IE 0015: Introduction to Data Analytics

Lecture 7: Applying Python

Housekeeping

- Assignments
    - Lab 4 is due by midnight this Friday
    - Homework 1 due 2/19
    - Midterm is 2/19
    - Review lecture on 2/17

    NO LAB THIS WEEK!

Announcements

Final project 1:

Deliverable 1 (3/17):
- Data exploration, hypothesis + abstract

Deliverable 2:
- Final report

Final project 2:

Deliverable 1 (3/17):
- Idea + plan, hypothesis + abstract

Deliverable 2:
- Final report + code (any attempt!)

Today

- Define python modules and packages
- Use python for a data pipeline

Packages, modules, libraries, and objects

What is the scope of a variable?

Packages, modules, libraries, and objects

What is the scope of a variable?

…region of code where variable is available

Packages, modules, libraries, and objects

- These all refer to importing someone else's code
- (bringing their variables into scope)

	Example from nunpy	Example from pandas
package	numpy	pandas
module	numpy or numpy.linalg	pandas.core.dtypes
library	Numpy.linalg/linalg.py	Pandas.core.dtypes/dtypes.py
object	tensorsolve()	Categorical

Packages, modules, libraries, and objects

- ***Package:*** *organized collection of modules / libraries (__init__.py)*
- ***Library:*** *collection of modules (typically one folder)*

	Example from nunpy	Example from pandas
package	numpy	pandas
library	numpy or numpy.linalg	pandas.core.dtypes
module	Numpy.linalg/linalg.py	Pandas.core.dtypes/dtypes.py
object	tensorsolve()	Categorical

Packages, modules, libraries, and objects

- ***Module:*** *single .py file to be imported*
- ***Object:*** *variable defined in module to be exported*

	Example from nunpy	Example from pandas
package	numpy	pandas
library	numpy or numpy.linalg	pandas.core.dtypes
module	Numpy.linalg/linalg.py	Pandas.core.dtypes/dtypes.py
object	tensorsolve()	Categorical

Import syntax

import pandas as pd

import <module>

import <module> as <alias>

from <module> import <object>

Import syntax

import pandas as pd

pd.<object>

import <module>

<module>.<object>

import <module> as <alias>

<alias>.<object>

from <module> import <object>

<object>

For big packages

```
import pandas as pd

import numpy as np
```

A new package

```
import requests
```

Problem statement

Data Science Process
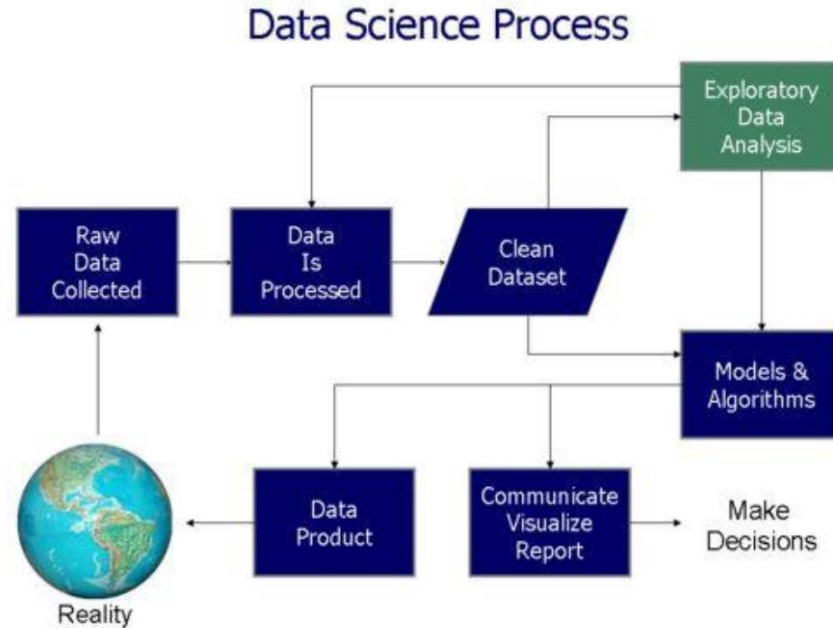
Problem statement

- A think-tank has approached IE0015 Spring 2025

- They have a new research direction: the future of social media

- They want our help to complete a more comprehensive social media research project

On to live coding

Let's see if we can make some progress…