

“Long Live *Keju*! The Persistent Effects of China’s Civil Examination System”

Description of the replication package:

Please note that a few of the software packages employed to produce the results in the paper require Internet connection to install and run. And ALL paths in the used code files must be changed manually to exactly where the downloaded files are stored. The replication package contains the following files.

The STATA data files are as follows:

census10_final.dta is for replicating Tables 1-9 (but not Table 6), A1-A3. For Tables 6 and A4, we use the confidential surname and location information in the *China Family Panel Studies* for construction. According to the data access policy of the Institute of Social Science Survey, such data could only be accessed via the computer facilities available at the Institute.

census05_final.dta is for replicating Table 10

figure2.dta and figure4.dta are for replicating Figure 2 and Figure 4, respectively.

The following do-file replicates all the tables (except for Tables 6 and A4) and figures (except for Figure 3) in the paper:

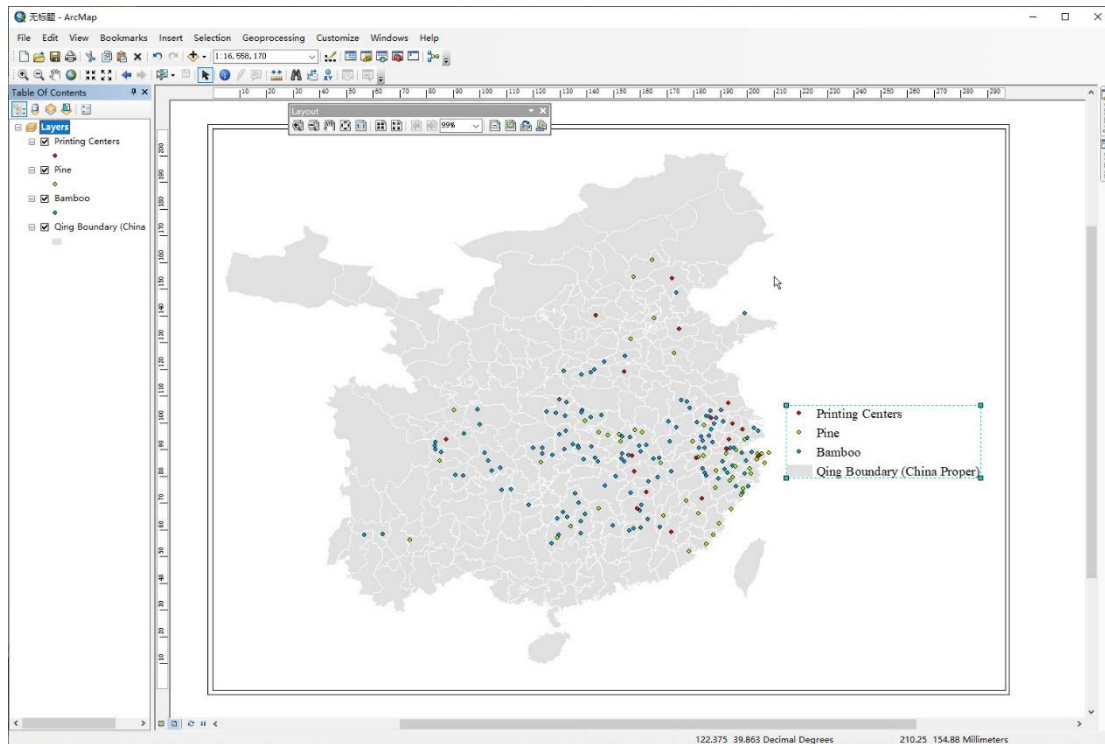
keju_final.do

The do-file was run on Stata/MP version 14.0 for Windows (64-bit x86-64).

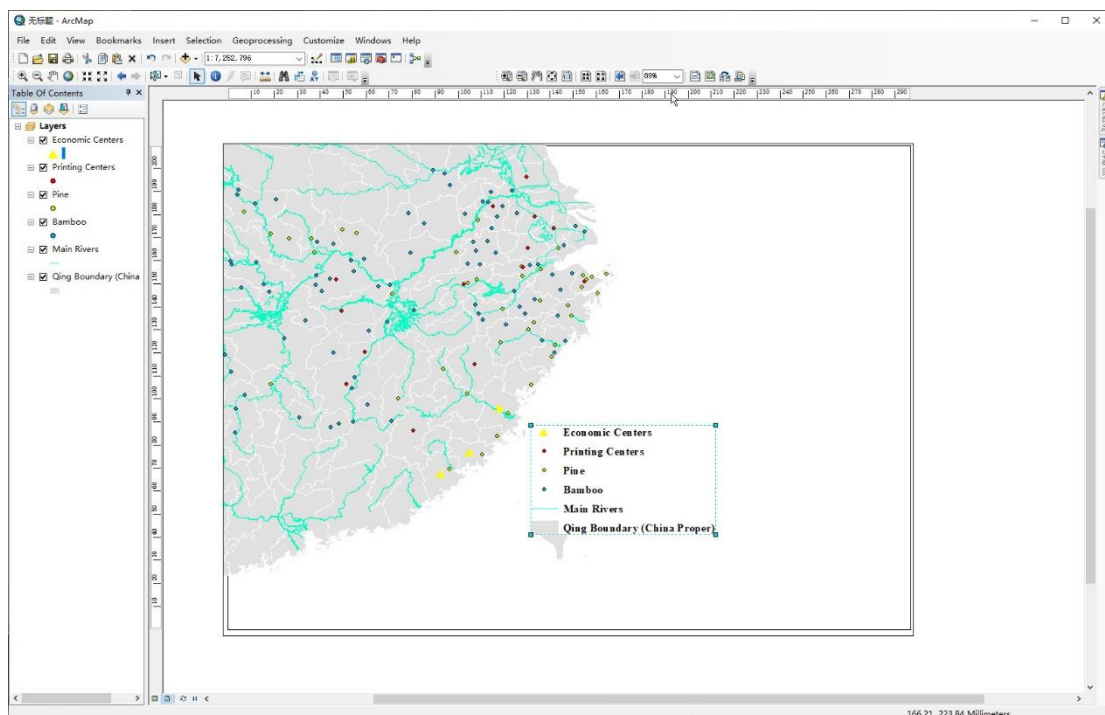
The GIS shapefiles are as follows:

Figure 3: The zipped file “figure3” contains the shapefiles for reproducing Figure 3.

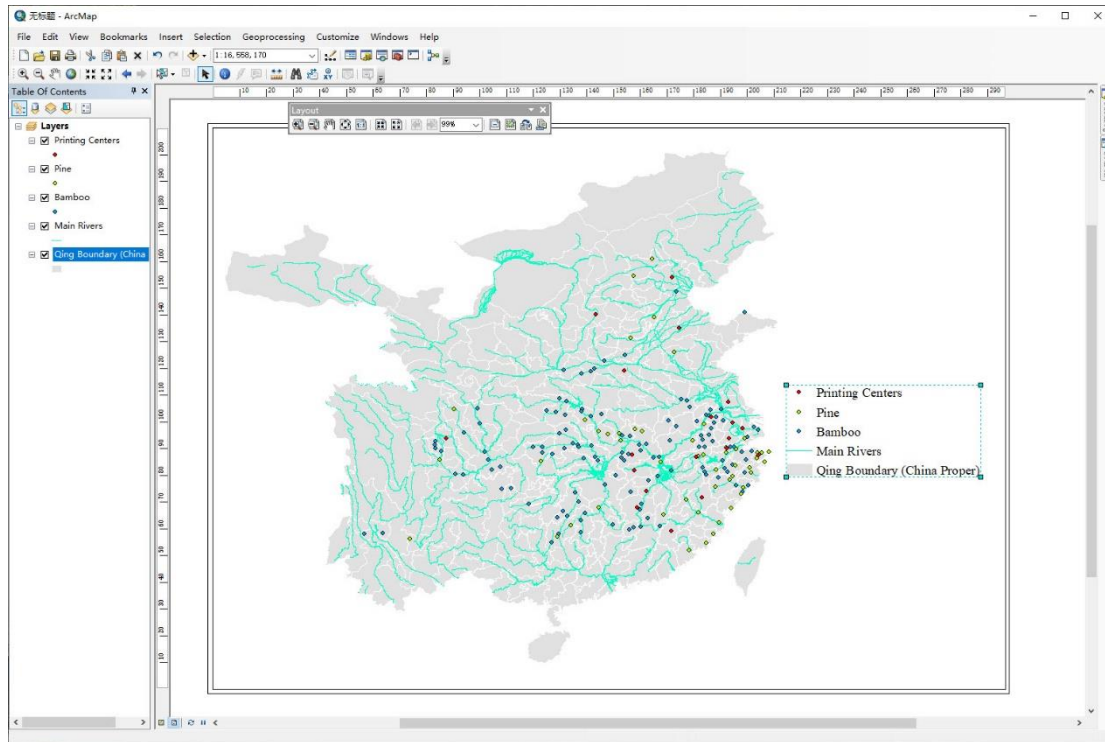
- Please refer to the following screen shot of ArcGIS for generating Figure 3A:
Step 1) Add the qing_china_proper.shp, printing_centers.shp, pine.shp, and bamboo.shp to the ArcMap;
Step 2) Adjust the colors;
Step 3) Finish by adding the legend and export map accordingly.



- Please refer to the following screen shot for generating Figure 3B:
 Step 1) Add the qing_china_proper.shp, printing_centers.shp, pine.shp, and bamboo.shp, MainRiver.shp, and Fujian_econ_centers.shp to the ArcMap;
 Step 2) Zoom in the Fujian Province (located in southeastern China);
 Step 3) Adjust the colors;
 Step 4) Add the legend and export map;
 Step 5) Finish by adding the arrows for illustration manually.

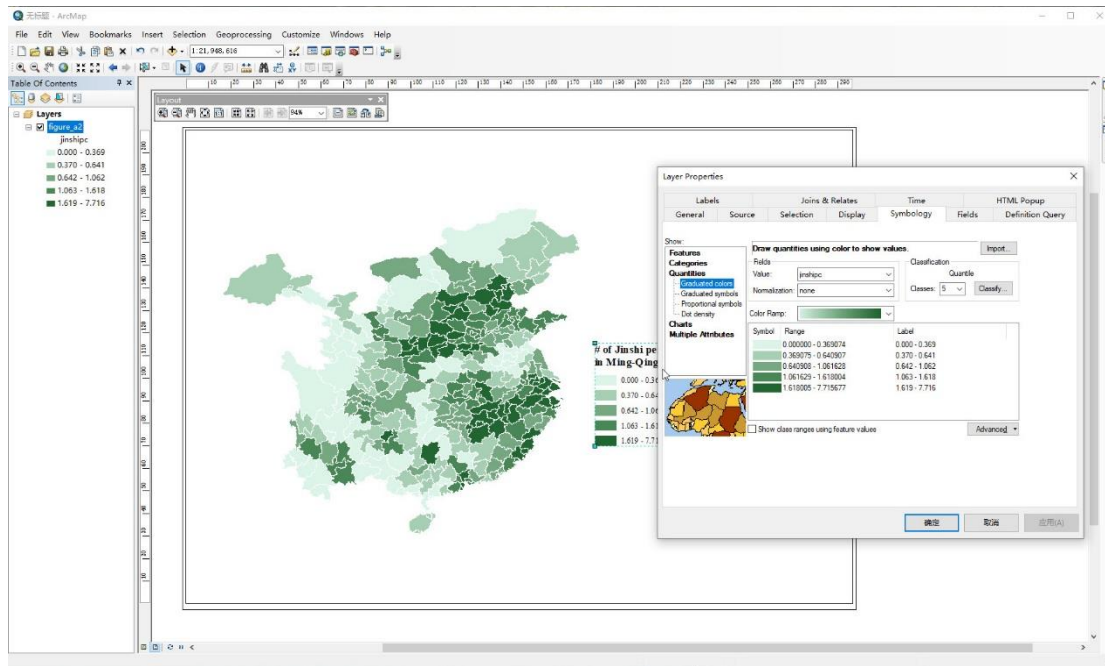


- Please refer to the following screen shot of ArcGIS for Figure 3C:
 Step 1) Add the qing_china_proper.shp, printing_centers.shp, pine.shp, and bamboo.shp and MainRiver.shp to the ArcMap;
 Step 2) Adjust the colors;
 Step 3) Finish by adding the legend and export map accordingly.

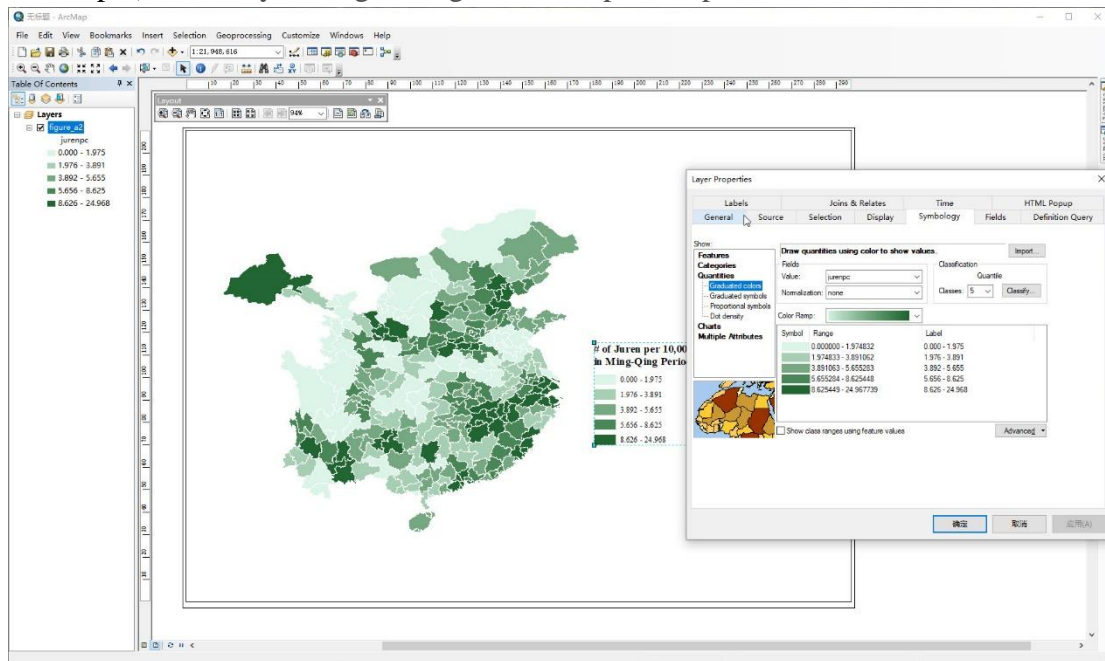


For Figure A2: The zipped file “figure_a2” contains the Figure_a2 shapefile, which is used for reproducing Figure A2 in the Appendix.

- Please refer to the following screen shot and the following procedure for reproducing the upper left panel of jinshi density in Figure A2:
 Step 1) Select the value “jinshipc” in the “Graduate Colors” type of Quantities map in the Symbology tab of the Layer Properties dialog box;
 Step 2) Use gradient green for Color Ramp;
 Step 3) Select “Quantile” method and Classes: 5 for Classification;
 Step 4) Format the Label selecting 3 for “Number of decimal places”
 Step 5) Finish by adding the legend and export map.

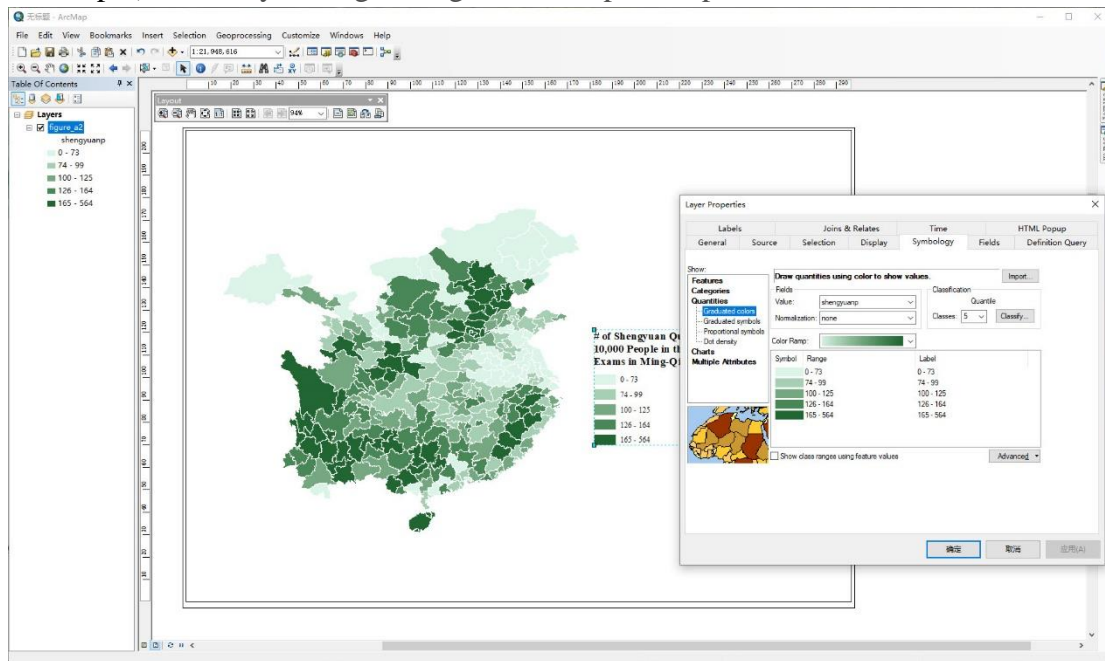


- Please refer to the following screen shot and the following procedure for reproducing the upper right panel of juren density in Figure A2:
 Step 1) Select the value “jurenpc” in the “Graduate Colors” type of Quantities map in the Symbology tab of the Layer Properties dialog box;
 Step 2) Use gradient green for Color Ramp;
 Step 3) Select “Quantile” method and Classes: 5 for Classification;
 Step 4) Format the Label selecting 3 for “Number of decimal places”
 Step 5) Finish by adding the legend and export map.



- Please refer to the following screen shot and the following procedure for reproducing the lower left panel for shengyuan quota density in Figure A2:

- Step 1) Select the value “shengyuanp” in the “Graduate Colors” type of Quantities map in the Symbology tab of the Layer Properties dialog box;
- Step 2) Use gradient green for Color Ramp;
- Step 3) Select “Quantile” method and Classes: 5 for Classification;
- Step 4) Format the Label selecting 3 for “Number of decimal places”
- Step 5) Finish by adding the legend and export map.

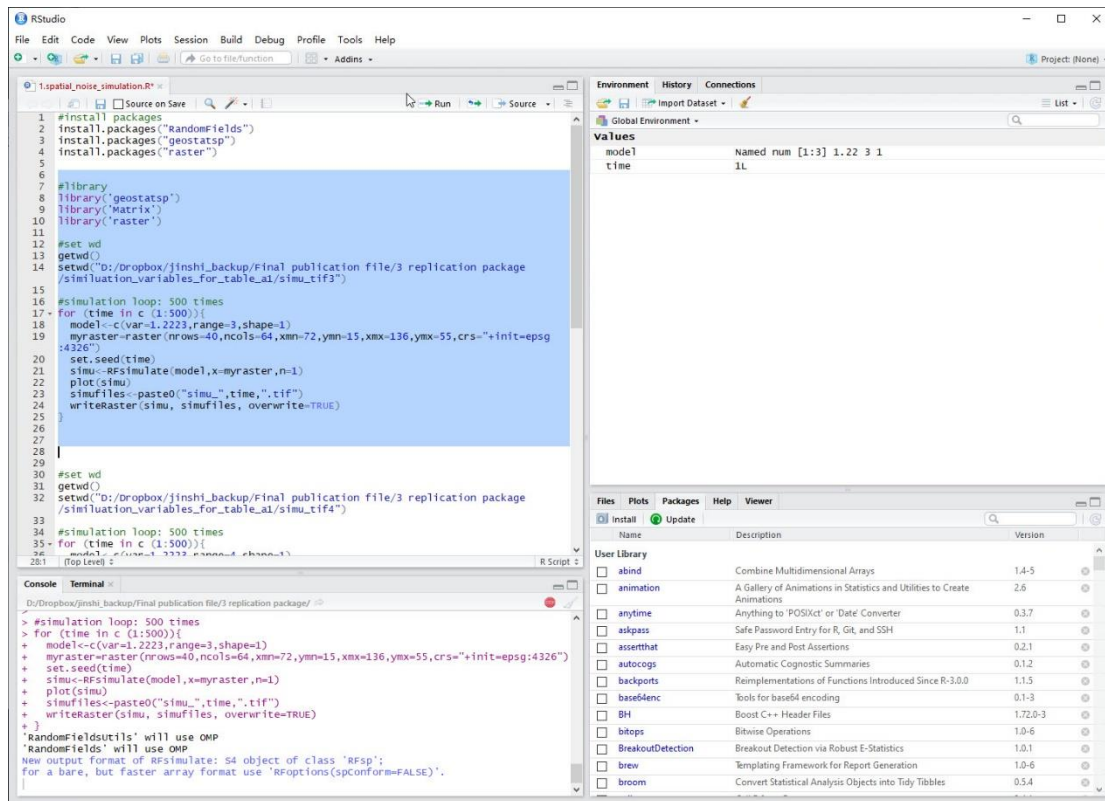


The R-scripts, Python .py and Stata .do files are used for generating the spatial noise variables and merge them to census10_final.dta used in Table A1:

The zipped file “simulation_variables_for_figure_a1” contains the following files. The R and Python scripts were run on R version i386 3.6.1 and ArcGIS Desktop version 10.6.1, respectively. The steps employed to generate and merge the spatial noise variables are as follows:

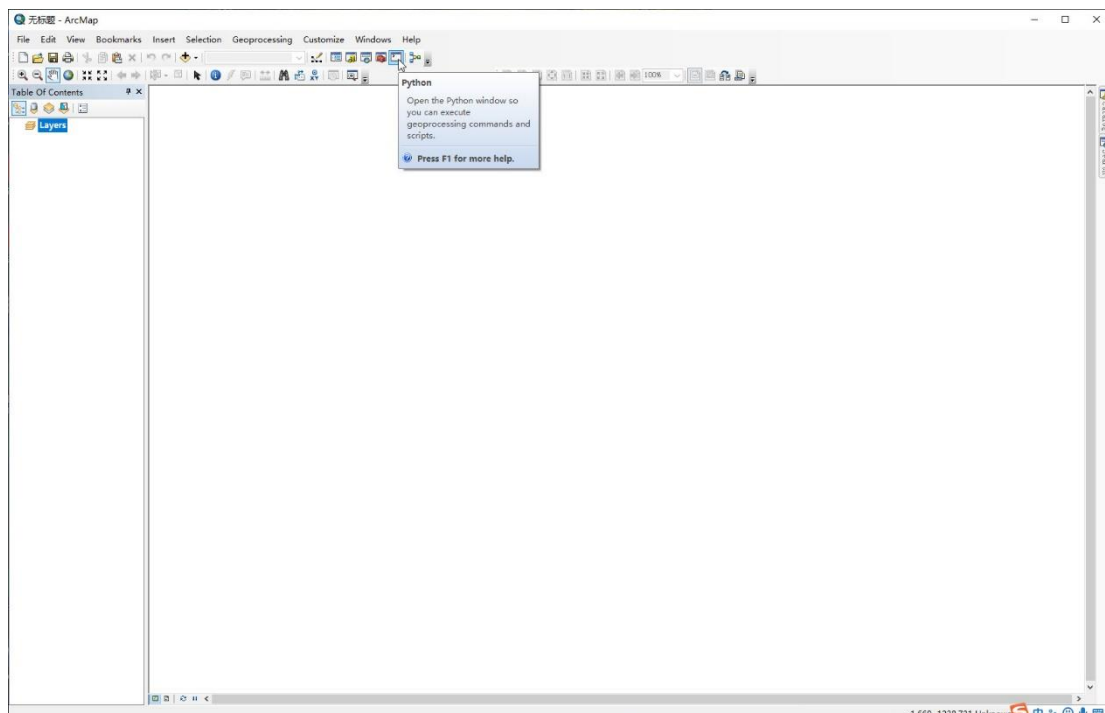
Step 1) Use the R-script “1.spatial_noise_simulation.R” to simulate the spatial noises on a 1*1 degree raster in R. For each ϕ ranges between 3 and 5, we simulated the noises 500 times. The output was thus 500*3 raster files. It may take up to 15 to 20 minutes to finish the simulation.

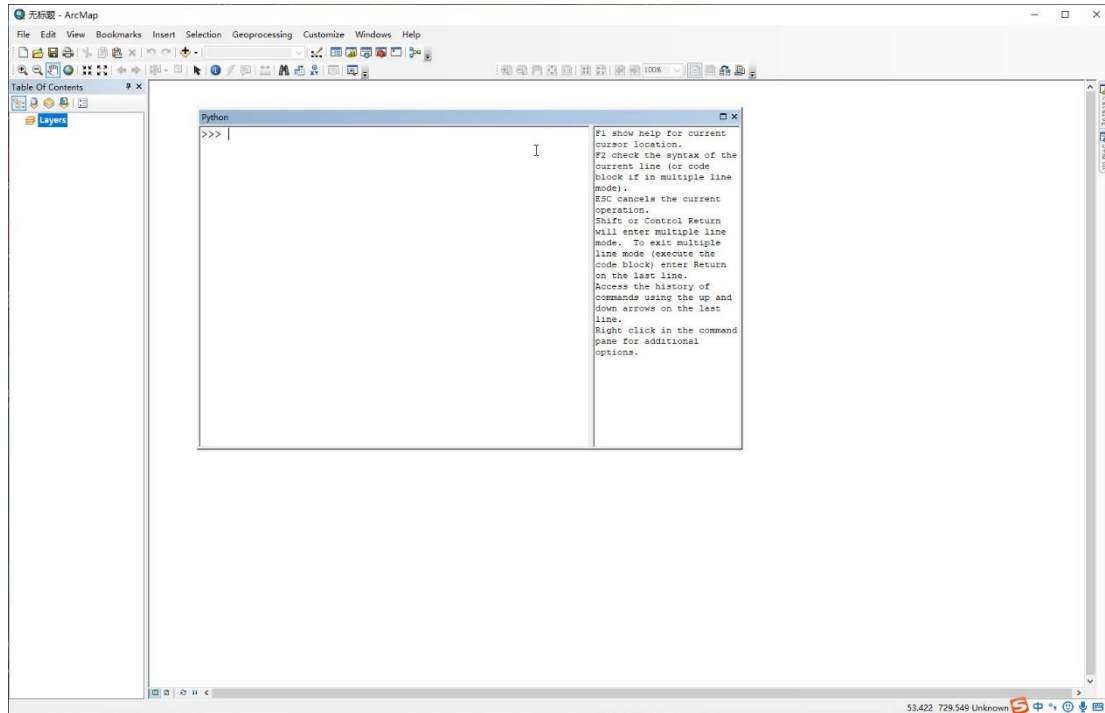
If the following is shown in your RStudio Console, it means the code has run smoothly:



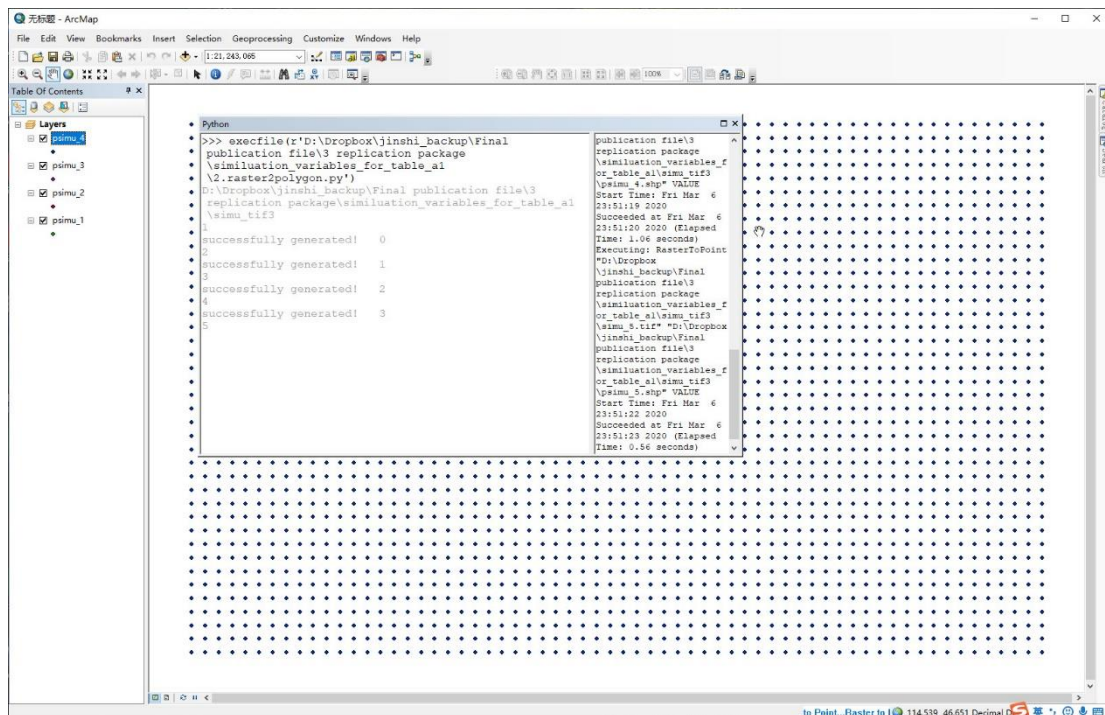
Step 2) Use the Python-script "2.raster2polygon.py" for matching the simulated raster files to the prefectural centroid points in ArcGIS, followed by exporting the attribute table of each shapefiles to the excel files. The procedure is as follows:

2.1) Click open the python window button  in the ArcGIS Desktop application;





2.2) Type the following code in the python console and run:
`exec(r'path to working directory\2.raster2polygon.py').read()`



If the above is shown in ArcGIS, it implies that the code has run smoothly. It will likely take 9-10 hours to finish.

Note: If the error message “Runtime error...IOError: [Errno 9] Bad file descriptor” appears, restart by typing the same code from step 2.2 in the python console. The procedure will resume where it was interrupted.

Step 3) After merging the tif file to the prefecture boundary and output it to excel, use

the STATA do-file “3.merge_to_prefectures.do” to import the excel files into STATA and merge them to the census10_final.dta using the variable “prefid”. This step may take approximately 15-20 minutes to run.