Data Dictionary

This activity uses a dataset called [marketing\_sales\_data.csv](https://www.coursera.org/learn/regression-analysis-simplify-complex-data-relationships/resources/fVqKH). It represents the amount of money spent on TV, radio, and social media promotions, as well as the corresponding sales. It is a fictional dataset that was created for learning purposes and made available on [Kaggle](https://www.kaggle.com/datasets/harrimansaragih/dummy-advertising-and-sales-data). The data has been modified for this activity.

The dataset contains:

**572 rows** – each row represents a different promotion budget

**5 columns**

|  |  |  |
| --- | --- | --- |
| **Column name** | **Type** | **Description** |
| TV | str | Television promotion budget, represented as low, medium, or high |
| Radio | int | Radio promotion budget (in millions) |
| Social Media | int | Social media budget (in millions) |
| Influencer | str | What type of influencer the promotion collaborated with (Mega, Macro, Nano, or Micro) |
| Sales | int | Total of sales (in millions) |

Remember, you can access and download the data for any Jupyter notebook activity from within the notebook itself by navigating to the **Lab Files** dropdown menu at the top of the page, clicking into the **/home/jovyan/work** folder, selecting the relevant data file, and clicking **Download**.