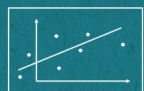














































	Data		Data Science		
	TRADITIONAL	BIG	BUSINESS INTELLIGENCE	TRADITIONAL METHODS	MACHINE LEARNING
WHEN it is applied	At the beginning of your analysis		After the data has been gathered & organized	After BI reports have been created and discussed	
	PAST		NOW	FUTURE	
WHY you need it	data-driven decisions require well-organized and relevant raw data stored in a digital format		use data to create reports and dashboards to gain business insights	Predictive Analytics	
WHAT techniques are involved	DATA COLLECTION PREPROCESSING <ul style="list-style-type: none"> class labeling (categorical vs numerical) data cleansing dealing with missing values CASE SPECIFIC <ul style="list-style-type: none"> e.g. balancing & shuffling datasets 		ANALYZE THE DATA EXTRACT INFO AND PRESENT IT IN THE FORM OF: <ul style="list-style-type: none"> metrics KPIs reports dashboards 	REGRESSION  LOGISTIC REGRESSION	SUPERVISED LEARNING <ul style="list-style-type: none"> SVMs NNs deep learning random forests bayesian networks
	DATA COLLECTION PREPROCESSING <ul style="list-style-type: none"> class labeling (number, text, digital images, digital video data, digital audio data) data cleansing dealing with missing values CASE SPECIFIC <ul style="list-style-type: none"> text data mining, confidentiality - preserving data mining techniques 		  	CLUSTERING  FACTOR ANALYSIS TIME SERIES 	UNSUPERVISED LEARNING <ul style="list-style-type: none"> k-means deep learning REINFORCEMENT LEARNING similar to supervised learning, but instead of minimizing the loss, one maximizes reward
WHERE	BASIC CUSTOMER DATA HISTORICAL STOCK PRICE DATA	SOCIAL MEDIA FINANCIAL TRADING DATA	PRICE OPTIMIZATION INVENTORY MANAGEMENT	USER EXPERIENCE (UX) SALES FORECASTING	FRAUD DETECTION CLIENT RETENTION
HOW using what tools	PROGRAMMING LANGUAGES     SOFTWARE  	PROGRAMMING LANGUAGES     SOFTWARE   	PROGRAMMING LANGUAGES     SOFTWARE     	PROGRAMMING LANGUAGES      SOFTWARE    	PROGRAMMING LANGUAGES      SOFTWARE   
WHO	DATA ARCHITECT DATABASE ENGINEER DATABASE ADMINISTRATOR	BIG DATA ARCHITECT BIG DATA ENGINEER	BI ANALYST BI CONSULTANT BI DEVELOPER	DATA SCIENTIST DATA ANALYST	DATA SCIENTIST MACHINE LEARNING ENGINEER
ARE YOU AWARE	200,000 lines of data is not necessarily big data. It is not just volume that defines a data set as 'big' - variety, variability, velocity, veracity, and other characteristics are determinative as well.		Qualitative analysis tools such as SWOT are not used for quantitative analysis. Hence, they are not part of business intelligence.	Software like Excel, SPSS, and Stata, can be successfully used by data science teams in many companies.	In deep learning, there is still a debate on WHY the algorithms used outperform all conventional methods.