Lesson 3:
Assessing Data

SEARCH

RESOURCES

CONCEPTS

Mentor Help
Ask a mentor on our Q&A platform

Peer Chat   3
Chat with peers and alumni

Lesson Outline

> Data wrangling process:
> - Gather
> - Assess (this lesson)
> - Clean

Assessing your data is the second step in data wrangling. When assessing, you
work, inspecting your dataset for two things: data quality issues (*i.e. content issu*
(*i.e. structural issues*).

Assessing is the precursor to cleaning. You can't clean something that you don'
lesson, you'll learn to identify and categorize common data quality and tidiness
the shortest and most "hands-off" code-wise of all four in the course because o
assessing relative to gathering and cleaning. We have tried to include quizzes w

This lesson will be structured as follows:

- You'll get motivated to assess (and later clean) the dataset for lessons 3 a
  trial data that compares the efficacy and safety of a new oral insulin to tre
- You'll learn to distinguish between dirty data and messy data
- You'll assess the data visually and programmatically to identify:
  - Data quality issues
  - Tidiness issues

- You'll learn about data quality dimensions and categorize each of the data
  identified above into its appropriate dimension

To begin, I want to introduce you to the dataset you will be assessing in this les