
Supply Chain Analytics

Kurt Y. Liu

Supply Chain Analytics

Concepts, Techniques and Applications

palgrave
macmillan

Kurt Y. Liu
Adam Smith Business School
University of Glasgow
Glasgow, UK

ISBN 978-3-030-92223-8 ISBN 978-3-030-92224-5 (eBook)
<https://doi.org/10.1007/978-3-030-92224-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission. No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6-10 Kirby Street, London EC1N 8TS. Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages

-

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To my family, who have always encircled me
with love, patience, and support—my
parents, my wife, and my sons.*

—K.Y.L

Preface

Supply chain analytics is an emerging topic in both academia and industries as using data to get useful insights into effective decision making in supply chain and logistics management field is gaining strong momentum in recent years. However, there is a lack of textbooks that specifically focus on applying data analytics in supply chain management, while majority of available texts on the market are written for subjects such as business analytics and marketing analytics.

The main objective of this book is to introduce how data analytics and machines learning can be applied in the supply chain management field to provide meaningful insights for supply chain and logistics managers in their daily decision making. To achieve this objective, the book is structured in such a way that the readers can learn not only the fundamental concepts of supply chain management, but also the relevant data analytics techniques and the applicable machine learning algorithms for solving practical problems.

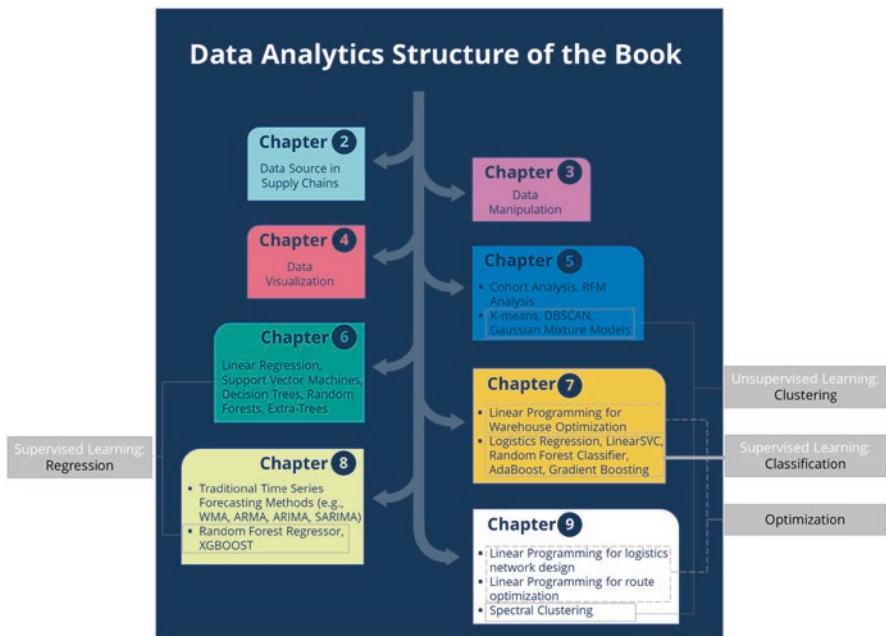
A particular focus of this book is on *application* rather than developing the various data analytics and machine learning algorithms in that the intended audience of this book is within the business management domain. Though having a computer science or mathematics background may have an advantage in understanding the content covered in this book, it is not essentially required for readers to follow through the book.

When writing this book, I have also tried to draw a clear line between supply chain analytics and operations research (OR) although these two subjects have many overlaps. However, OR is a well-established research domain which primarily involves the construction of mathematical models for optimization problems; whereas supply chain analytics focuses on the use of massive data and learning from them to improve decision making for effective supply chain management. We have only slightly touched upon linear programming for solving optimization problems in warehousing, logistics network design, and route optimization; other than that, no further topics from OR are included in the book.

This book is aimed at introductory level of supply chain analytics for those who have limited background in supply chain management, data analytics, and machine learning. It does not contain advanced machine learning algorithms and deep learning techniques and should be appropriate for postgraduate students in business and management subjects, MBA students, and senior undergraduate students, who have

interests in both supply chain management and data analytics. The book could also be useful as a reference book for practitioners in relevant industries looking for learning analytics and digital transformation in their supply chain and logistics operations.

Each chapter in this book generally contains key learning outcomes, concepts, techniques, and algorithms as well as practical examples, followed by discussion questions and keyboard exercise in the end. All the practical examples and analytics models introduced in this book are developed with Python; however, it does not prevent interested learners from applying alternative programming languages to solve these problems. A detailed structure of the data analytics introduced in this book is revealed in the diagram given below:



For Instructors

All the instructor resources can be found at Palgrave Macmillan's resource centre. The additional Python Jupyter notebooks, exercises, and solutions for each related chapter can be found at our online GitHub repository: <https://github.com/pysca/SCA>.

For Students

This book can be used by students who have no programming background; however, it does not intend to cover all Python basics in programming. Students are highly recommended to use complementary textbooks while studying this book to

gain a better understanding of Python and its various libraries and packages. In addition to the online GitHub repository provided above, where students can find all the Jupyter notebooks, exercises, and solutions for each chapter, there are numerous free online resources available (e.g., GitHub and Kaggle) for students to practise their coding and programming skills.

Glasgow, UK

Kurt Y. Liu

Acknowledgements

I would like to thank the reviewers whose suggestions significantly helped shape the structure and contents of the book. I would also like to thank the editors and the publisher for their efforts and support throughout the process. I am forever grateful to my former professor and friend at ASU, Prof. Sue P. Siferd, who has always believed in me and encouraged me to reach higher in my career. She will be deeply missed.

“The textbook Supply Chain Analytics by Kurt Liu provides an excellent hands-on introduction to supply chain management and logistics and associated quantitative methods from a data-science perspective using Python. The book first covers supply chain management basics and how to get started in data science using Python including data handling, pre-processing and visualization. Fundamental supply chain activities, data analysis, forecasting and decision making on the customer and supplier side, in inventory management, and for location and routing in logistics are introduced. The didactical concept to integrate topical basics and to present fundamental data analysis and decision support methods with examples coded in Python gives students a unique introduction to both important areas and the required knowledge and skills in supply chain analytics. I highly recommend this book for class adoption.”

—**Professor Dr Stefan Minner** Chair of Logistics and Supply Chain Management,
TUM School of Management, Technical University of Munich, Germany

“Supply Chain Analytics by Kurt Liu combines the essence of supply chain strategy with decisions taken at strategic, tactical, and operational levels and how data can support these decisions. Dr Liu makes this title appealing to supply chain students and researchers by integrating analytics themes with their execution in Python, which is among the top preferred data science languages. The treatment of data manipulation before any modelling exercise is commendable since the data from the real world are hardly ever ready to use. It is exciting to find the state-of-the-art prediction algorithms applied to supply chain problems. This book will join the reading lists of many supply chain programmes.”

—**Prof Dr Emel Aktas** Chair of Supply Chain Analytics, Cranfield School
of Management, Cranfield University, United Kingdom

“Data analytics is transforming the practice of supply chain management. This excellent primer provides an essential introduction of supply chain analytics that can benefit all interested learners, from novices to experienced ones. It offers an insightful synthesis of supply chain management and data analytics.

A truly impressive book on the topic that is well developed, very readable, and full of real-life examples that illuminate the underlying principles and emerging trends. Worth reading and highly recommended.”

—**Benjamin B. M. Shao** Professor of Information Systems, Director of the
Digital Society Initiative, W. P. Carey School of Business,
Arizona State University, United States of America

Contents

1	Introduction	1
1.1	What Is a Supply Chain?	1
1.1.1	Why Do We Need a Supply Chain?	2
1.1.2	Structure of a Supply Chain	2
1.1.3	Supply Chain Processes	4
1.1.4	Supply Chain Flows	6
1.2	Supply Chain Management	7
1.3	Business Analytics	9
1.4	Supply Chain Analytics	12
1.4.1	SMART Goals of SCA	13
1.4.2	SCA	15
	References	19
2	Data-Driven Supply Chains and Intro to Python	21
2.1	Data and Its Value in SCM	21
2.2	Data Source in Supply Chains	24
2.3	Big Data	30
2.4	Introduction to Python	31
2.4.1	Python Downloads and Installation	33
2.4.2	Python IDE and Jupyter Notebook	33
2.4.3	Essential Python Libraries	34
2.4.4	Jupyter Notebook Optimization	38
	References	43
3	Data Manipulation	45
3.1	What Is Data Manipulation?	45
3.2	Data Loading and Writing	47
3.3	Data Indexing and Selection	50
3.4	Data Merging and Combination	55
3.5	Data Cleaning and Preparation	62
3.6	Data Computation and Aggregation	67
3.7	Working with Text and Datetime Data	75

- 4 Data Visualization 83**
 - 4.1 Data Visualization in Python 83
 - 4.2 Creating a Figure in Python 85
 - 4.3 Formatting a Figure 91
 - 4.4 Plotting Simple Charts 93
 - 4.5 Plotting with Seaborn 100
 - 4.6 Geographic Mapping with Basemap 103
 - 4.7 Visualizing Starbucks Locations 106

- 5 Customer Management 113**
 - 5.1 Customers in Supply Chains 114
 - 5.2 Understanding Customers 115
 - 5.2.1 Benefits of a Customer-Centric Supply Chain 116
 - 5.3 How to Build a Customer-Centric SC 117
 - 5.3.1 Step One: Define Customers 118
 - 5.3.2 Step Two: Understand Customers’ Real Needs 118
 - 5.3.3 Step Three: Translate Needs into Product Features 119
 - 5.3.4 Step Four: Design Supply Chain Processes 119
 - 5.3.5 Step Five: Build Efficient Logistics Systems 120
 - 5.4 Cohort Analysis 121
 - 5.4.1 What Is a Cohort? 121
 - 5.4.2 What Is Cohort Analysis? 121
 - 5.4.3 Steps for Cohort Analysis 122
 - 5.4.4 Cohort Analysis Example in Python 123
 - 5.5 RFM Analysis 129
 - 5.5.1 What Is RFM? 129
 - 5.5.2 What Is RFM Analysis? 130
 - 5.5.3 Steps for RFM Analysis 131
 - 5.5.4 RFM Analysis Example in Python 133
 - 5.6 Clustering Algorithms 139
 - 5.6.1 K-Means Algorithm 140
 - 5.6.2 Customer Segmentation with K-Means 145
 - 5.6.3 DBSCAN 151
 - 5.6.4 Gaussian Mixture Model 155
 - References 162

- 6 Supply Management 163**
 - 6.1 Procurement in Supply Chains 164
 - 6.1.1 Vertical Integration 164
 - 6.1.2 Outsourcing 165
 - 6.2 Supplier Selection 168
 - 6.3 Supplier Evaluation 171
 - 6.3.1 Supplier Capability Assessment 173

6.4	Supplier Relationship Management	174
6.4.1	Managing Tiered Supply Network	175
6.5	Supply Risk Management	178
6.5.1	Step One: Risk Identification	179
6.5.2	Step Two: Risk Assessment	179
6.5.3	Step Three: Develop Risk Response Strategies	180
6.5.4	Step Four: Continuous Monitoring and Regular Review	180
6.6	Supplier Selection Examples	182
6.6.1	Coffee Quality Example	182
6.7	Regression Algorithms	192
6.7.1	Linear Regression	192
6.7.2	Support Vector Machines	196
6.7.3	Decision Trees	202
6.7.4	Cross-Validation	206
6.7.5	Random Forests	208
6.7.6	Model Fine-Tuning	210
6.7.7	Extra-Trees	214
	References	218
7	Warehouse and Inventory Management	219
7.1	Warehouse Management	220
7.1.1	What Is Warehouse Management?	220
7.1.2	Warehouse Management System	222
7.1.3	Benefits of WMS	223
7.1.4	Warehouse Management Performance Measurement	224
7.2	Inventory Management	225
7.2.1	What Is Inventory?	225
7.2.2	The Purpose of Holding Inventory	226
7.2.3	What Is Inventory Management?	228
7.2.4	Inventory Management Methods	229
7.3	Warehouse Optimization	236
7.3.1	Introduction to PuLP	236
7.3.2	Warehouse Optimization with PuLP	240
7.4	Classification Algorithms	243
7.4.1	Logistic Regression	243
7.4.2	Classification Performance Measurement	248
7.4.3	Dealing with Imbalanced Dataset	253
7.4.4	Linear Support Vector Classifier	258
7.4.5	Random Forest Classifier	260
7.4.6	Boosting Methods for Classification	262
	References	269

8	Demand Management	271
8.1	Demand Management	272
8.2	Demand Forecasting	273
8.3	Time Series Forecasting	275
8.3.1	Time Series Components	275
8.3.2	Traditional Time Series Forecasting Methods	277
8.4	Machine Learning Methods	302
8.4.1	Univariate vs Multivariate Time Series	302
8.4.2	Random Forest Regression	303
8.4.3	XGBoost	308
8.4.4	Learning Rate	311
	References	318
9	Logistics Management	319
9.1	Logistics Management	320
9.1.1	What Is Logistics Management?	320
9.1.2	Main Logistics Management Activities	322
9.2	Modes of Transport in Logistics	323
9.2.1	Chargeable Weight	323
9.2.2	Product Value Density	325
9.3	Logistics Service Providers	327
9.3.1	Freight Companies	327
9.3.2	Freight Carriers	328
9.3.3	Freight Forwarders	328
9.3.4	Third Party Logistics (3PL) providers	328
9.3.5	Fourth Party Logistics (4PL) Companies	329
9.4	Global Logistics Management	329
9.4.1	Incoterms	331
9.5	Logistics Network Design	332
9.5.1	Location Decisions	333
9.5.2	Centralization vs De-centralization	335
9.5.3	Logistics Network Design Example with PuLP	337
9.6	Route Optimization	342
9.6.1	Travelling Salesman Problem	342
9.6.2	Vehicle Routing Problem	348
9.6.3	Route Optimization Example: The Cambridge Pizza Store	350
	References	369
	Index	371

About the Author

Kurt Y. Liu is an Associate Professor of Supply Chain Analytics at the Adam Smith Business School, University of Glasgow. He teaches Executive, MBA, graduate, and undergraduate students on supply chain, logistics, and operations management. His research focuses on supply chain analytics, supply network configuration, and sustainable supply chain management. He studied at Waikato Management School, University of Waikato, New Zealand, and W.P. Carey School of Business, Arizona State University, USA, for his bachelor's degree in Supply Chain Management. He has Master's in Logistics Management and Master's in Transport Management from the Institute of Transport and Logistics Studies (ITLS), University of Sydney, Australia, and a PhD in Engineering from the University of Cambridge, UK. He is a Fellow of the Higher Education Academy (FHEA) and a Chartered Member of The Chartered Institute of Logistics and Transport (CILT), UK.