

Assignment 7: Working with Data, Part 1

Due Wednesday by 11:59pm **Points** 100 **Submitting** an external tool
Available Oct 31 at 12am - Nov 10 at 11:59pm


Assignment Goals


The goals of this assignment are to:

- Import data using APIs
- Create nested dictionaries
- Learn how to account for imperfections in data

Getting Started

This assignment is the first part of a two-part project where we explore and visualize the relationship across data.

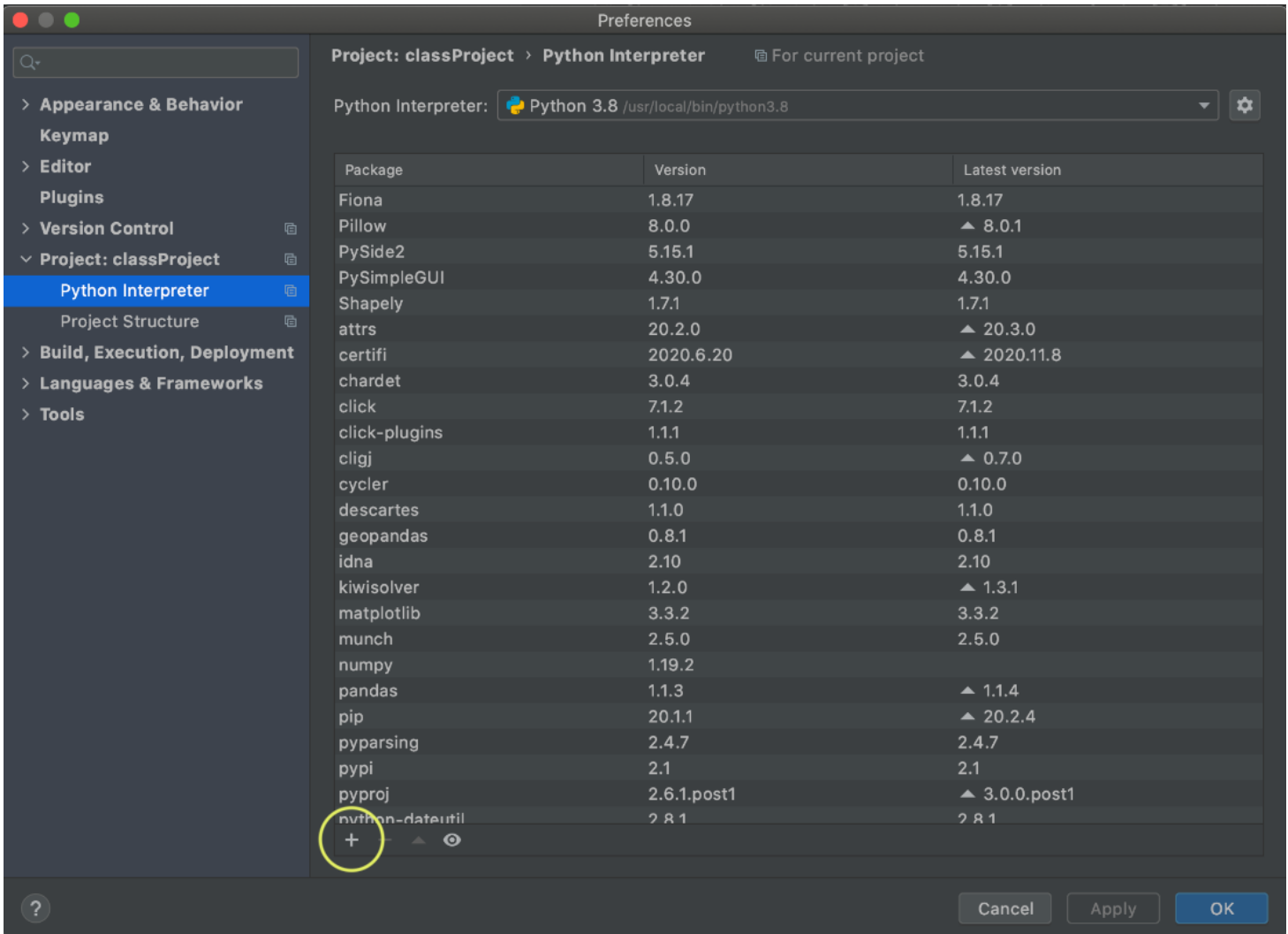
For this assignment, you will be retrieving three data sets from opendata.utah.gov  (<https://opendata.utah.gov/>) API: COVID cases, hospital characteristics, and average home prices. The three data sets will be combined into a nested dictionary where county names are the keys. In Part 2 (the next assignment) you will be visualizing this data in an interactive user interface that makes it easier to explore potential patterns across the data sets.

Start by downloading the starter code: [A7.zip](#) 

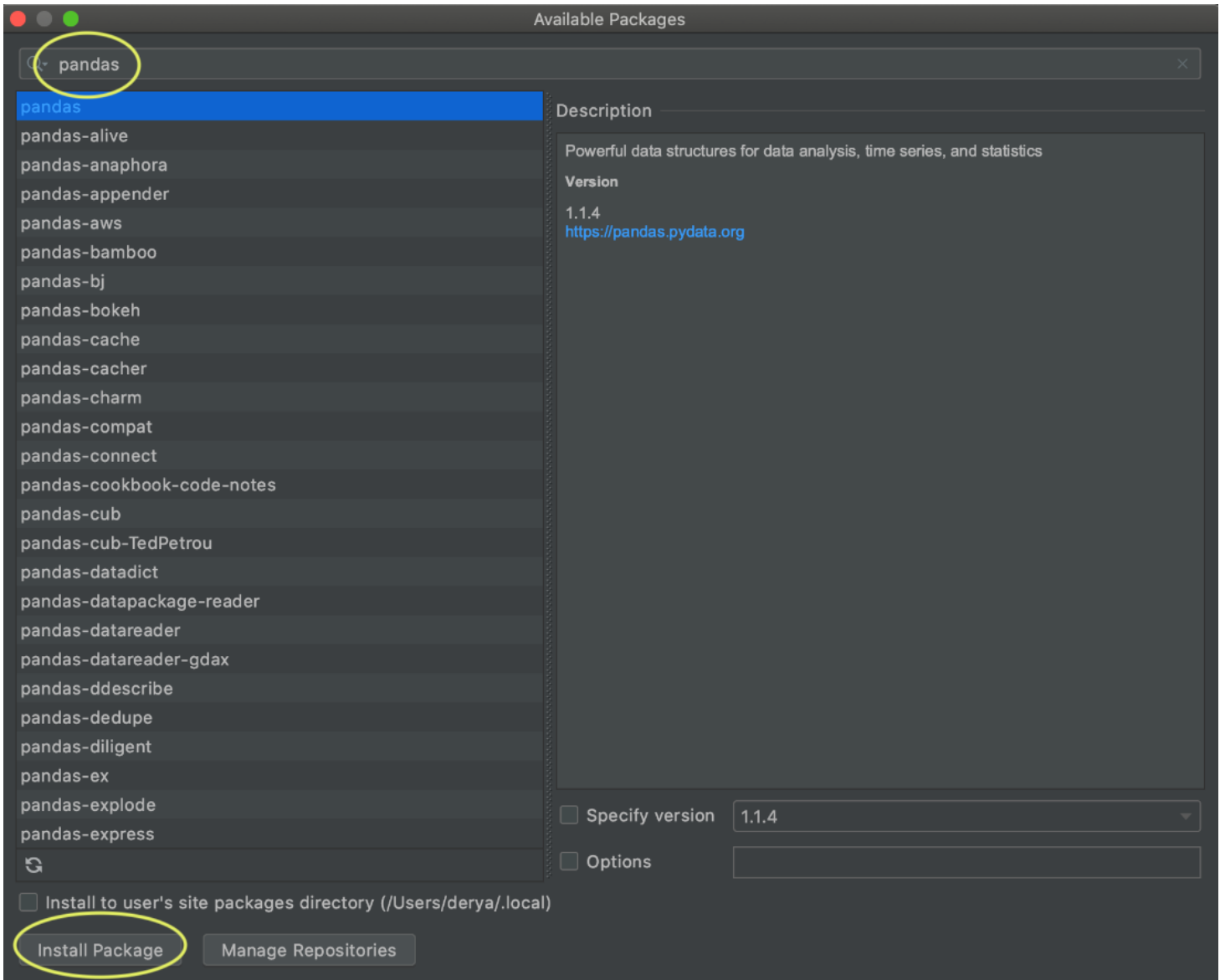
(https://utah.instructure.com/courses/796373/files/134138688/download?download_frd=1) . This project has a A7.py file where you will find starter code for your assignment and another python file, TableViewer.py, which you can call in your function.

Installing modules: You will need to install two Python modules for this assignment (requests and pandas). These are not standard parts of Python, so you will have to tell PyCharm to use it.

To do this, you go to Preferences --> Project Interpreter and then click the + button.



From here you will need to search for **requests** and click Install Package. You will need to do the same for the **pandas** module.



The A7.py file is broken into several sections:

- Data Retrieval: Importing data from the Utah Open Data portal using APIs
- Data Printing: Looking at the data is necessary for understanding what you have imported
- Dictionary Creation: You will be combining the three data sets into one dictionary.

Read through the starter code carefully. We include code for reading in one of the data sets into the complete_dict dictionary, but you will need to write your own code to read in the other two data sets. If you run the starter code, you will be able to see a print out of the first five rows of the initial COVID data set and complete_dict filled with only the COVID data.

Assignment Activities

Data Retrieval & Data Printing

We have provided one example of how to import data using the API. Using the `view_table()` function you can look at the first five lines of what you have imported.

You are responsible for importing the next two datasets:

1. Utah Hospital Characteristics: <https://opendata.utah.gov/Health/Utah-Hospital-Characteristics/ierb-h3t5> ➞ <https://opendata.utah.gov/Health/Utah-Hospital-Characteristics/ierb-h3t5>
2. Average Home Price By County In Utah 1996-2015: <https://opendata.utah.gov/Social-Services/Average-Home-Price-By-County-In-Utah-1996-2015/sma7-tpu2> ➞ <https://opendata.utah.gov/Social-Services/Average-Home-Price-By-County-In-Utah-1996-2015/sma7-tpu2>

In the example we provided, we are pulling the “Utah COVID Cases by County Map”:
<https://opendata.utah.gov/Health/Utah-COVID-Cases-by-County-Map/y4r8-7n5m> ➞ <https://opendata.utah.gov/Health/Utah-COVID-Cases-by-County-Map/y4r8-7n5m>

You can find the URLs for the APIs by following each link and copying the API endpoint.

Utah COVID Cases by County Map Health

This map represents the most recent COVID cases by County in Utah.

About this Dataset

Updated **April 15, 2020**

Data Last Updated: April 15, 2020 | Metadata Last Updated: April 15, 2020

Date Created: April 15, 2020

Views: **197** | Downloads: **25**

Data Provided by: Utah Dept of Health | Dataset Owner: Utah Open Data Portal Kung-Fu Master

[Contact Dataset Owner](#)

Dataset Summary

Organization: Utah Dept of Health

Microsites

Microsites: Department of Health

Topics

Category: Health

Tags: This dataset does not have any tags

Licensing and Attribution

License: Public Domain

Source Link: <https://coronavirus.utah.gov/>

Access this Dataset via SODA API

The Socrata Open Data API (SODA) provides programmatic access to this dataset including the ability to filter, query, and aggregate data.

[API Docs](#) [Developer Portal](#)

API Endpoint: <https://opendata.utah.gov/resource/y4r8-7n5/> JSON [Copy](#)

Dictionary Creation

You will create a nested dictionary that will combine all of the information from the data sets and organize them by county.

The final dictionary will have a structure similar to this:

```
{'county1': {'beds': 1, 'cases': 2, 'housing': 3}, 'county2': {'beds': 1, 'cases': 2, 'housing': 3}}
```

As you are constructing the dictionary pay attention to the following things:

- How will you handle incomplete entries?
- Are column names consistent across data sets?
- What is the type of the data (ex. *integer*, *string*, *float*) you want to store?

Your final nested dictionary should look something like this (remember dictionaries are *unordered*, so your dictionary might not look exactly similar)

```
>>> print(complete_dict)
{'BEAVER': {'beds': 72}, 'BOX ELDER': {'beds': 74, 'cases': 13}, 'CACHE': {'beds': 170, 'cases': 36}, 'CARBON': {'beds': 39, 'cases': 2}, 'DAVIS': {'beds': 481, 'cases': 21}, 'housing': 75800.0}, 'DUCHESNE': {'beds': 49, 'cases': 3}, 'GARFIELD': {'beds': 36, 'cases': 1}, 'GRAND': {'beds': 17, 'cases': 1}, 'IRON': {'beds': 50, 'cases': 15}, 'JUNEAU': {'beds': 25}, 'KANE': {'beds': 25, 'cases': 3}, 'MILLARD': {'beds': 37}, 'SALT LAKE': {'beds': 3607, 'cases': 1187, 'housing': 89900.0}, 'SAN JUAN': {'beds': 36, 'cases': 9}, 'SANPETE': {'beds': 43}, 'SEVIER': {'beds': 29}, 'TOOELE': {'beds': 47, 'cases': 34, 'housing': 61600.0}, 'UINTAH': {'beds': 39, 'cases': 6}, 'UTAH': {'beds': 1372, 'cases': 335, 'housing': 85600.0}, 'WASATCH': {'beds': 19, 'cases': 102, 'housing': 138500.0}, 'WASHINGTON': {'beds': 270, 'cases': 45, 'housing': 90200.0}, 'WEBER': {'beds': 606, 'cases': 105, 'housing': 52500.0}, 'SUMMIT': {'beds': 40, 'cases': 306}, 'EMERY': {'cases': 3}}
```

Submitting

Submit your A7.py file by the due date.

Extra Food for Thought

If you have submitted your assignment, great, you are done! If you want more practice, try the following (note this will not be graded and please do not submit this work):

1. Can you download other data sets? Can you write the data to a file?
2. Could you nest your dictionary using different attributes?
3. How can you use this dictionary to explore the relationship between COVID and socio-economic factors in Utah? (though feel free to come to TA office hours and talk about it)

