

8/21/23

## Data Science Lifecycle

1. Data scientists use techniques to transform data into a visual representation that can be easier to understand by humans
2. Data science platform market is expected to grow by upwards of 20% annually.
3. Data science generally falls under math, statistics, and computer science.
4. The Life Cycle
  - a. Question
  - b. Collect Data
  - c. Wrangle Data
  - d. Analyze Data
  - e. Visualize Information
  - f. Communicate Information
5. All steps in the life cycle are all fluid

9/5/23

## Python Fundamentals

1. Datasets
  - a. The collection of data
  - b. Types of datasets
    - i. Lists
      1. Ordered, changeable, duplicates allowed
    - ii. Dictionaries
      1. Ordered, changeable, duplicates not allowed
    - iii. Sets
      1. Unordered, unchangeable\*, duplicates not allowed
    - iv. Tuples
      1. Unordered, unchangeable, duplicates allowed
2. Representing datasets with code
  - a. Column-oriented
    - i. Grouping by features
  - b. Row-oriented
    - i. Grouping by a single observation
3. Indexing

- a. Used to access values of a collection type
  - b. Python syntax to access values
    - i. List
      - 1. name[index]
    - ii. Dictionary
      - 1. name[key]
    - iii. Set
      - 1. for loop
    - iv. Tuple
      - 1. Name[index]
4. Iteration
- a. Can repeat processes with loops or recursion in Python
  - b. Python loop types
    - i. While loop
      - 1. while condition: statements
    - ii. For loop
      - 1. for thing in collection: statements
5. Useful methods
- a. [Dictionaries](#)
    - i. values()
    - ii. items()
    - iii. keys()
  - b. [Lists](#)
    - i. len()
    - ii. append()
    - iii. sort()
  - c. [Other](#)
    - i. range()
    - ii. print()
    - iii. split()
    - iv. type()
    - v. int()
    - vi. str()