DAFTAR PUSTAKA

[1] C. Subramanian, S. Karthikeyan, and V. Sarala Devi, "A STUDY ON BIG DATA," *Int. J. Appl. Environ. Sci.*, vol. 10, p. 123, 2015.

[2] S. Mujawar and A. Joshi, "Data Analytics Type, Tools,and Their Comparison," *Int. J. Adv. Res. Comput. Commun. Eng.*, vol. 4, no. 2, pp. 488–491, 2015.

[3] R. V Zicari, "Big Data: Challenges and Opportunities," *Big Data Comput.*, pp. 103–128, 2013.

[4] Badan Pusat Statistik Provinsi DKI Jakarta, "Statistik Transportasi DKI Jakarta 2015," *Badan Pus. Stat. Provinsi DKI Jakarta*, 2015.

[5] Navint, "Why is BIG Data Important?" Navint, 2012.

[6] D. Sarkar, *Pro Microsoft HDInsight*. Apress, 2016.

[7] K. Santhiya, "Map Reduce Programming Model : Construction of Inverted Index for Automated Document Clustering," pp. 308–312, 2016.

[8] "Welcome To Hadoop." [Online]. Available: hadoop.apache.org. [Accessed: 03-Apr-2017].

[9] C.R. Rao, *HANDBOOK OF STATISTICS VOLUME 30*. 2012.

[10] P. J. Brockwell and R. A. Davis, *Time Series and Forecasting Methods*. 1991.

[11] A. Konwinski, Z. Holden Karau, P. Wendell, and M. Zaharia, *Learning Spark Spark LIGHTNING -FAST DATA ANALYSIS*, 1st ed. by O'Reilly Media, 2015.

[12] S. Gopalani and Rohan Arora, "Comparing Apache Spark and Map Reduce with Performance Analysis using K-Means," *Int. J. Comput. Appl.*, vol. 113, no. 1, p. 8887, 2015.

[13] G. Geethakumari and A. Srivatsava, "Big Data Analysis for Implementation of Enterprise Data Security," *Iracst.Org*, vol. 2, no. 4, pp. 742–746, 2012.

[14] M. Ahmed, "The Effect Of In-Memory Computing and Big Data on

Enterprise Software," Masarykova Univerzita, 2016.

[15] B. Y. Mike, "Big Data and Business Intelligence : a data-driven strategy for e-commerce organizations in the hotel industry Big Data and Business Intelligence : a data-driven strategy for e-commerce organizations in the hotel industry," University of Twente, 2015.

[16] Z. Han and Y. Zhang, "Spark: A Big Data Processing Platform Based on Memory Computing," *Proc. - Int. Symp. Parallel Archit. Algorithms Program. PAAP*, vol. 2016–Janua, pp. 172–176, 2016.

[17] M. Armbrust *et al.*, "Spark SQL: Relational Data Processing in Spark," 2015.

[18] D. Namiot, "Time series databases," in *CEUR Workshop Proceedings*, 2015.

[19] C. J. Wild and G. a. F. Seber, "Time series," *Chance Encount. A First Course Data Anal. Inference*, 1999.

[20] U. Jugel, Z. Jerzak, G. Hackenbroich, and V. Markl, "M4: a visualization-oriented time series data aggregation," *Proc. VLDB Endow.*, vol. 7, no. 10, pp. 797–808, 2014.

[21] M. Schaefer, "A Novel Explorative Visualization Tool for Financial Time Series Data Analysis," no. September, pp. 1–4, 2011.

[22] M. Cox and D. Ellsworth, "Application-controlled demand paging for out-of-core visualization," *Proceedings. Vis. '97 (Cat. No. 97CB36155)*, no. July, p. 235–244, 1997.

[23] R. Bryant, R. Katz, and E. Lazowska, "Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science and Society," *Comput. Res. Assoc.*, pp. 1–15, 2008.

[24] J. Anuradha, "A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology," *Procedia - Procedia Comput. Sci.*, vol. 48, pp. 319–324, 2015.

[25] Guru99, "What Is Big Data?," 2015. [Online]. Available: http://www.guru99.com/what-is-big-data.html.

[26] M. Ferguson, "Architecting A Big Data Platform for Analytics," *A*

Whitepaper Prep. IBM, no. October, pp. 1–36, 2012.

[27] M. Mayo, "Top Big Data Processing Framework," 2016. [Online]. Available: http://www.kdnuggets.com/2016/03/top-big-data-processing-frameworks.html. [Accessed: 30-Apr-2017].

[28] T. White, *Hadoop, The Definitive Guide*, 4th ed. California: O'Reilly, 2015.

[29] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, "The Hadoop distributed file system," *2010 IEEE 26th Symp. Mass Storage Syst. Technol. MSST2010*, pp. 1–10, 2010.

[30] Apache Software Foundation, "Sqoop User Guide." [Online]. Available: https://sqoop.apache.org/docs/1.4.1-incubating/SqoopUserGuide.html. [Accessed: 13-May-2017].

[31] "Streaming Programming Guide." [Online]. Available: http://spark.apache.org/docs/latest/streaming-programming-guide.html. [Accessed: 03-Apr-2017].

[32] D. Inc, "Monitoring and Instrumentation." [Online]. Available: http://spark.apache.org/docs/latest/monitoring.html. [Accessed: 13-May-2017].

[33] J. A. Scott, *Getting Started With Apache Spark*, 1st ed. San Jose: MapRTechnologies, Inc, 2015.

[34] H. Garcia-Molina *et al.*, *Database Systems: A Practical Approach to Design, Implementation, and Management*, vol. 49, no. 4. 2010.

[35] J. G. Politz *et al.*, "Python: The Full Monty A Tested Semantics for the Python Programming Language," *Oopsla*, 2013.

[36] M. Makai, "Why Use Python." [Online]. Available: https://www.fullstackpython.com/why-use-python.html.

[37] D. Inc, "Spark-SQL." [Online]. Available: https://github.com/databricks/spark-csv. [Accessed: 13-May-2017].

[38] D. Inc, "Spark-SQL API documentation." [Online]. Available: https://spark.apache.org/docs/1.6.2/api/java/org/apache/spark/sql/functions.html. [Accessed: 13-May-2017].

[39] J. Mitlohner, S. Neumaier, J. Umbrich, and A. Polleres, "Characteristics of open data CSV files," *Proc. - 2016 2nd Int. Conf. Open Big Data, OBD 2016*, vol. 838, pp. 72–79, 2016.

[40] Wikipedia, "Comma-Separated Values," 2005. [Online]. Available: https://en.wikipedia.org/wiki/Comma-separated_values. [Accessed: 29-Apr-2017].

[41] Apache Software Foundation, "Apache Parquet," 2014. [Online]. Available: https://parquet.apache.org/. [Accessed: 29-Apr-2017].

[42] "SQL Programming Guide," 2015. [Online]. Available: http://spark.apache.org/docs/latest/sql-programming-guide.html. [Accessed: 03-Apr-2017].

[43] D. J. Abadi, S. R. Madden, and N. Hachem, "Column-Stores vs . Row-Stores : How Different Are They Really ? Categories and Subject Descriptors," *Sigmod*, vol. June 9-12, pp. 967–980, 2008.

[44] S. Farooqui, "Building a modern Application with DataFrames." Databricks, p. 17.

[45] T. Unter, D. Liu, H. Von Hovell, R. Xin, and M. Zaharia, *Mastering Apache Spark*. .

[46] J. Wiley, *Data Visualization For Dummies*. John Wiley & Sons, Inc, 2014.

[47] T. Point, "Tableau - Overview." [Online]. Available: https://www.tutorialspoint.com/tableau/tableau_overview.htm. [Accessed: 13-May-2017].

[48] R. Shumway and D. Stoffer, *Time Series Analysis and Applications*, 4th ed. FREE DOG PUBLISHING, 2016.

[49] G. Casella, S. Fienberg, and I. Olkin, "Springer Texts in Statistics."

[50] Rob J Hyndman, "Forecasting: Forecasting: Principles & Practice," no. September, p. 138, 2014.

[51] W. Doble, "Automating Analytic Workflows on AWS," 2015. [Online]. Available: https://aws.amazon.com/blogs/big-data/automating-analytic-workflows-on-aws/. [Accessed: 07-May-2017].

[52] J. Wiley, *Big Data For Dummies*, 1st ed. New Jersey: John Wiley & Sons,

Inc, 2013.

[53] R. S. Pressman, *Software Engineering: A Practioner's Approach*. 2009.

[54] S. Liang, Z. Wang, and A. Fan, "Hadoop versus Spark," pp. 1–2, 2015.

[55] D. Anderson, "Column Oriented Database Technologies," 2012. [Online]. Available: https://www.dbbest.com/blog/column-oriented-database-technologies/. [Accessed: 29-Apr-2017].

[56] M. Yuk and S. Diamond, "Data Visualization: Choosing Simple and Effective Charts." [Online]. Available: http://www.dummies.com/programming/big-data/big-data-visualization/data-visualization-choosing-simple-and-effective-charts/. [Accessed: 13-May-2017].