**Data Definitions**

The goal of building data definitions is to describe the features in your dataset with enough detail to:

* Identify any issues that will require cleaning
* Identify the features of your dataset
* Begin to understand how those features will support the data science project you have in mind

This subunit will teach you how to do all that and more.

1

**pandas Profiling Module Overivew**

Save



30 - 50 Minutes

11 Points

This package automates many of the tasks required to develop data tables and profiles. In addition to speeding up your data definitions, the process outputs a beautiful html file for fast reference. Read the main page here and install the module on your machine — you’ll put this tool to use shortly.

2

**NASA Meteorite Data Exercise**

Save



1 - 2 Hours

27 Points

Ready to practice building data definitions?

**Steps:**

1. Download the Jupyter Notebook file.
2. Run through the notebook to develop the pandas proflie html output for the Meteorite data.
3. Jot down any questions you have as you work and bring them up during your next mentor call.