

19 April, 2021

# 15 Minutes Introduction to Data Analytics

Dr. Kevin Kam Fung Yuen  
Algorithm Artist

Email: [kevinkf.yuen@gmail.com](mailto:kevinkf.yuen@gmail.com)

# Learning outcomes

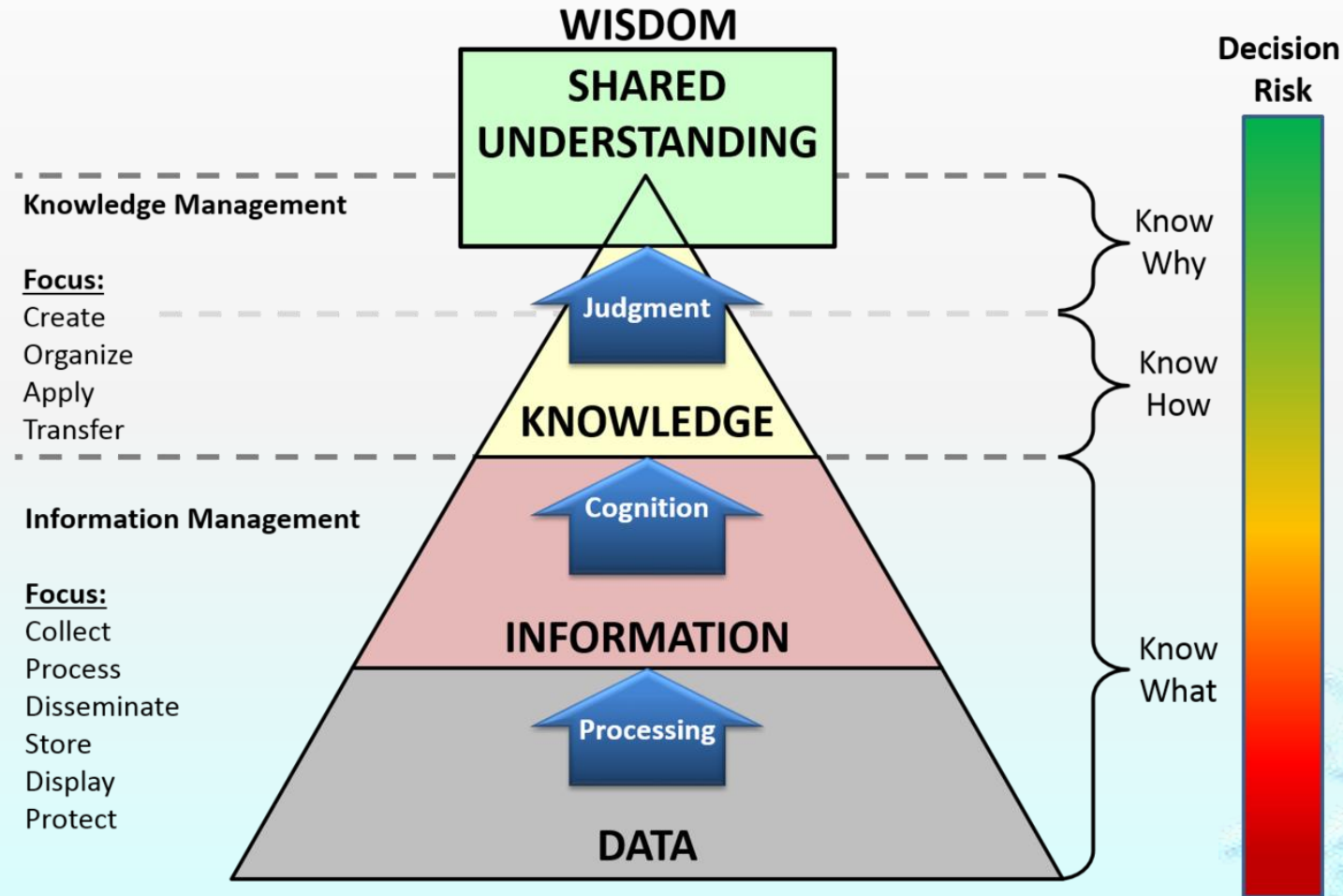
- ◆ Understand data analytics vs. Data science;
- ◆ Understand DIKW pyramid (Data, Information, Knowledge, and Wisdom);
- ◆ Understand the different steps in the advanced analytics: Descriptive, Diagnostic, Predictive, and Prescriptive Analytics;
- ◆ General process in Knowledge Discovery from Database (KDD);
- ◆ Understand relationship of Training and Prediction in Data Analytics;
- ◆ Understand the nexus among Artificial Intelligence, Machine Learning and Deep Learning for Data analytics.
- ◆ Demonstrate the concept with example how to use R in Insurance analysis

# Data analytics vs. Data science

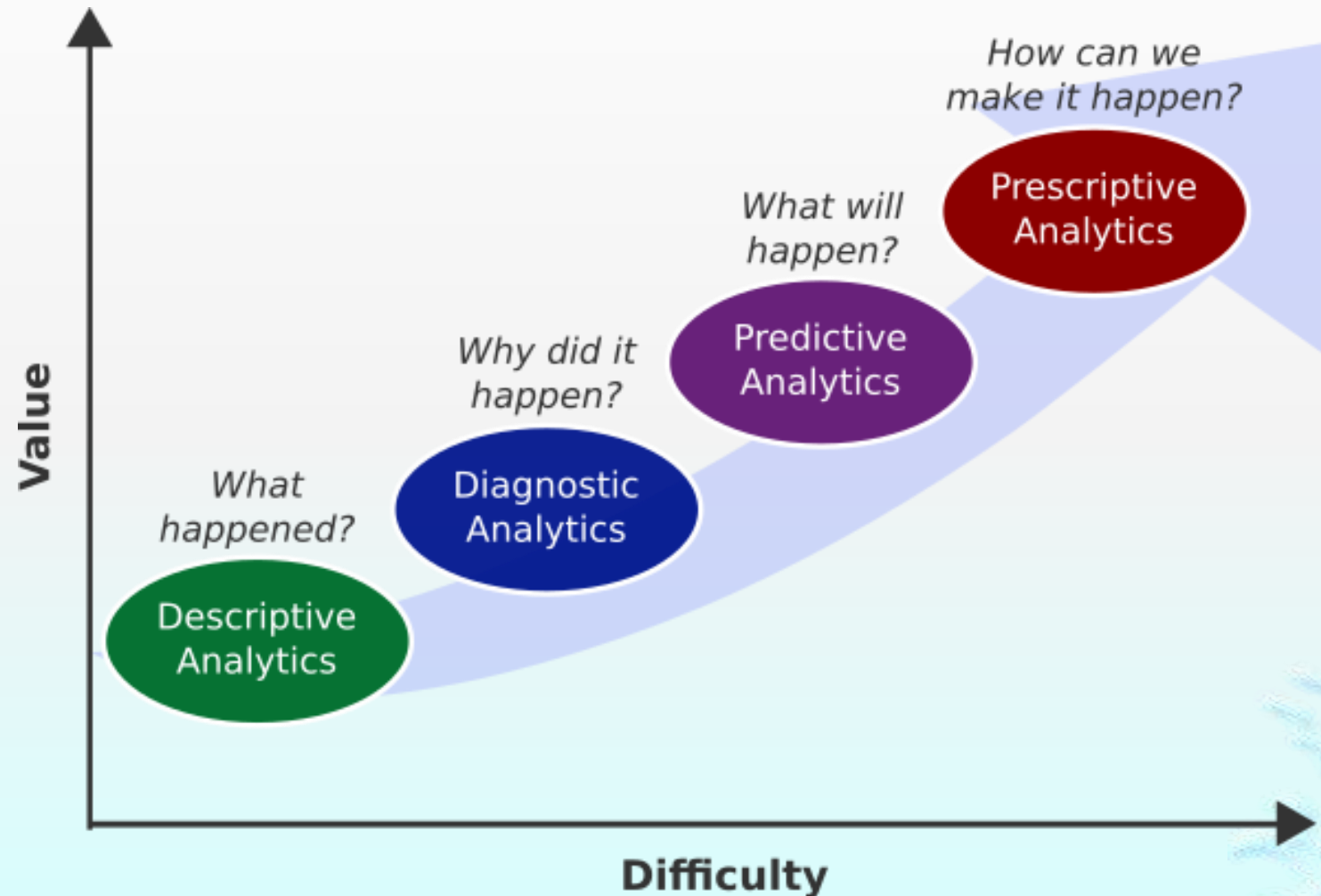
- ◆ **Data analytics** refers to the process and practice of analyzing data to answer questions, extract insights, and identify trends.
- ◆ This is done using an array of tools, techniques, and frameworks.
- ◆ **Data science** is centered on building, cleaning, and organizing datasets.
- ◆ Data scientists create and leverage algorithms, statistical models, and their own custom analyses to collect and shape raw data into something that can be more easily understood.

# Why analyse Data? DIKW pyramid

## Knowledge Management Cognitive Pyramid

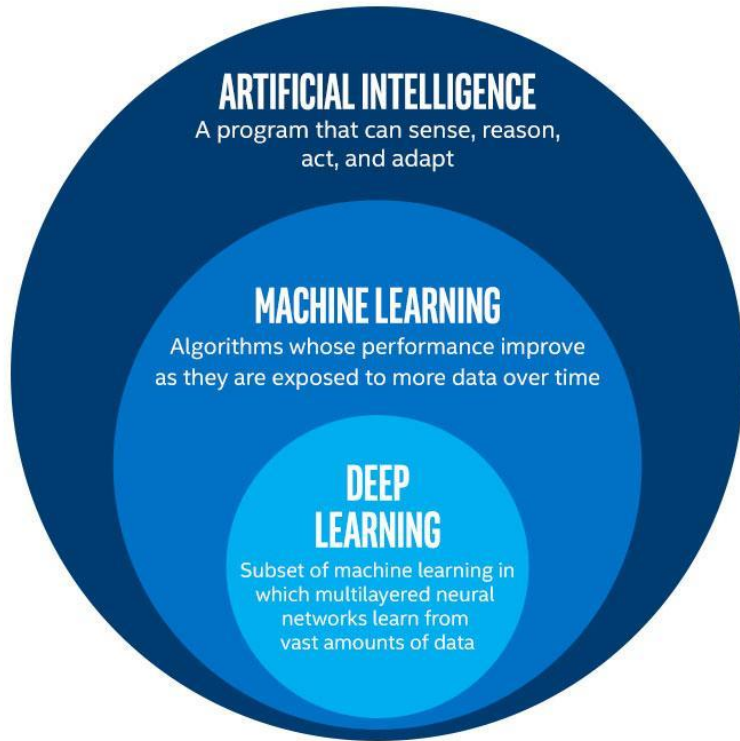


# The different steps in the advanced analytics

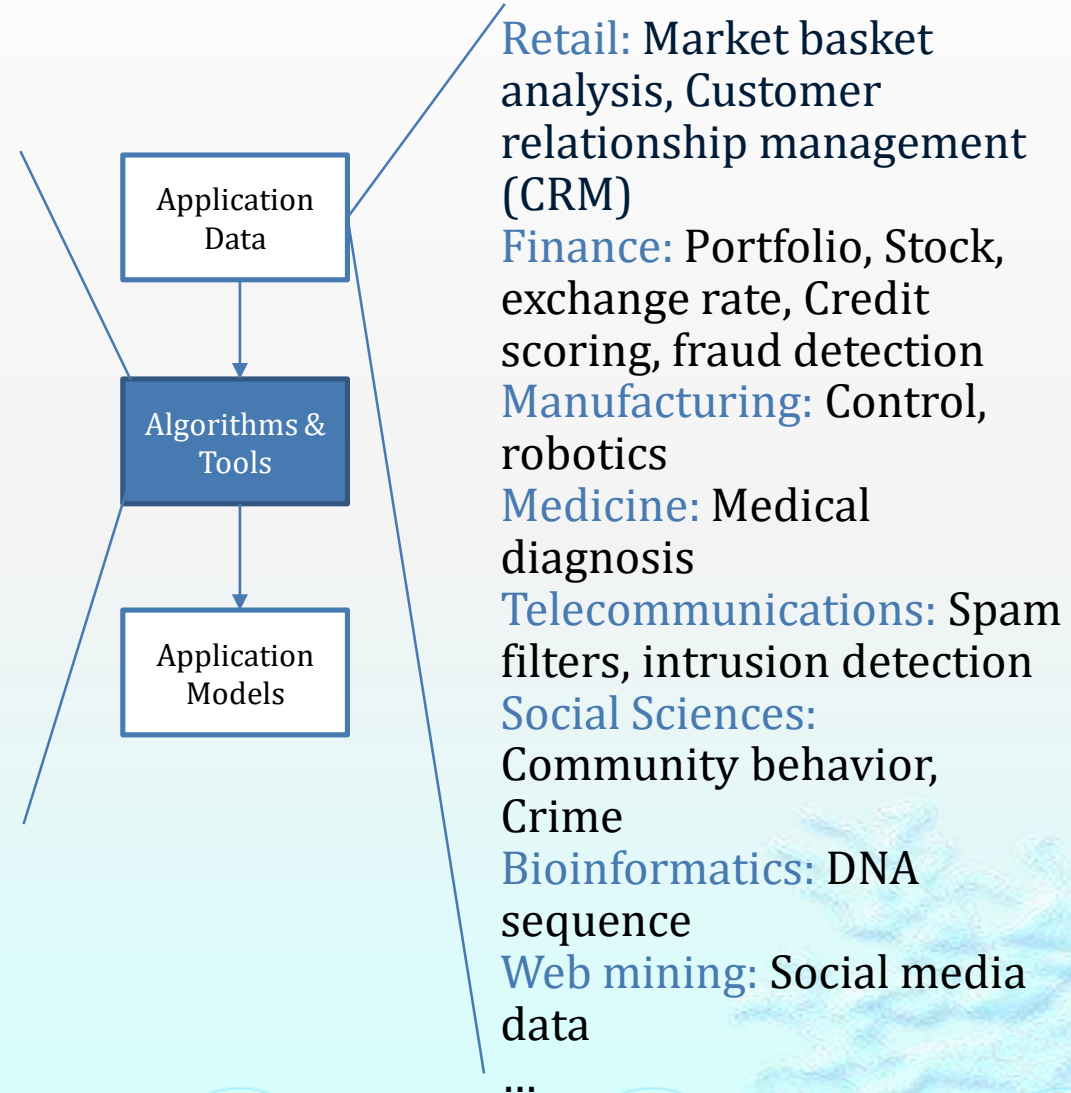




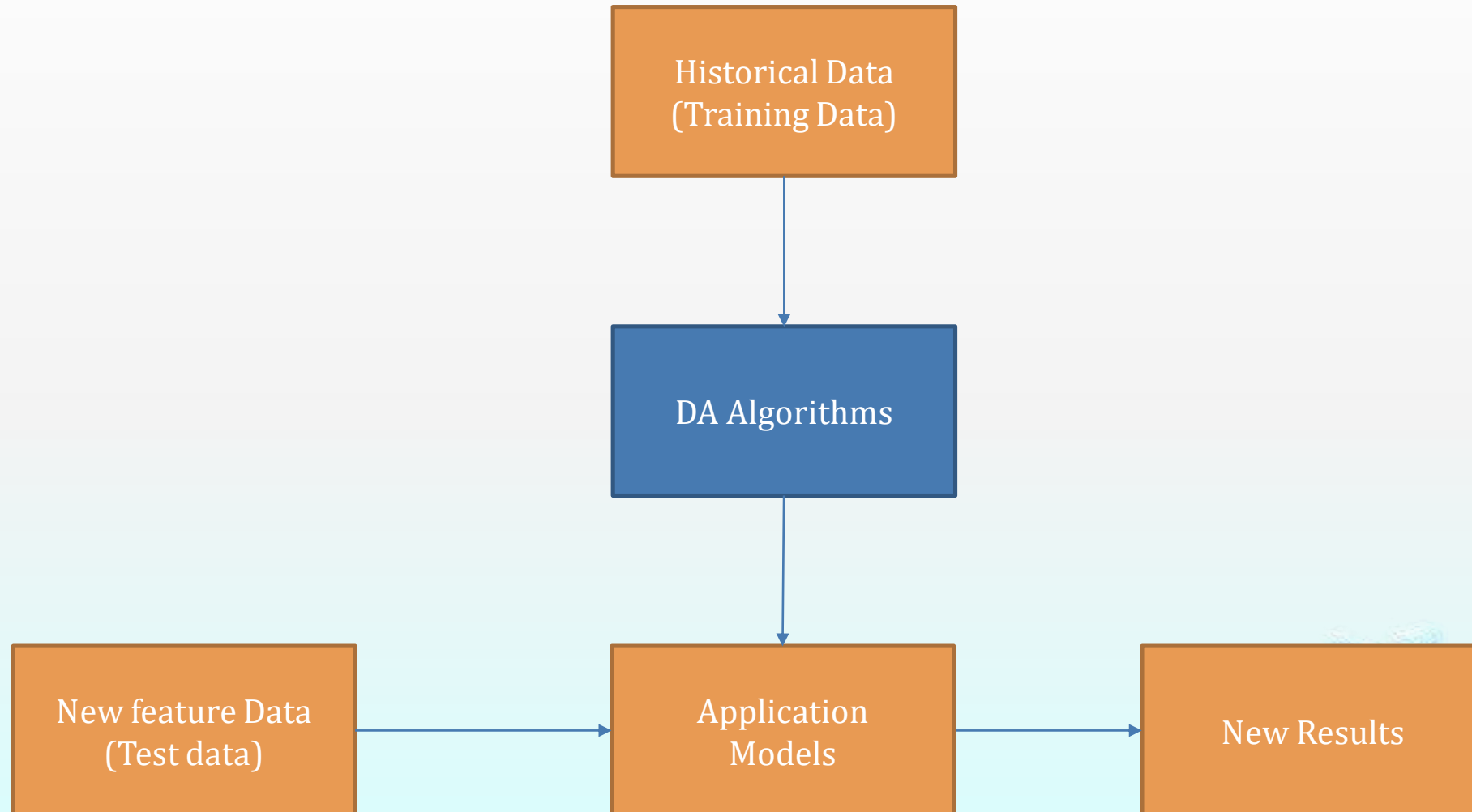
# AI, ML and DL for Data Analytics



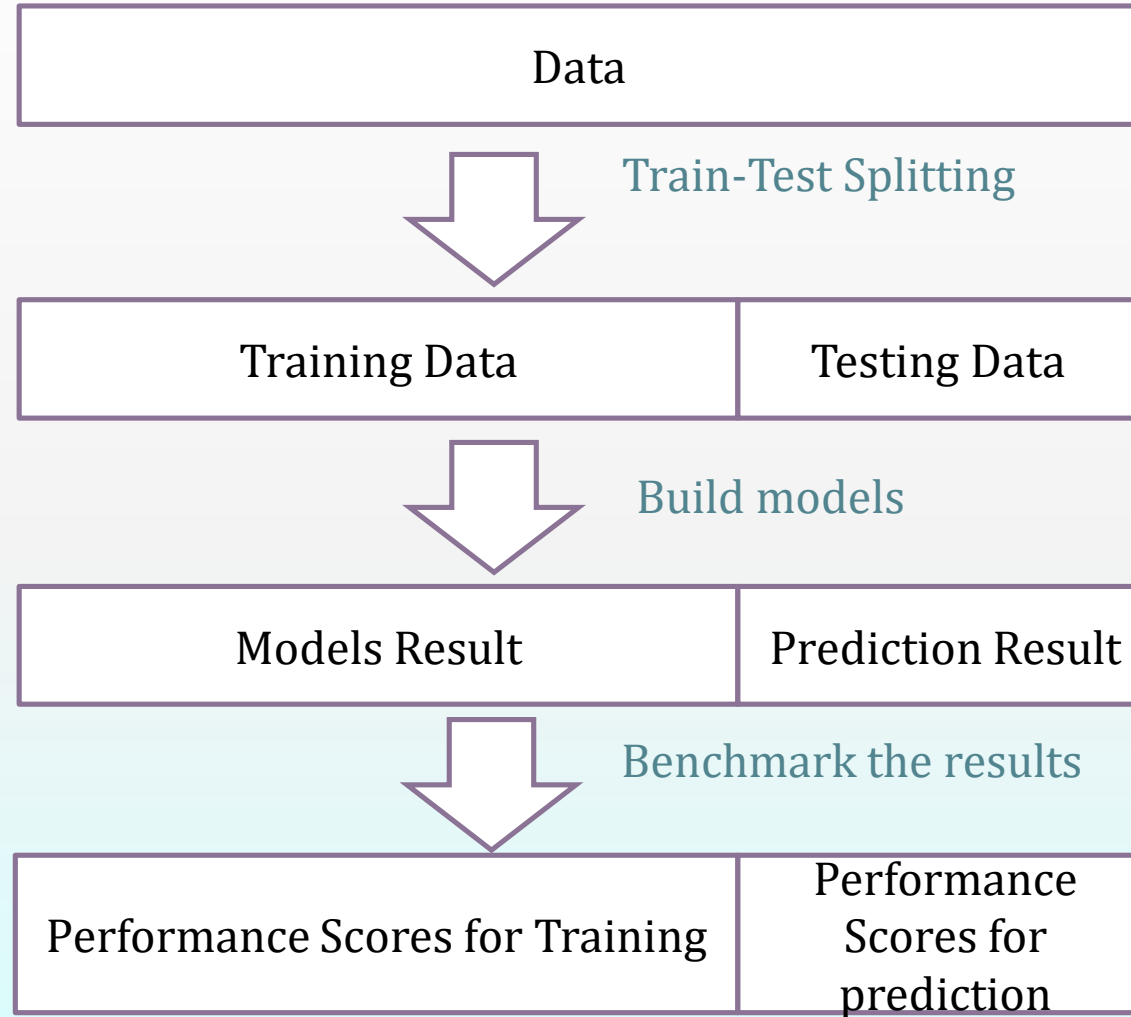
Mathematics is the queen of the sciences.  
(Carl Friedrich Gauss)



# Training and Testing in Data Analytics

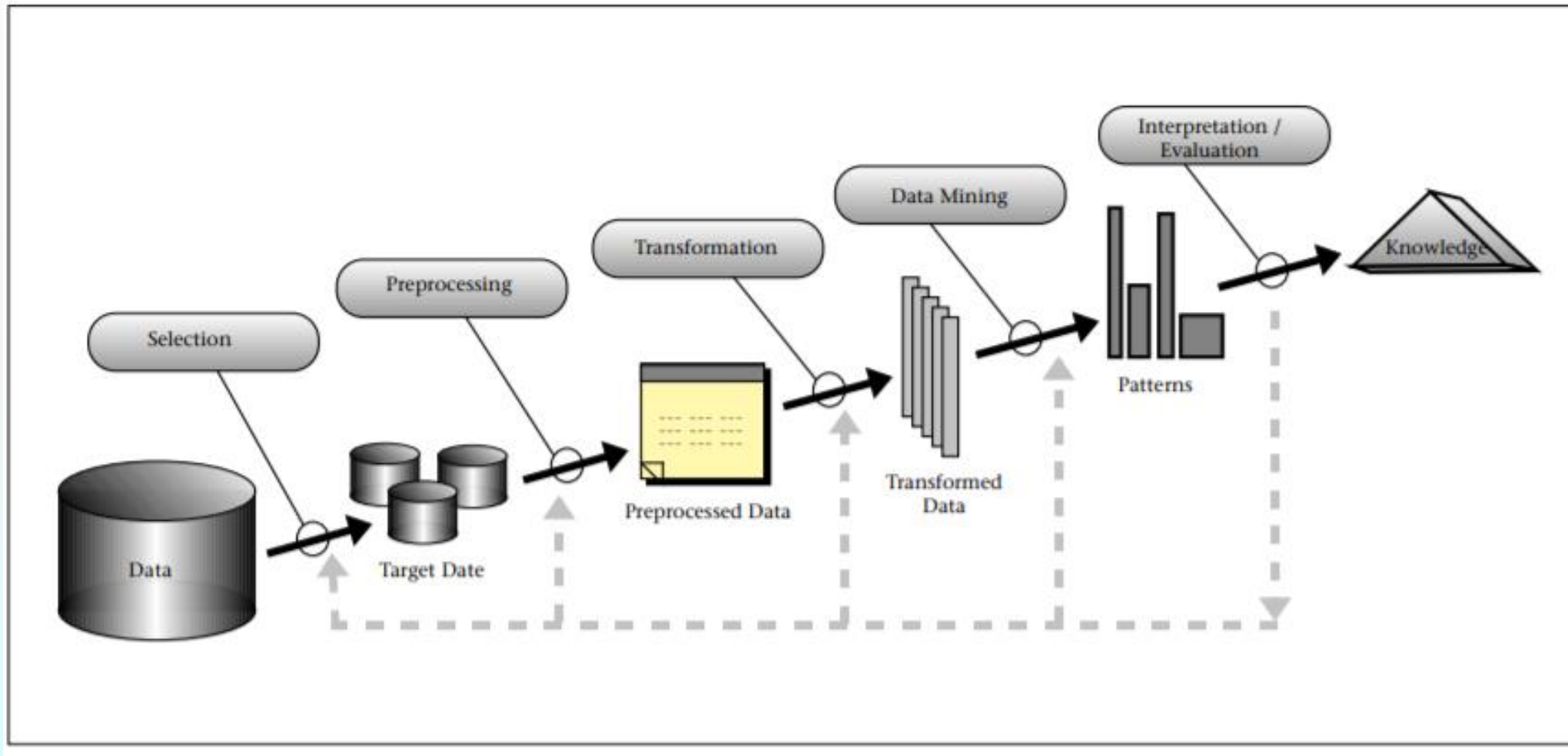


# Train-Test Splitting for Data





# Overview of the Knowledge Discovery in Databases (KDD) Process



# DEMO: INSURANCE ANALYSIS AND PREDICTION IN R

# References

- ◆ Wiki, *DIKW pyramid*, [https://en.wikipedia.org/wiki/DIKW\\_pyramid](https://en.wikipedia.org/wiki/DIKW_pyramid)
- ◆ Harward Business Online, *What's The Difference Between Data Analytics & Data Science?*, <https://online.hbs.edu/blog/post/data-analytics-vs-data-science>
- ◆ The different steps in the advanced analytics, [https://www.governanceanalytics.org/knowledge-base/Main\\_Tools/Data\\_classification\\_and\\_analysis](https://www.governanceanalytics.org/knowledge-base/Main_Tools/Data_classification_and_analysis)
- ◆ Fayyad, Piatetsky-Shapiro, Smyth, "From Data Mining to Knowledge Discovery: An Overview", in Fayyad, Piatetsky-Shapiro, Smyth, Uthurusamy, *Advances in Knowledge Discovery and Data Mining*, AAAI Press / The MIT Press, Menlo Park, CA, 1996, pp.1-34  
<https://www.aaai.org/ojs/index.php/aimagazine/article/view/1230/1131>
- ◆ K.K.F. Yuen, 2019, R fundamentals, <https://github.com/kkfyuen/Rbootcamp1>
- ◆ K.K.F. Yuen, 2019, Data Analytics with R, <https://github.com/kkfyuen/Rbootcamp2>
- ◆ KKF Yuen, Towards multiple regression analyses for relationships of air quality and weather, *Journal of Advances in Information Technology* Vol, 2017,  
<http://www.jait.us/uploadfile/2017/0927/20170927055511447.pdf>

# THANK YOU Q&A

