# **East West University**

### CI 280Data Analysis with Python

## Course Credit: 4

Prerequisite: CI 221

Instructor: Julie Zhang

Phone: 312 939-0111 ext 2211

Office: E211

E-mail: [jzhang@eastwest.edu](mailto:jzhang@eastwest.edu)

## Class Schedule:

**Tuesday 12:00 – 1:40PM – Remote Learning by Zoom**

Join Zoom Meeting

https://us02web.zoom.us/j/85737038447?pwd=MGVFTkc4SUlRaXR4V1VyTHlOL3Y0UT09

Meeting ID: 857 3703 8447

Passcode: dnu5LR

**Thursday 12:00 – 1:40 PM – Classroom E214**

## Course Description

This course covers the fundamentals of data Analysis by using Python.

## Topics include:

1. Fundamentals of Python programming
2. Structures of pandas
3. Data grouping and aggregation,
4. Data visualization
5. Data transformation

## Software

* Use the IPython interactive shell as your primary development environment. This course will use Jupyter Notebooks, Python 3 and various python modules.
* Students need to install the Anaconda platform ( Anaconda: <https://www.anaconda.com/>), which includes Jupyter/IPython as well all packages that will be required for the course.

## Textbook

**Python for Data Analysis, 2nd Edition**

**(https://www.amazon.com/Python-Data-Analysis-Wrangling-IPython/dp/1491957662)**

By [William McKinney](http://www.oreillynet.com/pub/au/5166)

Publisher: [O'Reilly Media](https://learning.oreilly.com/library/publisher/oreilly-media-inc/?utm_medium=referral&utm_campaign=publisher&utm_source=oreilly&utm_content=catalog&utm_content=catalog)

|  |  |
| --- | --- |
| https://covers.oreillystatic.com/images/0636920050896/lrg.jpg | * Use the IPython shell and Jupyter notebook for exploratory computing * Learn basic and advanced features in NumPy (Numerical Python) * Get started with data analysis tools in the pandas library * Use flexible tools to load, clean, transform, merge, and reshape data * Create informative visualizations with matplotlib * Apply the pandas group by facility to slice, dice, and summarize datasets * Analyze and manipulate regular and irregular time series data * Learn how to solve real-world data analysis problems with thorough, detailed examples. |

## Method of Instruction

Lectures and Labs

## Grading

#### Quizzes 10%

#### Labs 20%

#### Assignments 15%

Midterm Exam 20%

Final Exam 35%

A Above 90% of total points

B 89% - 80% of total points

C 79% - 70% of total points

D 69% - 60% of total points

F Below 60% of total points

|  |
| --- |
| **Late Work:** |

* All labs are collected during the classes.
* Assignments must be handed before the deadline, no points for late assignments

**Cheating**:

If cheating is confirmed on any assignment, no points. Cheating is a very serious offense. Cheating is defined as:

* Copying of another student's assignment or test,
* Allowing your own work to be copied, OR
* Developing a joint solution for an individual project.

**Make-up Exams**:

* No make-up exam for final exam.
* Any midterm make-up exam will lose 10 points of the midterm exam points, and make-up exam can only be taken within one week.

**Food and Drink**:

No food or drink is allowed in the computer labs.

**Class Schedule:**

|  |  |
| --- | --- |
| Date | Topic |
| Week 1 | * Introduction to the class * Set Jupyter Notebook * Python: Input, Processing and Output |
|  | Ch 2: [Python Language Basics, IPython, and Jupyter Notebooks](https://learning.oreilly.com/library/view/python-for-data/9781491957653/ch02.html#intro-python-environment)   * Python: Decision Structures * Repetition Structures * Functions |
| Week 2 | Ch 3: [Built-in Data Structures, Functions, and Files](https://learning.oreilly.com/library/view/python-for-data/9781491957653/ch03.html#intro-python-stdlib)   * Read from Write to Files * List, * Tuples, dictionary |
| Week 3 | Ch4 : NumPy Basics: Arrays and Vectorized Computation |
| Week 4 | Ch5 Pandas   * Pandas Data Structures * Series |
| Week 5 | Review Ch2 – Ch5  Midterm |
| Week 6 | Ch5 Pandas   * DataFrame * Essential Functions |
| Week 7 | Ch6 Data Loading  Ch7 Data Cleaning and Preparation |
| Week 8 | Ch8 Data Wrangling: Join, Combine, and Reshape |
| Week 9 | Ch9 Data Visualization  Plotting with pandas and seaborn |
| Week 10 | Ch10 Data Aggregation and Grouping   * Group By Mechanics * Data Aggregation |
| Week 11 | Final Exam |

**Useful Website for programmer:**

[**https://stackoverflow.com**](https://stackoverflow.com)

**About Python and other libraries**

Python: <https://www.python.org/>,

The Python Tutorial: <https://docs.python.org/3/tutorial/>

Pandas (Python Data Analysis Library) : <https://pandas.pydata.org/>

Numpy: <https://docs.scipy.org/doc/numpy/user/quickstart.html>

<https://matplotlib.org/>

<https://scikit-learn.org/>

**Dataset**

[https://www.kaggle.com](https://www.kaggle.com/)

**Cell Phone Usage**

Please turn off your cell phone during class time or exam. If a cell phone is used during exam, the student runs the risk of failing the exam. The instructor reserves the right to assign a penalty for cell phone use in class.