

**7**

# **Data Cleaning Hacks in SQL**

# 1

## Remove Duplicates

Hack:

Use the ***ROW\_NUMBER()*** window function to identify and remove duplicate records.

# 2

## Handling Missing Data

Hack:

Replace missing values using the ***COALESCE()*** function, which returns the first non null value.

# 3

## Correct Date Formats

Hack:

Use ***CAST()*** or ***CONVERT()***  
to transform date strings into  
proper date formats

# 4

## Fix Inconsistent Data

Hack:

Use a combination of ***CASE*** statements to correct inconsistent values in a column.

# 5

## Remove Unwanted Characters

Hack:

Use ***REPLACE()*** to strip out unwanted characters from a string.

# 6

## Split and Extract Data

Hack:

Use functions like

***SUBSTRING\_INDEX()*** or

***SPLIT\_PART()*** to extract parts of a string.

# 7

## Group and Aggregate for Cleaning

Hack:

Use ***GROUP BY*** with aggregation functions (SUM, AVG, MIN, MAX, COUNT) to clean data, like removing duplicates



Thanks for taking time to  
read through this.

I hope these explanations  
are helpful to you.

Follow

Sylvia Wutche

For More Data Tips

