



Bookmarks

- ▶ Before You Start
- ▶ Module 1: Introduction to Data Science
- ▶ Module 2: Probability and Statistics for Data Science
- ▶ Module 3: Simulation and Hypothesis Testing
- ▶ Module 4: Exploring and Visualizing Data
- ▼ **Module 5: Data Cleansing and Manipulation**

Data Ingestion and Flow

Data Cleansing

**Lab**

Module 5: Data Cleansing and Manipulation &gt; Lab &gt; Lab Verification



Bookmark

## Cleaning Data

(1/1 point)

How many rows were removed from the automobile data as a result of treating missing values or duplicate rows?

**Answer: 23****23****Correct:**

To determine how many rows were removed, visualize the output of the module where you joined the two source datasets and note the total rows. Then visualize the output after the data has been cleansed to remove rows with missing values and duplicate rows and note the total rows.

**EXPLANATION**

There were 216 rows in the source data, and there should be 193 after the data is cleansed - which means 23 rows were removed.

*You have used 1 of 2 submissions*

Lab



- ▶ Module 6: Introduction to Machine Learning
- ▶ Final Exam and Survey

## Handling Outliers

(1/1 point)

How many rows were removed as a result of removing outliers?

✓ Answer: 19

**Correct:**

To determine how many rows were removed, visualize the output of the module after you cleaned the dataset and note the total rows. Then visualize the output after the outliers have been removed and note the total rows.

**EXPLANATION**

There were 193 rows after rows with missing values and duplicate rows were removed, and there should be 174 after outliers were removed - which means 19 rows containing outliers were removed.

*You have used 1 of 2 submissions*

© All Rights Reserved



© 2016 edX Inc. All rights reserved except where noted. EdX, Open edX and the edX and Open EdX logos are registered trademarks or trademarks of edX Inc.

POWERED BY  
**OPENedX**

