

Sign up for our newsletter and get the latest big data news and analysis.

Email Address

Subscribe

[Home](#) » [Book Review](#) » Book Review: Julia for Data Science by Zacharias Voulgaris, Ph.D.

## Book Review: Julia for Data Science by Zacharias Voulgaris, Ph.D.

January 10, 2017 by [Daniel Gutierrez](#)

[Leave a Comment](#)



25



41



5



71  
SHARES

Here's a useful new book for data scientists looking to approach the field from a unique perspective that doesn't include language heavyweights like R and Python. "Julia for Data Science," by Zacharias Voulgaris, Ph.D. from [Technics Publications](#), allows you to master the Julia language to solve business critical data science challenges. But why look to a relatively new language when you already have other commonly-used languages at your disposal? Noted data scientist John Myles White wrote a compelling blog post "[Julia, I Love You](#)" which presents a number of good reasons to give Julia a chance. Pitted against R for example, Julia is a lot faster and less quirky as a programming language.



If you want to give Julia a look, "Julia for Data Science" is a good place to start as a learning resource. After covering the importance of Julia to the data science community and several essential data science principles, the book starts with the basics including how to install Julia and its libraries. Many examples are provided as the author illustrates how to leverage each Julia command, data set, and function.

The book introduces and describes specialized script packages. Hands-on problems representative of those commonly encountered throughout the data science pipeline are provided, and the book guides you in the use of Julia in solving them using published data sets. Many of these scenarios make use of existing packages and built-in functions. Here is a breakdown of the chapters:

Chapter 1: Introducing Julia – sets the stage by discussing how Julia can be used for data science.

Chapter 2: Setting Up the Data Science Lab – talks about the Julia IDE and how to use the language in the solution of data science problems. As an example, the author implements a simple kNN (k nearest neighbors) algorithm in Julia.

Chapter 3: Learning the Ropes of Julia – covers basic Julia language constructs,

commands and functions.

Chapter 4: Going Beyond the Basics in Julia – covers more advanced Julia language components, and implements a simple "skewness" algorithm.

Chapter 5: Julia Goes All Data Science-y – discusses the "data science pipeline," nice touch if you want to get all "data science-y!"

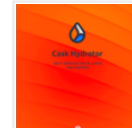
### INDUSTRY PERSPECTIVES

#### Swing and a Miss - How to Avoid Striking Out with the Wrong Data

In this special guest feature, Mark Budzinski, CEO at WhereScape, discusses how baseball can teach valuable data management lessons for enterprise by outlining how to get the right mix of automation and intelligent human analytical input. Context with any data set is just as critical to the result as is the collection of such data. [\[Read More...\]](#)



### WHITE PAPERS



#### Cask Hydrator: Self Service Data Lakes on Hadoop

The Cask Hydrator empowers organizations to blend, transform, and analyze data across multiple sources to create an on-demand analytical view across key customer interactions. This provides powerful insights into buyers, brand performance, products, and services.

[Download](#)

[See More White Papers »](#)

### LATEST VIDEO

[Video Archive](#)

FIND US ON:



Chapter 6: Julia The Data Engineer – shows how Julia can be used for typical data wrangling tasks.

Chapter 7: Exploring Data Sets – shows how Julia can be used for typical exploratory data analysis.

Chapter 8: Manipulating the Fabric of the Data Space – digs into dimensionality reduction techniques like principal component analysis (PCA), and feature evaluation.

Chapter 9: Sampling Data and Evaluating Results – reviews data sampling methods and model evaluation techniques like MSE, RMSE, and SSE metrics, plus cross validation.

Chapter 10: Unsupervised Machine Learning – focuses on unsupervised statistical learning techniques like k-means and hierarchical clustering.

Chapter 11: Supervised Machine Learning – focuses on supervised statistical learning techniques like: decision trees, regression trees, random forests, neural networks, etc.

Chapter 12: Graph Analysis – examines a different kind of modeling called graph analysis and uses two Julia packages, Graphs and LightGraphs.

Chapter 13: Reaching the Next Level – a number of tips for how to go beyond the book in using Julia for data science.

Each chapter concludes with a series of questions and exercises to reinforce what you learned.

I enjoyed this book: it explains exactly what the title says: Julia for Data Science! As outlined above, the first few chapters are sort of Julia 101, but then the book shifts gears and goes deeper into data science. As I longtime R developer, I appreciated the conciseness and the speed of Julia and the variety of libraries available for data science. There are other books available on Julia, but many of them are too CS-oriented or too general purpose. With the main focus on data science this one is different, allowing newbies or experienced data scientists to move into first steps and becoming proficient in Julia in a matter of weeks. A few chapters into the book, I'm sure most readers will be ready to experiment independently, effectively changing and shaping the provided code to test Julia on professional projects.

The author, Dr. Zacharias Voulgaris has worked at Georgia Tech as a Research Fellow, at an e-marketing startup in Cyprus as an SEO manager, and as a Data Scientist in both Elavon (GA) and G2 (WA). He also was a Program Manager at Microsoft, on a data analytics pipeline for Bing.



*Contributed by Daniel D. Gutierrez, Managing Editor of insideBIGDATA. In addition to being a tech journalist, Daniel also is a practicing data scientist, author, educator and sits on a number of advisory boards for various start-up companies.*

Sign up for the free insideBIGDATA [newsletter](#).



Filed Under: [Book Review](#), [Data Science](#), [Featured](#), [Google News Feed](#), [News / Analysis](#), [Uncategorized](#)

Tagged With: [data science](#), [Julia](#), [Weekly Newsletter Articles](#)

**Leave a Comment**

Name \* Email \* Website

- Notify me of follow-up comments by email.
- Notify me of new posts by email.



Copyright © 2017

[About insideBIGDATA](#)

[Contact](#)

[Advertise with insideBIGDATA](#)

[Visit Our Other Site - insideHPC](#)

[Terms of Service & Copyright](#)

[Privacy Policy](#)