



DO NOT COPY
bufetekaye.22@gmail.com

Cleaning data

- Data--entry errors
- Wrong data type
- Wrong code
- Illogical data
- Missing data
- Duplicate data
- Unacceptable characters in text
- Censored or truncated data
- Inconsistent data representation
- Data is not atomic
- Broken references



22

aws re/start

A database can become disorganized over time, and issues can occur. Some SQL string functions can be used to clean data. Trimming data by using LEFT, RIGHT, and TRIM are ways to select only certain elements of strings and to remove certain characters. You can use CONCAT to combine strings from several columns and put them together. Other statements like LOWER can be used to force every character in a string to lowercase. Similarly, UPPER can be used to make all the letters appear in uppercase.



Activity: Identifying Errors



24

Activity

- A database might encounter the following types of errors. Discuss what these errors might look like in a functional database. Also discuss possible techniques for addressing these errors.
 - Data-entry errors
 - Wrong data type
 - Wrong code
 - Illogical data
 - Missing data
 - Duplicate data
 - Unacceptable characters in text
 - Censored or truncated data
 - Inconsistent data representation
 - Data is not atomic
 - Broken references



Checkpoint questions



What are two ways to use the INSERT INTO statement?



What is the purpose of NULL?

25

aws re/start

Answers:

1. The **INSERT INTO** statement can be used to add new **data** to a **database**.
The **INSERT INTO** statement can also add a new record to a **table**.
2. The term **NULL** is used to represent a missing value.

Key takeaways



© 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.

26

- When you insert a row, you must specify a column where the data will go.
- When you test for a NULL value, use the IS NULL and IS NOT NULL commands.
- Databases can become disorganized over time, and issues can occur. Some SQL string functions can be used to clean data.



Some key takeaways from this lesson include:

- When you insert a row, you must specify a column where the data will go.
- When you test for a NULL value, use the IS NULL and IS NOT NULL commands.
- Databases can become messy over time, and issues can occur. Some SQL string functions can be used to clean data.