



# Becoming a Data Scientist

Teguh Nugraha  
Statistika Ria 12  
2017



**Josh Wills**

@josh\_wills

Follow

Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.

9:55 AM - 3 May 2012

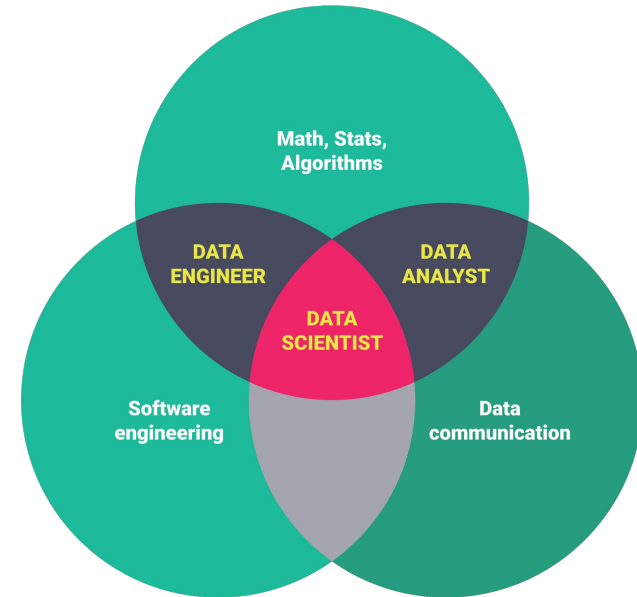
1,665 Retweets 1,346 Likes



52

1.7K

1.3K

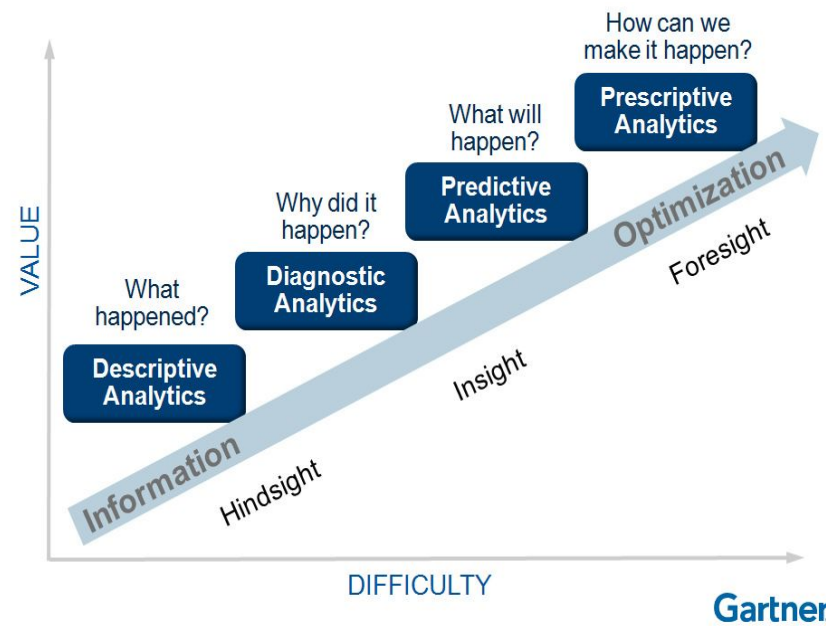


<https://www.springboard.com/blog/data-science-career-paths-different-roles-industry/>

# Data science goals and deliverables

## Analytics

- Descriptive Analytics (What happened?)
- Diagnostic Analytics (Why did it happen?)
- Predictive Analytics (What will happen?)
- Prescriptive Analytics (How can we make it happen?)








# Data science goals and deliverables

## Data Product/Artificial Intelligence

- Classification (e.g., spam or not spam)
- Recommendations (e.g., Amazon and Netflix recommendations)
- Anomaly detection (e.g., fraud detection)
- Pattern detection and grouping (e.g., classification without known classes)
- Automated processes and decision-making (e.g., credit card approval)
- etc

<https://www.innoarchitech.com/what-is-data-science-does-data-scientist-do>

Pembeli yang Melihat Barang Ini, Juga Tertarik dengan

 <p>Headphone Bluetooth Extra Bass Stereo <b>Rp300.000</b></p>	 <p>Jam Tangan Pria Skmei Original Model Casio <del>Rp380.000</del> <b>Rp285.000</b> ★★★★☆ 11 ulasan</p>	 <p>Jam Tangan Pria Skmei Tahan Air 50m Water <del>Rp260.000</del> <b>Rp182.000</b> ★★★★☆ 23 ulasan</p>	 <p>Headset Bluetooth Wireless Handsfree <b>Rp240.000</b> ★★★★★ 8 ulasan</p>	 <p>Headset Headphones Bass Beats With Mic HDF <b>Rp105.000</b> ★★★★☆ 6 ulasan</p>	 <p>Headphones Bluetooth Headphone Wireless <del>Rp300.000</del> <b>Rp225.000</b> ★★★★☆ 23 ulasan</p>
---	---	--	---	---	--



# Get Familiar with These

## Basic

- SQL
- Git and Github
- Command Line Interface
- R, Python
- Jupyter Notebook
- Google Analytics
- Statistical Inference and Probability
- AB Testing
- Machine Learning
- Cron Job

## Advanced

- NoSQL
- MapReduce
- Big Data Platforms (Hadoop, Hive)
- Deep Learning
- Parallel processing



---

# Self Learning

- [Coursera](#)
- [Datacamp](#)
- [edX](#)
- YouTube ([Data School](#), [Siraj Raval](#))
- [Udacity](#)
- [Data Science Specialization](#)
- [Analytics Academy](#)
- [SoloLearn](#) (SQL, Python)
- [Khan Academy](#)
- [Awesome Repository](#)
- Real cases: [Kaggle](#)



Thank You

teguh@bukalapak.com

