bConnected powered by Google

DANQING ZHANG <danqing0703@berkeley.edu>

Quick question about data-source

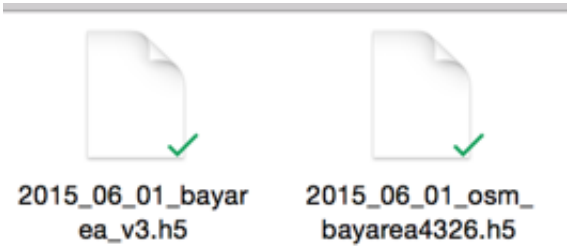
9 messages

DANQING ZHANG <danqing0703@berkeley.edu> Mon, Feb 22, 2016 at 6:04 PM
To: SAMUEL MARSH MAURER <maurer@berkeley.edu>

Hi Sam,

This is Danqing, really sorry to interrupt you, but I have a very quick question about h5 file.

I am wondering how are the below two files read, and then converted into orca tables? I went through the datasources.py file in bayarea_urbansim, but I could not figure out how this works.



Many thanks, Best, Danqing

--
Danqing Zhang

Ph. D Student
UC Berkeley, Civil & Environmental Engineering

DANQING ZHANG <danqing0703@berkeley.edu> Mon, Feb 22, 2016 at 6:09 PM
To: SAMUEL MARSH MAURER <maurer@berkeley.edu>

I am thinking I want to create a new h5 file that contains the data I needed, but I do not know where I can change the reading-h5 file path code. Or should I just load a new csv file or ? I am kind of looking for suggestions~~

Many thanks!

Best, Danqing
[Quoted text hidden]

Samuel Maurer <maurer@berkeley.edu> Mon, Feb 22, 2016 at 10:57 PM
To: DANQING ZHANG <danqing0703@berkeley.edu>

Hi Danqing!

Yes, that makes sense. Here's how the h5 files are loaded, I think.

In settings.yaml there's a line specifying the filename:
https://github.com/ual/bayarea_urbansim/blob/arb/configs/settings.yaml#L150

And you're right that it's datasources.py that loads the file, but inside urbansim_defaults:
https://github.com/ual/urbansim_defaults/blob/arb/urbansim_defaults/datasources.py#L42-L46

Then this line pulls the 'buildings' table out, for example:
https://github.com/ual/urbansim_defaults/blob/arb/urbansim_defaults/datasources.py#L84

It might be easier to load a CSV file, I'm not sure! That's what I did for the Craigslist data:
https://github.com/ual/bayarea_urbansim/blob/arb/datasources.py#L339-L341

Sam
[Quoted text hidden]

DANQING ZHANG <danqing0703@berkeley.edu> Tue, Feb 23, 2016 at 3:22 PM
To: Samuel Maurer <maurer@berkeley.edu>

Great, thanks! it helps!

I am also wondering about the pdtarget.csv file

https://github.com/ual/bayarea_urbansim/blob/arb/data/pdatargets.csv

I am wondering what exactly is it, and where the data come from..And I guess there is something wrong with the csv format.

Thanks so much again!

Best, Danqing

Samuel Maurer <maurer@berkeley.edu> Tue, Feb 23, 2016 at 3:32 PM
To: DANQING ZHANG <danqing0703@berkeley.edu>

Hah, yeah, that's not a csv. Maybe it's a tsv?

It has to do with regional land use planning scenarios. PDA stands for Priority Development Areas, which are districts near transit that are supposed to accommodate most of the region's future growth. I don't know whether the numbers in that file are used in a forecasting model, or just to evaluate results.

Here's some more info about the planning process. UrbanSim was used to create the scenario forecasts for the most recent regional plan. http://abag.ca.gov/priority/
[Quoted text hidden]

DANQING ZHANG <danqing0703@berkeley.edu> Tue, Feb 23, 2016 at 7:26 PM
To: Samuel Maurer <maurer@berkeley.edu>

Hi Sam,

Many thanks for the reply! I just have a few questions that I still cannot get the answer, and I am wondering could I ask you, really sorry for interrupting you.....I guess I can manage to make a wiki page also after I get these answers. Thanks!

【1】I found that I don't understand when these three files are called:
neighborhood_vars.yaml, price_vars.yaml and settings.yaml
I searched and got the idea rrh_estimate.yaml etc, is called through hedonic_estimate function of urbansim_defaults package, and is packaged as an orca step in the models.py file.

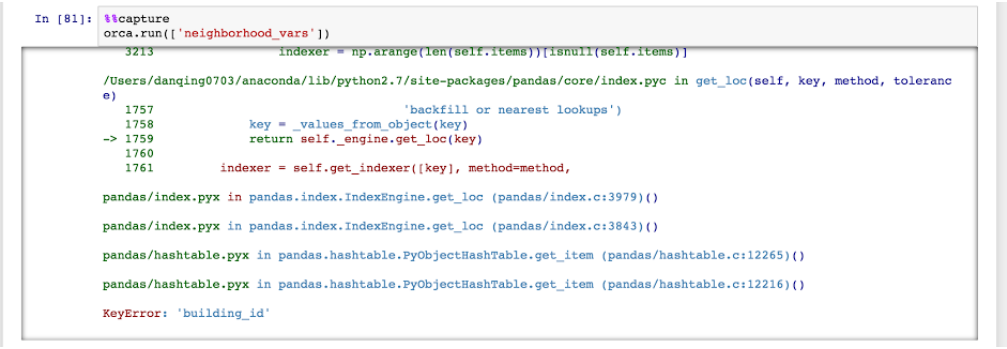
And I searched and got the idea that settings.yaml, everything is settled. like "aggregations" and "unit_aggregations" are defined. etc. But I do not know when this yaml is called.

Also, I also do not know when neighborhood_vars.yaml and prices_vars.yaml are used to update certain columns of "nodes" table from other dataframes like residential_unit and households etc. But again I failed to figure out how this is called. I can see this is an orca step, but I did not get the idea how this orca step is created.

【2】
some Key errros some times occur, like "unit_id" and "building_id" after I run the following orca steps(I just random try out everthing)

- orca.run(["travel_model_output"])

- orca.run(['neighborhood_vars']) in another ipython notebook



【3】
what data are actually used?
I know:
from HDF5 file that we mentioned "store", parcels, buildings,households,jobs,zones are created
and from below csv files certain tables or columns are added through datasources.py

But how are they linked? Like in postgresql , I can visualize the table relationships , to know how tables are connected and linked, do we have that for urbansim dataframes?
Also is pdatargets.csv actually used in current estimation & simulation?
And I guess I am getting errors like building_ids, because of the table relationships...
And I do not through what steps are multiple table joined?Like I can get this error when I randomly play with the

code.

```
Exception                                 Traceback (most recent call last)
<ipython-input-42-c0bfaa733c89> in <module>()
      5         geom_name='ZONE_ID', # from JSON file
      6         join_name='zone_id', # from data frames
----> 7         precision=2)

/Users/danqing0703/anaconda/lib/python2.7/site-packages/urbansim-3.1.dev0-py2.7.egg/urbansim/maps/dframe_explorer.pyc
in start(views, center, zoom, shape_json, geom_name, join_name, precision, port, host, testing)
    130         if join_name not in views[k].columns:
    131             raise Exception("Join name must be present on all dataframes - "
--> 132                             "'%s' not present on '%s'" % (join_name, k))
    133
    134     config['schema'] = simplejson.dumps(get_schema())

Exception: Join name must be present on all dataframes - 'zone_id' not present on 'zones'
```

【4】

visualization

I am trying to run this one:https://github.com/ual/bayarea_urbansim/blob/master/scripts/make_pda_map.py

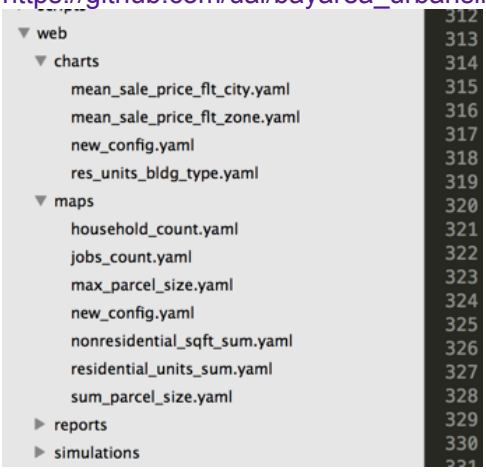
But got the below errors...I am wondering have we been changing urbansim_explorer that we are changing the visualization part?

```
In [82]: from urbansim_explorer import sim_explorer as se
         import sys

         runnum = int(sys.argv[1])

         parcel_output = 'runs/run%d_parcel_output.csv' % runnum
         zone_output = 'runs/run%d_simulation_output.json' % runnum
         outfile = '/var/www/html/sim_explorer%d.html' % runnum

         se.start(
             zone_output,
             parcel_output,
             port=8080,
             host='0.0.0.0',
             write_static_file=outfile
         )

         ---------------------------------------------------------------------------
         ImportError                               Traceback (most recent call last)
         <ipython-input-82-1628638ebd60> in <module>()
         ----> 1 from urbansim_explorer import sim_explorer as se
               2 import sys
               3
               4 runnum = int(sys.argv[1])
               5

         ImportError: No module named urbansim_explorer
```

【5】

in the web folder of bayarea_urbansim, there are many yaml files...I am wondering whether they are used and when and how shall they be used? A bit similar to question 1.

https://github.com/ual/bayarea_urbansim/tree/master/web



Many thanks!

I think these are currently all my questions for urbansim, and I am thinking I need to solve them so that I can write code that can be used inside urbansim. Many thanks and look forward to your reply!

Best, Danqing

DANQING ZHANG <danqing0703@berkeley.edu> Tue, Feb 23, 2016 at 7:27 PM
To: Samuel Maurer <maurer@berkeley.edu>

Hi Sam, I hear that you are preparing for qualifying exam and are very busy,...if these questions are too time-consuming to answer, do you have suggestions like who should I communicate with ...Many Thanks for the constant help!

Best, Danqing
[Quoted text hidden]
--
Danqing Zhang

Ph. D Student
UC Berkeley, Civil & Environmental Engineering

DANQING ZHANG <danqing0703@berkeley.edu> Tue, Feb 23, 2016 at 7:29 PM
To: Samuel Maurer <maurer@berkeley.edu>

Also, I am using "grep -R "neighborhood_vars.yaml" to find out certain matches in code. How do you find matches in multiple repos? I do find sometimes, I am just missing out something that is defined in either urbansim.defaults or urbansim.utils...

Thanks!
[Quoted text hidden]

DANQING ZHANG <danqing0703@berkeley.edu> Tue, Feb 23, 2016 at 7:34 PM
To: Samuel Maurer <maurer@berkeley.edu>

And I am kind of writing my own reference for urbansim review...but I find the above questions, especially the table relationships one, bother me a lot...
[Quoted text hidden]

---

urbansim.pdf
820K