Data Dictionary

This project uses a dataset called **HR\_capstone\_dataset.csv.** It represents 10 columns of self-reported information from employees of a fictitious multinational vehicle manufacturing corporation.

The dataset contains:

14,999 rows – each row is a different employee’s self-reported information

10 columns

|  |  |  |
| --- | --- | --- |
| **Column name** | **Type** | **Description** |
| satisfaction\_level | int64 | The employee’s self-reported satisfaction level [0-1] |
| last\_evaluation | int64 | Score of employee's last performance review [0–1] |
| number\_project | int64 | Number of projects employee contributes to |
| average\_monthly\_hours | int64 | Average number of hours employee worked per month |
| time\_spend\_company | int64 | How long the employee has been with the company (years) |
| work\_accident | int64 | Whether or not the employee experienced an accident while at work |
| left | int64 | Whether or not the employee left the company |
| promotion\_last\_5years | int64 | Whether or not the employee was promoted in the last 5 years |
| department | str | The employee's department |
| salary | str | The employee's salary (low, medium, or high) |

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**.

The dataset can be found on [Kaggle](https://www.kaggle.com/datasets/mfaisalqureshi/hr-analytics-and-job-prediction). It has been repurposed for this project.