

1. What does a typical data wrangling workflow include?

1 / 1 point

- ☐ Predicting probabilities
- ☒ Validating the quality of the transformed data
- ☐ Recognizing patterns
- ☐ Using mathematical techniques to identify correlations in data

✓ Correct

Validating the quality of transformed data is an essential step in a data wrangling workflow.

2. OpenRefine is an open-source tool that allows you to:

1 / 1 point

- ☒ Transform data into a variety of formats such as TSV, CSV, XLS, XML, and JSON
- ☐ Enforces applicable data governance policies automatically
- ☐ Use add-ins such as Microsoft Power Query to identify issues and clean data
- ☐ Automatically detect schemas, data types, and anomalies

✓ Correct

Using OpenRefine, you can transform data into a wide variety of formats such as TSV, CSV, XLS, XML, and JSON.

• OpenRefine is an open-source tool that allows you to import and export data in a wide variety of formats, such as TSV, CSV, XLS, XML, and JSON. Using OpenRefine, you can clean data, transform it from one format to another, and extend data with web services and external data.

- DataPrep is an intelligent cloud data service that allows you to visually explore, clean, and prepare both structured and unstructured data for analysis. DataPrep can automatically detect schemas, data types, and anomalies.
- Watson Studio Refinery, available via IBM Watson Studio, allows you to discover, cleanse, and transform data with built-in operations. It transforms large amounts of raw data into consumable, quality information that's ready for analytics.
- Data Refinery offers the flexibility of exploring data residing in a spectrum of data sources. It detects data types and classifications automatically and enforces applicable data governance policies automatically.

3. What is one of the steps in a typical data cleaning workflow?

1 / 1 point

- ☒ Inspecting data to detect issues and errors
- ☐ Building classification models
- ☐ Clustering data
- ☐ Establishing relationships between data events

✓ Correct

Inspecting data to detect issues and errors is one of the first steps in a typical data cleaning workflow.