

Data Wrangling Hands On

Software Craftsmanship New York

Code Kata

Our Format

— — —

- Pair Programming
 - One laptop per pair
 - Driver - Navigator or Ping Pong
- Test-Driven Development (TDD)
- Clean Code
- Pomodoro Sessions 30 minutes
- Small Discussion

Deliberate practice

— — —

Our Goals:

- Practice writing the best code you can.

Themes:

- File I/O
- String Parsing
- Regular Expressions (Regex)
- Refactoring

Ready?

Data Wrangling - Part 1

— — —

- Download <http://codekata.com/data/04/weather.dat>
 - Daily Weather data for Morristown, NJ for June 2002
- Output the **day number** (column one) with **the smallest temperature spread** (the max temperature is the second column, the min is in the third column).

What did you think?

Data Wrangling - Part 2

— — —

- Download <http://codekata.com/data/04/football.dat>
 - Results from the English Premier League for 2001/2.
- The columns labeled 'F' and 'A' contain the total number of goals scored for and against each team in that season (e.g. Arsenal scored 79 goals against opponents, and had 36 goals scored against them).
- **Print the name of the team with the smallest difference in 'for' and 'against' goals.**

What did you think?

Thank You!

Keep practicing!

Based on the **Data Munging**
kata found in

<http://codekata.com/>

— — —