Data Interpretation and Writing Exercise (PHL)

March 1, 2024

Pedro Huet López

Contents:

This is a short excercise, based on an unorganized xlsx file containing a relatively small proportion of information regarding a variety of variables related to housing information on four Indianapolis Neighborhoods. The first part of the code is employed to standardize the data structure and making it a workable dataframe, while the second part is a series of visualizations to explore the actual tendencies we can see among these neighborhoods. The third part is an interpretation based on the visualizations.

Stage 1) Pre-Processing: building working and replicable data for academic research

[1]:
```
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

file_path = r"C:\Users\pedro\Documents\Data Analysis Work\Job␣
↪Excercise\Interpretation and writing exercise for candidates (1).xlsx"
Datex = pd.read_excel(file_path, header=1)

def remove_total(cell):
    if isinstance(cell, str):
        return cell.replace(' Total', '')
    return cell

Datex = Datex.applymap(remove_total)
Datex = Datex[~Datex.apply(lambda row:
                            row.astype(str).str.contains('Source:').any() or
                            row.dropna().empty or
                            any(any(word in str(cell) for cell in row) for word␣
↪in ["Housing Cost", "Housing Market"]),
                            axis=1)]

Datex = Datex.reset_index(drop=True)
```

[2]:
```python
def transform_df(df, start_row, end_row):
    new_df = df.iloc[start_row:end_row]
    columns_to_drop = new_df.columns[12:21]
    new_df = new_df.drop(columns=columns_to_drop)

    vertical_df = new_df.transpose()
    vertical_df = vertical_df.fillna("Year")
    vertical_df.columns = vertical_df.iloc[0]
    vertical_df = vertical_df[1:]
    vertical_df.reset_index(drop=True, inplace=True)
    vertical_df = vertical_df.rename_axis(columns=None)

    # Correcting some of the headers
    vertical_df = vertical_df.rename(columns={'Mapleton-Fall Creak': 'Mapleton␣
↪Fall Creek','St. Clair': 'St. Clair Place'})
    return vertical_df

med_sales_pri_df = transform_df(Datex, 53, 59)
num_home_s_df = transform_df(Datex, 60, 66)
```

[3]:
```python
def nuanced_df(df, start_row, end_row):
    new_df = df.iloc[start_row:end_row]
    columns_to_drop = new_df.columns[6:21]
    polished_df = new_df.drop(columns=columns_to_drop)
    polished_df = polished_df.fillna("Year")
    polished_df.columns =  polished_df.iloc[0]
    polished_df = polished_df[1:]
    polished_df.reset_index(drop=True, inplace=True)
    polished_df = polished_df.rename_axis(columns=None)
    return polished_df

mortgage_loan_app_df = nuanced_df(Datex, 29, 36)
res_buil_per_df = nuanced_df(Datex, 37, 44)
med_a_val_df = nuanced_df(Datex, 45, 52)
```

[4]:
```python
resDatex = pd.concat([Datex.iloc[0:28], Datex.iloc[66:]], ignore_index=True)
selected_data = resDatex.iloc[20:29, 0:11]
header_texts = selected_data.iloc[0, [1, 3, 5, 7, 9]].tolist()
new_column_labels = ["Year"] + header_texts

dataframes = []
for i in range(2):
    cols = [0] + [j for j in range(1 + i, 11, 2)]
    df = selected_data.iloc[:, cols]
    df = df.iloc[2:]
    df.columns = new_column_labels
    df.reset_index(drop=True, inplace=True)
```

```
    dataframes.append(df)

med_mon_mort_df, med_mon_rent_df = dataframes
```

[5]:
```python
def process_dataframe(data):
    data.columns = data.iloc[0]
    data = data[1:]
    data = data.drop(data.index[1])

    header_texts = data.iloc[0, [1, 4, 7, 10, 13]].tolist()
    new_column_labels = ["Year"] + header_texts

    df = {}
    for i in range(1, 4):
        df[i] = data.iloc[:, [0, i + 0, i + 3, i + 6, i + 9, i + 12]]
        df[i] = df[i].drop(df[i].index[0])
        df[i].reset_index(drop=True, inplace=True)
        df[i].columns = new_column_labels
    return df[1], df[2], df[3]

selected_data_v2 = resDatex.iloc[15:19, 0:16]

cb_h_nm_df, cb_h_wm_df, cb_rent_df = process_dataframe(selected_data_v2)
```

[6]:
```python
new_data = resDatex.iloc[7:15]
header_texts = new_data.iloc[0, [1, 5, 9, 13, 17]].tolist()
new_column_labels = ["Year"] + header_texts

dataframes = []
for i in range(4):
    cols = [0] + [j+1 for j in range(i, 20, 4)]
    df = new_data.iloc[:, cols]

    df = df.replace('**', np.nan)

    df[df.columns[1]] = pd.to_numeric(df[df.columns[1]], errors='coerce') / 100

    df = df.iloc[2:]
    df.columns = new_column_labels
    df.reset_index(drop=True, inplace=True)
    dataframes.append(df)

asian_ho_df, afr_ho_df, lat_ho_df, white_ho_df = dataframes
```

[7]:
```python
last_data = resDatex.iloc[1:6, 1:21].drop(resDatex.index[2])
last_data = last_data.reset_index(drop=True)
last_data = last_data.drop(last_data.index[1])
```

```
def process_data(last_data):
    header_texts = last_data.iloc[0, [0, 4, 8, 12, 16]].tolist()
    new_column_labels = header_texts

    df = {}
    for i in range(4):
        columns = [i + j for j in range(0, 19, 4) if i + j < 21]
        df[i] = last_data.iloc[:, columns]
        df[i] = df[i].drop(df[i].index[0])
        df[i].reset_index(drop=True, inplace=True)
        df[i].columns = new_column_labels

    return df[0], df[1], df[2], df[3]

hholds_ten_own, hholds_ten_rent, perc_hholds_ten_own, perc_hholds_ten_rent = 
 ↪process_data(last_data)

hholds_ten_own.iloc[1], hholds_ten_rent.iloc[0] = hholds_ten_rent.iloc[0].
 ↪copy(), hholds_ten_own.iloc[1].copy()
perc_hholds_ten_own.iloc[1], perc_hholds_ten_rent.iloc[0] =
 ↪perc_hholds_ten_rent.iloc[0].copy(), perc_hholds_ten_own.iloc[1].copy()

def modify_dataframe(df):
    year_column = pd.DataFrame({'Year': [2012, 2017]})
    df_with_year = pd.concat([year_column, df], axis=1)
    return df_with_year

hholds_ten_own = modify_dataframe(hholds_ten_own)
hholds_ten_rent = modify_dataframe(hholds_ten_rent)
perc_hholds_ten_own = modify_dataframe(perc_hholds_ten_own)
perc_hholds_ten_rent = modify_dataframe(perc_hholds_ten_rent)
```

All dataframes in long form and ready for time series analysis

[8]:
```
dataframes = [hholds_ten_own, hholds_ten_rent, perc_hholds_ten_own,
 ↪perc_hholds_ten_rent, asian_ho_df, afr_ho_df, lat_ho_df, white_ho_df,
 ↪cb_h_nm_df, cb_h_wm_df, cb_rent_df, med_mon_mort_df, med_mon_rent_df,
 ↪mortgage_loan_app_df, res_buil_per_df, med_a_val_df, med_sales_pri_df,
 ↪num_home_s_df]

for df in dataframes:
    if 'Year' in df.columns:
        df['Year'] = pd.to_datetime(df['Year'], format='%Y')

print(hholds_ten_own)
print(hholds_ten_rent)
```

```python
print(perc_hholds_ten_own)
print(perc_hholds_ten_rent)
print(asian_ho_df)
print(afr_ho_df)
print(lat_ho_df)
print(white_ho_df)
print(cb_h_nm_df)
print(cb_h_wm_df)
print(cb_rent_df)
print(med_mon_mort_df)
print(med_mon_rent_df)
print(mortgage_loan_app_df)
print(res_buil_per_df)
print(med_a_val_df)
print(med_sales_pri_df)
print(num_home_s_df)
```

```
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01        204401        519        685                1359
1 2017-01-01        198434        555        976                1598

  St. Clair Place
0             480
1             751
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01        155037       1764       1365                3199
1 2017-01-01        168781       1762       1283                3018

  St. Clair Place
0            1027
1            1235
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01      0.568668   0.227332   0.334146            0.298157
1 2017-01-01      0.540376   0.239534    0.43205            0.346187

  St. Clair Place
0        0.318514
1        0.378147
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01      0.431332   0.772668   0.665854            0.701843
1 2017-01-01      0.459624   0.760466    0.56795            0.653813

  St. Clair Place
0        0.681486
1        0.621853
        Year  Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01       0.453566          0          0                   0
1 2013-01-01       0.434568          0   0.333333                   0
```

```
2 2014-01-01       0.457675          0          1                  0
3 2015-01-01       0.446326          0          1                  0
4 2016-01-01       0.392507          0        0.6                  0
5 2017-01-01       0.555920          0   0.611111                  0

  St. Clair Place
0             NaN
1             NaN
2             NaN
3               1
4        0.466667
5        0.416667
        Year  Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01       0.378957   0.217833   0.072327            0.300362
1 2013-01-01       0.366115   0.178912   0.069401            0.283486
2 2014-01-01       0.359220   0.190141   0.079903            0.282676
3 2015-01-01       0.352226   0.181598   0.071429             0.29519
4 2016-01-01       0.343028   0.205431   0.151667            0.293333
5 2017-01-01       0.339403   0.225066   0.141046            0.350245

  St. Clair Place
0        0.084233
1        0.120507
2        0.120968
3        0.040076
4        0.066451
5        0.090038
        Year  Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01       0.657522    0.25679   0.427372            0.315989
1 2013-01-01       0.651979   0.214286   0.421508            0.307985
2 2014-01-01       0.645810    0.26555   0.491228            0.324906
3 2015-01-01       0.635305   0.312044   0.534351            0.346364
4 2016-01-01       0.631702   0.302564   0.542799            0.352682
5 2017-01-01       0.637127   0.325153   0.564953            0.377656

  St. Clair Place
0        0.416366
1         0.42809
2        0.464522
3        0.464529
4        0.480806
5        0.502339
        Year  Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01       0.341612   0.674419   0.222222            0.366667
1 2013-01-01       0.326738     0.6875   0.144928            0.589041
2 2014-01-01       0.331888   0.693878   0.328244            0.597403
3 2015-01-01       0.328359          0   0.409091            0.149533
4 2016-01-01       0.326443          0    0.40678            0.070707
```

```
5 2017-01-01       0.352395          0   0.376147            0.119658

  St. Clair Place
0        0.320988
1        0.306604
2        0.291339
3        0.389105
4           0.375
5        0.424051
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2017-01-01          0.13   0.266393   0.264331           0.240642

  St. Clair Place
0            0.18
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2017-01-01          0.24   0.234892   0.257991            0.33313

  St. Clair Place
0            0.33
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2017-01-01           0.5   0.728682   0.619184           0.658448

  St. Clair Place
0             0.6
        Year Marion County   Crown Hill   Holy Cross Mapleton Fall Creek  \
0 2012-01-01          1172   937.664615         1343         1019.561026
1 2013-01-01          1148   972.482036         1126         1009.761555
2 2014-01-01          1129    830.92429         1142          904.612176
3 2015-01-01          1111   990.182724         1182            960.3878
4 2016-01-01          1111  1020.393939  1114.099275          997.458647
5 2017-01-01          1123      1011.41  1134.936589         1011.220829

  St. Clair Place
0             982
1             982
2             848
3             852
4             850
5             907
        Year Marion County  Crown Hill  Holy Cross Mapleton Fall Creek  \
0 2012-01-01           751  659.080288         623          714.082287
1 2013-01-01           768  681.540134         641          718.240488
2 2014-01-01           781  690.624771         692           689.11477
3 2015-01-01           788  630.867588         721          710.016728
4 2016-01-01           806  639.299131         749          700.061211
5 2017-01-01           836  661.199773  778.571317          658.203575

  St. Clair Place
```

```
0             671
1             714
2             711
3             730
4             751
5             791
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01     26.236624       32.4       69.2                38.3
1 2013-01-01      28.36917       27.9       62.5                37.9
2 2014-01-01     30.322532       36.8       58.7                39.8
3 2015-01-01     36.263376       58.8      111.5                64.3
4 2016-01-01     40.753079       76.5      118.3                76.5
5 2017-01-01     44.172724        100      126.9                91.8

  St. Clair Place
0       41.666667
1       31.944444
2       33.333333
3       73.611111
4       83.333333
5      123.611111
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01     11.097826       58.8       60.6                37.8
1 2013-01-01      9.744565       51.5       60.6                45.4
2 2014-01-01        11.875       58.8       92.3                40.8
3 2015-01-01     10.964674       64.7       85.6                37.8
4 2016-01-01     10.980978       33.8       81.7                33.2
5 2017-01-01     11.358696       54.4       94.2                55.6

  St. Clair Place
0            77.8
1            43.1
2            76.4
3           106.9
4            80.6
5           112.5
        Year Marion County Crown Hill Holy Cross Mapleton Fall Creek  \
0 2012-01-01         91500      51655      56400               63897
1 2013-01-01         88900      49501      54528               63687
2 2014-01-01         90400      51109      57768               60461
3 2015-01-01         93300      51776      60089               61676
4 2016-01-01         95700      52606      62843               62678
5 2017-01-01        100200      54871      70077               65832

  St. Clair Place
0           28013
1           31716
2           32267
```

```
3          35022
4          37467
5          44452
         Year Crown Hill Holy Cross Mapleton Fall Creek St. Clair Place  \
0  2008-01-01    13412.5      32500                19500            8750
1  2009-01-01      12500      19000                28750           10000
2  2010-01-01      15700      91500                29450           10000
3  2011-01-01      18900      75500                64500           11000
4  2012-01-01      24900     145000                77000           12600
5  2013-01-01      28500     122000                55275           19375
6  2014-01-01      20400     155000               145000           22405
7  2015-01-01      31375     170100               109000           30000
8  2016-01-01      53500     214000               105000           28040
9  2017-01-01      44000     165000               136750           56000
10 2018-01-01      84750     210000               142500           79000

   Marion County
0          89000
1          89999
2          93000
3          91000
4          97000
5         107000
6         115500
7         121000
8         126000
9         135000
10        148050
         Year Crown Hill Holy Cross Mapleton Fall Creek St. Clair Place  \
0  2008-01-01         50         36                  137             115
1  2009-01-01         37         42                  121              96
2  2010-01-01         27         45                   98              76
3  2011-01-01         29         23                   82              64
4  2012-01-01         25         37                  111              64
5  2013-01-01         32         30                   98              56
6  2014-01-01         23         37                   93              56
7  2015-01-01         44         44                  142              67
8  2016-01-01         40         50                  148              64
9  2017-01-01         45         69                  166             113
10 2018-01-01         66         56                  203             140

   Marion County
0          11654
1          10457
2           9172
3           9063
4          10527
5          12160
```

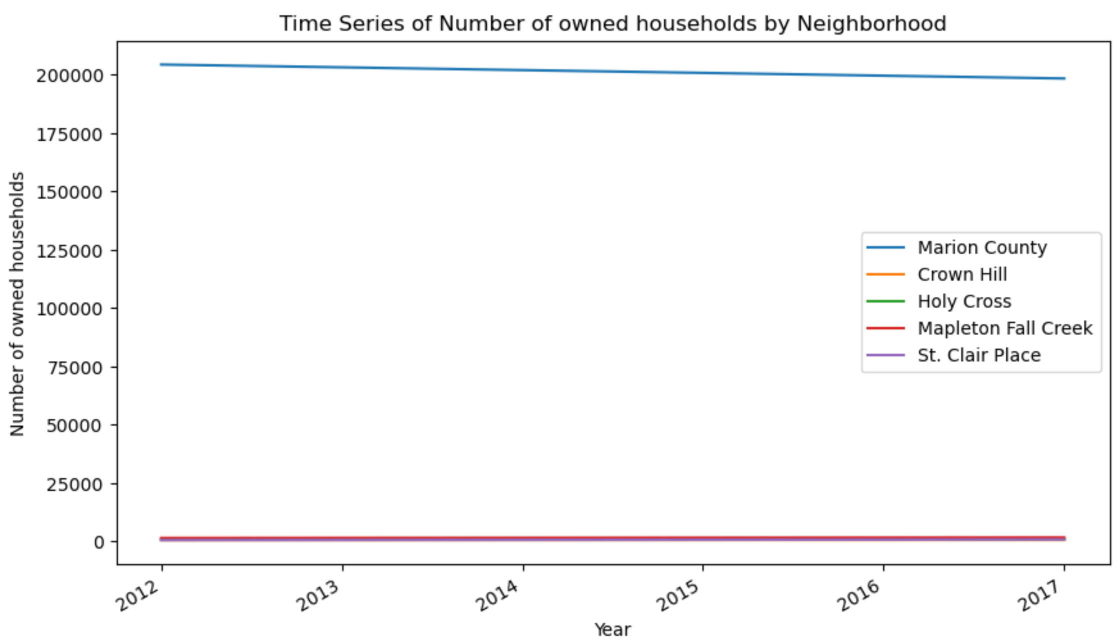
```
6          11796
7          12814
8          14083
9          14875
10         14844
```

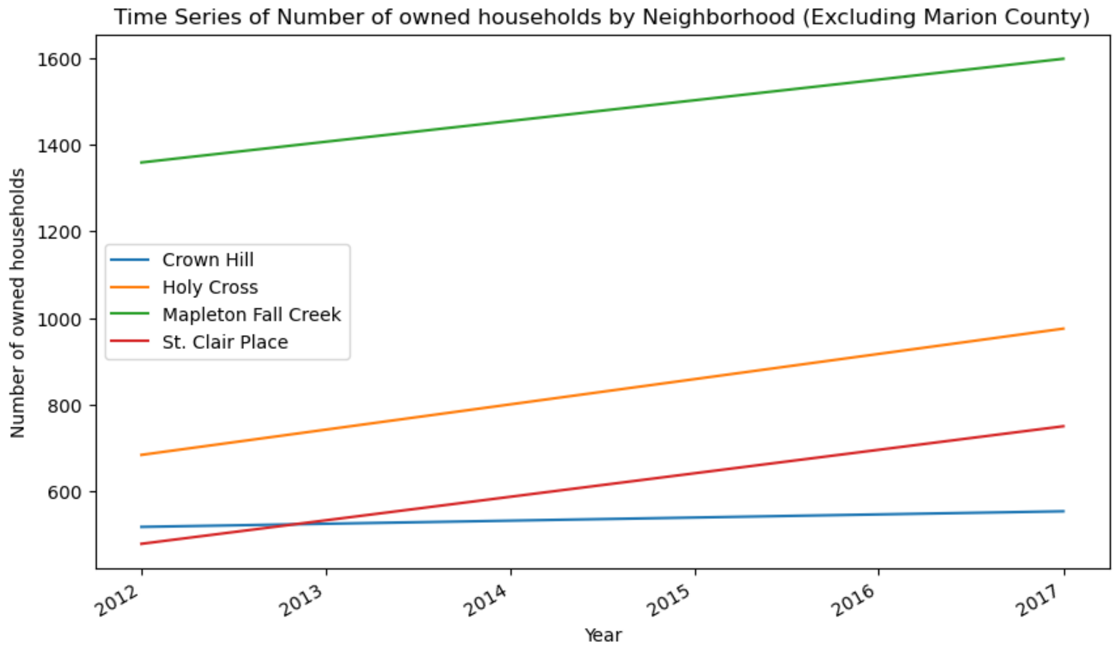
Stage 2) Data Modelling and Visualization

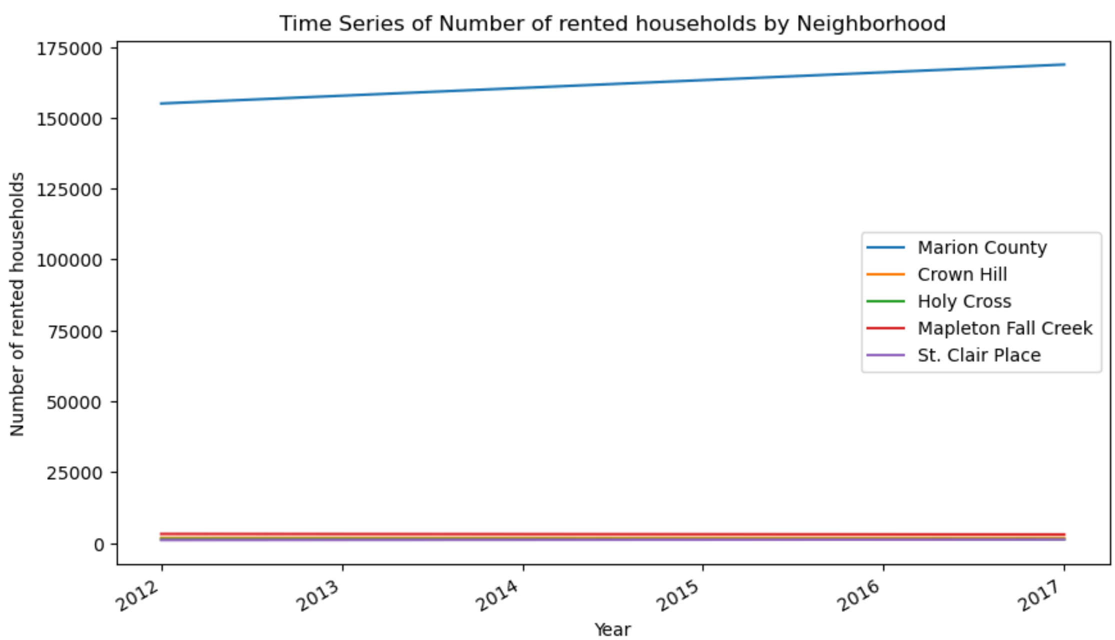
[9]:
```python
#Homeownership

#Households by tenure

def plot_time_series(df, ylabel, title, include_marion=True):
    neighborhoods = ['Crown Hill', 'Holy Cross', 'Mapleton Fall Creek', 'St.␣
 ↪Clair Place']
    if include_marion:
        neighborhoods.insert(0, 'Marion County')
    df.plot(kind='line', x='Year', y=neighborhoods, figsize=(10, 6),␣
 ↪title=title)
    plt.xlabel('Year')
    plt.ylabel(ylabel)
    plt.grid(False)
    plt.show()

#Number of owned households
plot_time_series(hholds_ten_own, 'Number of owned households', 'Time Series of␣
 ↪Number of owned households by Neighborhood')
plot_time_series(hholds_ten_own, 'Number of owned households', 'Time Series of␣
 ↪Number of owned households by Neighborhood (Excluding Marion County)',␣
 ↪include_marion=False)

#Number of rented households
plot_time_series(hholds_ten_rent, 'Number of rented households', 'Time Series␣
 ↪of Number of rented households by Neighborhood')
plot_time_series(hholds_ten_rent, 'Number of rented households', 'Time Series␣
 ↪of Number of rented households by Neighborhood (Excluding Marion County)',␣
 ↪include_marion=False)

#Percentage of owned households
plot_time_series(perc_hholds_ten_own, 'Percentage of owned households', 'Time␣
 ↪Series of Percentage of owned households by Neighborhood')
plot_time_series(perc_hholds_ten_own, 'Percentage of owned households', 'Time␣
 ↪Series of Percentage of owned households by Neighborhood (Excluding Marion␣
 ↪County)', include_marion=False)

#Percentage of rented households
```

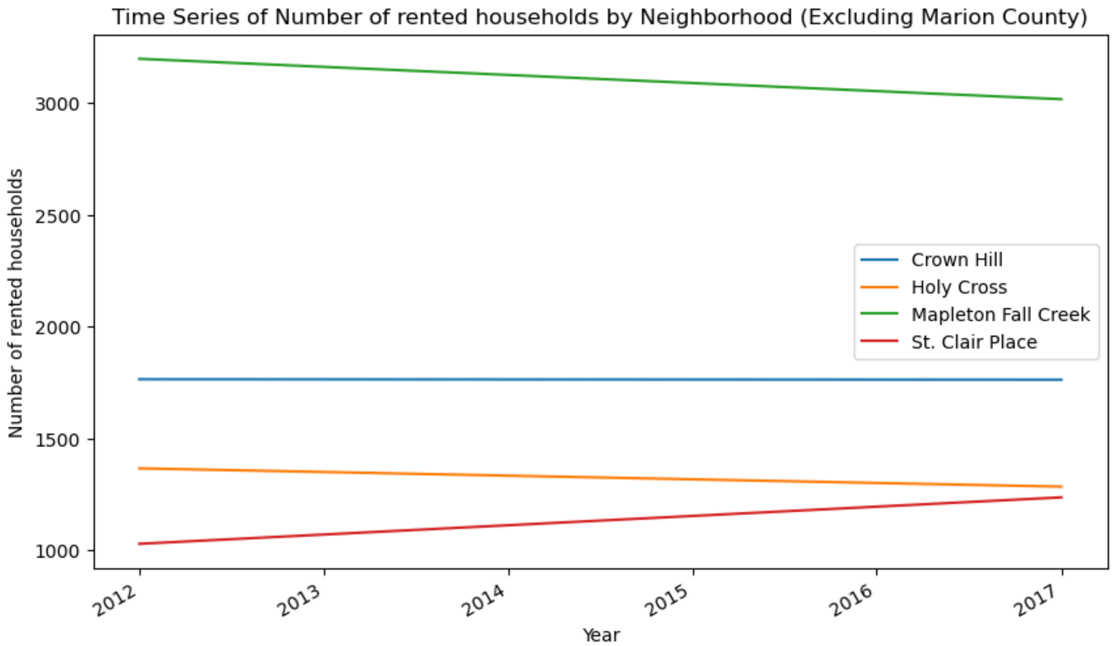
```python
plot_time_series(perc_hholds_ten_rent, 'Percentage of rent households', 'Time␣
 ↪Series of Percentage of rented households by Neighborhood')
plot_time_series(perc_hholds_ten_rent, 'Percentage of rent households', 'Time␣
 ↪Series of Percentage of rented households by Neighborhood (Excluding Marion␣
 ↪County)', include_marion=False)

#Homeownership Rate by Race (Group-Specific)

def plot_homeownership(df, demographic, title):
    df.plot(kind='line', x='Year', y=['Marion County', 'Crown Hill', 'Holy␣
 ↪Cross', 'Mapleton Fall Creek', 'St. Clair Place'], figsize=(10, 6),␣
 ↪title=title)
    plt.xlabel('Year')
    plt.ylabel('Percentage of Home Ownership')
    plt.grid(False)
    plt.show()

plot_homeownership(asian_ho_df, 'Asian', 'Time Series of Asian Homeownership␣
 ↪Rate by Neighborhood')
plot_homeownership(afr_ho_df, 'African-American', 'Time Series of␣
 ↪African-American Homeownership Rate by Neighborhood')
plot_homeownership(lat_ho_df, 'Latina/o', 'Time Series of Latina/o␣
 ↪Homeownership Rate by Neighborhood')
plot_homeownership(white_ho_df, 'Caucasian', 'Time Series of Caucasian␣
 ↪Homeownership Rate by Neighborhood')

#Homeownership Rate by Race (Neighborhood-Specific)

ethnicities = ['Asian', 'African-American', 'Caucasian', 'Latina/o']

dataframes = [asian_ho_df, afr_ho_df, lat_ho_df, white_ho_df]

neighborhoods = ['Marion County', 'Crown Hill', 'Holy Cross', 'Mapleton Fall␣
 ↪Creek', 'St. Clair Place']

fig, axes = plt.subplots(nrows=len(neighborhoods), ncols=1, figsize=(10, 24))

for i, neighborhood in enumerate(neighborhoods):
    ax = axes[i]
    ax.set_title(f'Homeownership Rate by Ethnicity in {neighborhood}')
    ax.set_xlabel('Year')
    ax.set_ylabel('Percentage of Home Ownership')
    ax.grid(False)

    for ethnicity, df in zip(ethnicities, dataframes):
        df.plot(kind='line', x='Year', y=neighborhood, ax=ax, label=ethnicity)
```

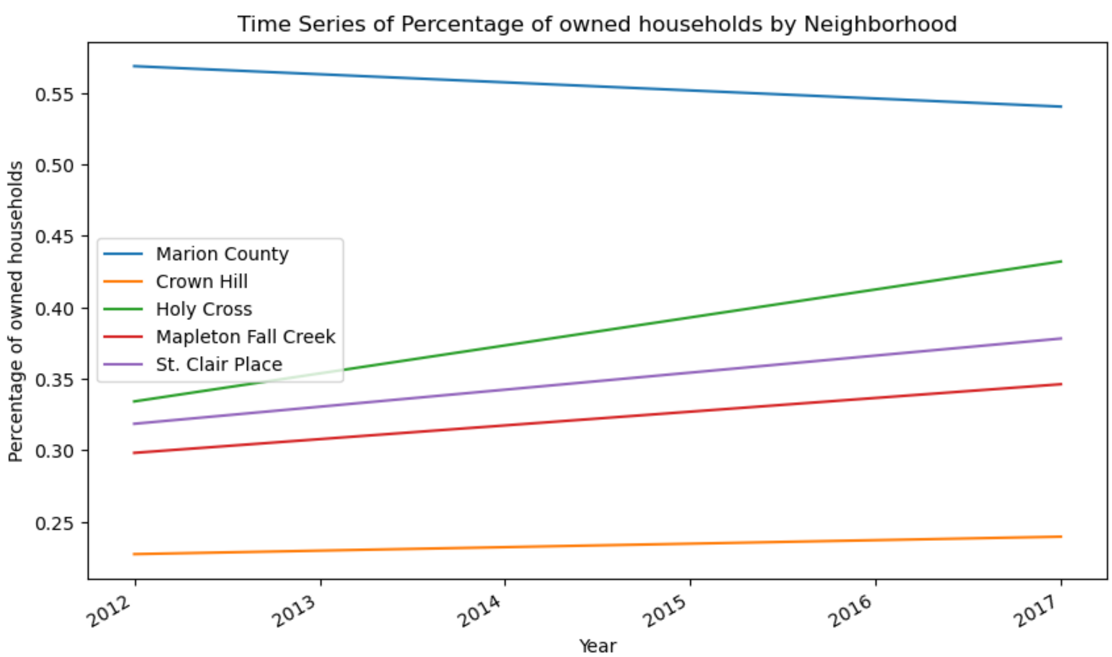
```
plt.tight_layout()
plt.legend()
plt.show()

#Homeownership by Race Visualization

fig, axes = plt.subplots(nrows=2, ncols=2, figsize=(15, 10))

axes = axes.flatten()

for i, neighborhood in enumerate(neighborhoods[1:]):
    ax = axes[i]
    ax.set_title(f'Homeownership Rate by Ethnicity in {neighborhood}')
    ax.set_xlabel('Year')
    ax.set_ylabel('Percentage of Home Ownership')
    ax.grid(False)

    for ethnicity, df in zip(ethnicities, dataframes):
        df.plot(kind='line', x='Year', y=neighborhood, ax=ax, label=ethnicity)

    ax.legend(loc='upper left')

plt.tight_layout()

plt.show()
```

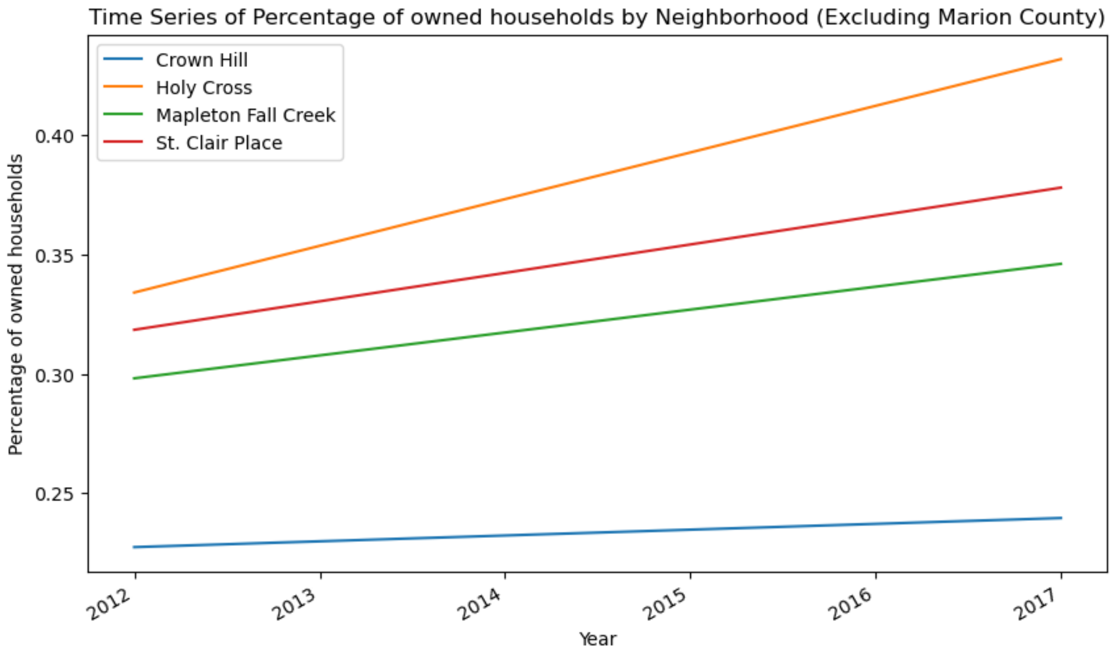


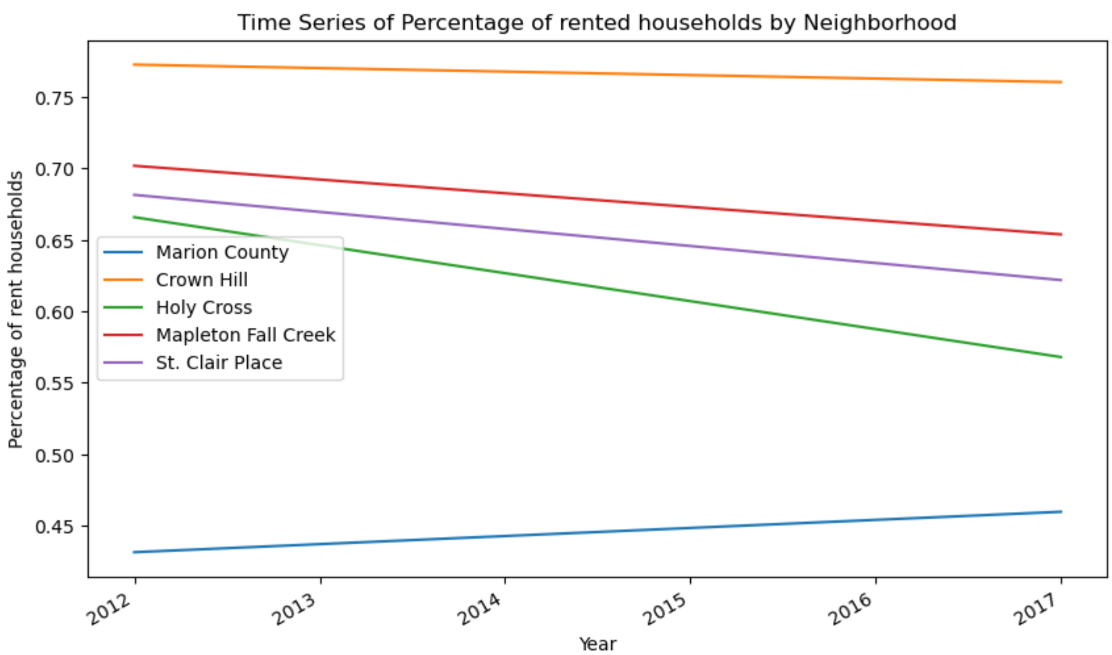

Time Series of Number of owned households by Neighborhood (Excluding Marion County)
Number of owned households
Year
1600
1400
1200
1000
800
600
2012
2013
2014
2015
2016
2017
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place

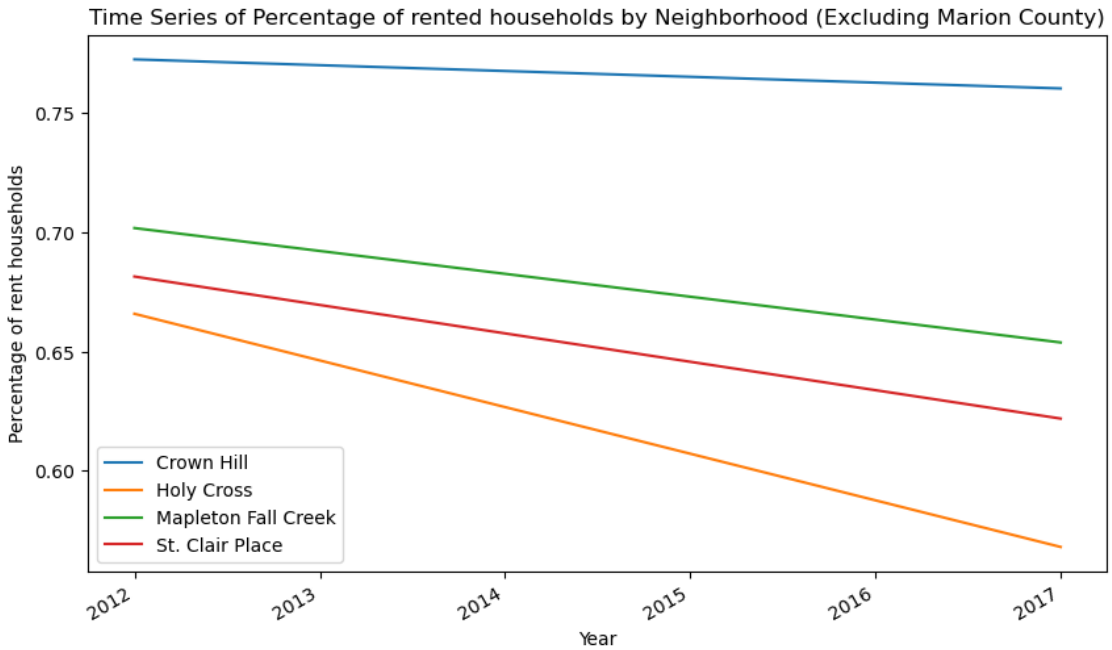

Time Series of Number of rented households by Neighborhood
Number of rented households
Year
175000
150000
125000
100000
75000
50000
25000
0
2012
2013
2014
2015
2016
2017
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place

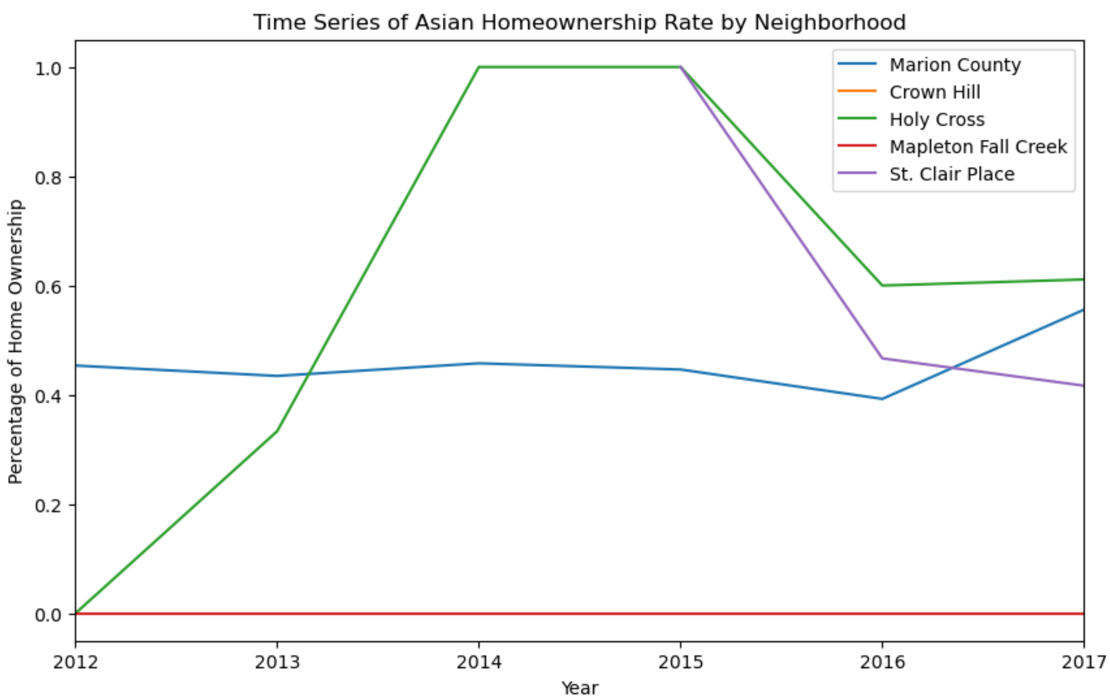

Time Series of Number of rented households by Neighborhood (Excluding Marion County)
Number of rented households
3000
2500
2000
1500
1000
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
2012
2013
2014
2015
2016
2017
Year

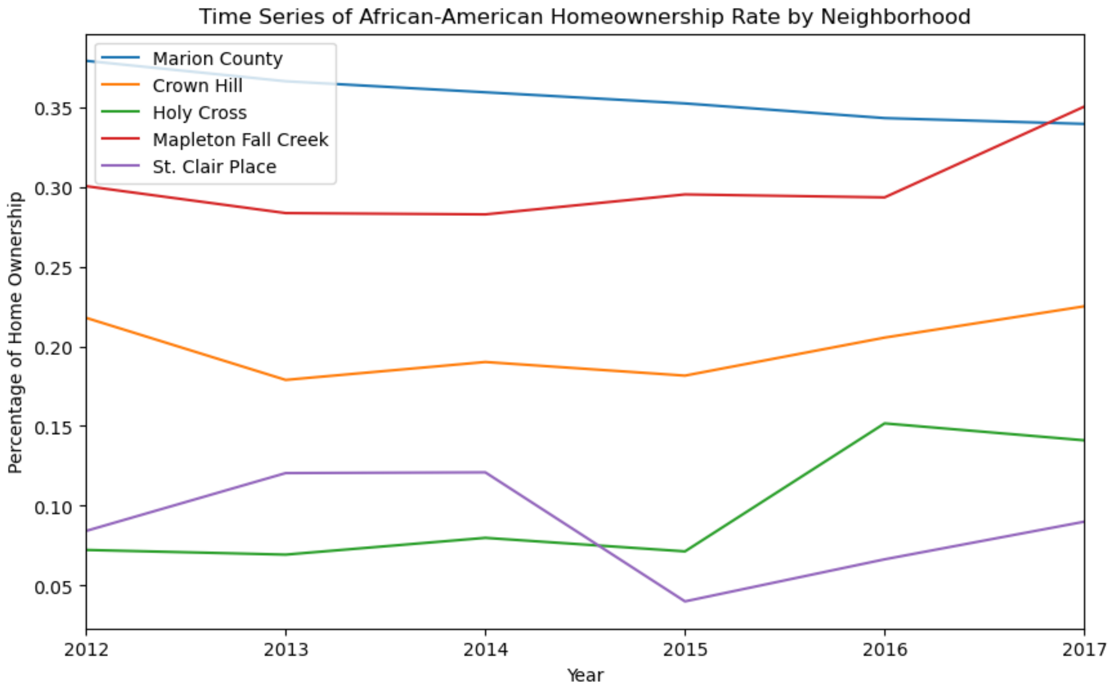

Time Series of Percentage of owned households by Neighborhood
Percentage of owned households
0.55
0.50
0.45
0.40
0.35
0.30
0.25
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
2012
2013
2014
2015
2016
2017
Year

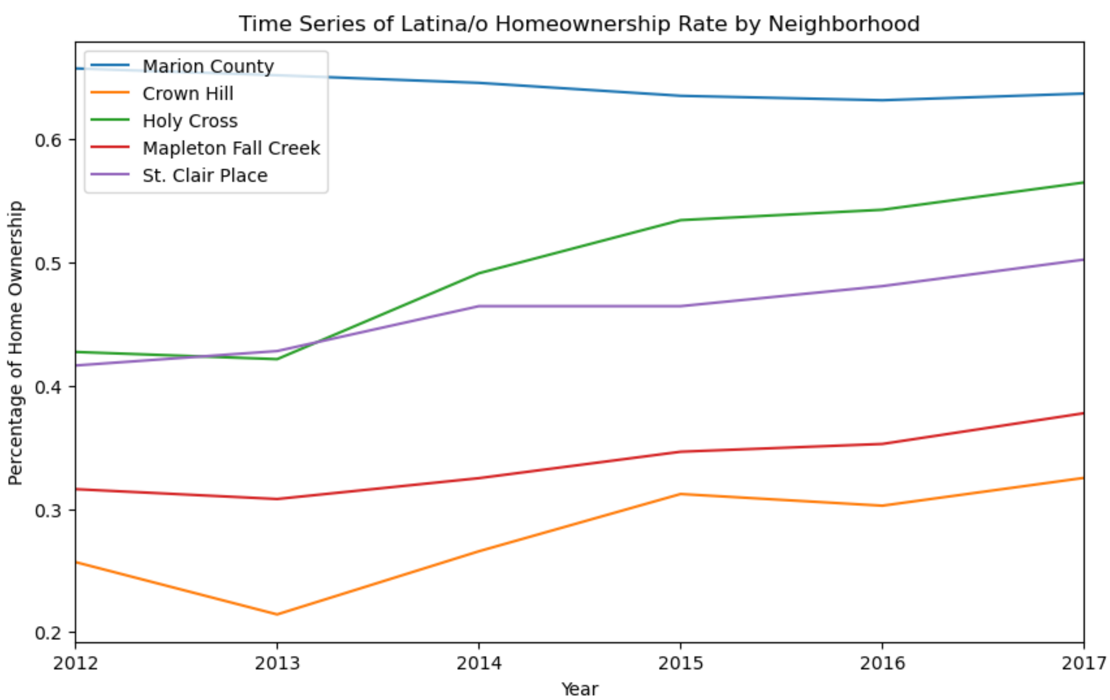

Time Series of Percentage of owned households by Neighborhood (Excluding Marion County)
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of owned households
0.40
0.35
0.30
0.25
2012
2013
2014
2015
2016
2017
Year

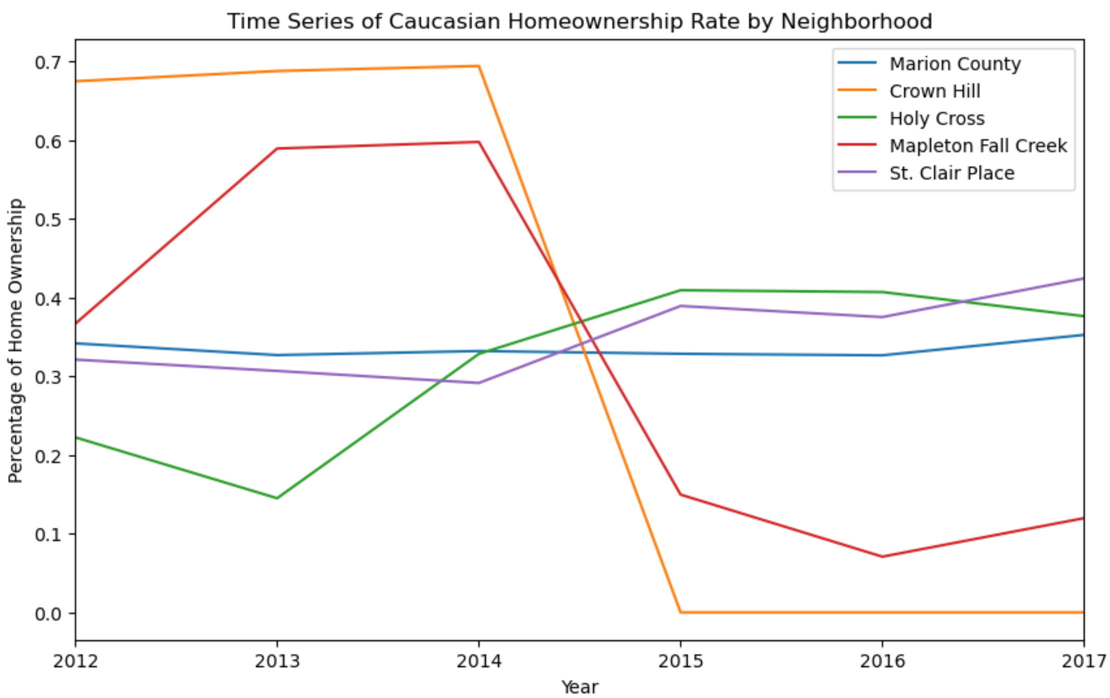

Time Series of Percentage of rented households by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of rent households
0.75
0.70
0.65
0.60
0.55
0.50
0.45
2012
2013
2014
2015
2016
2017
Year


Time Series of Percentage of rented households by Neighborhood (Excluding Marion County)
Percentage of rent households
0.75
0.70
0.65
0.60
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
2012
2013
2014
2015
2016
2017
Year

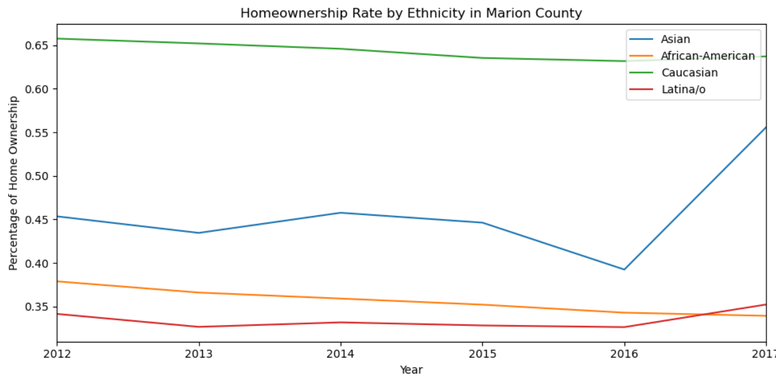

Time Series of Asian Homeownership Rate by Neighborhood
Percentage of Home Ownership
1.0
0.8
0.6
0.4
0.2
0.0
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
2012
2013
2014
2015
2016
2017
Year

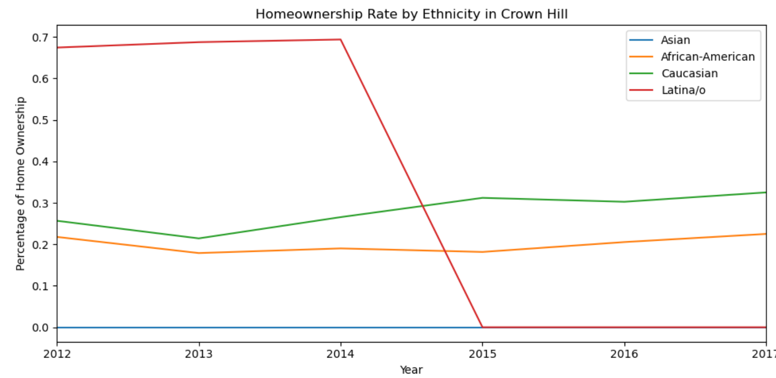

Time Series of African-American Homeownership Rate by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of Home Ownership
0.35
0.30
0.25
0.20
0.15
0.10
0.05
2012
2013
2014
2015
2016
2017
Year

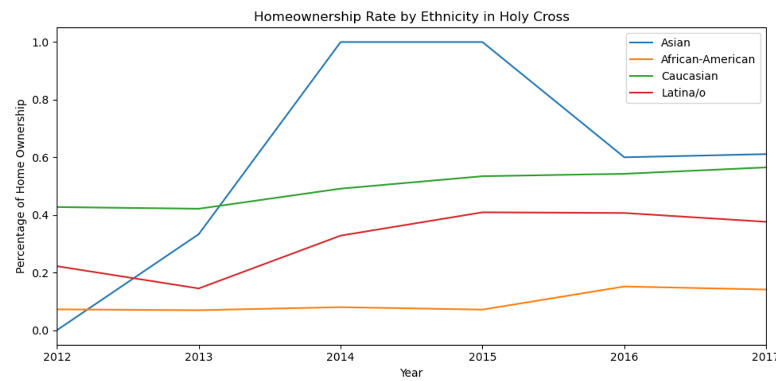

Time Series of Latina/o Homeownership Rate by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of Home Ownership
0.6
0.5
0.4
0.3
0.2
2012
2013
2014
2015
2016
2017
Year

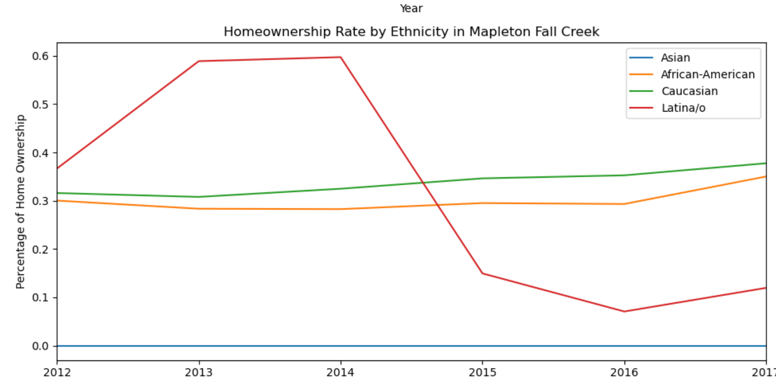


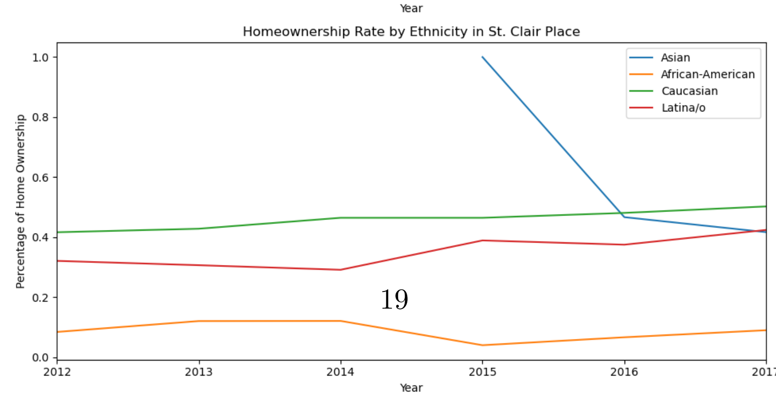
Time Series of Caucasian Homeownership Rate by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of Home Ownership
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
2012
2013
2014
2015
2016
2017
Year

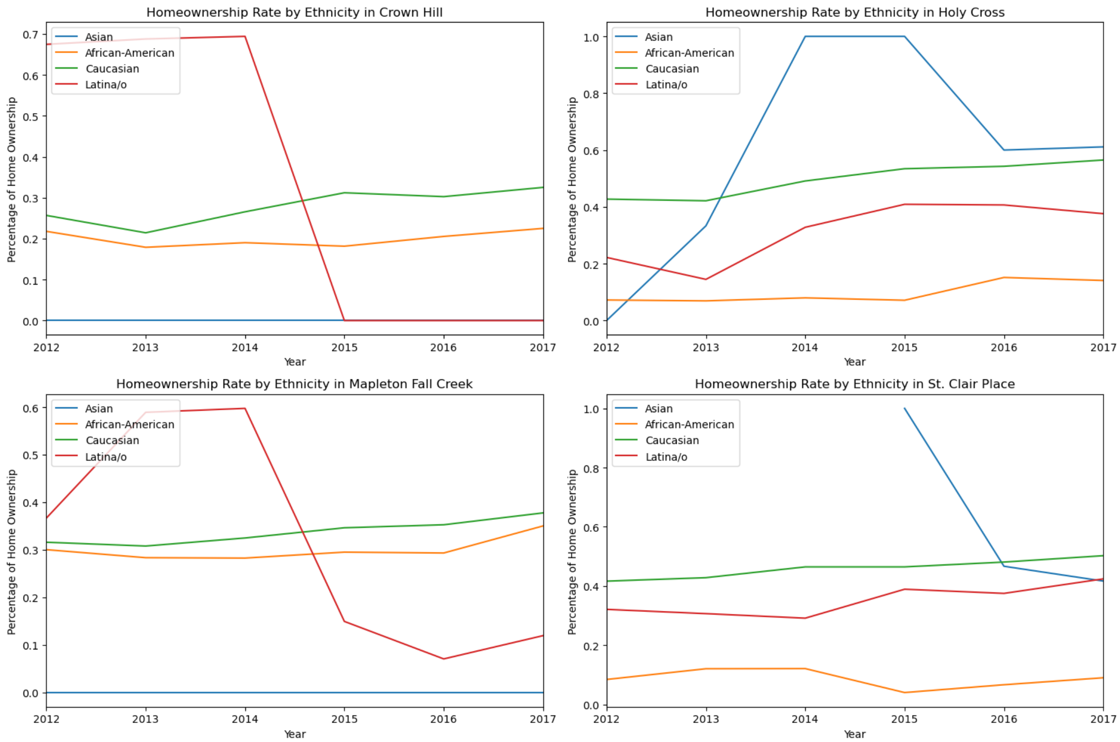

Homeownership Rate by Ethnicity in Marion County
Percentage of Home Ownership
Year
2012 2013 2014 2015 2016 2017
0.35 0.40 0.45 0.50 0.55 0.60 0.65
Asian
African-American
Caucasian
Latina/o


Homeownership Rate by Ethnicity in Crown Hill
Percentage of Home Ownership
Year
2012 2013 2014 2015 2016 2017
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7
Asian
African-American
Caucasian
Latina/o


Homeownership Rate by Ethnicity in Holy Cross
Percentage of Home Ownership
Year
2012 2013 2014 2015 2016 2017
0.0 0.2 0.4 0.6 0.8 1.0
Asian
African-American
Caucasian
Latina/o

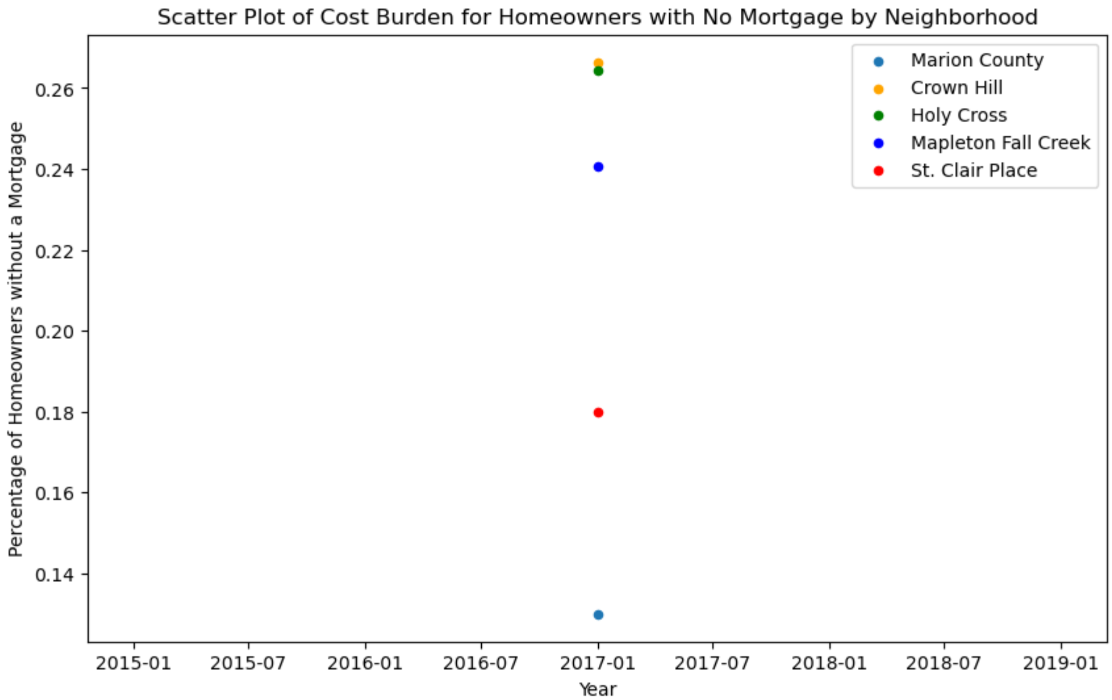

Homeownership Rate by Ethnicity in Mapleton Fall Creek
Percentage of Home Ownership
Year
2012 2013 2014 2015 2016 2017
0.0 0.1 0.2 0.3 0.4 0.5 0.6
Asian
African-American
Caucasian
Latina/o

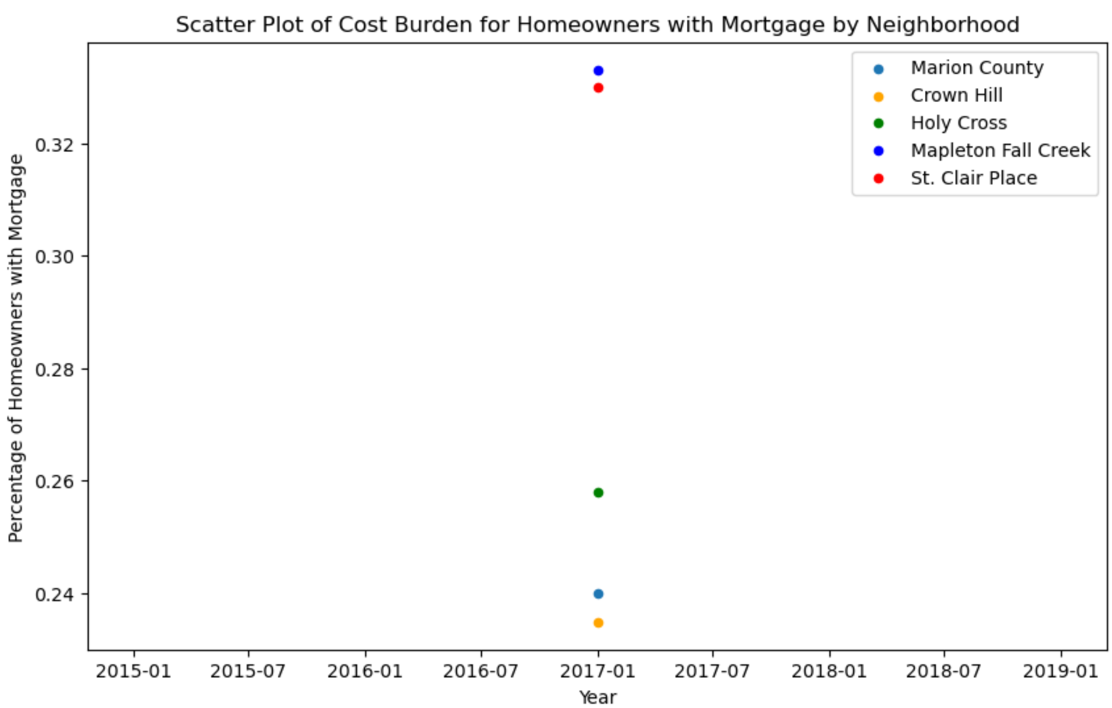

Homeownership Rate by Ethnicity in St. Clair Place
Percentage of Home Ownership
Year
2012 2013 2014 2015 2016 2017
0.0 0.2 0.4 0.6 0.8 1.0
Asian
African-American
Caucasian
Latina/o

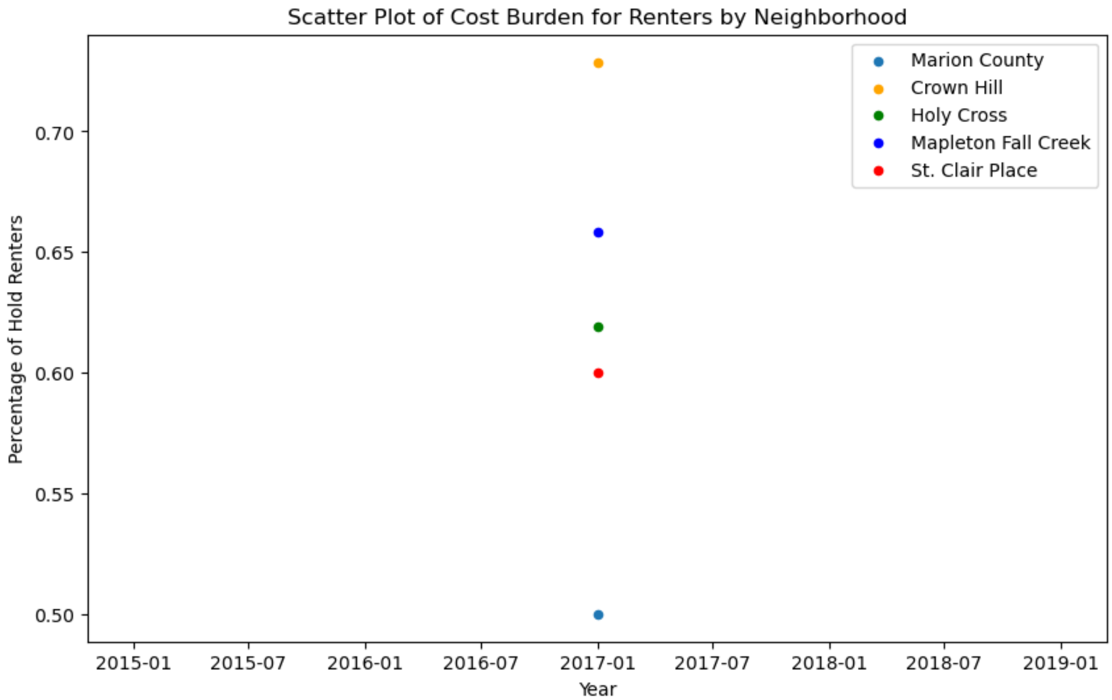


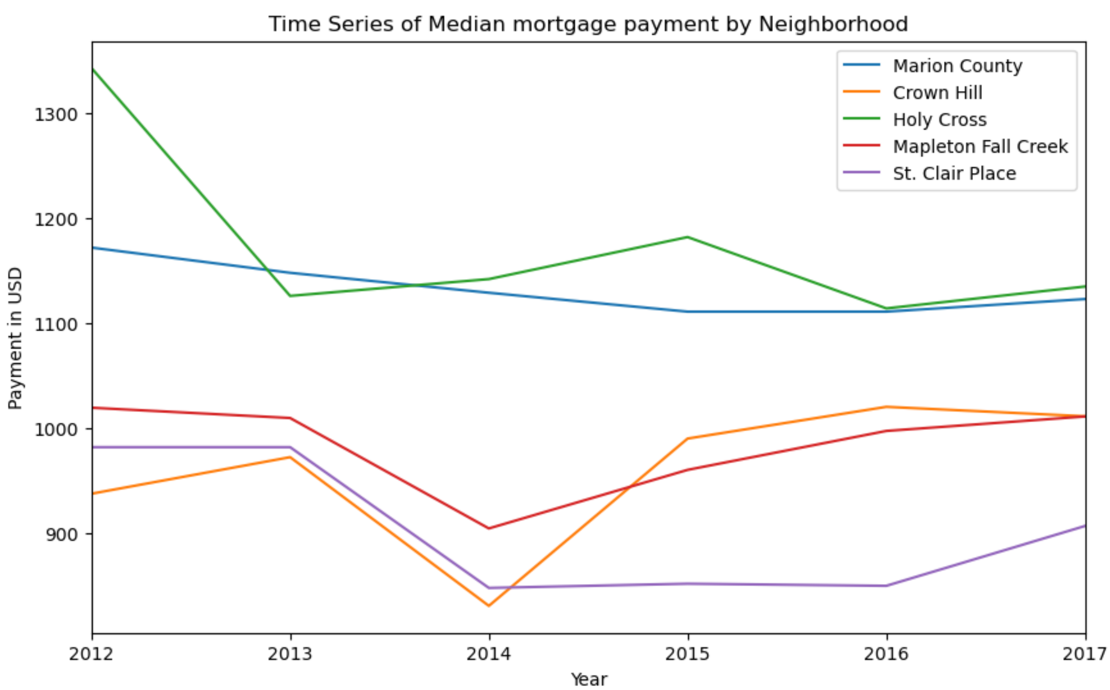
[10]:
```python
#Housing Cost

#Cost burden

def plot_scatter(df, ylabel, title):
    df.plot(kind='scatter', x='Year', y='Marion County', figsize=(10, 6), title=title)
    df.plot(kind='scatter', x='Year', y='Crown Hill', color='orange', ax=plt.gca())
    df.plot(kind='scatter', x='Year', y='Holy Cross', color='green', ax=plt.gca())
    df.plot(kind='scatter', x='Year', y='Mapleton Fall Creek', color='blue', ax=plt.gca())
    df.plot(kind='scatter', x='Year', y='St. Clair Place', color='red', ax=plt.gca())

    plt.xlabel('Year')
    plt.ylabel(ylabel)
    plt.grid(False)
```

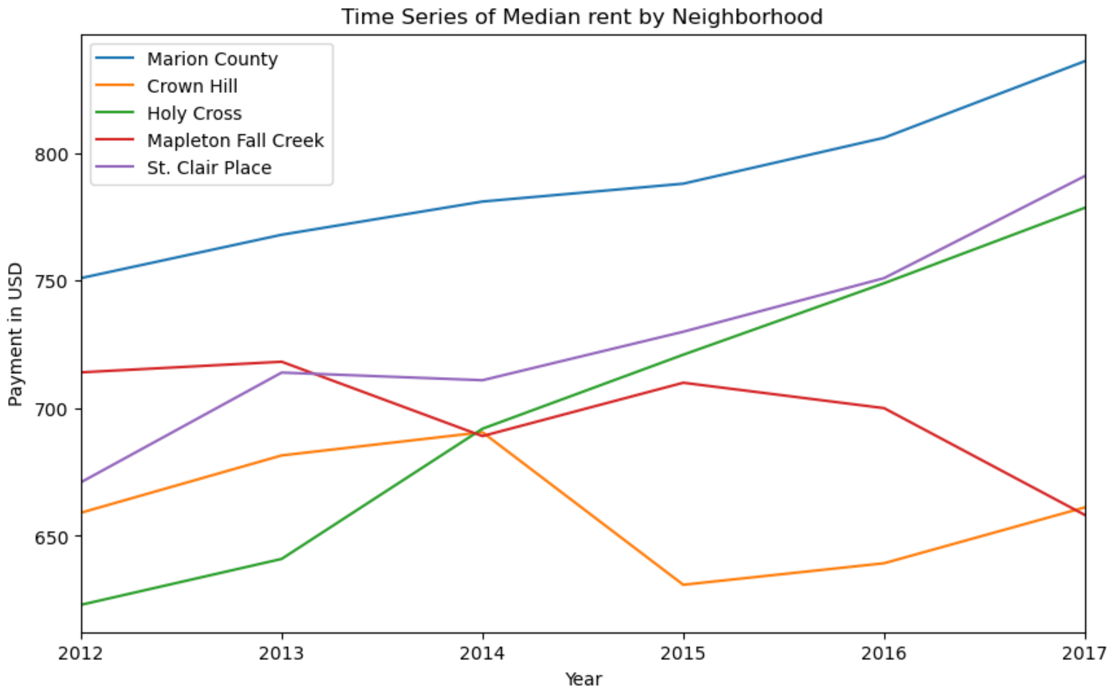
```
    plt.legend(['Marion County', 'Crown Hill', 'Holy Cross', 'Mapleton Fall␣
 ↪Creek', 'St. Clair Place'])
    plt.show()

plot_scatter(cb_h_nm_df, 'Percentage of Homeowners without a Mortgage',␣
 ↪'Scatter Plot of Cost Burden for Homeowners with No Mortgage by␣
 ↪Neighborhood')
plot_scatter(cb_h_wm_df, 'Percentage of Homeowners with Mortgage', 'Scatter␣
 ↪Plot of Cost Burden for Homeowners with Mortgage by Neighborhood')
plot_scatter(cb_rent_df, 'Percentage of Hold Renters', 'Scatter Plot of Cost␣
 ↪Burden for Renters by Neighborhood')

# Median monthly rent and mortgage payment amounts

def plot_time_series(df, ylabel, title):
    df.plot(kind='line', x='Year', y=['Marion County', 'Crown Hill', 'Holy␣
 ↪Cross', 'Mapleton Fall Creek', 'St. Clair Place'], figsize=(10, 6),␣
 ↪title=title)
    plt.xlabel('Year')
    plt.ylabel(ylabel)
    plt.grid(False)
    plt.show()

plot_time_series(med_mon_mort_df, 'Payment in USD', 'Time Series of Median␣
 ↪mortgage payment by Neighborhood')
plot_time_series(med_mon_rent_df, 'Payment in USD', 'Time Series of Median rent␣
 ↪by Neighborhood')
```


Scatter Plot of Cost Burden for Homeowners with No Mortgage by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of Homeowners without a Mortgage
0.26
0.24
0.22
0.20
0.18
0.16
0.14
2015-01
2015-07
2016-01
2016-07
2017-01
2017-07
2018-01
2018-07
2019-01
Year


Scatter Plot of Cost Burden for Homeowners with Mortgage by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of Homeowners with Mortgage
0.32
0.30
0.28
0.26
0.24
2015-01
2015-07
2016-01
2016-07
2017-01
2017-07
2018-01
2018-07
2019-01
Year

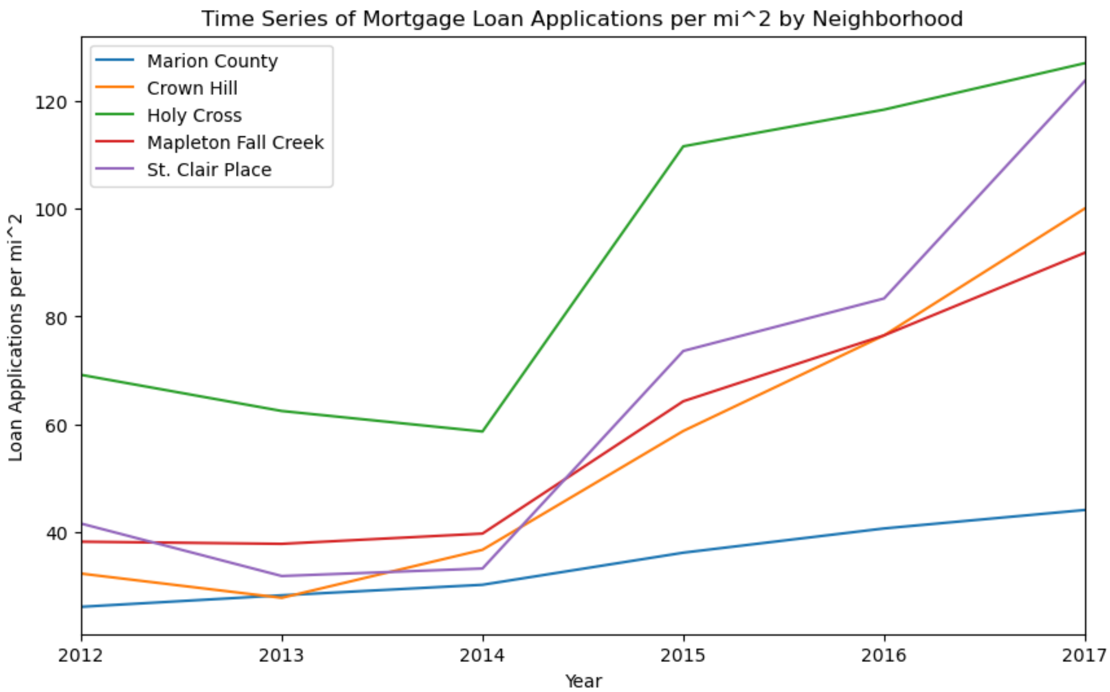

Scatter Plot of Cost Burden for Renters by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Percentage of Hold Renters
0.70
0.65
0.60
0.55
0.50
2015-01
2015-07
2016-01
2016-07
2017-01
2017-07
2018-01
2018-07
2019-01
Year

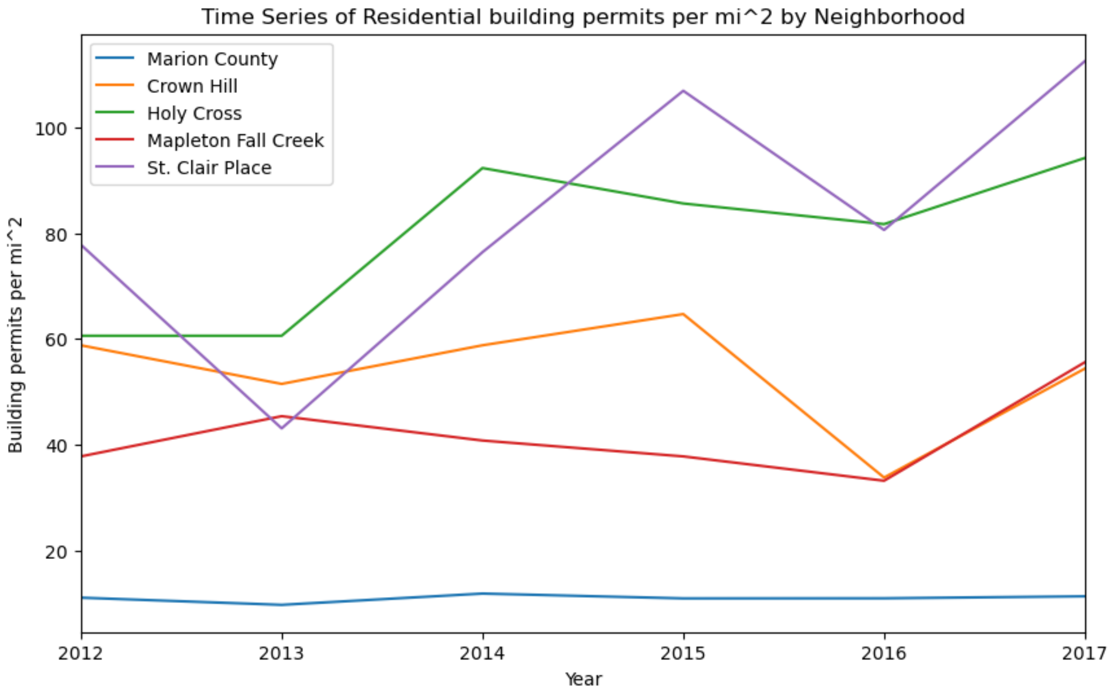

Time Series of Median mortgage payment by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Payment in USD
1300
1200
1100
1000
900
2012
2013
2014
2015
2016
2017
Year

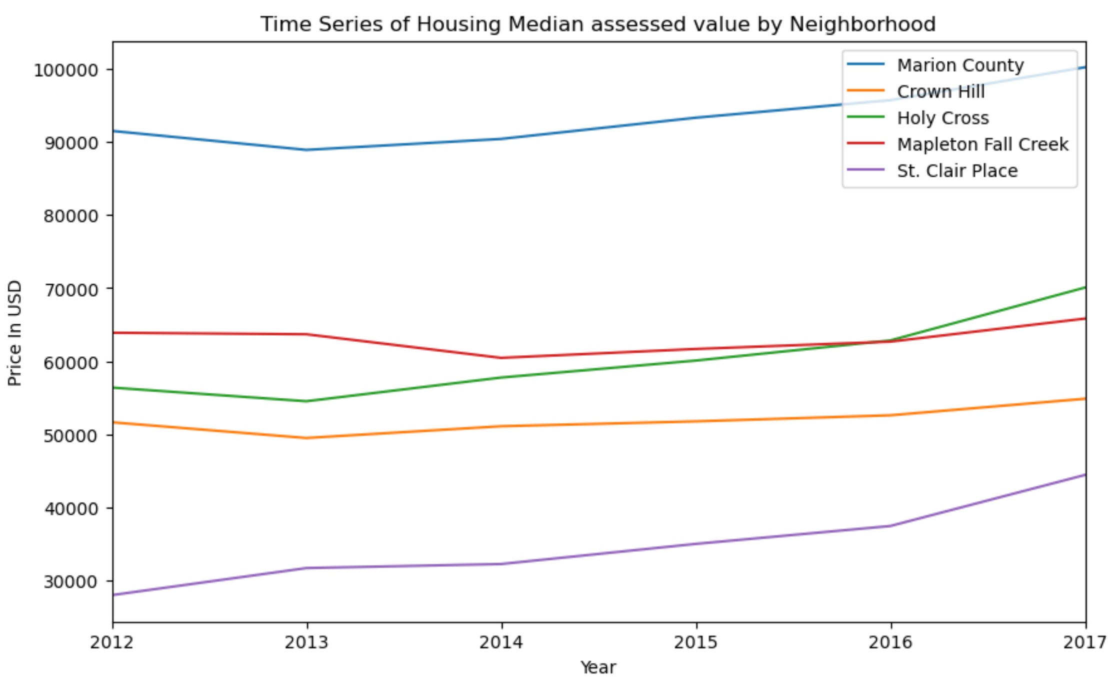


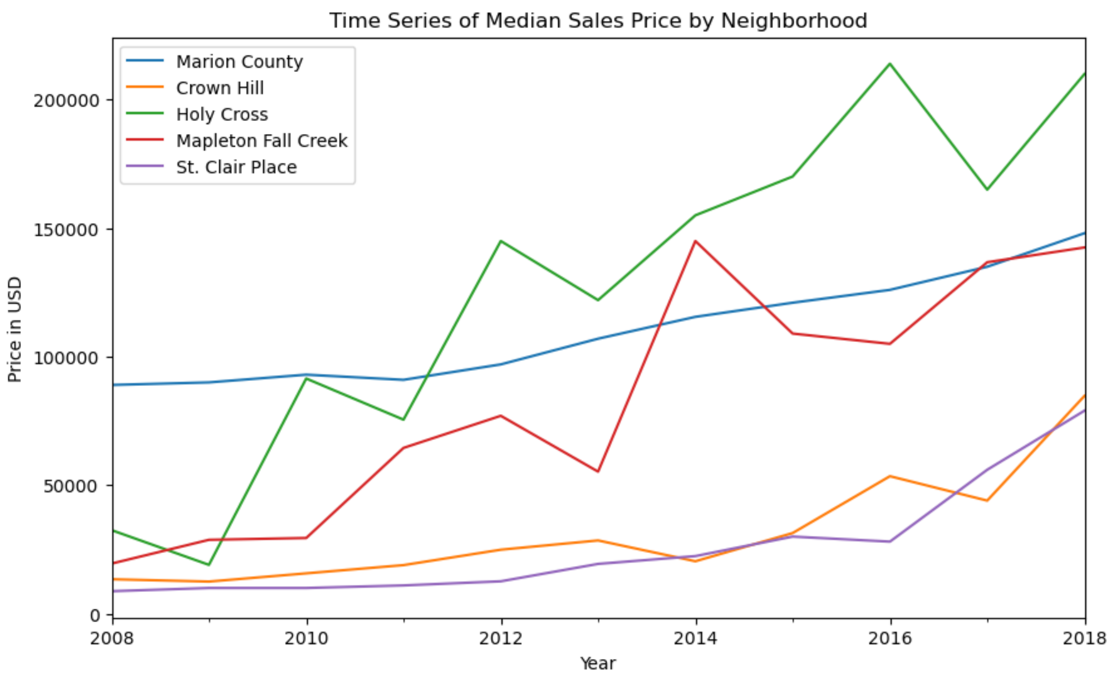
[11]:
```python
def plot_housing_metric(df, ylabel, title, include_marion=True):
    neighborhoods = ['Marion County', 'Crown Hill', 'Holy Cross', 'Mapleton Fall Creek', 'St. Clair Place']
    if not include_marion:
        neighborhoods.remove('Marion County')
    df.plot(kind='line', x='Year', y=neighborhoods, figsize=(10, 6), title=title)
    plt.xlabel('Year')
    plt.ylabel(ylabel)
    plt.grid(False)
    plt.show()

# Mortgage loan applications per sq mile
plot_housing_metric(mortgage_loan_app_df, 'Loan Applications per mi^2', 'Time Series of Mortgage Loan Applications per mi^2 by Neighborhood')

# Residential building permits per square mile
plot_housing_metric(res_buil_per_df, 'Building permits per mi^2', 'Time Series of Residential building permits per mi^2 by Neighborhood')

# Median assessed value
plot_housing_metric(med_a_val_df, 'Price In USD', 'Time Series of Housing Median assessed value by Neighborhood')
```

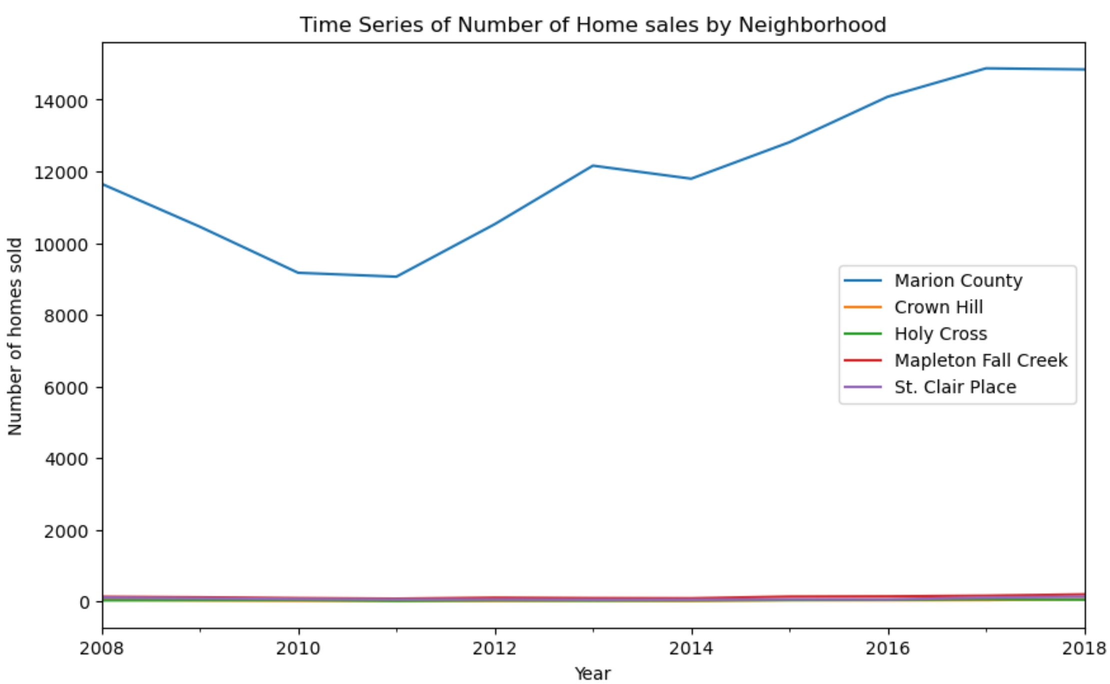
```python
# Median sales price
plot_housing_metric(med_sales_pri_df, 'Price in USD', 'Time Series of Median␣
↪Sales Price by Neighborhood')

# Number of home sales
plot_housing_metric(num_home_s_df, 'Number of homes sold', 'Time Series of␣
↪Number of Home sales by Neighborhood')
plot_housing_metric(num_home_s_df, 'Number of homes sold', 'Time Series of␣
↪Number of Home sales by Neighborhood', include_marion=False)
```

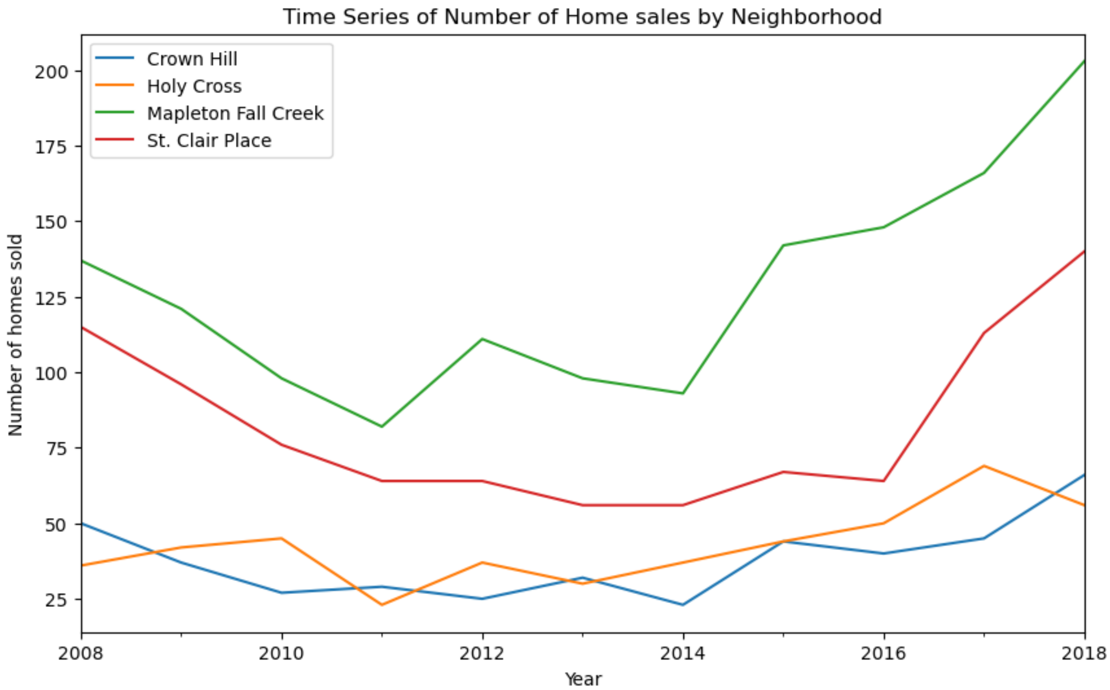



Time Series of Residential building permits per mi^2 by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Building permits per mi^2
20
40
60
80
100
2012
2013
2014
2015
2016
2017
Year


Time Series of Housing Median assessed value by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Price In USD
30000
40000
50000
60000
70000
80000
90000
100000
2012
2013
2014
2015
2016
2017
Year


Time Series of Median Sales Price by Neighborhood
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Price in USD
200000
150000
100000
50000
0
2008
2010
2012
2014
2016
2018
Year


Time Series of Number of Home sales by Neighborhood
Number of homes sold
14000
12000
10000
8000
6000
4000
2000
0
Marion County
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
2008
2010
2012
2014
2016
2018
Year

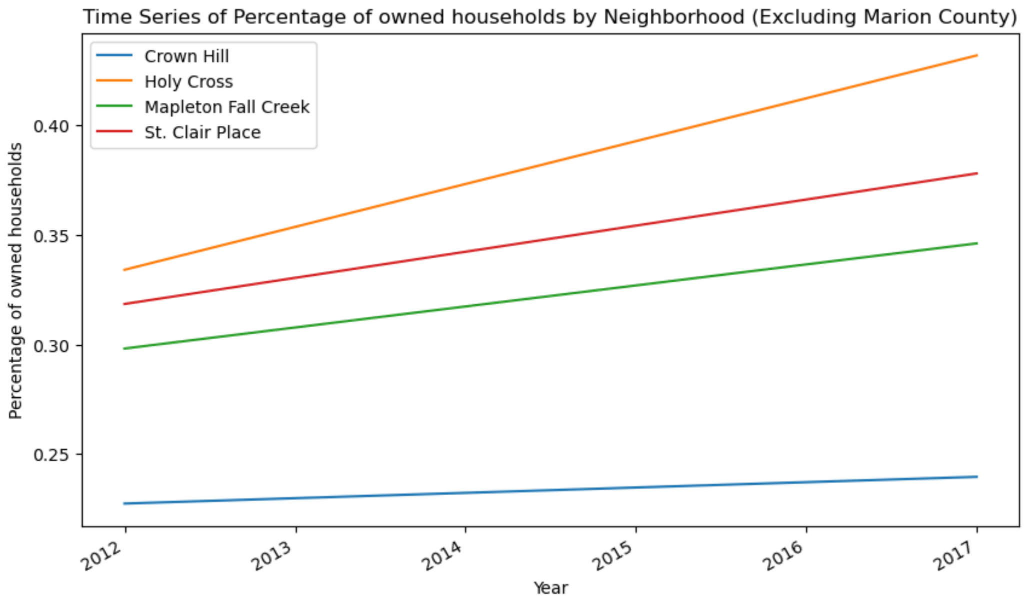


Time Series of Number of Home sales by Neighborhood
Crown Hill
Holy Cross
Mapleton Fall Creek
St. Clair Place
Number of homes sold
200
175
150
125
100
75
50
25
2008
2010
2012
2014
2016
2018
Year

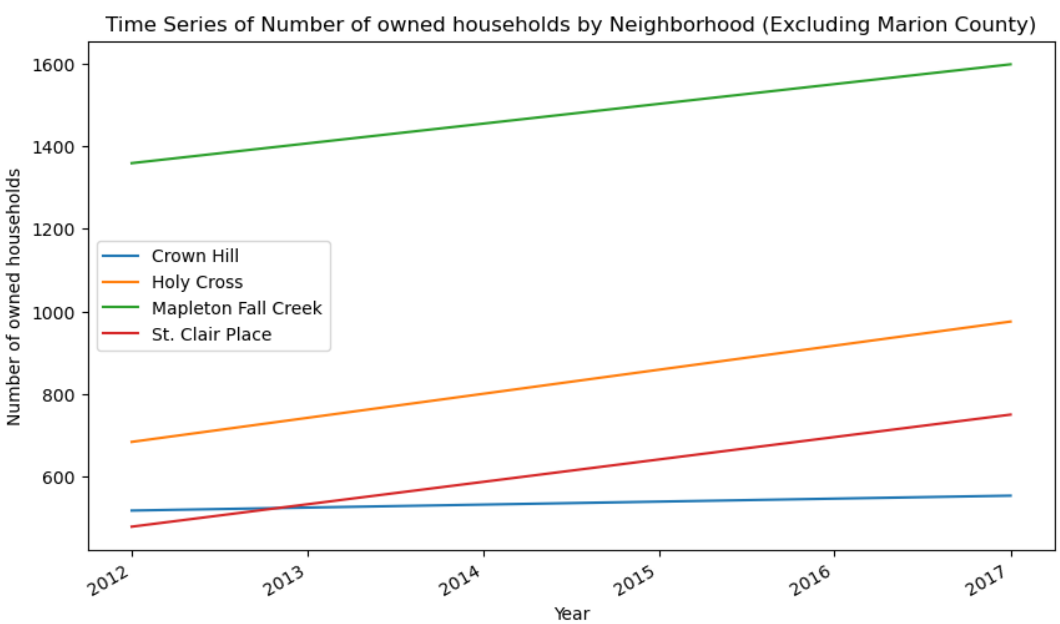
Stage 3) Interpretations, Analysis and Conclusions

Based on the data, it is possible that increasing building permit issuance in some Indianapolis neighborhoods could have a beneficial impact for its communities.

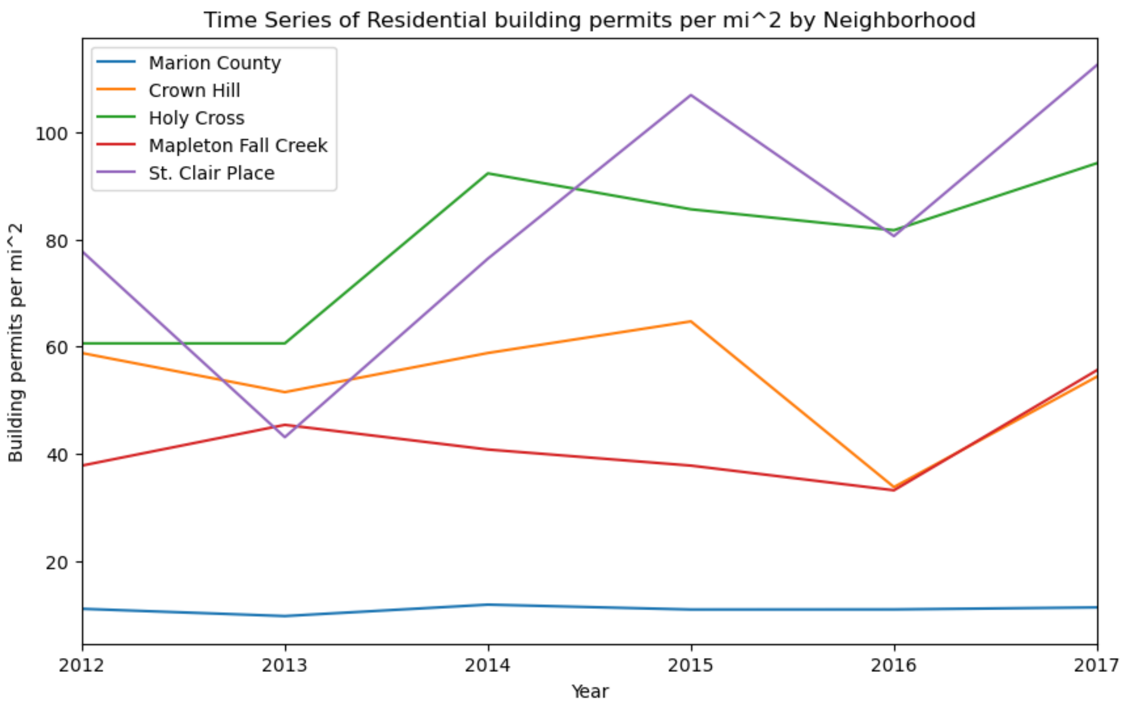
When observing the household ownership trends among some of Indianapolis neighborhoods from 2012 to 2017, it seems that the two neighborhoods from the Eastside area, Holy Cross and St. Clair Place, have had a larger growth (10% and 6%) when compared to two neighborhoods in the Mid-North area, Mapleton Fall Creek and Crown Hill (5% and 1%).



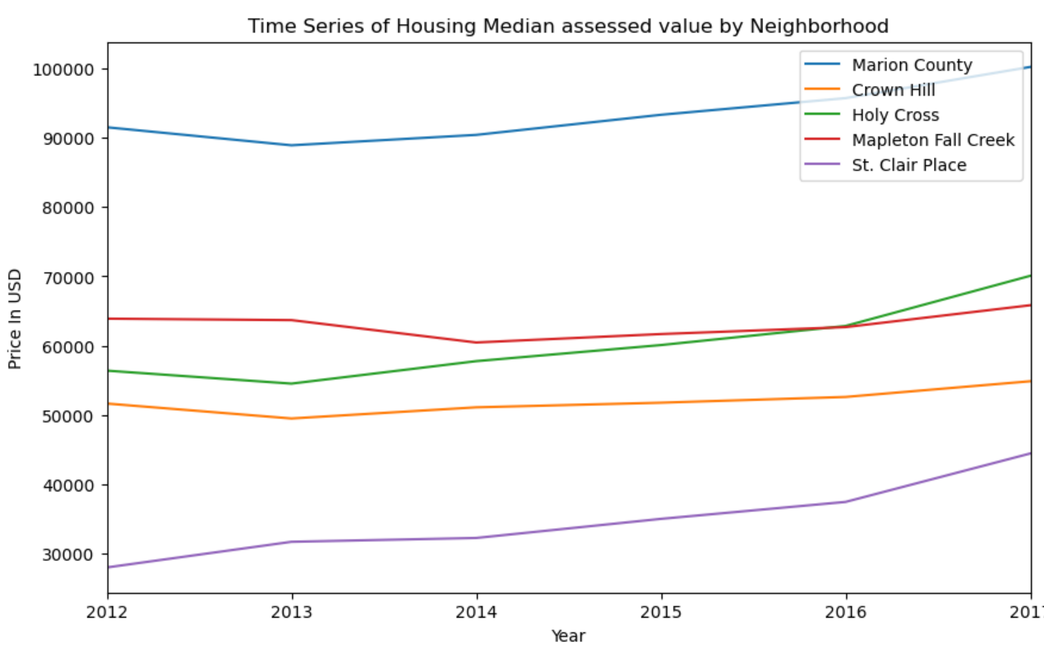
This difference seems real, as the actual numbers of household owners stagnated in Crown Hill (38 increase) and Mapleton Fall Creek had a relatively small growth (239 increase), despite starting with more than double the number of homes to its Eastside counterparts, which both experienced an almost 300 increase during this period.



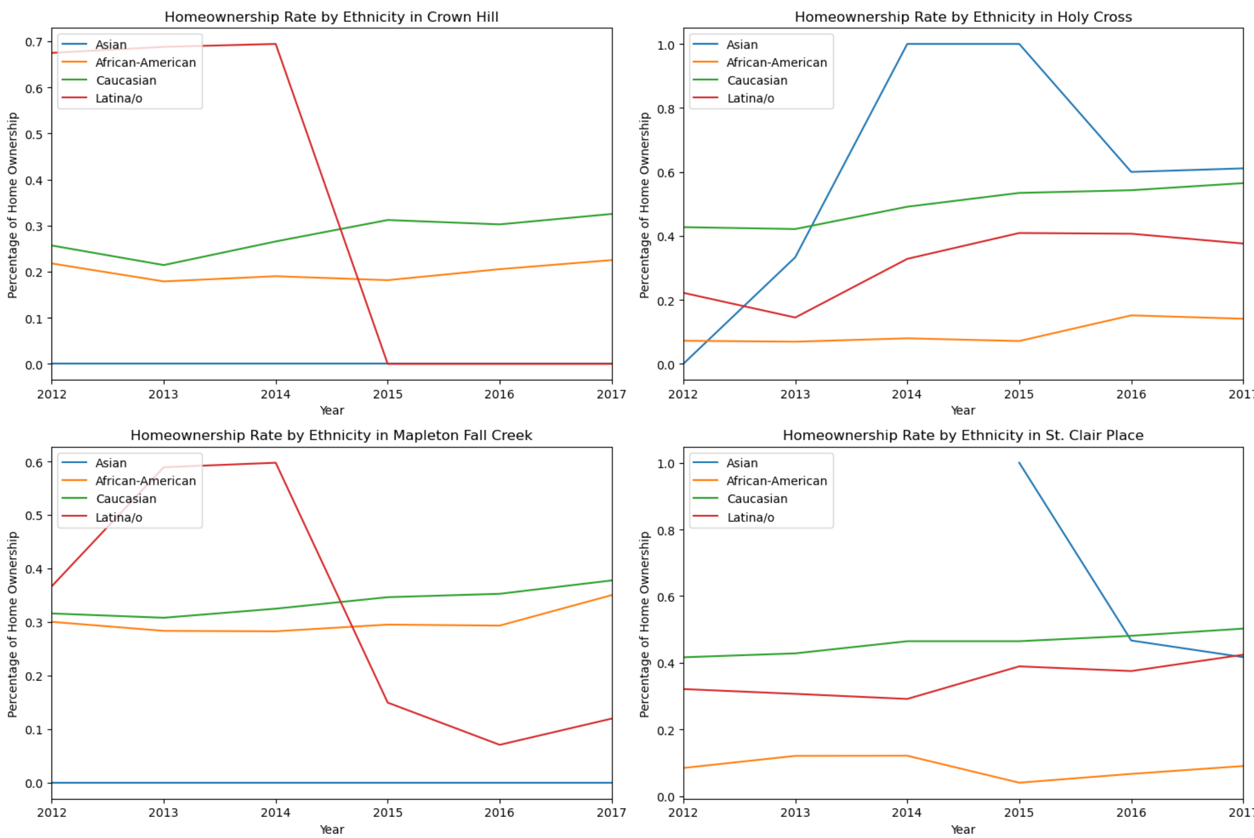
One of the possible variables that could be explaining this difference in trends between these two pairs of neighborhoods could be the unique and prolonged increase of residential building permits issued in these Eastside neighborhoods starting in 2013. More residential building permits result in more households and can allow a locality to meet the housing demands of the population.



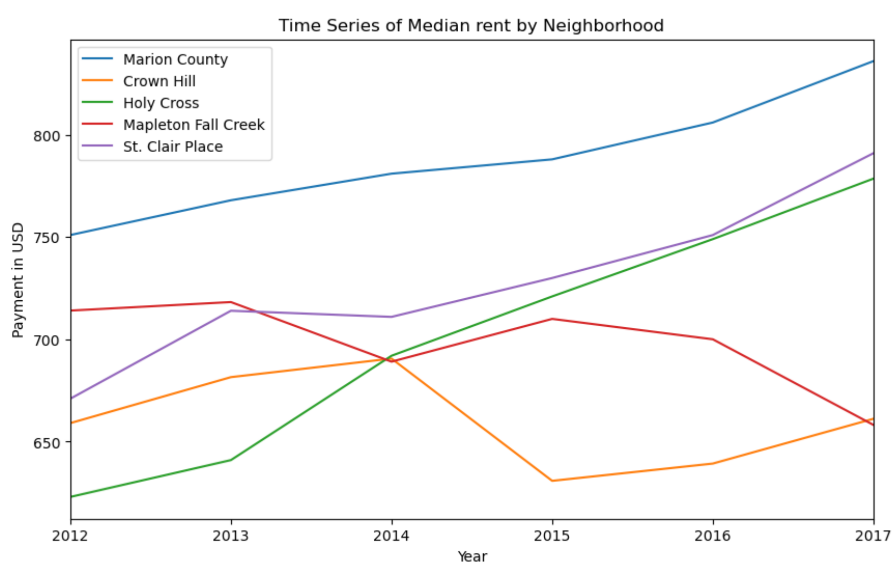
This permit expansion could also be affecting the local median assessed property values, as the Eastside neighborhoods had a noticeably greater growth during this time period.

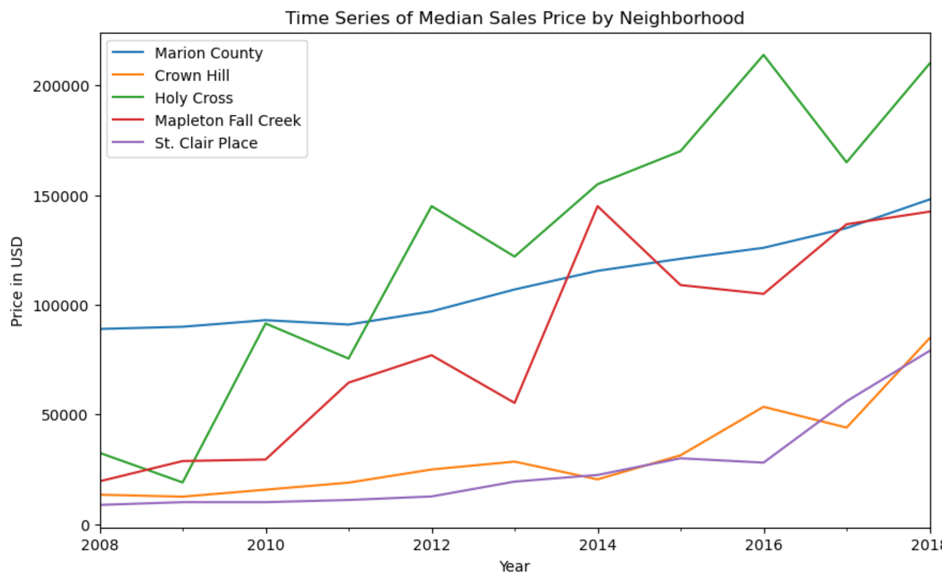


These two Eastside neighborhoods could also be benefiting minorities, as these two localities seem to be experiencing a higher Non-Caucasian homeownership rate during this period.



The increase in building permits could also be beneficial to current homeowners, as the median rent and sales prices of households in these two neighborhoods is greater than the other two.





As a nuance, expanding residential building permits in this locality should be cautious, as intensifying the issuance of these permits also has risks. It could induce an excess supply of real estate in neighborhoods, leading to insurmountable household prices that stagnate sales, followed by an economic downturn due to massive losses from investments.

This was a cursory analysis, as it used only five cases with a few observations (between 1 – 11 for each). For future research, I propose nurturing this data by making a larger database with more variables (median income, number of registered businesses) and cases (neighborhoods) that could allow performing thorough econometric (cointegration tests, difference in difference estimation, fixed effects panel models) and machine learning (K-Nearest Neighbors algorithm) techniques to arrive to robust findings.