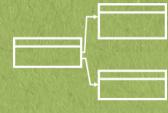
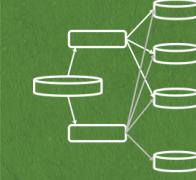


Data		Data Science		
WHEN it is applied	At the beginning of your analysis	BIG	BUSINESS INTELLIGENCE	TRADITIONAL METHODS
WHY you need it	data-driven decisions require well-organized and relevant raw data stored in a digital format		After the data has been gathered & organized	After BI reports have been created and discussed
PAST	NOW	FUTURE		
WHAT techniques are involved	<b>DATA COLLECTION</b> <b>PREPROCESSING</b> <ul style="list-style-type: none"><li>• class labeling (categorical vs numerical)</li><li>• data cleansing</li><li>• dealing with missing values</li></ul> <b>CASE SPECIFIC</b> <ul style="list-style-type: none"><li>• e.g. balancing &amp; shuffling datasets</li></ul>  	<b>DATA COLLECTION</b> <b>PREPROCESSING</b> <ul style="list-style-type: none"><li>• class labeling (number, text, digital images, digital video data, digital audio data)</li><li>• data cleansing</li><li>• dealing with missing values</li></ul> <b>CASE SPECIFIC</b> <ul style="list-style-type: none"><li>• text data mining, confidentiality-preserving data mining techniques</li></ul> 	<b>ANALYZE THE DATA</b> <b>EXTRACT INFO AND PRESENT IT IN THE FORM OF:</b> <ul style="list-style-type: none"><li>• metrics</li><li>• KPIs</li><li>• reports</li><li>• dashboards</li></ul>   	<b>PREDICTIVE ANALYTICS</b> <ul style="list-style-type: none"><li>use data to create reports and dashboards to gain business insights</li></ul> <b>REGRESSION</b>  <b>LOGISTIC REGRESSION</b>  <b>CLUSTERING</b>  <b>FACTOR ANALYSIS</b>  <b>TIME SERIES</b>  <b>SUPERVISED LEARNING</b> <ul style="list-style-type: none"><li>• SVMs</li><li>• NNs</li><li>• deep learning</li><li>• random forests</li><li>• bayesian networks</li></ul> <b>UNSUPERVISED LEARNING</b> <ul style="list-style-type: none"><li>• k-means</li><li>• deep learning</li></ul> <b>ML</b>  <b>REINFORCEMENT LEARNING</b> <p>similar to supervised learning, but instead of minimizing the loss, one maximizes reward</p>
WHERE	<b>BASIC CUSTOMER DATA</b> HISTORICAL STOCK PRICE DATA	<b>SOCIAL MEDIA</b> FINANCIAL TRADING DATA	<b>PRICE OPTIMIZATION</b> INVENTORY MANAGEMENT	<b>USER EXPERIENCE (UX)</b> SALES FORECASTING
HOW using what tools	<b>PROGRAMMING LANGUAGES</b>   <b>SOFTWARE</b>  	<b>PROGRAMMING LANGUAGES</b>   <b>SOFTWARE</b>  	<b>PROGRAMMING LANGUAGES</b>   <b>SOFTWARE</b>  	<b>PROGRAMMING LANGUAGES</b>   <b>SOFTWARE</b>     <b>DATA SCIENTIST</b>  
WHO	<b>DATA ARCHITECT</b> <b>DATA ENGINEER</b> <b>DATABASE ADMINISTRATOR</b>	<b>BIG DATA ARCHITECT</b> <b>BIG DATA ENGINEER</b>	<b>BI ANALYST</b> <b>BI CONSULTANT</b> <b>BI DEVELOPER</b>	<b>DATA SCIENTIST</b> <b>DATA ANALYST</b> <b>MACHINE LEARNING ENGINEER</b>
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.	In deep learning, there is still a debate on WHY the algorithms used outperform all conventional methods.