Recap: Week 6 Data Wrangling and 7 Data Cleaning

In Week 6 we covered Data Wrangling
which is the most time-consuming phase of Data Analytics.

In Week 7 we covered Data Cleaning
a major step in Data Wrangling.

It is very important to clean and organize your data.

Remember GIGO (Garbage-In, Garbage-Out)

READ the files marked READ.

For these topics, it is very worthwhile to also read/watch the supplementary material.

But not the supplementary material on Entity Resolution!

Go back to the video of the POI system development several times
to see each step of Data Wrangling in action.

Chapter 7 of the `pandas` book is on Data Wrangling

Chapter 5 of the `pandas` book has a section on *Handling Missing Values*

and Chapter 9 covers more advanced features for *Data Aggregation and Group Operations*