COLUMBIA | ENGINEERING
EXECUTIVE EDUCATION

Week 2

Intermediate Python — Data Structures for your Analysis

Applied Data Science

Columbia University - Columbia Engineering

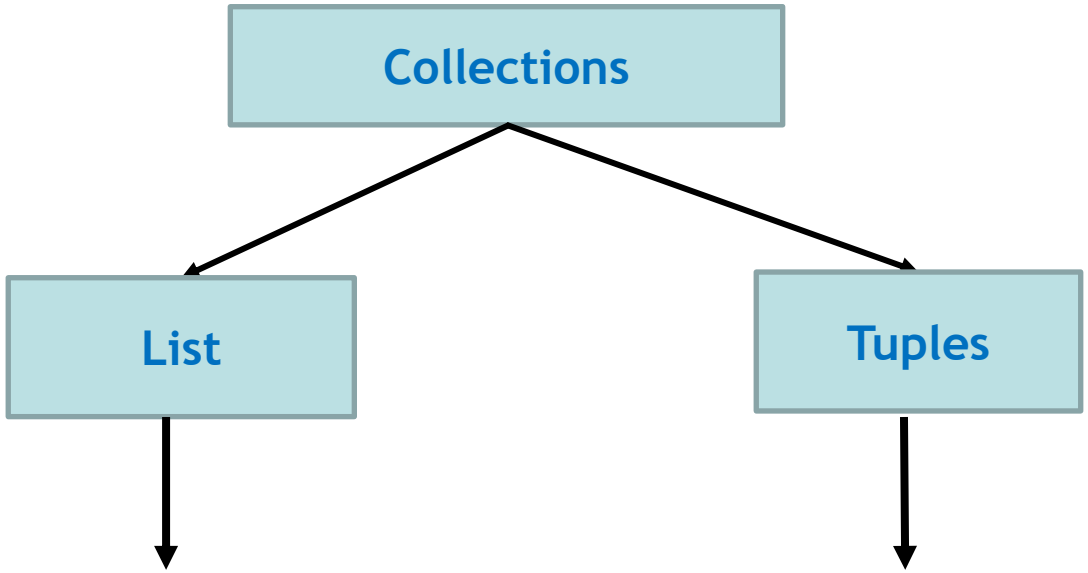
Course Agenda

COLUMBIA | ENGINEERING
EXECUTIVE EDUCATION

- Week 1: Python Basics: How to translate procedures into codes
- **Week 2: Intermediate Python — Data Structures for Your Analysis**
- Week 3: Relational Databases — Where Big Data is Typically Stored
- Week 4: SQL — Ubiquitous Database Format/Language
- Week 5: Statistical Distributions — The Shape of Data
- Week 6: Sampling — When You Can't or Won't Have ALL the Data
- Week 7: Hypothesis Testing — Answering Questions about Your Data
- Week 8: Data Analysis and Visualization — Using Python's NumPy for Analysis
- Week 9: Data Analysis and Visualization — Using Python's Pandas for Data Wrangling
- Week 10: Text Mining — Automatic Understanding of Text
- Week 11: Machine Learning — Basic Regression and Classification
- Week 12: Machine learning — Decision Trees and Clustering

Python Data Structures- Collections

In addition to basic data types like strings and floating integers, there are collections.



Lists are collections, that are;

- Sequential
- Ordered
- Mutable

Tuples are collections, that are;

- Sequential
- Ordered
- Immutable

Dictionaries are collections that are;

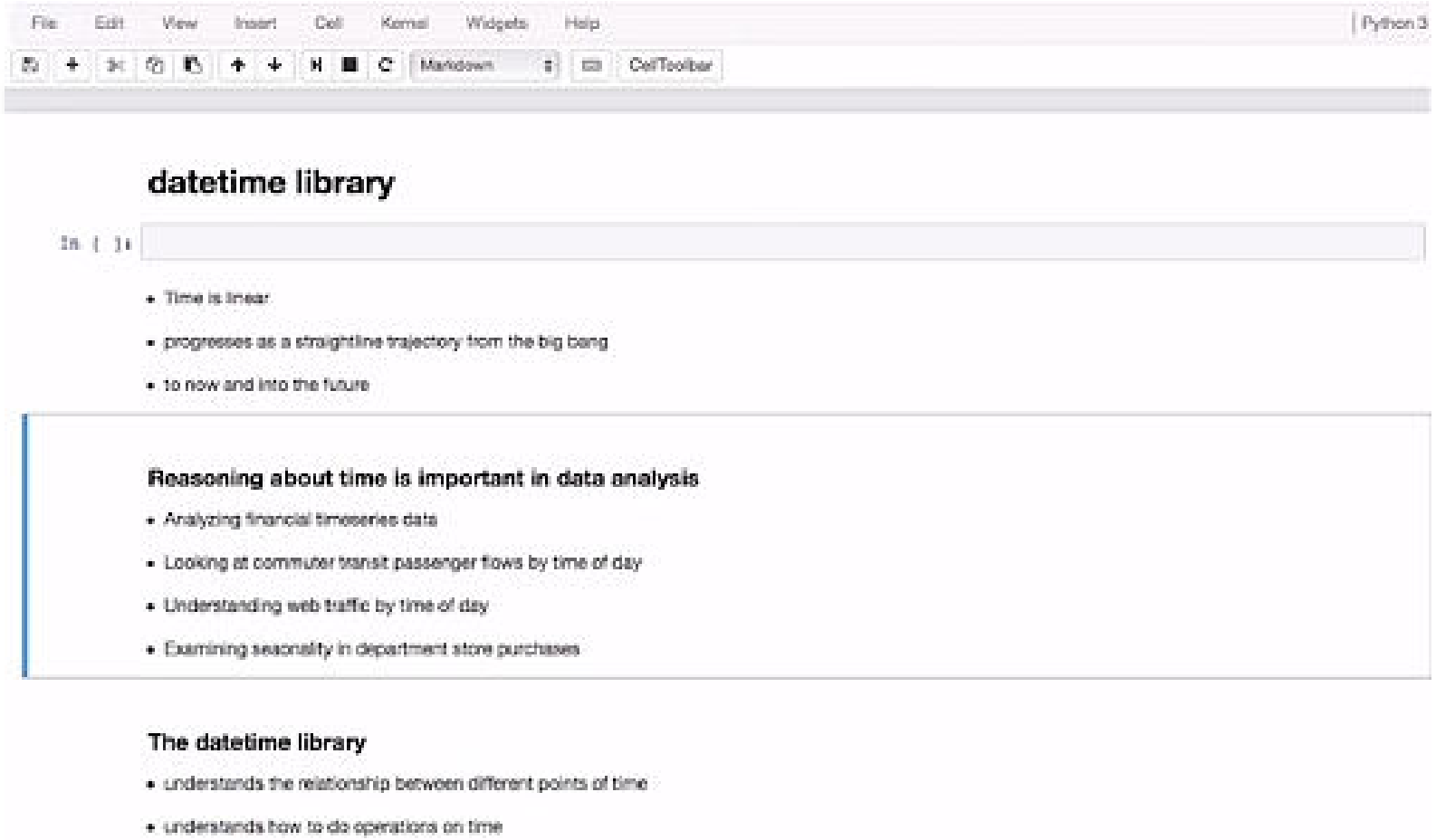
- Unordered
- Pair of elements with a key and a value
- Access values through keys
- Keys are immutable

Sets are collections that are;

- Unordered
- Collection of unique elements
- Does not contain key-value pairs
- Values are immutable

The Datetime Library

In Python, the ‘datetime' library is an extremely useful library for data analysis because time is a critical data element.



Using the Datetime Library

The datetime library helps us understand the relationship between different points of time and how to perform operations on time.

Components of a Datetime library

- Date: Stores the date (days, months, years)
- Time: Stores the time (hours, minutes, seconds)
- Datetime: Stores the date as well as time (days, months, years, hours, minutes, seconds)
- Timedelta: Duration between two datetime and date objects

www.emeritus.org