
# CSE 6040/x Bootcamp

Nested Data

NBA Data Example

# Nested Data

- Why so much emphasis on nested data structures?
- Let's look at the notebooks leading up to MT1
  - NB 1 – Intro to variables and nesting (student grades)
  - NB 2 – Association rule mining (in a nested data structure)
  - NB 4 – Underlying number data storage (numbers as strings in a complex data structure)
  - NB 5 – String manipulation/regex (in a complex and nested HTML data structure)
- What is the common theme here?
  - We are doing more complex coding, within the framework of the data structure.
  - So, we must understand how to address the underlying data structures, to perform the required operations/manipulations.
  - With practice, directly addressing each element of the data structure should become "2nd nature", so that you can focus on the data manipulations of the exercise requirements.

Georgia Tech
CREATING THE NEXT

# Nested Data – 2

- As we move forward in the course, the emphasis will go from basic data manipulations to applying these techniques to solve “real world” data scenarios.
- You will need to understand which data structures are being used and how to manipulate them.
- With practice, directly addressing each element of the data structure must become “2nd nature”, so that you can focus on the algorithmic/logic requirements of the exercises.
- So, what we want to do is provide you with some tools and techniques to help understand how to read, understand, and work with nested data structures.

Georgia Tech
CREATING THE NEXT

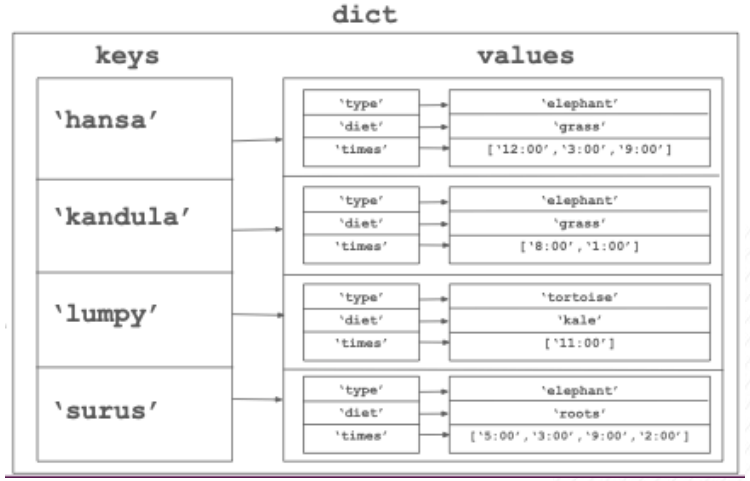
# Nested Data – 3

- We will use the NBA data as an example to work through.
- All of this is in Google Colab (links in Canvas/edX)
- For those interested, the below are links to the data and NBA API:
  - https://github.com/swar/nba_api/blob/master/docs/table_of_contents.md
  - https://www.playingnumbers.com/2019/12/how-to-get-nba-data-using-the-nba_api-python-module-beginner/
  - https://pypi.org/project/nba-api/
- This was the data source for a project in CSE6242.

Georgia Tech
CREATING THE NEXT

# Troubleshooting Nested Data

- Don’t try to print the whole data (it might be large)
- Use diagnostic functions like **len(), type()** to understand the structure of the nested data
- Usually, the nested data is logically structured. We recommend you create a mental map of how the data is organized - It will enable you to write code and extract the required data easily
- Drop it into Python Tutor to visualize.
- Develop the syntax of addressing each individual element of the data structure.




Georgia Tech
CREATING THE NEXT

# Open Discussion / Q&A

Georgia Tech
CREATING THE NEXT