



UNIVERSIDAD  
CENTRAL

LIBROS

Referencias sugeridas

---

Jesús Alfredo Sanabria Mejía

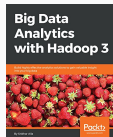
0060 Bases de Datos

24 de septiembre de 2022

# Libros Observados

1. Big Data Analytics with Hadoop 3.
2. Hadoop in Practice 2nd Edition
3. Hands-On Python Natural Language Processing
4. Next generation databases: nosql and big data
5. SQL NoSQL Databases: Models Languages

# Big Data Analytics with Hadoop 3



[Alla, 2018]

- Editorial Packt Publishing (472 paginas)
- Edición (29/05/2018)
- Idioma Inglés
- ISBN 781788628846



Sridhar Alla: es un Licenciado en Informática Indio, experto en big data que ayuda a las empresas a resolver problemas complejos en sistemas de computo distribuidos, ciencia de datos a gran escala y práctica de análisis.

El libro es una guía práctica de Apache Hadoop para el procesamiento de big data y enseña a combinar con una gran cantidad de otras herramientas de big data para crear soluciones de análisis. cuenta con la orientación mediante una ayuda de ejemplos prácticos. con códigos gratuitos que se puede descargar desde github

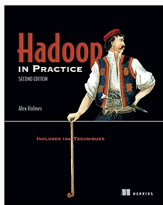
<https://github.com/packtpublishing/big-data-analytics-with-hadoop-3>

# Big Data Analytics with Hadoop 3



- 1. Introducción a Hadoop
- 2. Descripción general de Big Data Analytics
- 3. Procesamiento de Big Data con Mapreduce
- 4. Computación Científica y Grandes Análisis de datos con Python y Hadoop
- 5. Informática estadística de grandes datos con R y Hadoop
- 6. Análisis por lotes con Apache Spark
- 7. Análisis en tiempo real con Apache Spark
- 8. Análisis por lotes con Apache Flink
- 9. Procesamiento de Stream con Apache Flink
- 10. Visualización de Big Data (Herramientas)
- 11. Introducción al Computación en la Nube
- 12. Utilización de Servicios Web en Amazon.

# Hadoop in Practice 2nd Edition



[Holmes, 2015]

- MANNING Shelter Island(513 paginas)
- Edición (2015)
- Idioma Inglés
- ISBN 9781617292224



Alex Holme: es un ingeniero de software con experiencia en la resolución de problemas de BigData mediante Hadoop. Ha presentado en multiples congresos y es líder técnico en VeriSign (empresa USA de infraestructuras de red).

El libro presenta una guía en hadoop práctica donde recopila cerca de 85 ejemplos de Hadoop y los presenta en un formato de problema/solución. Cada técnica aborda una tarea específica, tratando de presentar un problema paso a paso, aprendiendo cómo construir e implementar esa solución específica junto con el pensamiento que se utilizó en su diseño

# Hadoop in Practice 2nd Edition(5 Partes 13 capitulos)



Los códigos de Ejemplo son de acceso libre por github <https://github.com/alexholmes/hiped2>

- Antecedentes y fundamentos
- logística de datos
- Patrones de Big data
- Ciencia de datos
- Domar al elefante

en el Libro se encuentra un Resumen conceptual de Hadoop y MapReduce  
85 técnicas prácticas y probadas acerca de Problemas reales, soluciones reales y con ejemplos muestra como integrar MapReduce y R

# Hands-On Python Natural Language Processing



[Kedia and Rasu, 2020]

- Editorial BIRMINGHAM - MUMBAI (304 paginas)
- Edición (2020)
- Idioma Inglés
- ISBN 978-1-83898-959-0



Mayank Rasu Vicepresidente, Análisis Cuantitativo - Desarrollo de Modelos en Barclays Investment Bank tiene más de 12 años de experiencia global como científico de datos y analista cuantitativo en la industria de banca de inversión.

Aman Kedia es promotor de la Inteligencia artificial (IA) y los algoritmos que la soportan. Ha trabajado en problemas del procesamiento del lenguaje natural (NLP), que abarcan la comparación de currículums y asistentes digitales, entre otros. Ha trabajado en Oracle y SAP, tratando de resolver problemas que aprovechan los avances en IA.

Este libro presentará información del Procesamiento de Lenguaje Natural PNL y sus aplicaciones, junto con modernas bibliotecas de Python que sirven de ayuda a ejemplos prácticos, para crear aplicaciones de NLP. El libro cubre varias metodologías y desafíos en la implementación de aplicaciones de NLP. Cubre tareas clave de NLP, como la clasificación de texto, la incrustación semántica, el análisis de sentimientos, la traducción automática y el desarrollo de un chatbot utilizando técnicas de aprendizaje automático y aprendizaje profundo.

# Hands-On Python

## Natural Language Processing (11 capítulos)

Los códigos de Ejemplo son de acceso libre por github

<https://github.com/packt publishing/hands-on-python-natural-language-processing>

- 1. conceptos básicos de PNL
- 2. PNL usando Python
- 3. Construyendo vocabulario de PNL
- 4. Transformar texto en Estructuras de datos
- 5. incrustaciones de palabras y Mediciones para Textos
- 6. Exploración de sentencias en documentos e incrustaciones a nivel de documento y carácter.
- 7. Identificación de patrones en textos usando machine learnig
- 8. De las neuronas humanas a Neuronas artificiales para Comprender el texto
- 9. Aplicando Convoluciones l texto
- 10. Capturando relaciones temporales en textos
- 11. Estado del Arte en NLP



# Next generation databases:

## nosql and big data



[Harrison, 2020]

- Editorial APRESS / Springer (381 paginas)
- Edición (2020)
- Idioma Inglés
- ISBN 978-1-4842-1330-8



Guy Harrison (Australia) comenzó a trabajar como desarrollador de bases de datos a mediados de la década de 1980. Él ha escrito numerosos libros sobre desarrollo y rendimiento de bases de datos. Se unió a Quest Software (ahora parte de Dell) en 2000 y actualmente lidera el equipo que desarrolla Toad, Spotlight y Shareplex familias de productos.

Este libro describe para qué se diseñó cada tecnología de Base de Datos. Muestra cómo se puede usar cada tecnología para resolver problemas reales de aplicaciones y negocios. Lo que es más importante, este libro destaca las diferencias arquitectónicas entre las tecnologías que son los factores críticos a considerar al elegir una plataforma de base de datos para proyectos nuevos y futuros.

# Next generation databases:

nosql and big data (13 capítulos)

- Bases de datos de próxima generación
- Tres revoluciones de bases de datos
- Google, grandes datos y Hadoop
- Sharding(fragmentación), Amazon y el nacimiento de NoSQL
- Bases de datos de documentos
- Las tablas no son tus amigas: bases de datos de gráficos
- Bases de datos de columnas
- ¿El fin del disco? SSD y bases de datos en memoria
- Patrones de bases de datos distribuidas
- Modelos de consistencia
- Modelos de datos y almacenamiento
- Lenguajes e Interfaces de Programación

# SQL NoSQL Databases:

## Models, Languages



[Meier and Kaufmann, 2019]

- Springer Vieweg (238 paginas)
- Edición (2019)
- Idioma Inglés
- ISBN 9783658245481



Michael Kaufmann (Alemania) es profesor de Ciencia de Datos y Big Data en la escuela de tecnologías de la información de la Universidad en Lucerna, Suiza, donde como investigador ha fundado y conseguido financiar el equipo de investigación de inteligencia de datos. Andreas Meier fue miembro de la Facultad de Economía y Ciencias Sociales de Friburgo donde fue profesor de Tecnologías de la Información. Se especializa en empresa electrónica, gobierno electrónico y gestión de la información.

Este libro incluye más de 100 tablas, ejemplos e ilustraciones, y cada capítulo ofrece una lista de recursos para lecturas adicionales. En el libro se explican las fortalezas y debilidades de los enfoques relacionales y no relacionales, y muestra cómo emprender el desarrollo de aplicaciones de big data.

# SQL NoSQL Databases:

## Models, Languages(7 capítulos)

- 1. Administración de Datos
- 2. Modelado de datos
- 3. Lenguajes de Bases de Datos
- 4. Garantizando la consistencia de base de  
\_ datos
- 5. Arquitectura del Sistema
- 6. Bases de datos posrelacionales
- 7. Bases de datos NoSQL

Todos los capítulos finalizan con una lista comentada de sugerencias y referencias bibliográficas.

# Bibliografía i



Alla, S. (2018).

***Big Data Analytics with Hadoop 3: Build highly effective analytics solutions to gain valuable insight into your big data.***

Packt Publishing, Birmingham, England.



Harrison, G. (2020).

***Next generation databases: Nosql, NewSQL, and Big Data.***

Apress.



Holmes, A. (2015).

***Hadoop in Practice.***

Manning Publications, New York, NY, 2 edition.



Kedia, A. and Rasu, M. (2020).

***Hands-On Python Natural Language Processing: Explore tools and techniques to analyze and process text with a view to building real-world NLP applications.***

Packt Publishing, Birmingham, England.

# Bibliografía ii



Meier, A. and Kaufmann, M. (2019).  
*SQL NoSQL databases.*  
Springer Vieweg.

Thanks!