

DATA ANALYTICS FOR ORGANIZATION

Prepared by:
Dr. Nor Azizah Ali
Dr. Noorfa Haszlinna Mustaffa
Faculty of Computing, UTM

2022/2023 – Sem. (1)

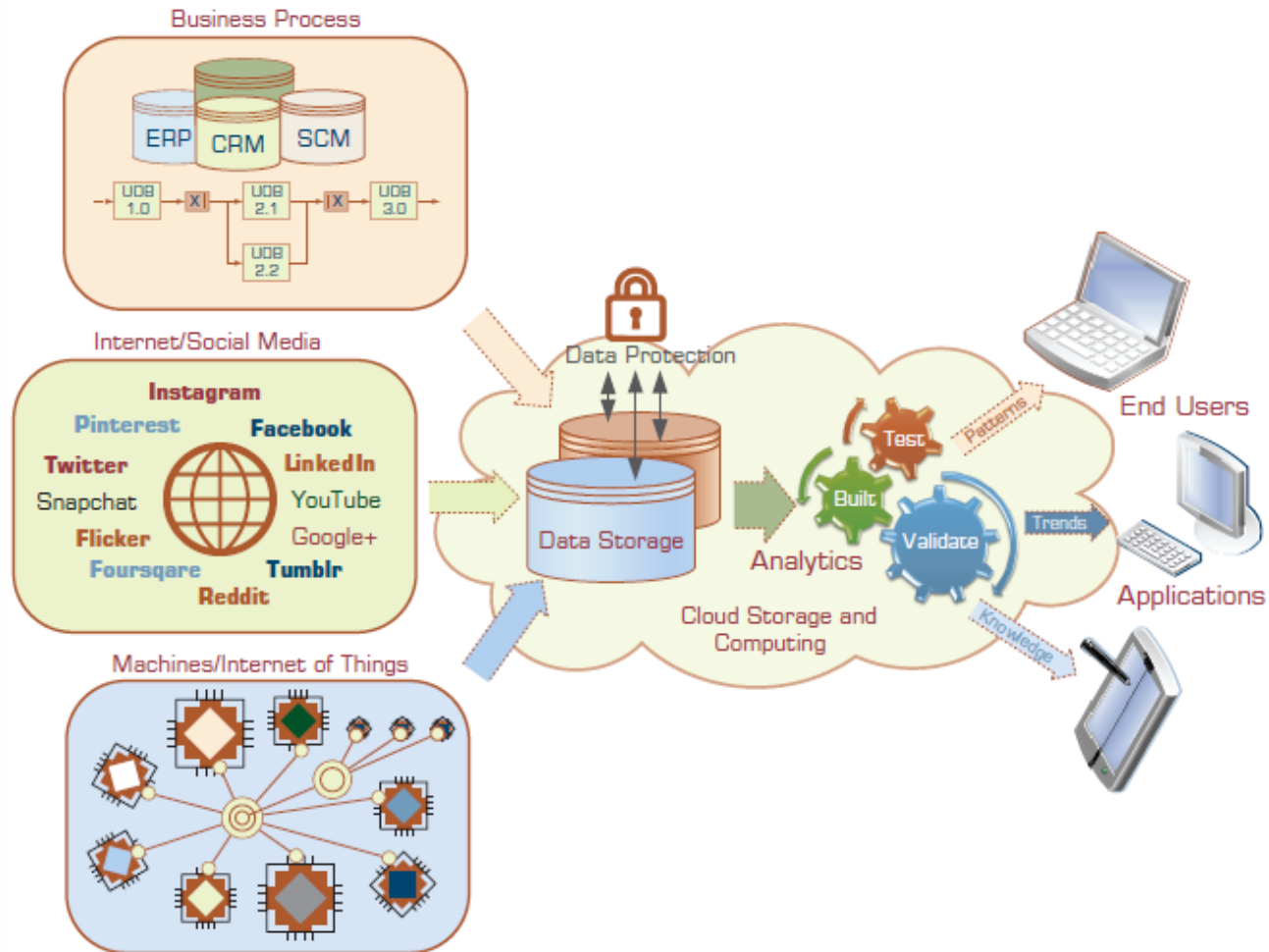
Course Outline:

Session	Topic
Session 1	Fundamental Concept of Data Analytics
Session 2	Applications of Data Analytics with Tableau – Part 1
Session 3	Applications of Data Analytics with Tableau – Part 2
Session 4	Hands-On Project (Dashboard Development)
Session 5	Competency Test

Session 1:

Fundamental Concepts of Data Analytics

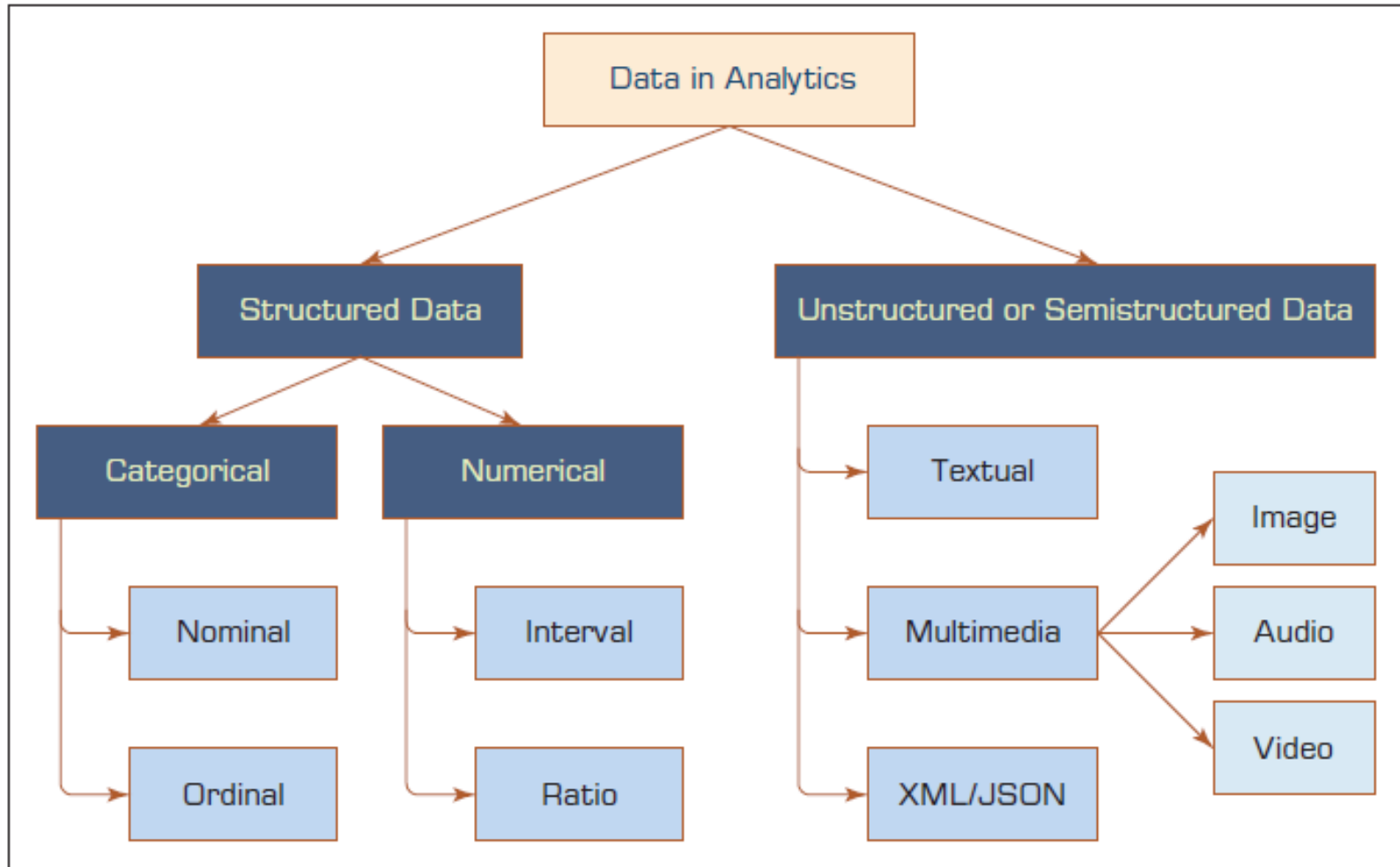
The Nature of Data



A Simple Taxonomy of Data

- Data (datum—singular form of data): **facts**
- Types of Data:
 - Structured data
 - Targeted for computers to process
 - Numeric versus nominal
 - Unstructured/textual data
 - Targeted for humans to process/digest
 - Semi-structured data
 - XML, HTML, Log files, etc.

A Simple Taxonomy of Data (cont'd)



Common Data Problems

■ Structured, unstructured, semi-structured:

- Information and knowledge management is the management of both structured data (15% of information) and unstructured data (85% of information), according to the Butler Group.
- 80 percent of business is conducted on unstructured information (Gartner Group).

■ Information overloading:

- too much data and information with varied formats and structure
- difficulty of data organization for effective access and retrieval
- difficult to find useful information (knowledge) from them
- Multiple copies of data exists sometimes with conflicts

■ Big data:

- Variety, Velocity, Volume, Veracity

(Source: <https://www.ibmbigdatahub.com/infographic/four-vs-big-data>)

Common Data Problems (cont'd)

■ Data everywhere:

- Data in separate systems and different sources; internal and external
- Problem of spreadmart (<http://en.wikipedia.org/wiki/Spreadmart>)
- Over 43 percent of organizations have more than six content stores. (Forrester Research).

■ Difficulty of access:

- We may have that data but we cannot access it (or difficult to get it), because of technical issues or administrative issues.

■ Don't have that data:

- The data is simply not available.
- The collection of data may need additional process and is costly.

Decision Making in Business Processes

- The decision-making process, in a business context, is a set of steps taken by managers in an enterprise to determine the planned path for business initiatives and to set specific actions in motion.
- Ideally, business decisions are based on an analysis of objective facts, aided by the use of **Business Intelligence** (BI) and analytics tools.

What is Business Intelligence?

Technology that Allows:

- Gathering, storing, accessing & analyzing data to help business users make better decisions

Set of Applications that Allow:

- Decision support systems
- Query and reporting
- online analytical processing (OLAP)
- Statistical analysis, forecasting, and data mining

Help in analyzing business performance through data-driven insight:

- Understand the past & predict the future

Importance of Business Intelligence

- Overall, the role of business intelligence is to improve all parts of a company by improving access to the firm's data and then using that data to **increase profitability**.
- Companies that employ BI practices can **translate their collected data into insights of their business process**.
- The insights can then be used to **create strategic business decisions** that improve productivity, increase revenue and accelerate growth.

Challenges in Decision Making Processes

- Problems in decision making:
 - A gap between data and knowledge (useful information leading to a decision).
 - Management/operation by intuition.
 - Lack of effective feedback and alignment systems, no improvement cycles.
 - Need good analytical processing and models.
- Evolving analytical needs in decision support:
 - Real-time, most recent data.
 - Business user driven, agile, instant.
 - Exploratory and interactive.

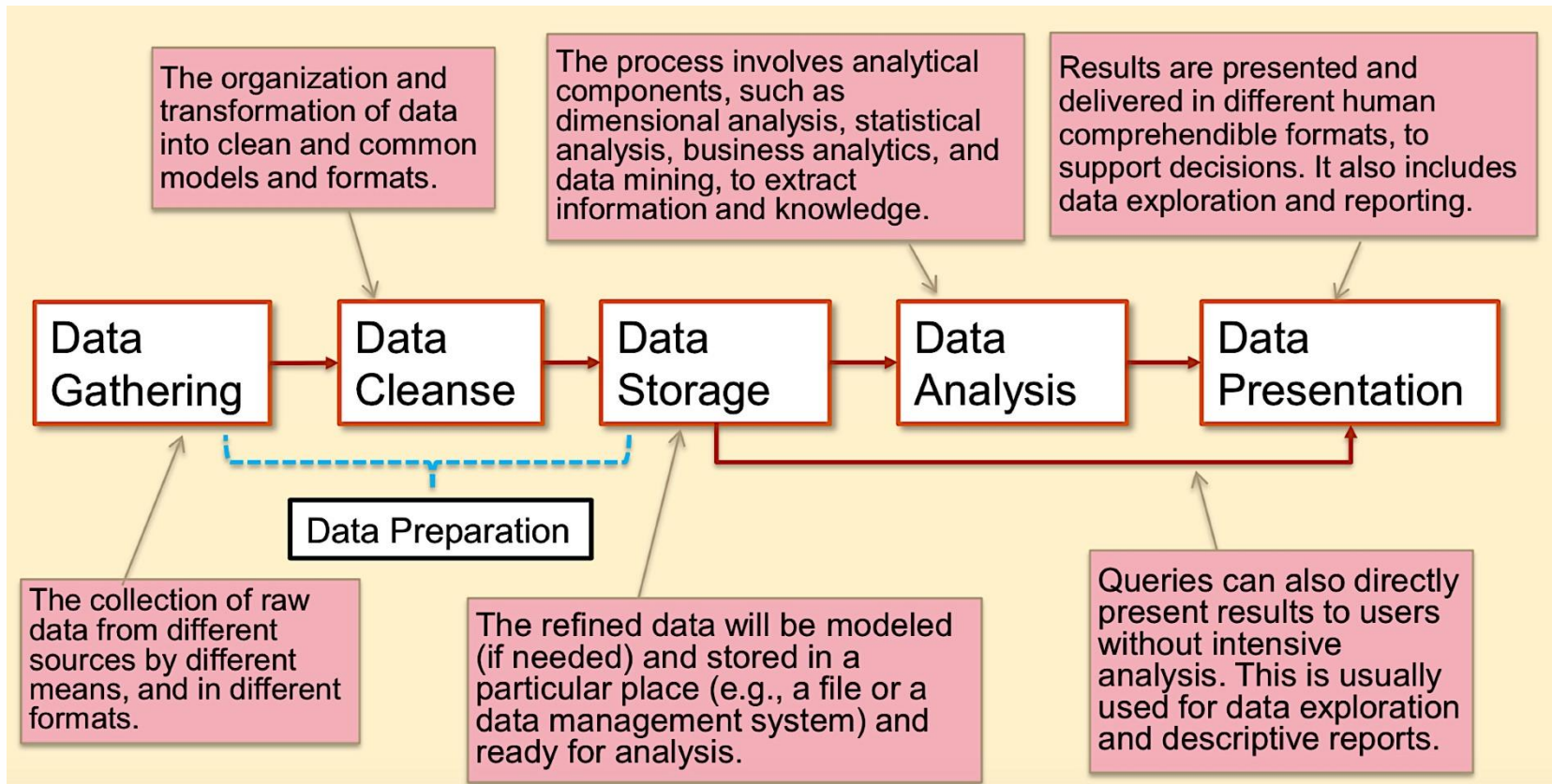
Who make a decision?

Decision making at different levels:

- Operational
 - Related to daily activities with short-term effect
 - Structured decisions taken by lower management
- Tactical
 - Semi-structured decisions taken by middle management
- Strategic
 - Long-term effect
 - Unstructured decisions taken by top management



General Process – Data Analytics in Organization



Analysis vs Analytics

The primary **difference** between analytics and analysis is **a matter of scale**, as data analytics is a broader term of which data analysis is a subcomponent. **Data analysis** refers to the process of examining, transforming and arranging a given data set in specific ways in order to study its individual parts and extract useful information. **Data analytics** is an overarching science or discipline that encompasses the complete management of data. This not only include analysis, but also data collection, organization, storage and all the tools and techniques used.

Source: www.getsmarter.com

Data Analytics Phases

Introduction



Source: Gartner (March 2012)

EXAMPLE

Phase 1: Data collection

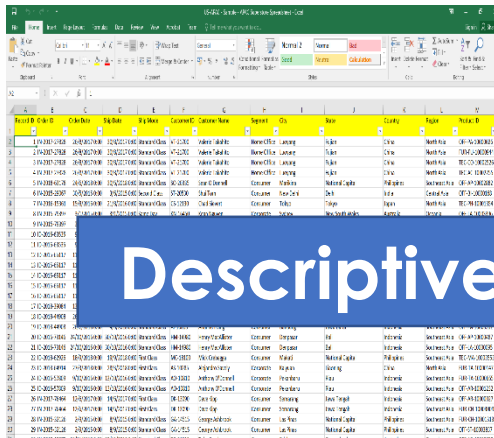
Data

US-APAC - Sample - APAC Superstore Spreadsheet - Excel													
Record ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	Country	Region	Product ID	
1	IN-2017-27828	26/8/2017 0:00	30/8/2017 0:00	Standard Class	VT-21700	Valerie Takahito	Home Office	Luoyang	Fujian	China	North Asia	OFF-PA-10000026	
2	IN-2017-27828	26/8/2017 0:00	30/8/2017 0:00	Standard Class	VT-21700	Valerie Takahito	Home Office	Luoyang	Fujian	China	North Asia	FUR-FU-10000944	
3	IN-2017-27828	26/8/2017 0:00	30/8/2017 0:00	Standard Class	VT-21700	Valerie Takahito	Home Office	Luoyang	Fujian	China	North Asia	TEC-CO-10002526	
4	IN-2017-27828	26/8/2017 0:00	30/8/2017 0:00	Standard Class	VT-21700	Valerie Takahito	Home Office	Luoyang	Fujian	China	North Asia	TEC-AC-10002255	
5	IN-2018-63178	24/6/2018 0:00	30/6/2018 0:00	Standard Class	SO-20335	Sean O'Donnell	Consumer	Marikina	National Capital	Philippines	Southeast Asia	OFF-AP-10002882	
6	IN-2015-29067	30/5/2015 0:00	3/6/2015 0:00	Second Class	ST-20530	Shui Tom	Consumer	New Delhi	Delhi	India	Central Asia	OFF-BI-10000168	
7	IN-2016-15361	15/8/2016 0:00	21/8/2016 0:00	Standard Class	CS-12130	Chad Sievert	Consumer	Tokyo	Tokyo	Japan	North Asia	TEC-PH-10001354	
8	IN-2015-79397	3/1/2015 0:00	3/1/2015 0:00	Same Day	KN-16450	Kean Nguyen	Corporate	Sydney	New South Wales	Australia	Oceania	OFF-LA-10003396	
9	IN-2015-79397	3/1/2015 0:00	3/1/2015 0:00	Same Day	KN-16450	Kean Nguyen	Corporate	Sydney	New South Wales	Australia	Oceania	OFF-AP-10000304	
10	ID-2016-63535	9/5/2016 0:00	15/5/2016 0:00	Standard Class	ML-17410	Maris LaWare	Consumer	Hobart	Tasmania	Australia	Oceania	TEC-PH-10003062	
11	ID-2016-63535	9/5/2016 0:00	15/5/2016 0:00	Standard Class	ML-17410	Maris LaWare	Consumer	Hobart	Tasmania	Australia	Oceania	OFF-AR-10001134	
12	ID-2015-65117	11/7/2015 0:00	15/7/2015 0:00	Standard Class	SW-20245	Scot Wooten	Consumer	Manila	National Capital	Philippines	Southeast Asia	OFF-SU-10004177	
13	ID-2015-65117	11/7/2015 0:00	15/7/2015 0:00	Standard Class	SW-20245	Scot Wooten	Consumer	Manila	National Capital	Philippines	Southeast Asia	TEC-PH-10001354	
14	ID-2015-65117	11/7/2015 0:00	15/7/2015 0:00	Standard Class	SW-20245	Scot Wooten	Consumer	Manila	National Capital	Philippines	Southeast Asia	OFF-BI-10000089	
15	ID-2015-65117	11/7/2015 0:00	15/7/2015 0:00	Standard Class	SW-20245	Scot Wooten	Consumer	Manila	National Capital	Philippines	Southeast Asia	OFF-EN-10003463	
16	ID-2015-65117	11/7/2015 0:00	15/7/2015 0:00	Standard Class	SW-20245	Scot Wooten	Consumer	Manila	National Capital	Philippines	Southeast Asia	FUR-CH-10001871	
17	ID-2018-39084	13/6/2018 0:00	17/6/2018 0:00	Standard Class	PO-18865	Patrick O'Donnell	Consumer	Valencia	Northern Mindanao	Philippines	Southeast Asia	TEC-CO-10000666	
18	ID-2018-44908	26/2/2018 0:00	4/3/2018 0:00	Standard Class	AI-10855	Arianne Irving	Consumer	Bandung	Jawa Barat	Indonesia	Southeast Asia	OFF-SU-10004848	
19	ID-2018-44908	26/2/2018 0:00	4/3/2018 0:00	Standard Class	AI-10855	Arianne Irving	Consumer	Bandung	Jawa Barat	Indonesia	Southeast Asia	OFF-PA-10003332	
20	ID-2016-73146	24/10/2016 0:00	30/10/2016 0:00	Standard Class	HM-14980	Henry MacAllister	Consumer	Denpasar	Bali	Indonesia	Southeast Asia	OFF-PA-10003487	
21	ID-2016-73146	24/10/2016 0:00	30/10/2016 0:00	Standard Class	HM-14980	Henry MacAllister	Consumer	Denpasar	Bali	Indonesia	Southeast Asia	OFF-LA-10000095	
22	ID-2018-62926	18/9/2018 0:00	19/9/2018 0:00	First Class	MC-18100	Mick Crebagga	Consumer	Makati	National Capital	Philippines	Southeast Asia	TEC-MA-10003553	
23	ID-2018-64914	27/3/2018 0:00	28/3/2018 0:00	First Class	AS-10285	Alejandro Savely	Corporate	Kaiyuan	Liaoning	China	North Asia	FUR-TA-10000147	
24	ID-2016-57809	9/10/2016 0:00	13/10/2016 0:00	Standard Class	AO-10810	Anthony O'Donnell	Corporate	Pekanbaru	Riau	Indonesia	Southeast Asia	FUR-TA-10000665	
25	ID-2016-57809	9/10/2016 0:00	13/10/2016 0:00	Standard Class	AO-10810	Anthony O'Donnell	Corporate	Pekanbaru	Riau	Indonesia	Southeast Asia	OFF-AR-10001232	
26	IN-2017-76464	12/6/2017 0:00	14/6/2017 0:00	First Class	DK-13090	Dave Kipp	Consumer	Semarang	Jawa Tengah	Indonesia	Southeast Asia	OFF-AR-10000027	
27	IN-2017-76464	12/6/2017 0:00	14/6/2017 0:00	First Class	DK-13090	Dave Kipp	Consumer	Semarang	Jawa Tengah	Indonesia	Southeast Asia	FUR-CH-10003009	
28	IN-2015-32126	2/9/2015 0:00	8/9/2015 0:00	Standard Class	GA-14515	George Ashbrook	Consumer	Las Pinas	National Capital	Philippines	Southeast Asia	FUR-CH-10001313	
29	IN-2015-32126	2/9/2015 0:00	8/9/2015 0:00	Standard Class	GA-14515	George Ashbrook	Consumer	Las Pinas	National Capital	Philippines	Southeast Asia	OFF-ST-10003837	
30	IN-2018-13575	24/11/2018 0:00	20/11/2018 0:00	Standard Class	TC-31210	Toby Gonda	Consumer	Brisbane	Queensland	Australia	Oceania	TEC-AR-10000059	

EXAMPLE

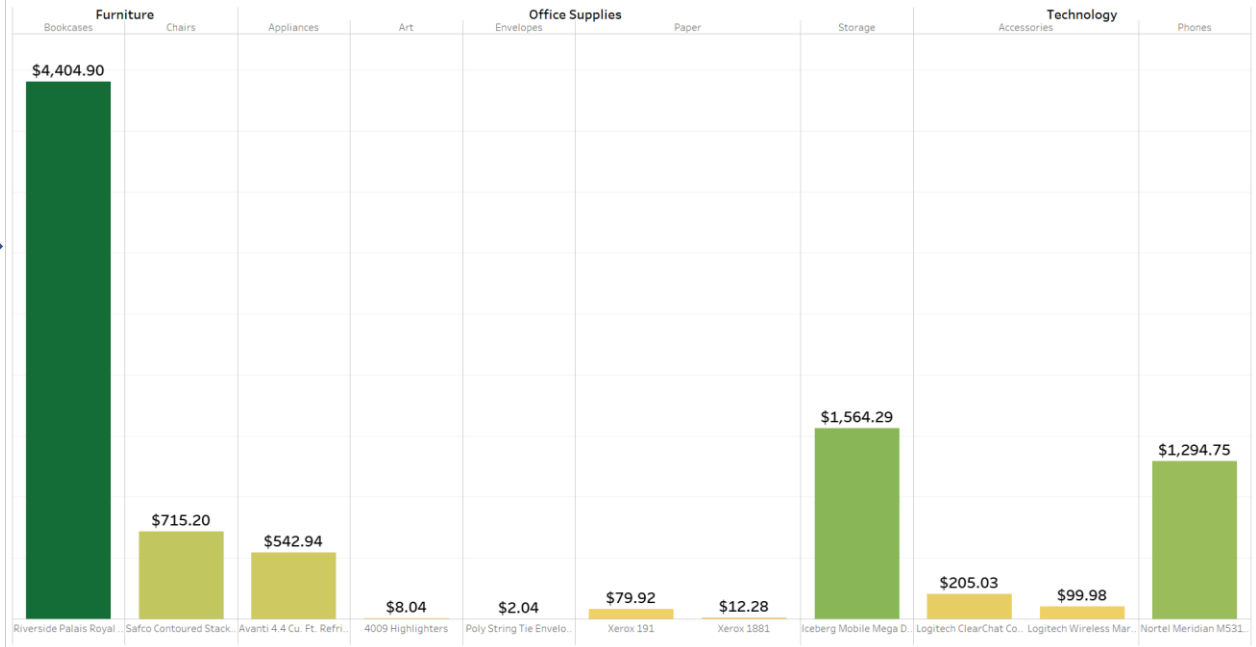
Phase 2: Descriptive

- ✓ Clean
- ✓ Relate
- ✓ Summarize
- ✓ Visualize



Descriptive

Sales by Category

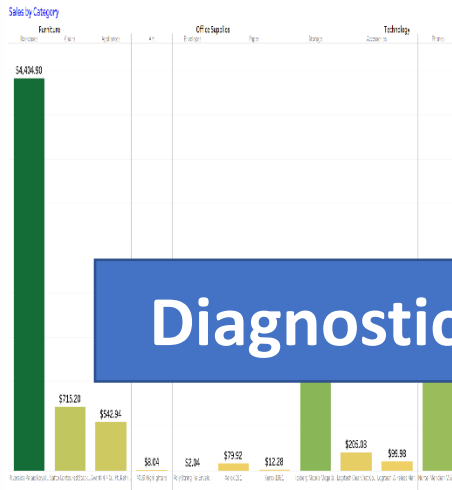
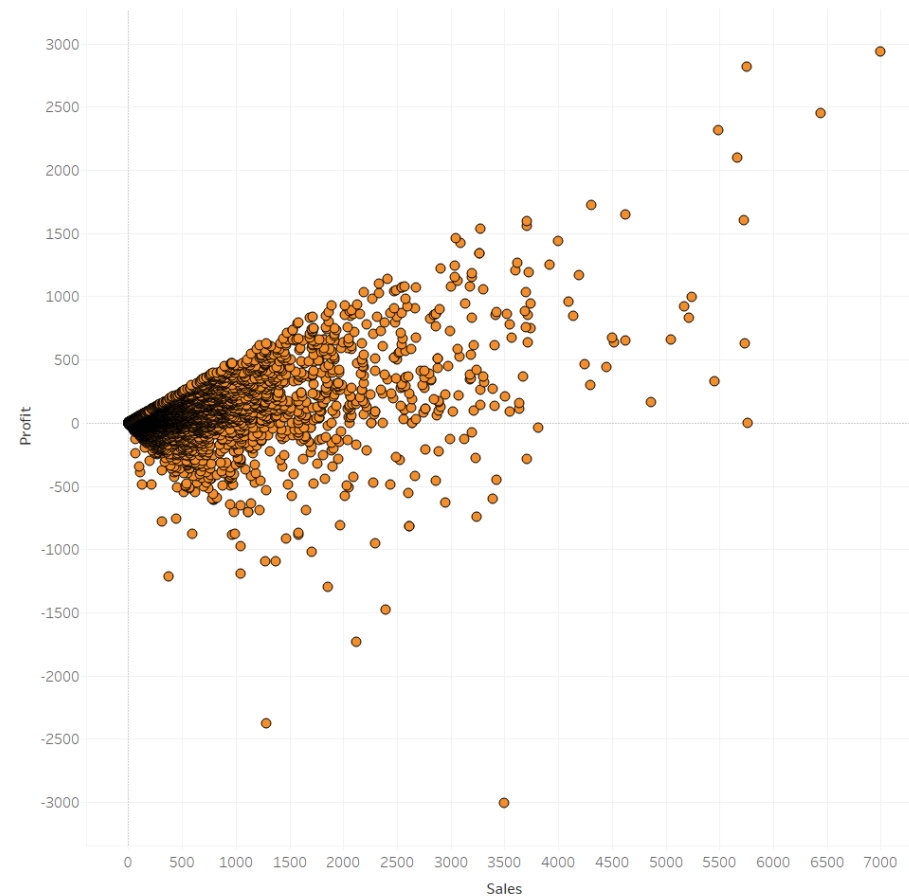


EXAMPLE

Phase 3: Diagnostic

- ✓ Drill - down
- ✓ Data Discovery
- ✓ Data Mining
- ✓ Correlation
- ✓ Causation

Scatter Plot: Profit vs Sales



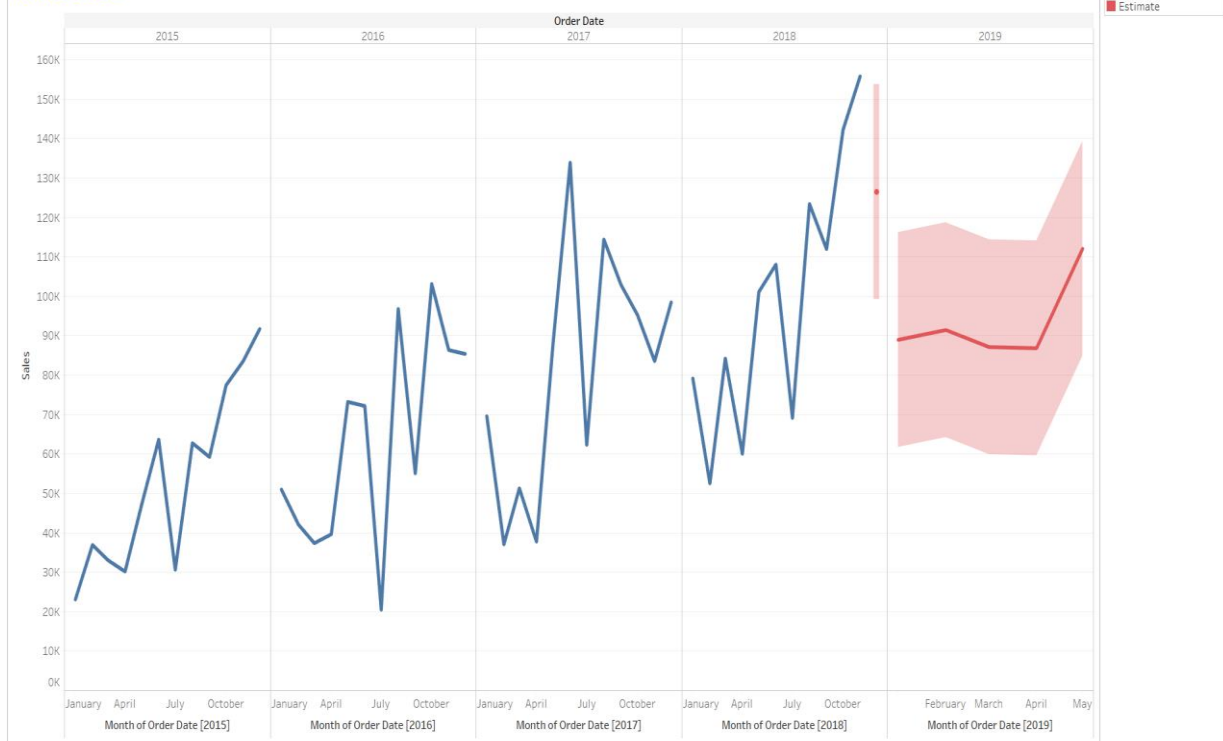
EXAMPLE

Phase 4: Predictive

Scatter Plot: Profit vs Sales



Forecast



Magic Quadrant for Analytics and Business Intelligence Platforms.

Gartner is the world's leading research and advisory company.



Source: Gartner (February 2020)

End of Session 1