Data Analytics

Lesson 02.

Foundation of Business Analytics

Dr. Hai Tran

hai.tran@sbsuni.edu.vn

Scholar: https://scholar.google.com/citations?user=kHZvlTkAAAAJ&hl=en&oi=ao

Co-Founder: XAI - https://xai.foo/

Saigon Business School

In partnership with

UNIVERSITY CANADA WEST

MACQUARIE University SYDNEY·AUSTRALIA

University of Northampton

AMITY GLOBAL INSTITUTE

Saigon Business School

Learning materials

- Textbook
  - Evans, J. (2016) Business Analytics. 2nd edn. Pearson.
  - Runkler, T. (2016) Data Analytics: Models and Algorithms for Intelligent Data Analysis. 2nd edn. Vieweg+Teubner Verlag.

- Online reference materials
  - archive.ics.uci.edu/ml/
  - powerbi.microsoft.com
  - https://github.com/topics/data-analysis-python
  - https://media.pearsoncmg.com/ph/esm/esm_evans_eba3e_20/tools/eba3e_analytic_solver.html
  - https://data.imf.org/

sbsedu.vn

Saigon Business School

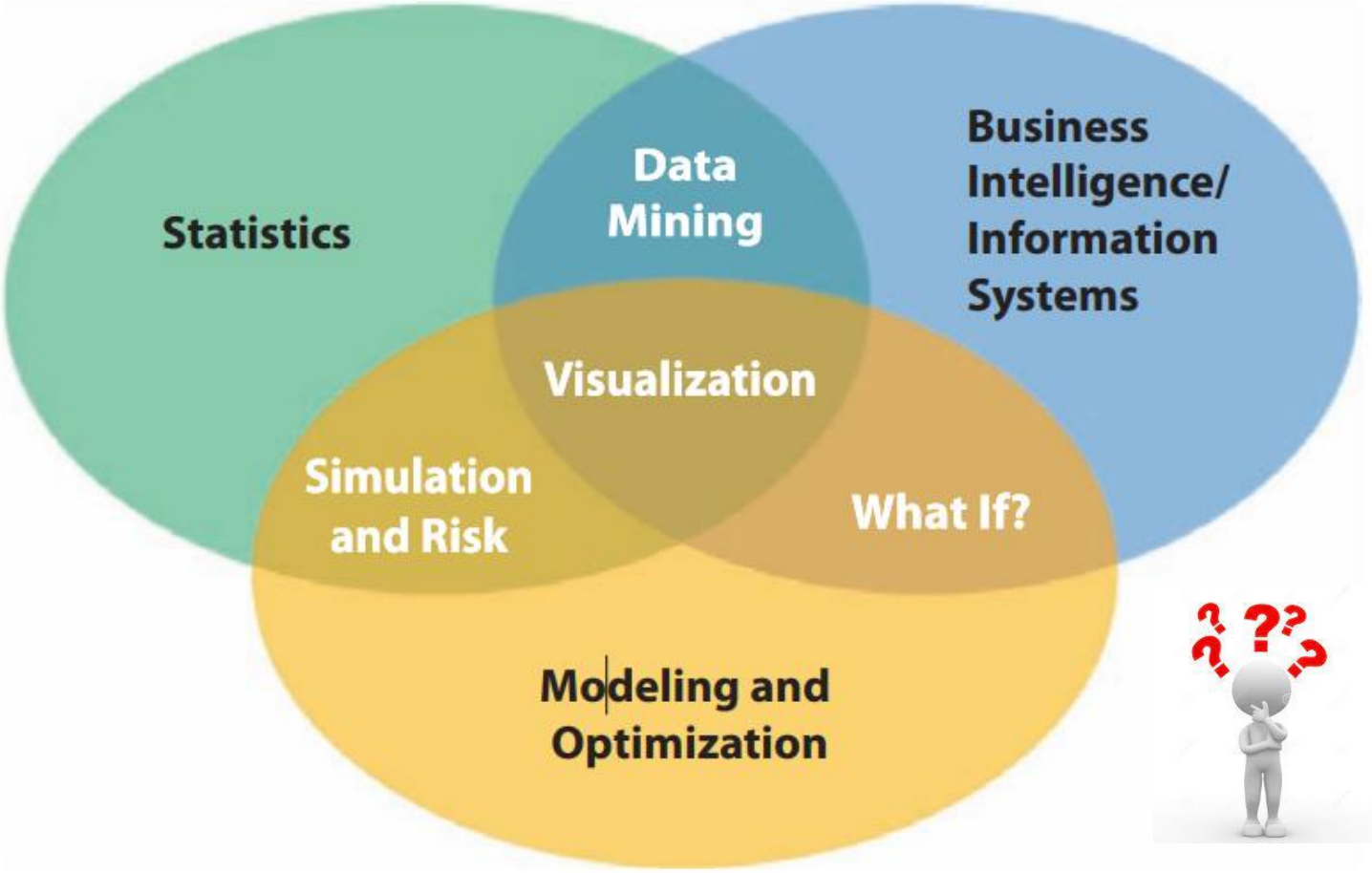
Agenda

- Lesson 1: Understanding Data Analytics Terminologies.
- Lesson 2: Foundation of Business Analytics
- Lesson 3: Visualizing and Exploring data
- Lesson 4: Applying Descriptive Analytic Techniques
- Lesson 5: Data Modeling
- Lesson 6: Predictive Analytics
- Lesson 7: Regression, Classification and Clustering
- Lesson 8: Forecasting Techniques
- Lesson 9: Investigating Predictive Analytic Techniques
- Lesson 10: Introduction to Data Mining
- Lesson 11: Demonstrating Prescriptive Analytic Methods
- Lesson 12: Recap and advanced topics

sbsedu.vn

Saigon Business School

Foundation of Business Analytics



- **Business analytics**, or simply **analytics**, is the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight about their business operations and make better, fact-based decisions. Business analytics is "a process of transforming data into actions through analysis and insights in the context of organizational decision making and problem solving."3 Business analytics is supported by various tools such as Microsoft Excel and various Excel add-ins, commercial statistical software packages such as SAS or Minitab, and more complex business intelligence suites that integrate data with analytical software.

Saigon Business School

Foundation of Business Analytics

- Algorithm
- Big data
- Business analytics (analytics)
- Business intelligence (BI)
- Categorical (nominal) data
- Constraint Continuous metric
- Data mining, Data set
- Database Decision model
- Decision support systems (DSS)
- Descriptive analytics
- Modeling and optimization
- Objective function
- Operations Research/Management Science (OR/MS)
- Search algorithm
- Simulation and risk analysis
- Statistics Stochastic model
- Visualization
- What-if analysis

sbsedu.vn

Saigon Business School

Foundation of Business Analytics

sbsedu.vn

- A bank developed a model for predicting the average checking and savings account balance as balance = -17,732 + 367 * age + 1,300 * years education + 0.116 * household wealth.
- a. Explain how to interpret the numbers in this model.
- b. Suppose that a customer is 32 years old, is a college graduate (so that years education = 16), and has a household wealth of $150,000. What is the predicted bank balance?

- a. To interpret the numbers in the given model:
- The intercept, -17,732, represents the predicted balance when all other variables (age, years of education, and household wealth) are equal to zero. In practical terms, this means that if a person had zero years of education, zero wealth, and was just born (age = 0), their predicted bank balance would be -17,732. The negative sign suggests that the model starts with a negative balance, which may not be meaningful in a real-world context.
- The coefficient 367 for "age" indicates how the bank balance is expected to change for each additional year of age. In this case, for each year older a customer is, their predicted bank balance is expected to increase by 367 units (currency).
- The coefficient 1,300 for "years of education" suggests that for each additional year of education, the predicted bank balance is expected to increase by 1,300 units.
- The coefficient 0.116 for "household wealth" implies that for each unit increase in household wealth (measured in the same currency), the predicted bank balance is expected to increase by 0.116 units.
-

Saigon Business School

Foundation of Business Analytics

- A bank developed a model for predicting the average checking and savings account balance as balance = -17,732 + 367 * age + 1,300 * years education + 0.116 * household wealth.
- a. Explain how to interpret the numbers in this model.
- b. Suppose that a customer is 32 years old, is a college graduate (so that years education = 16), and has a household wealth of $150,000. What is the predicted bank balance?

- b. To calculate the predicted bank balance for a customer who is 32 years old, is a college graduate (16 years of education), and has a household wealth of $150,000, you can use the provided model:
- Balance = -17,732 + (367 * age) + (1,300 * years of education) + (0.116 * household wealth)
- Balance = -17,732 + (367 * 32) + (1,300 * 16) + (0.116 * 150,000)
- Now, calculate the values:
- Balance = -17,732 + 11,744 + 20,800 + 17,400
- Balance = 32,212
- So, the predicted bank balance for a customer with these characteristics would be $32,212.

sbsedu.vn

Saigon Business School

Foundation of Business Analytics

- Example: Python script for histogram analytic

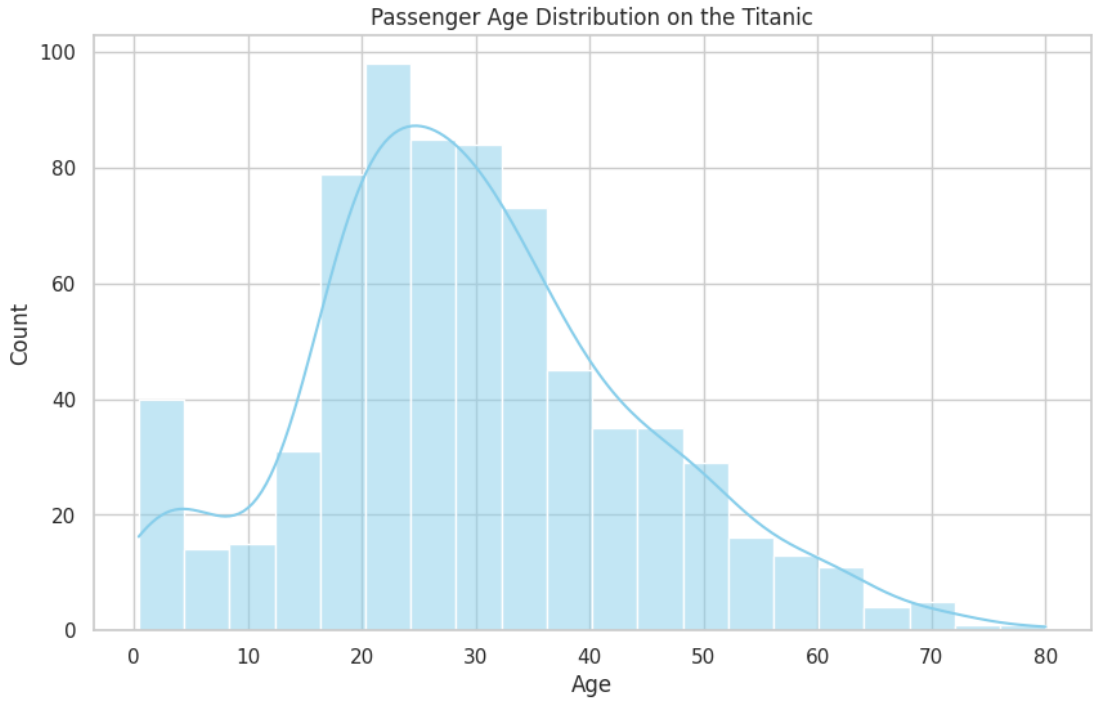
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the Titanic dataset
titanic_data = sns.load_dataset("titanic")

# Plot a histogram for passenger ages
sns.set(style="whitegrid")
plt.figure(figsize=(10, 6))
sns.histplot(data=titanic_data, x="age", kde=True,
bins=20, color="skyblue")

# Customize the plot
plt.title("Passenger Age Distribution on the
Titanic")
plt.xlabel("Age")
plt.ylabel("Count")

# Show the histogram
plt.show()
```



sbsedu.vn

Saigon Business School

Learning Mission

30 mins

Reading Business Analytics textbook: Example 1.2 A Sales Transaction Database File, page 14 – 17.

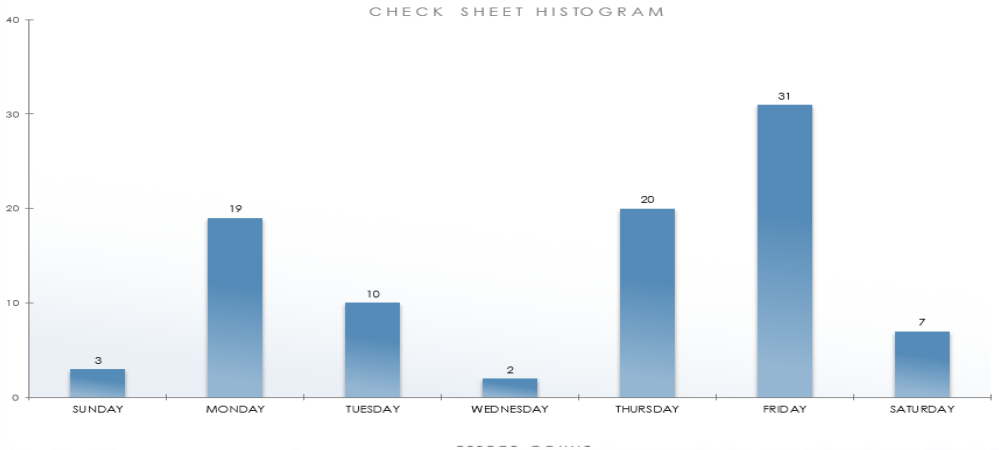
Discussion and answer:

- What are the key characteristics that IBM uses to define "big data" in the context of analytics, and why are these characteristics important for organizations?
- Discuss the different types of data classification based on measurement scales. Provide examples of each type and explain their characteristics.

sbsedu.vn

Saigon Business School

Excel Visualization



	A	B	C	D	E	F
1	Month	Product A	Product B	Product C	Product D	Product E
2	January	7792	5554	3105	3168	10350
3	February	7268	3024	3228	3751	8965
4	March	7049	5543	2147	3319	6827
5	April	7560	5232	2636	4057	8544
6	May	8233	5450	2726	3837	7535
7	June	8629	3943	2705	4664	9070
8	July	8702	5991	2891	5418	8389
9	August	9215	3920	2782	4085	7367
10	September	8986	4753	2524	5575	5377
11	October	8654	4746	3258	5333	7645
12	November	8315	3566	2144	4924	8173
13	December	7978	5670	3071	6563	6088

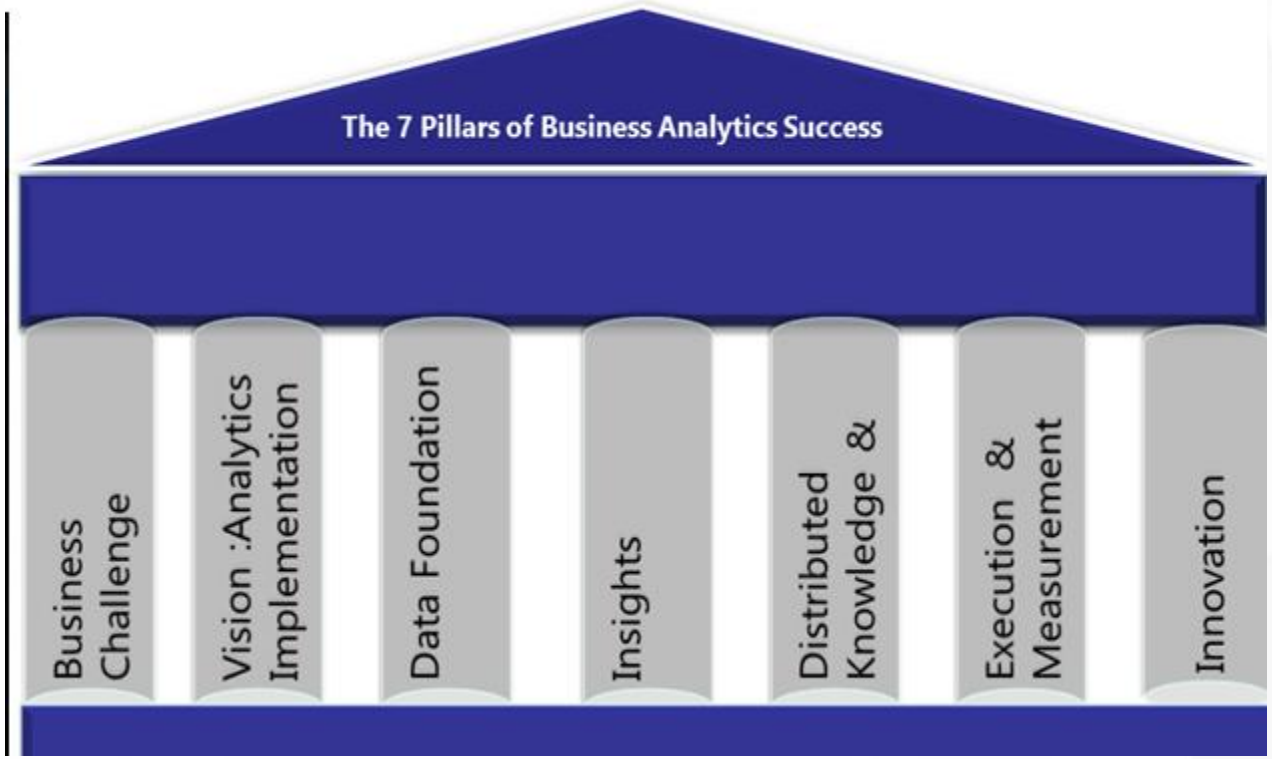
sbsedu.vn

Saigon Business School

Conclusion and Questions

- Foundation of Business Analytics

- Business analytics begins with the collection, organization, and manipulation of data and is supported by three major components.



sbsedu.vn

Saigon Business School

Thank you

sbsedu.vn