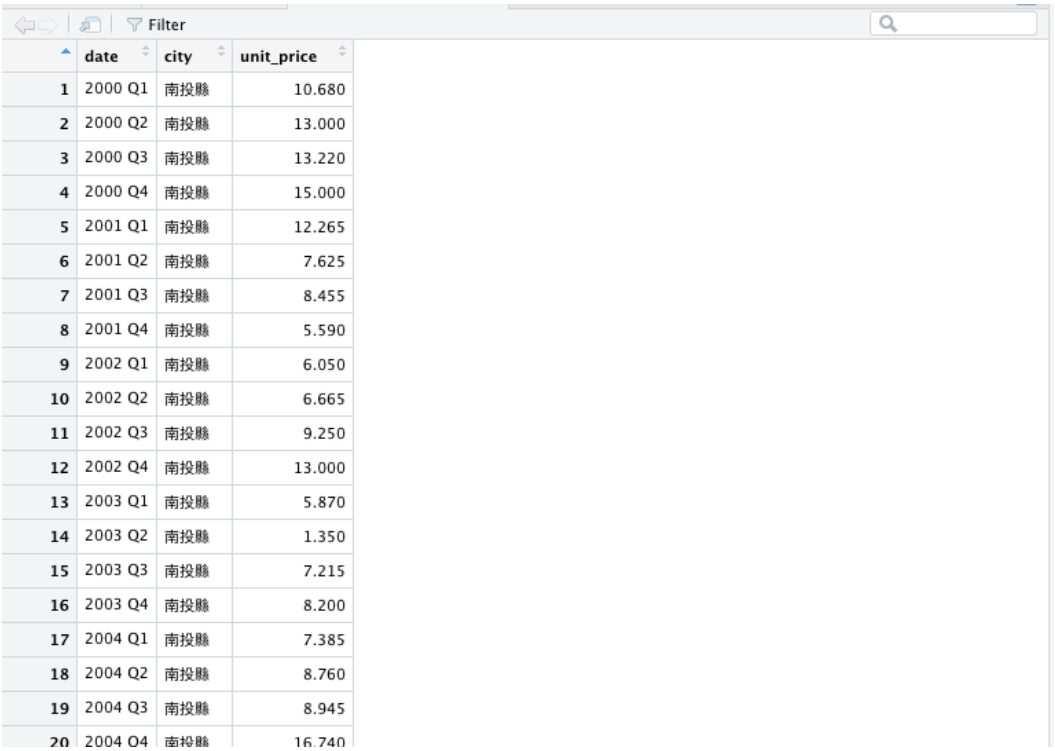
Before running my script of R, please change the first line of my R script "setwd("~/Desktop/Assignment") to any working directory you used to save the two raw files. Meanwhile, my R script will create a file called "results" in the working directory, storing the results (graphs, data, regression summary, etc.) of the seven questions. I have attached the "results" file with the email too, so you can see the results without executing my R script.

1. **Using local land raw.dta, compute the median unit price for each quarter-year and city combination.**

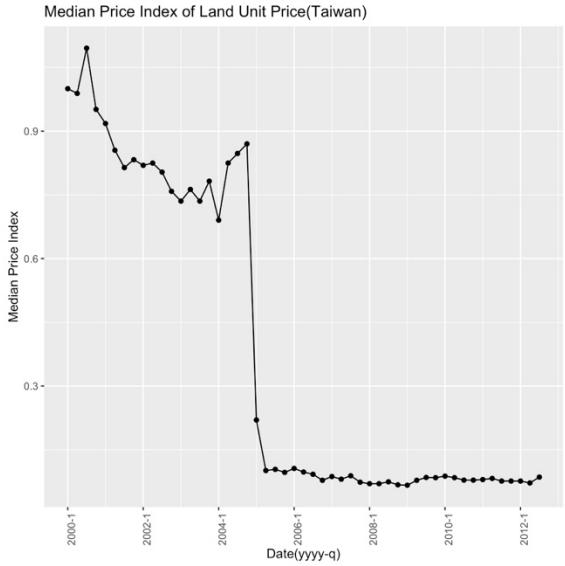
I have saved the data result into the "results" file. The following screenshot is a glance of the data. Please visit the "results" file for full data.

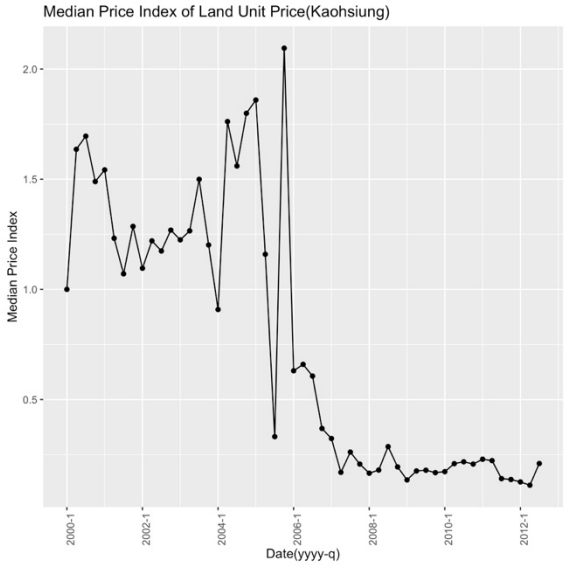

Filter

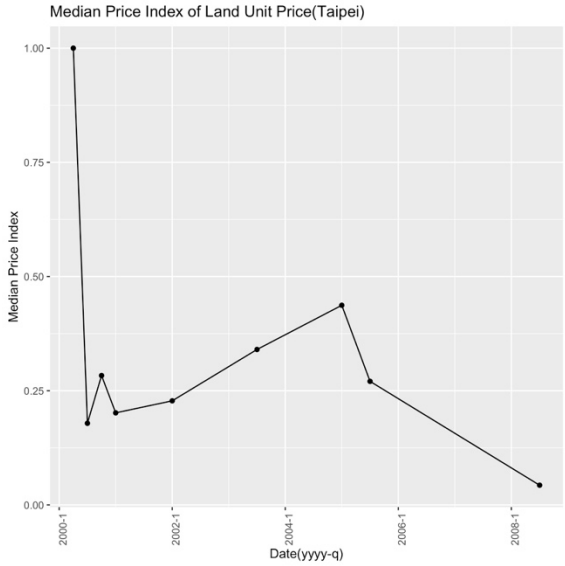
	date	city	unit_price
1	2000 Q1	南投縣	10.680
2	2000 Q2	南投縣	13.000
3	2000 Q3	南投縣	13.220
4	2000 Q4	南投縣	15.000
5	2001 Q1	南投縣	12.265
6	2001 Q2	南投縣	7.625
7	2001 Q3	南投縣	8.455
8	2001 Q4	南投縣	5.590
9	2002 Q1	南投縣	6.050
10	2002 Q2	南投縣	6.665
11	2002 Q3	南投縣	9.250
12	2002 Q4	南投縣	13.000
13	2003 Q1	南投縣	5.870
14	2003 Q2	南投縣	1.350
15	2003 Q3	南投縣	7.215
16	2003 Q4	南投縣	8.200
17	2004 Q1	南投縣	7.385
18	2004 Q2	南投縣	8.760
19	2004 Q3	南投縣	8.945
20	2004 Q4	南投縣	16.740

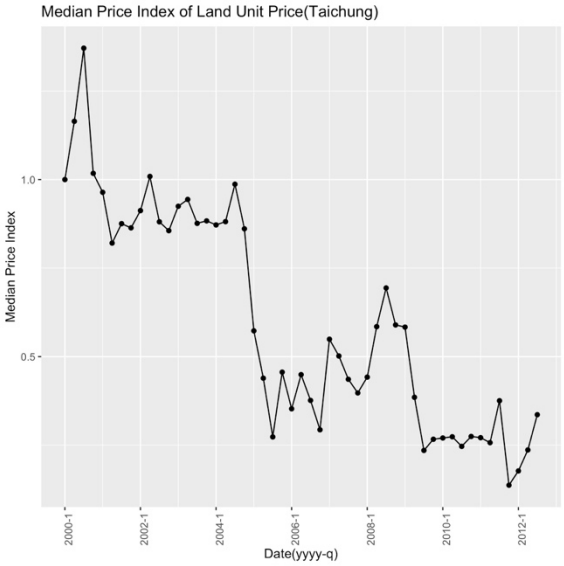
2. **Create a time series graph which plots the median unit price over time for all of Taiwan (including all 22 regions) alongside separate series for each of Taiwan's six major cities (Taipei, New Taipei, Taoyuan, Taichung, Tainan, and Kaohsiung). Do this for the full sample time period of 2000Q1 to 2012Q3. A common procedure when making this type of plot is to "normalize" each series by dividing prices by the value in a base period. Use 2000Q1 as the base period. The resulting series are sometimes called median price indices. Save the above time series graph in .gph and .eps formats.**

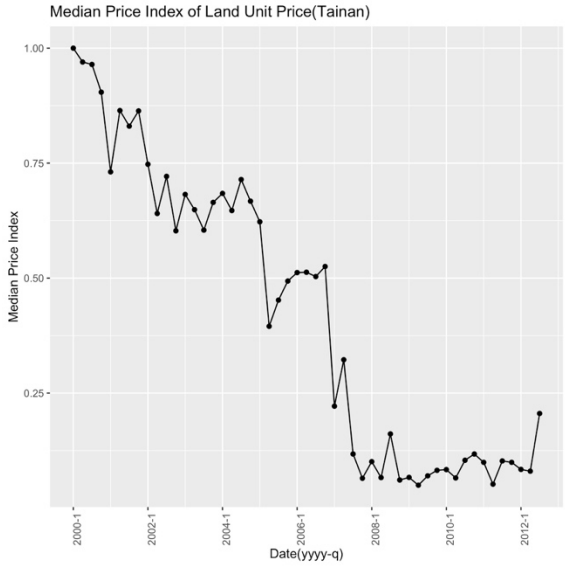
I have created the seven charts, and saved them into the "results" file. It is worth noting that since Taipei did not have land trading data in 2000 Q1, I used 2000 Q2 as the base period for Taipei while calculating the median price index. The followings are the charts, you can see bigger versions in the "results" file.


Median Price Index of Land Unit Price(Taiwan)
Median Price Index
0.9
0.6
0.3
2000-1
2002-1
2004-1
2006-1
2008-1
2010-1
2012-1
Date(yyyy-q)

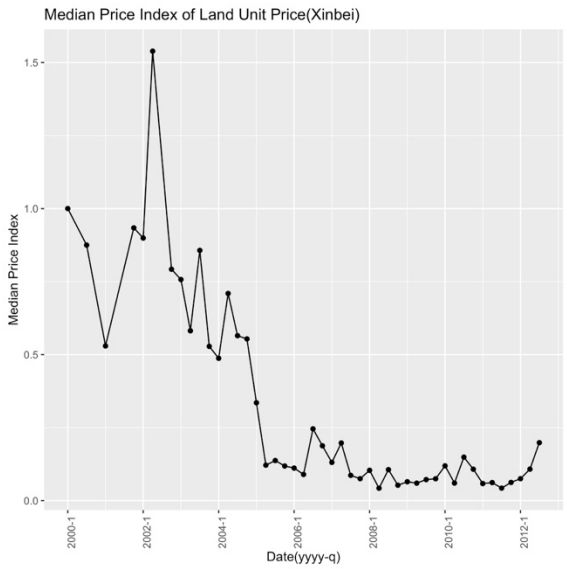

Median Price Index of Land Unit Price(Kaohsiung)
Median Price Index
2.0
1.5
1.0
0.5
2000-1
2002-1
2004-1
2006-1
2008-1
2010-1
2012-1
Date(yyyy-q)


Median Price Index of Land Unit Price(Taipei)
Median Price Index
1.00
0.75
0.50
0.25
0.00
2000-1
2002-1
2004-1
2006-1
2008-1
Date(yyyy-q)


Median Price Index of Land Unit Price(Taichung)
Median Price Index
1.0
0.5
2000-1
2002-1
2004-1
2006-1
2008-1
2010-1
2012-1
Date(yyyy-q)

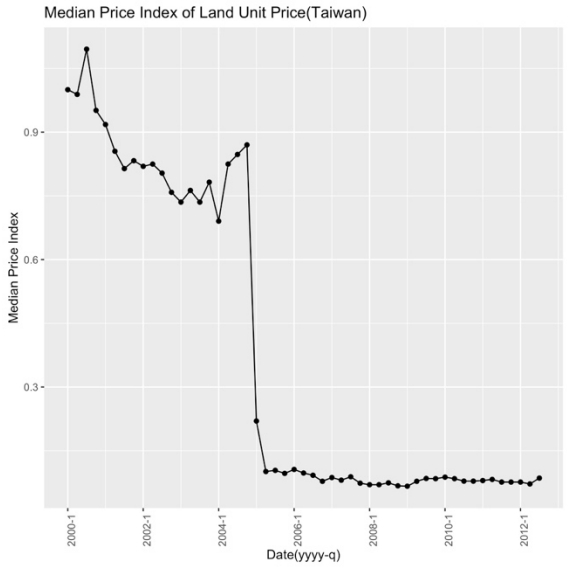

Median Price Index of Land Unit Price(Tainan)
Median Price Index
1.00
0.75
0.50
0.25
2000-1
2002-1
2004-1
2006-1
2008-1
2010-1
2012-1
Date(yyyy-q)

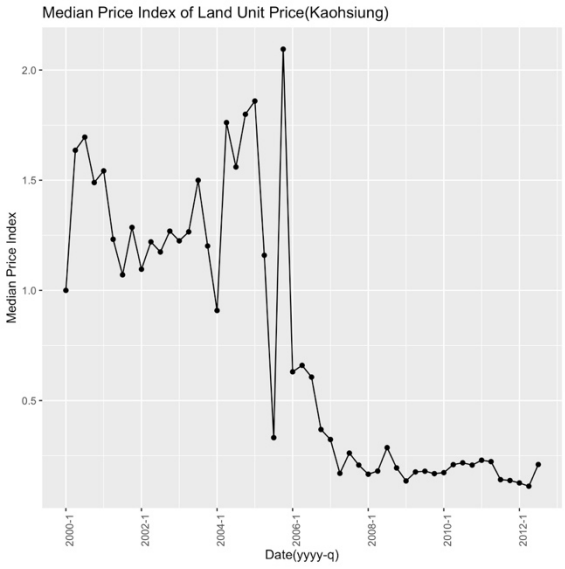

Median Price Index of Land Unit Price(Taoyuan)
Median Price Index
1.5
1.0
0.5
2000-1
2002-1
2004-1
2006-1
2008-1
2010-1
2012-1
Date(yyyy-q)

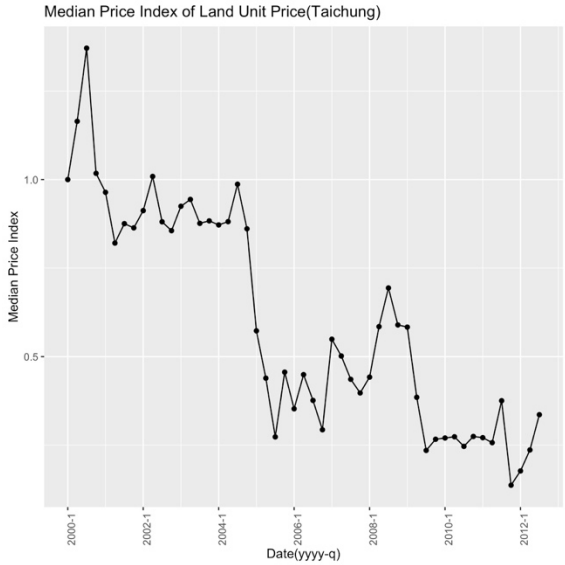


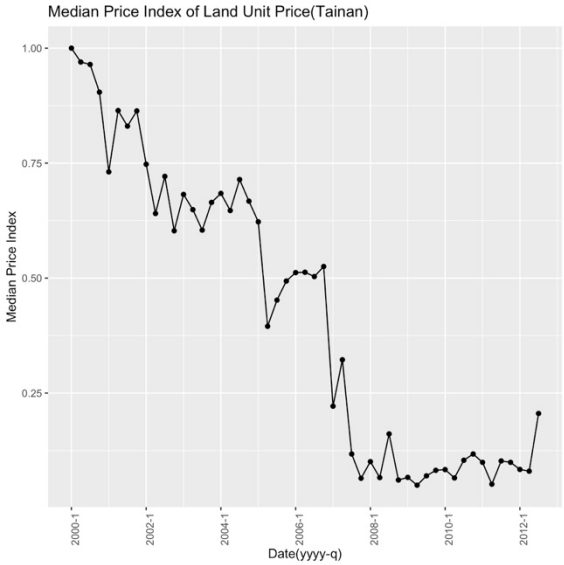
3. Redo the first two steps, restricting to transactions of land parcels within the top 20% by transaction value.

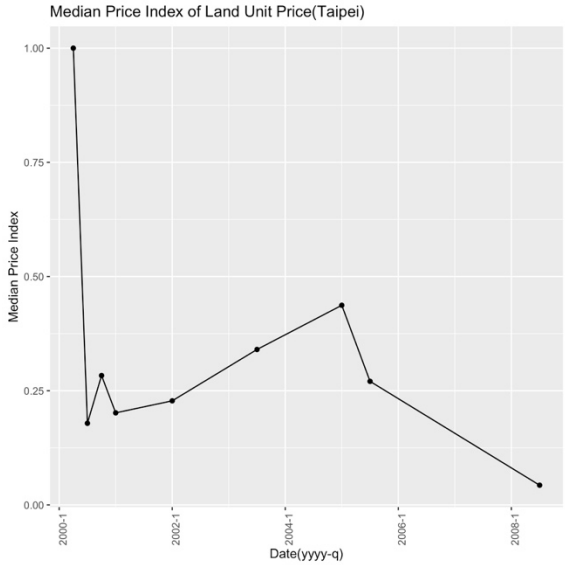
I have limited the data base to data points within top 20% of transaction value. The results, including the data and the seven charts, were also saved into the "results" files. Here are some quick views of the charts. Similar to question 2, Taipei did not have land trading data in 2000 Q1, thus I used 2000 Q2 as the base period for Taipei while calculating the median price index.

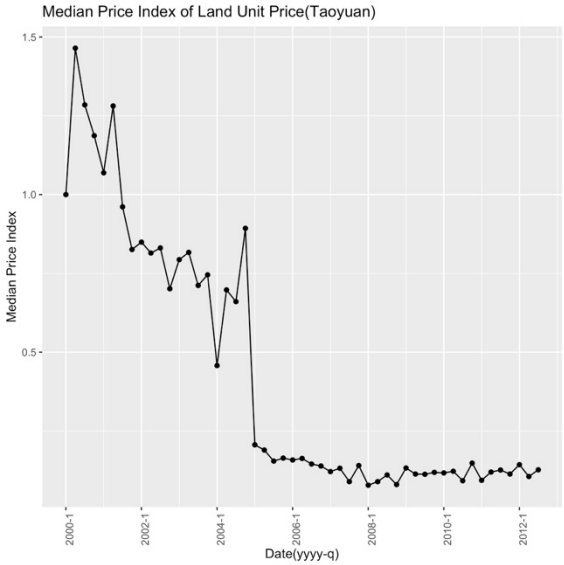


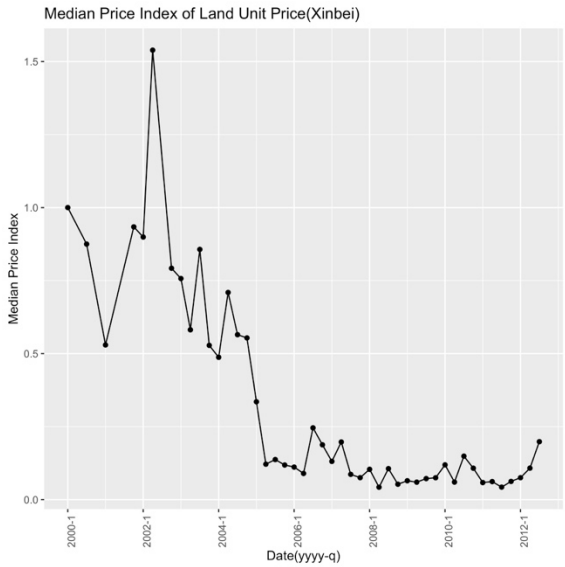












4. Now combine the two raw data files into one file. Create a variable(s) that allows you to keep track of which transactions are (i) land-only, (ii) building-only, and (iii) a building-land bundle. You may need to combine several variables to identify these transaction types.

I merged the two data sets with the following keys: city, township, address, date, roadwidth, street_section. I leave out several other common variables, including transact_value, unit_price, land_area and land_use. I tried to use land_area and land_use as keys. However, this resulted in almost no data matched between the two raw files. Thus, I decided to take them out.

Meanwhile, I used two columns to decide whether the transaction is (i) land-only, (ii) building-only, or (iii) a building-land. I discovered that the "shape" variable is exclusive in land-trade, while the "build_mat" variable is only seen in building-trade. Therefore, I used these two to decide the transaction type. The merged data is also in the "results" file.

5. **In addition, create the following variables which can be used to help determine why one property may be more/less valuable:**
    - **The age of the building being sold in years (if applicable).**
    - **Convert num floor and property floor to numeric variables.**
    - **Sort the properties into five categories based on use: single-family housing, apartments, commercial (offices/stores), industrial, and agricultural. Call this variable "broad use code." You may find it helpful to combine the existing variables "land use" and "main use."**
    - **A single numeric variable which identifies distinct neighborhoods within the same region.**
      Save the resulting file as local alltrans clean.dta. All numerical variables should be saved as numeric type, not as strings.

I trim the data and get the year of construction of the buildings, and then calculated the age of the building at the time of transaction.

I set property_floor equal to num_floor when there is "全" in order to transfer property_floor to numeric data.

I used main_use to determine the "broad_use_code" of the buildings. I tried, but failed, to take land_use into account, as the land_use variable is too diverse.

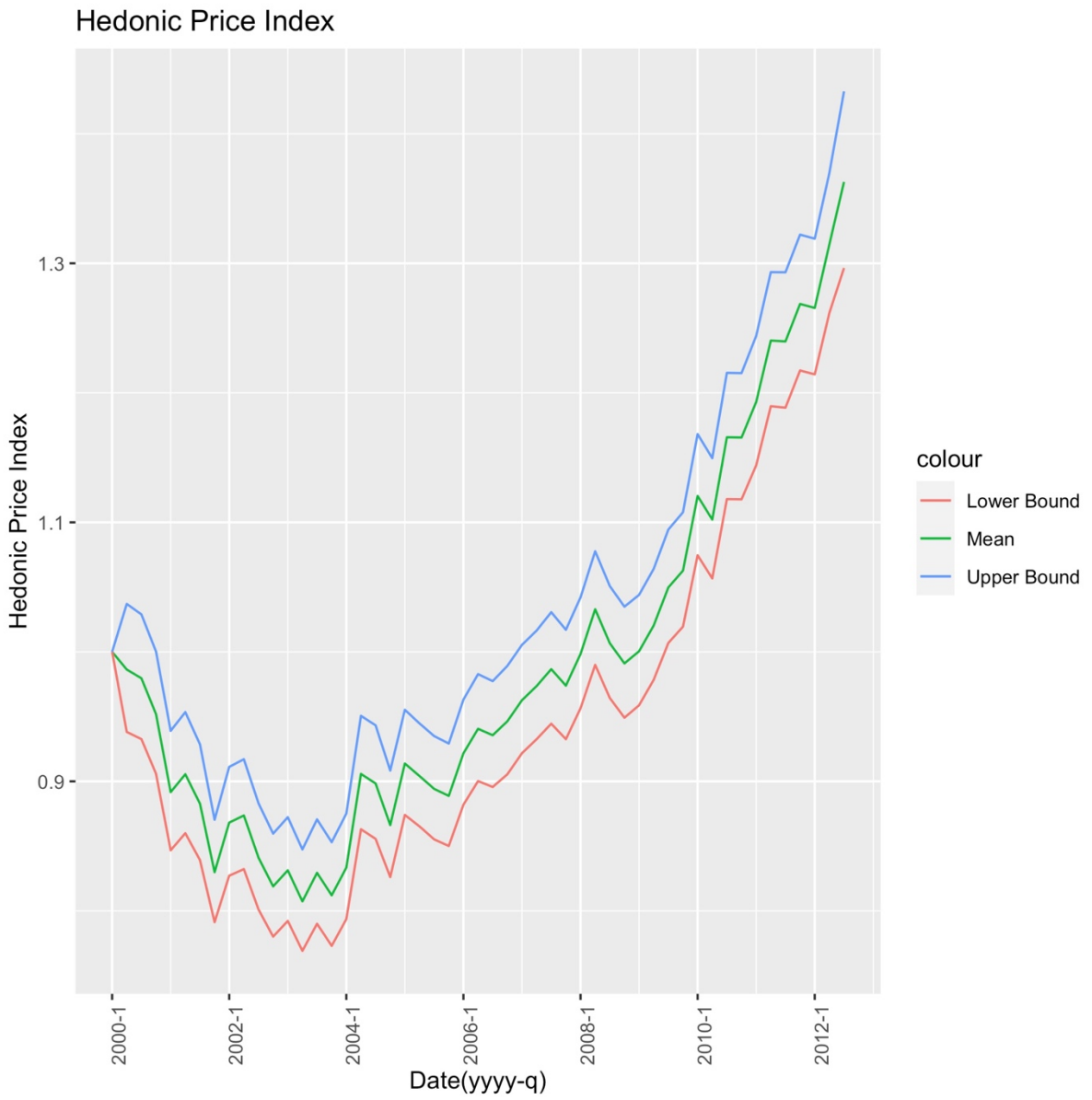
I could not understand "A single numeric variable which identifies distinct neighborhoods within the same region". Hence, I simply combine city and township to create a new variable, noting that many cities have townships with the same name. Creating a variable of city-township could then separate the townships with the same name.

The clean data is saved in the "results" file.

6. **Pool all of the transactions together to estimate the following regression model:**

   where i refers to individual transactions, and t refers to a quarter-year. Use the log transaction value as the dependent variable. The vector $X_{i,t}$ in the above model should include: road width, dummies for "broad use code," dummies for the number of floors in the unit, dummies for neighborhood, a quadratic in age, a quadratic in floor space, and a quadratic in land area (see step 5 for details).
   Transform the coefficient estimates exp(β_t) and plot them as a time series with the 95% confidence intervals. As in step 2, normalize the coefficients so that β_t is equal to 1 in 2000Q1. This step creates a hedonic price index.[1]

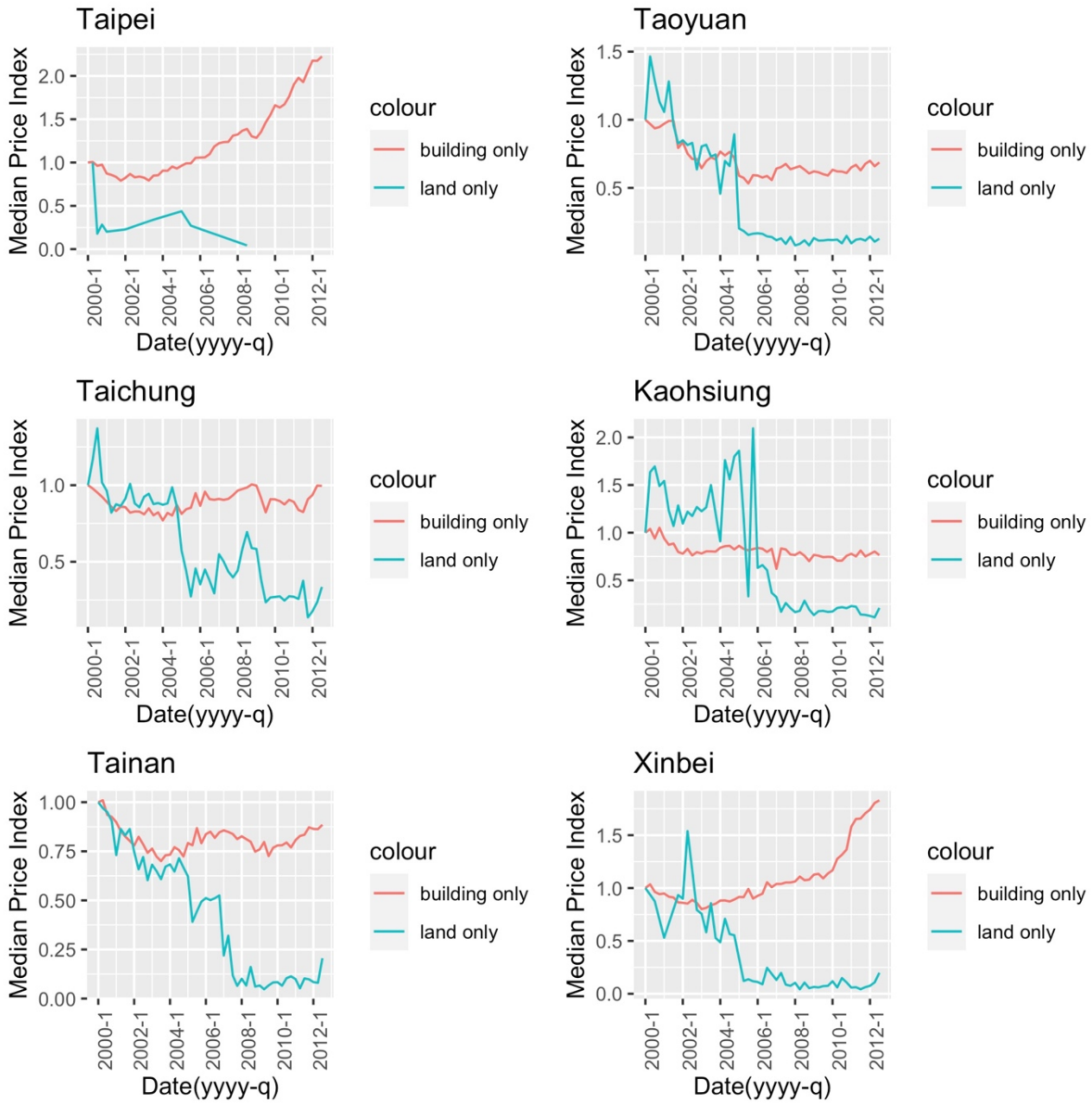
I used the variables given to run a regression model, and I saved the results in the "results" file. Meanwhile, I get the coefficients of the date variable. Then I use the standard errors of each coefficients to calculate the 95% confidence intervals of each coefficient. Afterwards, I use exp() to transfer the coefficients, as well as their upper and lower bounds within the 95% confidence intervals. It is worth noting that I manually add the coefficient of 2000 Q1 as 0, since it is used as the reference category of the dummy variable, and thus is not shown in the regression report.



7. **Finally, use the variable you created in step 4 to do the following:**
   - **(i) Create six time series graphs, one for each of Taiwan's six major cities. Each graph should plot the median unit price index over time for the three transaction types.**
   - **(ii) Combine the six time series graphs into a single figure displayed as two rows of three graphs (a 2 × 3 grid).**
   - **(iii) Based on the series in these graphs, generate a simple table which computes growth rates for each region-transaction type combination between three periods: 2000Q1-2007Q4, 2008Q1-2010Q4, and 2011Q1-2012Q3.**

I used the merged data in step 4 to plot the median unit price index over time for the three transaction types. Nonetheless, the six cities all lack the “land and building” transaction type, which might result from my mistakes of choosing the merging keys in step 4 (but I couldn’t think of other more reasonable keys to merge the two raw files). Hence the charts would only show two of the transaction types, “land only” and “building only”.

Meanwhile, for the “land only” type of Taipei, there isn’t data for 2000 Q1, and thus I have to use 2000 Q2 as the base period. Furthermore, there isn’t data after 2008 Q3 for Taipei “land only” either.



Here is the calculation of the growth rate of the three periods given. Since all cities lack “land and building” type, the growth rates of “land and building” couldn’t be calculated. Meanwhile, Taipei lacks data for 2000 Q1 and 2007Q4, thus I used 2000 Q2 and 2008Q3 to calculate the growth rate between 2000-2007 for Taipei. The following is the summary of the growth rate. You can also retrieve the same data in the “results” file.

	city	transaction_type	2000-2007(%)	2008-2010(%)	2011-2012(%)
4	Kaohsiung City	building only	-0.22898230	0.01251204	0.01893938
3	Taichung City	building only	-0.06531091	0.06204379	-0.11813784
5	Tainan	building only	-0.18771925	0.06847134	-0.09733556
2	Taipei City	building only	0.31070502	-0.33146997	-0.17046360
6	Taoyuan	building only	-0.36290858	0.06069961	-0.05586884
1	Xinbei City	building only	0.05275233	-0.28417268	-0.15820031
16	Kaohsiung	land and building	NA	NA	NA
17	Taichung	land and building	NA	NA	NA
18	Tainan	land and building	NA	NA	NA
14	Taipei	land and building	NA	NA	NA
13	Taoyuan	land and building	NA	NA	NA
15	Xinbei	land and building	NA	NA	NA
10	Kaohsiung	land only	-0.79261179	-0.25000005	0.08474575
9	Taichung	land only	-0.60263155	0.37896823	-0.23948221
11	Tainan	land only	-0.93519216	-0.11940296	-1.06818184
8	Taipei	land only	-0.95688084	NA	NA
12	Taoyuan	land only	-0.85977337	-0.89999988	-0.35606052
7	Xinbei City	land only	-0.92486532	-0.03551912	-2.39805835