by 2025, the volume of data created, captured, copied, and consumed globally is projected to double compared to 2021, reaching 181 zettabytes (**Source**: Statista, **https://www.statista.com/statistics/871513/worldwide-data-created/**). Thus, a mere momentary lapse in financial decision-making could translate to substantial losses or foregone profits in this high-velocity digital age. Even a slight delay or a small mistake when making financial decisions can lead to lost profits or significant financial losses. Consequently, using tools offered by modern high-level programming languages (R, Python, Java, etc) has become indispensable for finance, and even parts of most analytical software products.

The conceptual view of gaining profit from investing in financial assets using Python (the general logic of mastering the material of this book) can be illustrated in *Figure 1.1*:

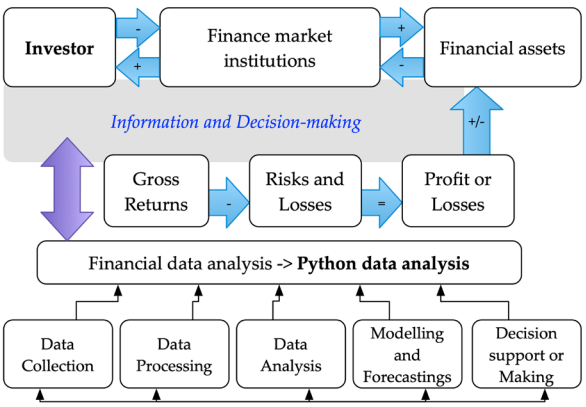

***Figure 1.1**: Outcomes workflow for Python for finance*

Figure 1.1 depicts both the participants in the financial markets and the tools utilized for decision-making in the management of the portfolio of financial assets:

Financial investor

A participant in financial relations with liquidity (financial assets that can easily be exchanged for other assets or services, such as cash). Think of Bob, a software developer who just received his year-end bonus. He is an individual investor looking to grow that bonus by investing in the financial market. An investor can be an individual or a financial institution making a profit in financial markets. The distinguishing characteristic of an investor in the turnover process of financial assets is the owner of the initial capital for investment. Active investors might decide to conduct all financial calculations and forecasts to manage their financial investment portfolio using all analytical tools, including financial data analysis tools from Python. Passive investors delegate some or all analytical and decision-making functions to financial institutions. However, both use financial data analysis tools to maximize profit and minimize losses.

Financial market institutions

This category comprises financial and service (consulting) institutions offering diverse services to access financial market opportunities. Such institutions include financial asset traders, stock exchanges, consulting firms, analytical bureaus, rating agencies, etc. Let us understand each, one by one:

- **Exchange institutions**: Participants in any financial market (stock market, bond market, options, or cryptocurrency) need an institution that provides legal compliance and a certain level of investor protection. For example, imagine a busy market where buyers and sellers come to exchange goods. Without the market administration, police, and even cleaners, this market could not exist. Likewise, the exchange institution (usually an exchange) helps financial investors buy, sell, or exchange financial assets. Modern financial exchanges (stock, foreign exchange, cryptocurrency, etc.) have expanded their role. They can also provide consulting and trading services. However, the primary function of the financial exchange institution remains to fulfil client's orders and conduct settlement transactions.
- **Financial market traders**: These are individuals or entities actively trading in the financial markets, either on their behalf or for clients (individual and corporate investors). Imagine someone at multiple computer screens, watching the numbers and graphs move, deciding when to buy or sell—that is a trader in action. They bring liquidity (supply and demand) to the markets, ensuring that securities can be bought or sold anytime. Thanks to traders, we can be sure that the valuable assets that we have in our possession can be sold to someone. Traders use various investment strategies based on analytical calculations and the preferences of their clients.

- **Consulting institutes**: These include rating and information agencies, trust companies, and mutual funds, which are key players in the financial sector. They offer various services, from one-time consultations for purchase and sale decisions to comprehensive management of investment portfolios. These institutes serve as financial guides, advising clients on their financial tasks and problems. Rating agencies show the creditworthiness of various organizations, from corporations to governments. News agencies provide essential analytical data for informed trading. Trust companies and mutual funds are responsible for professionally financial assessing and managing investors' resources. They devise and execute priority investment strategies for maximum profit. In essence, these institutes combine consulting, trading, and exchange functions.

Intertwining the needs of financial institutions and investors ensures a dynamic and fluid financial market. Therefore, analytical skills are critical for all participants in financial relations.

Two critical finance categories

For finance, as illustrated in *Figure 1.1*, two critical categories emerge: Return (profit, revenue) and Risk (potential losses). Imagine setting out on a sea voyage. The **Returns** are the treasures you hope to find, and the **Risks** are the turbulent waters and pirates you might encounter on the way. These twin pillars are the basis of any financial analysis.

When an analyst forecasts high returns based on financial computations and constructs a portfolio of financial assets without adequately accounting for risk, potential losses may culminate in eventual deficits. Conversely, the net investment profit will likely approximate zero if an investor's strategy is singularly oriented towards risk minimization. Further, when inflation's impact is not considered, this may even result in a capital reduction. Hence, the deployment of analytical tools, both within this book and in broader financial practice, usually aims at gauging and prognosticating return metrics from the utilization of financial assets and potential associated losses.

However, these categories can be assessed using various metrics or indicators. This significantly increases the volume of financial information and other data used in making management decisions in finance. Imagine a dashboard full of dials, each representing indicators such as speed, direction, and weather conditions. All the dashboard data is essential to reaching your destination port successfully. Financial indicators often include qualitative and quantitative measures. For example, knowledge of past market sentiment, prices, and trading volumes. They can be compared to a ship's compass (quantitative) and a sailor's intuition based on experience (qualitative). Balancing risk and return are the cornerstone of sound financial management. After all, the pursuit of profits can lead to bankruptcy. It is as if we are going around the reefs, trying to reach our destination port with the treasure faster than the competitors, but successfully.