Introduction to Business Analytics 1st edition Richardson Solutions Manual

SKU: 9781265454340-SOLUTIONS

TBEXAM.COM

Chapter 2 End-of-Chapter Assignment Solutions

Multiple Choice Questions

1. (LO 2.3) The text in blog posts is generally considered to be
   a. structured data
   b. interval data
   c. unstructured data
   d. ratio data

2. (LO 2.3) Data that identify letter grades for students (A, B, C, D, F) are
   a. Ratio data
   b. Ordinal data
   c. Numerical data
   d. Interval data

3. (LO 2.3) Employee salaries are examples of
   a. interval data
   b. ratio data
   c. ordinal data
   d. unstructured data

TBEXAM.COM

4. (LO 2.3) If a company has 20 return transactions and 50 sales transactions, what is the proportion of return transactions?
   a. =20 ÷ 70
   b. =20 ÷ 50
   c. =70 ÷ 20
   d. =50 ÷ 20

5. (LO 2.3) If there are five transactions on 12/1, four transactions on 12/2 and nine transactions on 12/3, what is the proportion of the transactions that occur during the first three days of December took place on 12/1?
   a. =18 ÷ 5
   b. =5 ÷ 12
   c. =5 ÷ 18
   d. =5 ÷ 9

6. (LO 2.3) The data in the Transaction_Type column in Exhibit 2.2 are
   a. ratio data
   b. nominal data
   c. numerical data
   d. ordinal data

7. (LO 2.3) Averaging is a summarization method for which type of data?

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES

Introduction to Business Analytics 1st edition Richardson Solutions Manual

TBEXAM.COM

a. Categorical
b. Numerical
c. Both Categorical and Numerical

8. (LO 2.3) A dimension in Tableau is which type(s) of data?
a. Categorical
b. Numerical
c. Both Categorical and Numerical

9. (LO 2.3) A measure in Tableau is which type(s) of data?
a. Categorical
b. Numerical
c. Both Categorical and Numerical

10. (LO 2.4) Which function in Excel or Power BI can be used to remove nonprintable characters?
a. TRIM()
b. CLEAN()
c. REMOVE()
d. CUT()

Discussion Questions

1. (LO 2.4) What is ETL and how does it relate to obtaining the data?

TBEXAM.COM

Suggested Solution:

ETL is an acronym for Extract, Transform, and Load. It relates to obtaining the data because frequently the data you wish to analyze needs to first be obtained from a different data source. For example, perhaps the data is stored in a database and you wish to analyze it in Tableau or Excel, so you need to extract that data from the database, transform it for analysis, and then load it into the appropriate tool.

2. (LO 2.4) Why is validating the data for completeness and integrity an important component of obtaining the data?

Suggested Solution:

Validating the data is important to ensure that the data you requested or retrieved is complete – as in every record you requested is represented, and that the data has integrity – as in it is accurate.

3. (LO 2.1) How could economic indicators, such as unemployment and GDP, help predict sales in the next quarter and the next year?

Suggested Solution:

Economic indicators are examples of external data types. Both internal and external data types can help predict sales, but the external data types are primarily areas that the company doesn't have control over. Recognizing external impacts on sales (such as unemployment and GDP) can help a company recognize what might impact consumer behavior.

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES
Introduction to Business Analytics 1st edition Richardson Solutions Manual

TBEXAM.COM

TBEXAM.COM

4. (LO 2.1) If you could have any internal data you wanted, which data would you use to predict **Walmart**'s sale of pecans at Thanksgiving time each year?

Suggested Solution:

Answers here will vary, but Walmart could certainly benefit from its own data reporting historic sales of pecans by day, particularly around the Thanksgiving season.

5. (LO 2.1) Referencing the previous question concerning the internal data that can be used to predict **Walmart**'s sale of pecans at Thanksgiving: Which external data would be helpful in predicting these sales?

Suggested Solution:

Answers here will vary, but Walmart could benefit from external data sources about the price of pecans, the sales of pecans at other stores around Thanksgiving and the relative supply of pecans.

6. (LO 2.3) Why is net income or net sales better defined as ratio data instead of interval data?

Suggested Solution:

Net income and net sales are ratio data because 0 is meaningful for these measures (as is any measure related to monetary amounts or ratios).

7. (LO 2.3) Why is temperature measured using a Fahrenheit scale not considered to be ratio data?

TBEXAM.COM

Suggested Solution:

Fahrenheit temperature is interval (not ratio) data because zero is not numerically meaningful in Fahrenheit temperature. 0 degrees does not indicate the absence of temperature, it simply indicates one degree colder than 1 degree.

8. (LO 2.3) Are exam grades classified as interval data or ratio data? Why? Keep in mind that in typical exams, a score of 0 means that you answered every question wrong (or didn't answer the question at all). Does this 0 score indicate an absolute absence of knowledge? If Student A earns 40% on an exam and Student B earns 80% on the same exam, can we assume that, compared to Student A, Student B knows twice as much about the subject?

Suggested Solution:

This can be a tricky one for students, but exam grades are classified as interval data. Earning a 0 on an exam does not represent the total absence of knowledge – it simply represents not getting the answers right on that particular exam. A student who earns an 80% does not have twice the knowledge as a student who earns a 40%.

9. (LO 2.4) Which step in the ETL process (ensuring data quality, validating the data, cleansing the data, or performing preliminary exploratory analysis) do you think is the most time consuming? Which do you think is the most important? Why?

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM
Introduction to Business Analytics 1st edition Richardson Solutions Manual
TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES

TBEXAM.COM

Suggested Solution:

Answers will vary here, and it will depend on the dataset you are working with and the goals of the analysis project, in general.

10. (LO 2.1) Why is an understanding of the way data are structured in a relational database (particularly regarding primary keys and foreign keys) important?

Suggested Solution:

Understanding the structure of data and how PK and FKs are related will help an analyst request the correct data and/or retrieve the data appropriately when it is stored in a relational database and they need data that is stored across multiple tables.

Brief Exercises

1. (LO 2.3) Match the data examples to its appropriate data type: categorical or numerical.

Data Examples	Data Type (Categorical or Numerical)
Fahrenheit scale	Numerical
Survey answers: agree, indifferent, disagree	Categorical
Advertising expense	Numerical
Employee choice of retirement plan vendor (Fidelity, Edward Jones, Betterment)	Categorical
Level of customer service: good, neutral, bad	Categorical
Total monthly sales	Numerical
Employee count	Numerical

TBEXAM.COM

2. (LO 2.3) Match each data example to its appropriate data type. Choose from structured nominal, structured ordinal, structured interval, structured ratio, or unstructured.

Data Example	Data Type (Structured Nominal, Structured Ordinal, Structured Interval, Structured Ratio, or Unstructured)
SAT scores and scores on other standardized exams	Structured Interval data
Total assets	Structured Ratio data
Gold, silver, and bronze medals at the Olympics	Structured Ordinal data
School grades (A, B, C, D, F)	Structured Ordinal data
Depreciation Method (declining balance, straight-Line, units of production, etc.)	Structured Nominal Data
Conference call transcript	Unstructured Data

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES
Introduction to Business Analytics 1st edition Richardson Solutions Manual

TBEXAM.COM

Data Example	Data Type (Structured Nominal, Structured Ordinal, Structured Interval, Structured Ratio, or Unstructured)
Education level of employees	Structured Nominal Data
Credit score	Structured Ratio Data
Height of WNBA basketball players	Structured Ratio Data
Sales	Structured Ratio Data
Hair color	Structured Nominal Data

3. (LO 2.1, 2.2, 2.3, 2.4) Match each term to its appropriate definition.

Term	Definition
Data dictionary	a centralized repository of information about data containing a separate record for each field/variable in the database.
Nominal data	categorical data that cannot be ranked
Ordinal data	categorical data with natural order that allows them to be ranked
Ratio data	numerical data with an equal and definitive ratio between each data point, and with a meaningful zero
Interval data	numerical data measured along with a scale with an equal and definitive distance between each data point, but no meaningful zero

TBEXAM.COM

TBEXAM.COM

4. (LO 2.3) Identify each summarization method as either numerical, categorical or both.

Summarization Method	Numerical, Categorical or Both
Counting and grouping	Both
Averaging	Numerical
Proportion	Both
Ranking	Categorical
Summing	Numerical

5. (LO 2.3) What is the data type for each of these variables in a products table? Choose from short text, long text, currency, Y/N flag or number.

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES
Introduction to Business Analytics 1st edition Richardson Solutions Manual

TBEXAM.COM

Field Name	Data Type (Short Text, Long Text, Currency, Y/N flag, Number)
Product name	Short text
Product description	Long text
Sales price	Currency
Coupon available	Y/N flag
Prime shipping available	Y/N flag
Count on hand	Number

6. (LO 2.5) Match the software tool (Alteryx, Excel, Gretl, Power BI, Python, Tableau, Tableau Prep) with its description.

Description of Software Tool	Software Tool
Similar to Tableau Prep, but offers more file types for the output and more advanced analytics	Alteryx
Usually considered the best software tool for visualizations	Tableau
Multipurpose tool for basic data preparation, business analysis, and data visualizations	Excel
Requires knowledge of programming to use for business analysis	Python
Works with Tableau as a data-preparation tool	Tableau Prep
Intuitive open-source tool for users with a base level of statistics knowledge	Gretl
Microsoft product that goes beyond Microsoft Excel for reporting and visualizing data	Power BI

TBEXAM.COM

TBEXAM.COM

Problems

1. (LO 2.3) **Answers will vary but presented below is a possible solution.**

Product Name	Price	Pick-Up Availability Location	Primary Vendor	Review
Toy Story 4 Jessie	**$22.37**	**Fayetteville 2875 W. Martin Luther King**	**Walmart**	**N/A**

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES
Introduction to Business Analytics 1st edition Richardson Solutions Manual

TBEXAM.COM

2. (LO 2.3)

a. UnitPrice, UnitsinStock, UnitsonOrder, and ReorderLevel

b. No – it would not be a useful value because SupplierID is not numerical

c. Yes – UnitPrice is a numerical value, so the average would be a useful number.

3. (LO 2.3)

a. SupplierID is categorical and represents the unique identifier for each supplier. In this table, it is a foreign key (it would exist as a primary key in the Suppliers table), and indicates which supplier is associated with the given product.

b. The foreign keys in this table are SupplierID and CategoryID. Each of these fields exist as the primary key (unique identifier) in their related tables of Suppliers and Categories, respectively, and exist in this Products table to indicate which supplier and which category each Product is associated with. If one wanted to learn more details about the Category or Supplier, they could use the Foreign Key field as a lookup value to find the associated details in the Suppliers or Categories tables.

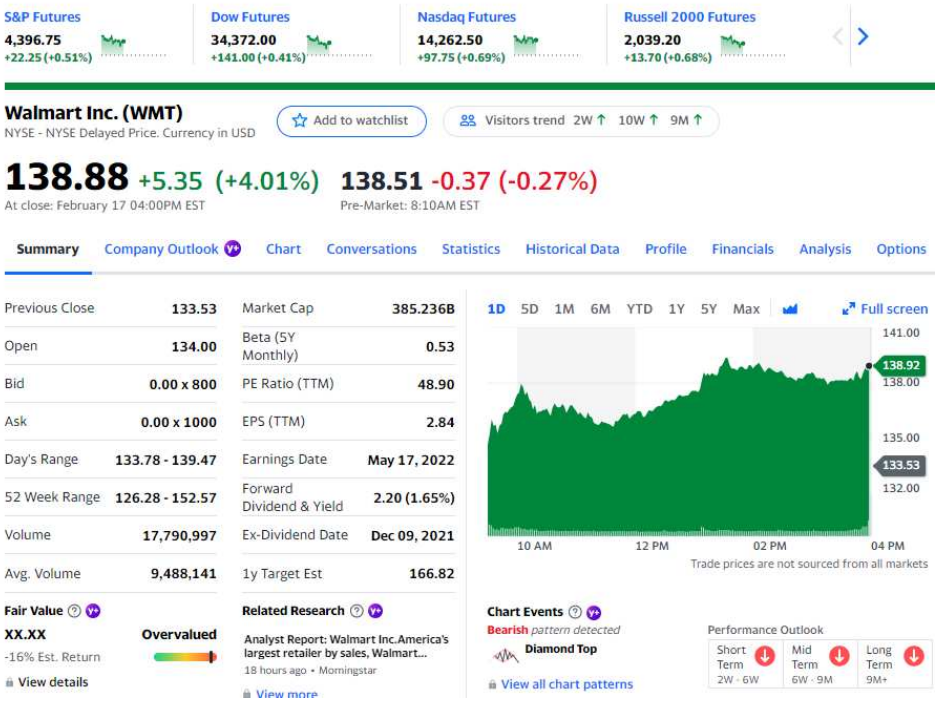
4. (LO 2.8)

a. Answers will vary, but the ProductName, ReorderLevel, and a calculated field of UnitsInStock + UnitsOnOrder. The answer to how often to run the report will vary, this is dependent on the business.

5. (LO 2.1, 2.2)

a.

TBEXAM.COM



Problem 2.5 jpg

b. 140.78

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES
Introduction to Business Analytics 1st edition Richardson Solutions Manual

TBEXAM.COM

c. 148.67

d. The adjusted closed price is Structured Data measured on the ratio scale because the data is meaningfully numeric.

e. Perhaps the company finance team is deciding if the valuation of the stock is related to the level of earnings or cash flows each period. That would require a combination of external data (stock price) to internal data (company's financial statements).

6. (LO 2.4)

a. The dates in the Due Date column are not uniform, some are written as long dates and some are formatted as short date.

b. Saved in the Excel file. Screenshot:

	A	B	C	D	E	F	G
1	**Customer**	**InvoiceAmount**	**Due Date**	**Days Past Due**		Today's Date	12/31/2025
2	IBM	406.57	10/3/2025	33			
3	PepsiCo	747.12	12/15/2025				
4	Smirnoff	373.59	9/12/2025	54			
5	3M	1415.76	11/27/2025				
6	Canon	308.63	11/30/2025				

7. (LO 2.4)

a. The count number does match. There are a variety of methods students could take, including just selecting the first column and viewing the summary statistic in the bottom right (it will read as 201, that is because it is including the header role in the count). They also could run a =count() function on the Customer column.

TBEXAM.COM

b. Saved in the Excel Solution file.

c. The students should choose to remove the records associated with null values in the Days Past Due field instead of replacing them with the average. The null values represent invoices that are not past due. Although, since the analysis will be performed in a PivotTable, those items that have null values will automatically be excluded from the PivotTable.

d. Solutions will vary. Students can talk about the advantages of using averages (using interpolation), or just removing the observation altogether. Either method is acceptable. Would you rather have an approximation (averaging) that might be biased, or no observation at all that would also remove an observation from the dataset that just happened to be missing important data.

TBEXAM.COM

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

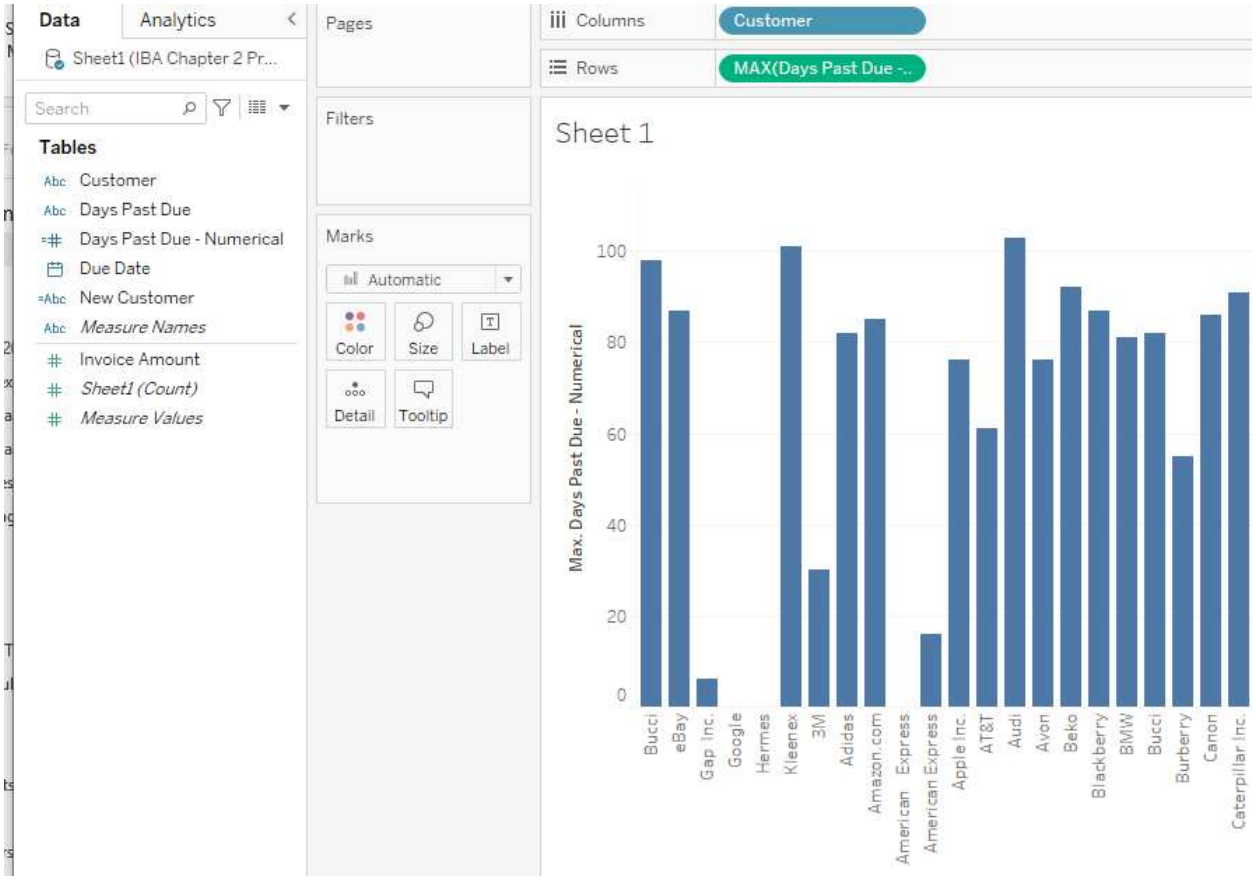
TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES
Introduction to Business Analytics 1st edition Richardson Solutions Manual

Data Visualization Exercises

1. (LO 2.4) Saved in the Power BI Solution file. Screenshot:

Customer	Max of Days Past Due
3M	30
Adidas	82
Amazon.com	85
American Express	16
Apple Inc.	76
AT&T	61
Audi	103
Avon	76
Beko	92

2. (LO 2.4)
   a. The data interpreter cleaned the extra columns associated with "Today's Date"
   b. Saved in the Tableau Solution file. Screenshot:



b. part IV: How does viewing these data in a visualization aid in understanding and spotting trends?

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES

Introduction to Business Analytics 1st edition Richardson Solutions Manual

TBEXAM.COM

TBEXAM.COM

The bar chart (or column chart) in the visualization helps see which companies have late invoices and which do not. They also allow a view of the diversity with some owing long periods of time, and others owing nothing at all (or very little).

TBEXAM.COM

TBEXAM.COM

© McGraw Hill LLC. All rights reserved. No reproduction or distribution without the prior written consent of McGraw Hill LLC.

TBEXAM.COM – THE WORLD'S LARGEST TEST BANK AND SOLUTIONS DATABASE WITH UNBEATABLE RATES
Introduction to Business Analytics 1st edition Richardson Solutions Manual