Quiz 2 - Data Collection and Data Wrangling

1. Web scraping is used to extract what type of data?   1 / 1 point

   ○ Text, videos, and data from relational databases

   ○ Images, videos, and data from NoSQL databases

   ◉ Text, videos, and images

   ○ Data from news sites and NoSQL databases

   ⊘ **Correct**
   Web scraping is used for extracting text, videos, and images from the web.

2. _____ focuses on cleaning the database of unused data and reducing redundancy and inconsistency.   1 / 1 point

   ◉ Normalization

   ○ Data Profiling

   ○ Data Visualization

   ○ Denormalization

   ⊘ **Correct**

3. OpenRefine is an open-source tool that allows you to: 1 / 1 point

   ○ Automatically detect schemas, data types, and anomalies

   ○ Use add-ins such as Microsoft Power Query to identify issues and clean data

   ◉ Transform data into a variety of formats such as TSV, CSV, XLS, XML, and JSON

   ○ Enforces applicable data governance policies automatically

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

4. When you're combining rows of data from multiple source tables into a single table, what kind of data transformation are you performing?   1 / 1 point

   ○ Normalization

   ◉ Unions

   ○ Denormalization

   ○ Joins

   ⊘ **Correct**
   Unions are a common structural transformation used for combining rows of data from multiple source tables.

5. When you detect a value in your data set that is vastly different from other observations in the same data set, what would you report that as?   1 / 1 point

- ⦿ Outlier
- ◯ Missing value
- ◯ Syntax error
- ◯ Irrelevant data

✓ **Correct**
Outliers are values in your data set that may be vastly different from other values in the same data field.