

rental_analysis

May 6, 2022

1 Toronto Dwellings Analysis

In this assignment, you will perform fundamental analysis for the Toronto dwellings market to allow potential real estate investors to choose rental investment properties.

```
[1]: # imports
import panel as pn
pn.extension('plotly')
import plotly.express as px
import pandas as pd
import hvplot.pandas
import matplotlib.pyplot as plt
import os
import random
from pathlib import Path
from dotenv import load_dotenv
```

```
[2]: # Read the Mapbox API key
load_dotenv()

map_box_api = os.getenv("map_box_api")
# Set the Mapbox API
px.set_mapbox_access_token(map_box_api)
```

```
[3]: type("mapbox")
```

```
[3]: str
```

1.1 Load Data

```
[4]: # Read the census data into a Pandas DataFrame
file_path = Path("Data/toronto_neighbourhoods_census_data.csv")
to_data = pd.read_csv(file_path, index_col="year")
to_data.head()
```

```
[4]:
```

	neighbourhood	single_detached_house	\
year			
2001	Agincourt North		3715

2001	Agincourt	South-Malvern West	3250
2001		Alderwood	3175
2001		Annex	1060
2001		Banbury-Don Mills	3615

	apartment_five_storeys_plus	movable_dwelling	semi_detached_house	\
year				
2001	1480	0	1055	
2001	1835	0	545	
2001	315	0	470	
2001	6090	5	1980	
2001	4465	0	240	

	row_house	duplex	apartment_five_storeys_less	other_house	\
year					
2001	1295	195	185	5	
2001	455	105	425	0	
2001	50	185	370	0	
2001	605	275	3710	165	
2001	380	15	1360	0	

	average_house_value	shelter_costs_owned	shelter_costs_rented
year			
2001	200388	810	870
2001	203047	806	892
2001	259998	817	924
2001	453850	1027	1378
2001	371864	1007	1163

1.2 Dwelling Types Per Year

In this section, you will calculate the number of dwelling types per year. Visualize the results using bar charts and the Pandas plot function.

Hint: Use the Pandas `groupby` function.

Optional challenge: Plot each bar chart in a different color.

```
[5]: # Calculate the sum number of dwelling types units per year (hint: use groupby)
dwelling_per_year = to_data.groupby("year").sum().
    ↪drop(columns=["average_house_value", "shelter_costs_owned",
    ↪"shelter_costs_rented" ])

dwelling_per_year
```

```
[5]:      single_detached_house  apartment_five_storeys_plus  movable_dwelling  \
year
2001                300930                355015                75
2006                266860                379400                165
2011                274940                429220                100
2016                269680                493270                95

      semi_detached_house  row_house  duplex  apartment_five_storeys_less  \
year
2001                90995        52355    23785                116900
2006                69430        54690    44095                162850
2011                72480        60355    44750                163895
2016                71200        61565    48585                165575

      other_house
year
2001            3040
2006            1335
2011            2165
2016            2845
```

```
[6]: # Save the dataframe as a csv file
dwelling_per_year.to_csv("Data/dwelling_types_per_year.csv")
```

```
[7]: # Helper create_bar_chart function
def create_bar_chart(data, title, xlabel, ylabel, color):
    """
    Create a barplot based in the data argument.
    Input:
    data = DataFrame to use for plotting the data
    title = Chart Title
    xlabel = Label for X Axis
    ylabel = Label for Y Axis
    color = Colour of the bar chart

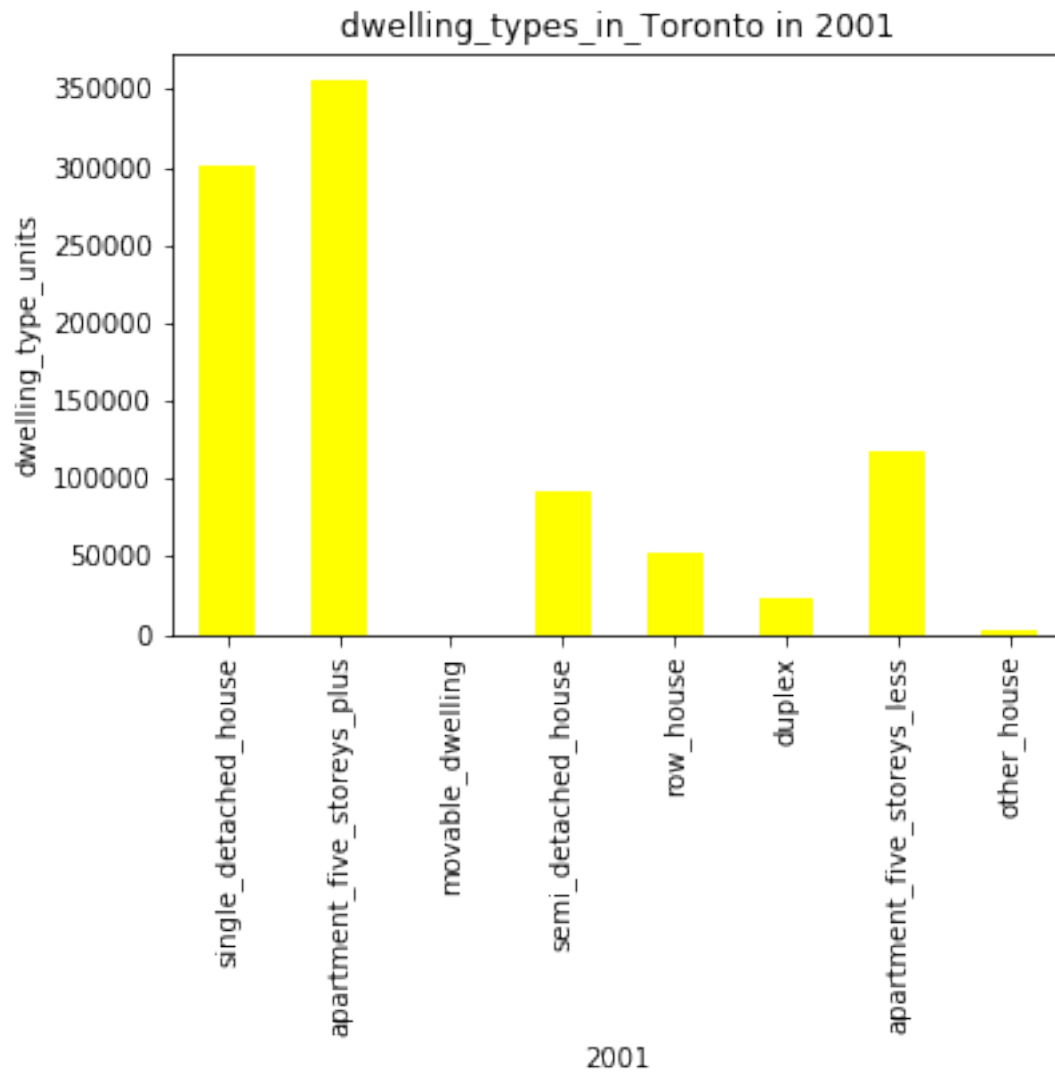
    """
    data.plot(kind="bar", xlabel=xlabel, ylabel=ylabel, color=color,
    ↪title=title)

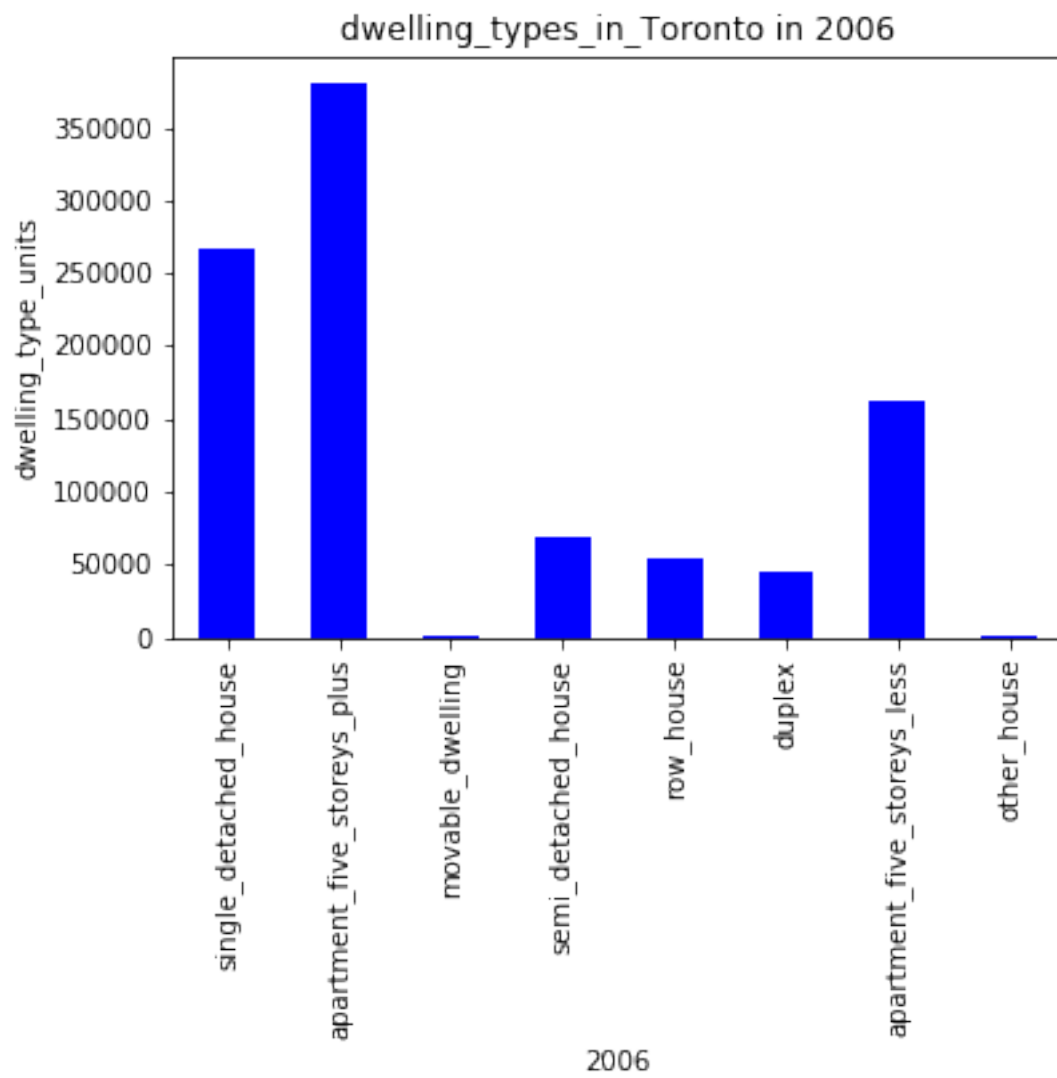
    # Show the plot after plotting it, to avoid overwrite the previously plotted
    ↪bar chart.
    plt.show()
```

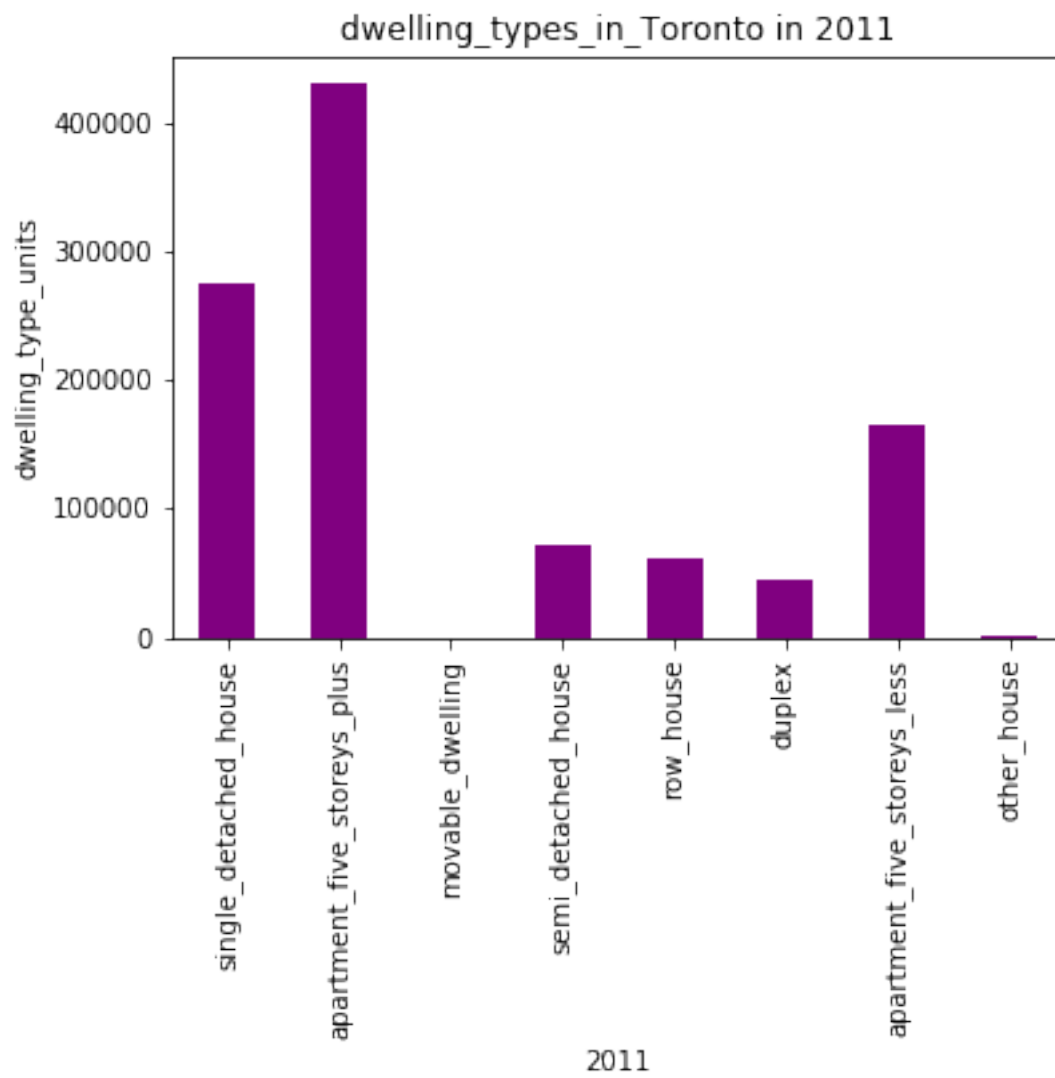
```
[31]: # Create a bar chart per year to show the number of dwelling types
colors = ["orange", "yellow", "blue", "purple", "green", "pink"]
```

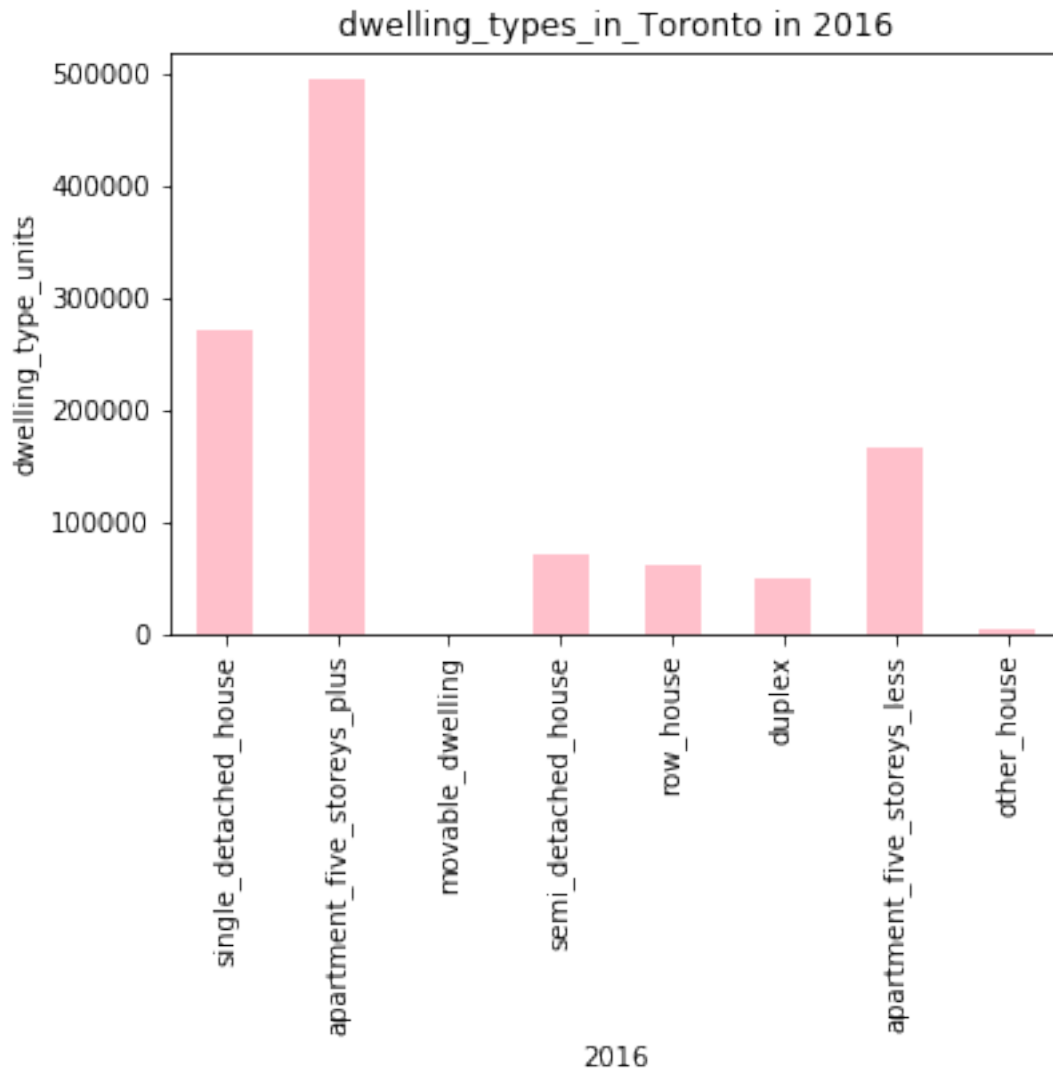
```
#For loop is used to do charts for each year using a random color from the
↳ colors list.

for index in dwelling_per_year.index:
    title = "dwelling_types_in_Toronto in " + str(index)
    ylabel = "dwelling_type_units"
    create_bar_chart(dwelling_per_year.loc[index], title=title, xlabel=index,
↳ ylabel=ylabel, color = random.choice(colors))
```









1.3 Average Monthly Shelter Costs in Toronto Per Year

In this section, you will calculate the average monthly shelter costs for owned and rented dwellings and the average house value for each year. Plot the results as a line chart.

Optional challenge: Plot each line chart in a different color.

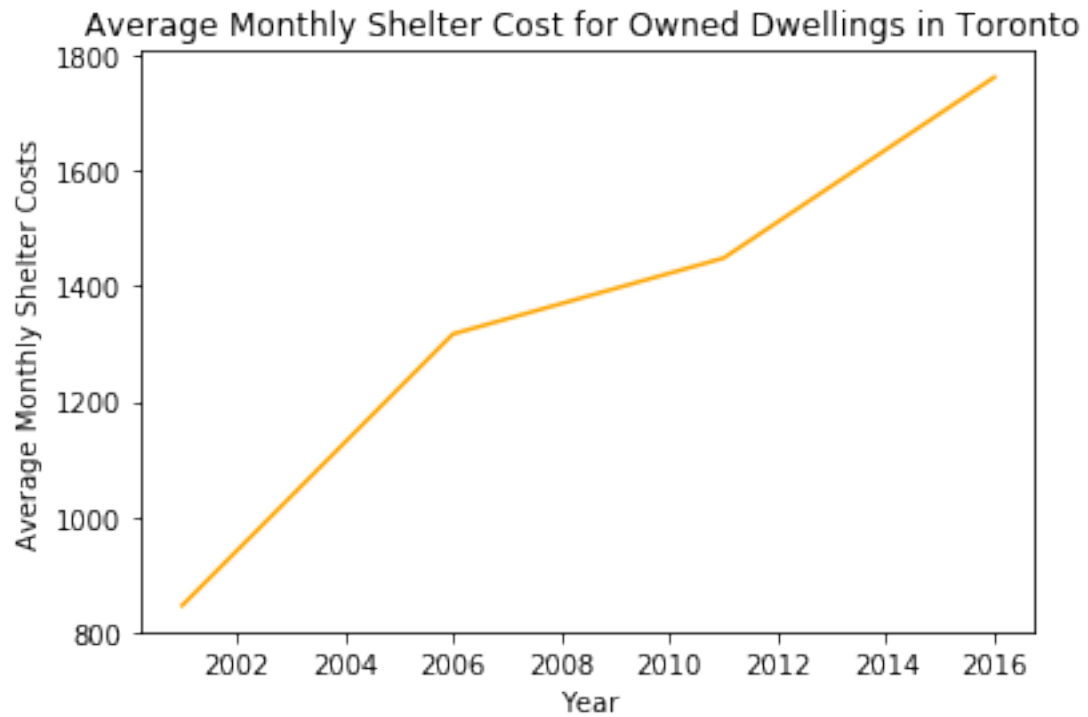
```
[9]: # Calculate the average monthly shelter costs for owned and rented dwellings
monthly_shelter_cost_per_year = □
    ↳ to_data[["shelter_costs_owned", "shelter_costs_rented"]].groupby("year").
    ↳ mean()
monthly_shelter_cost_per_year
```

```
[9]:      shelter_costs_owned  shelter_costs_rented
year
2001          846.878571          1085.935714
2006          1316.800000          925.414286
2011          1448.214286          1019.792857
2016          1761.314286          1256.321429
```

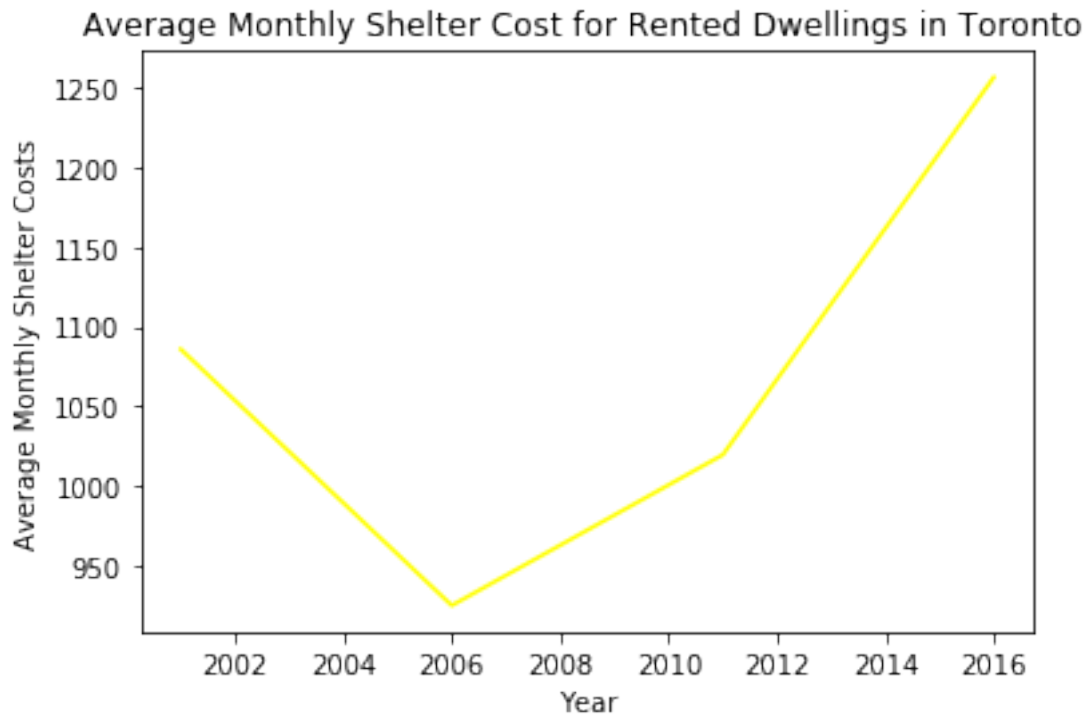
```
[10]: # Helper create_line_chart function
def create_line_chart(data, title, xlabel, ylabel, color):
    """
    Create a line chart based in the data argument.
    Input:
    data = DataFrame to use for plotting the data
    title = Chart Title
    xlabel = Label for X Axis
    ylabel = Label for Y Axis
    color = Colour of the line chart
    """
    data.plot(kind="line", xlabel=xlabel, ylabel=ylabel, color=color,
    ↪title=title)
    # Show the plot after plotting it, to avoid overwrite the previously plotted
    ↪bar chart.
    plt.show()
```

```
[11]: # Create two line charts, one to plot the monthly shelter costs for owned
    ↪dwelling and other for rented dwellings per year

# Line chart for owned dwellings
create_line_chart(data=monthly_shelter_cost_per_year["shelter_costs_owned"],
    ↪title="Average Monthly Shelter Cost for Owned Dwellings in Toronto ",
    xlabel="Year", ylabel="Average Monthly Shelter Costs",
    ↪color="orange")
```

```
[12]: # Line chart for rented dwellings
create_line_chart(data=monthly_shelter_cost_per_year["shelter_costs_rented"],
    title="Average Monthly Shelter Cost for Rented Dwellings in Toronto ",
    xlabel="Year", ylabel="Average Monthly Shelter Costs",
    color="yellow")
```



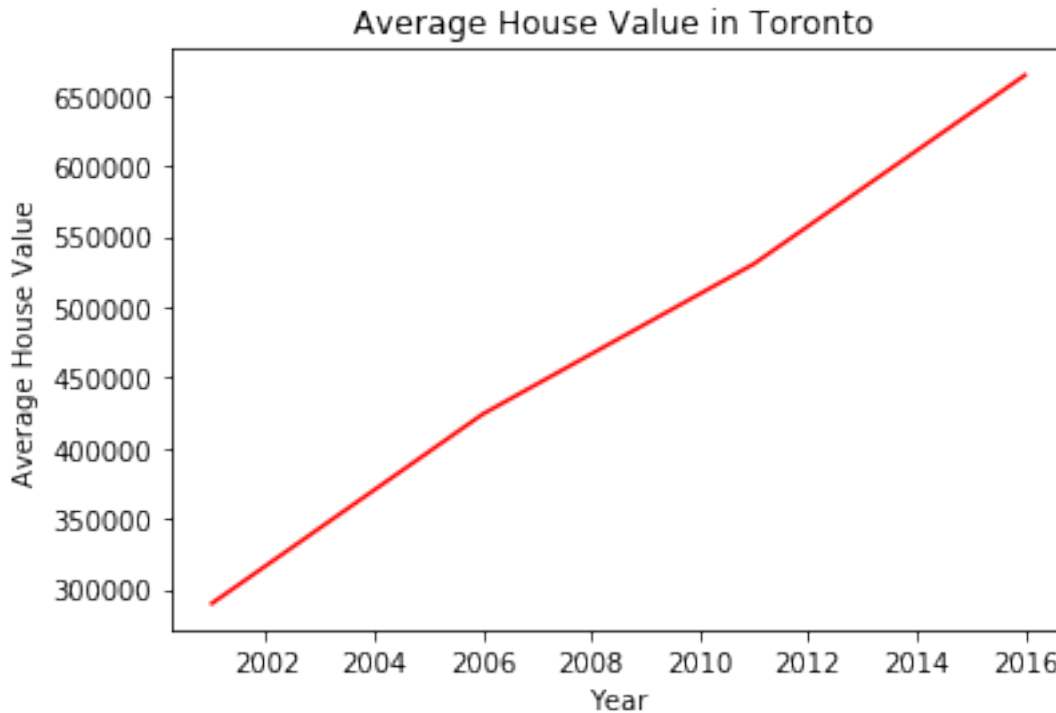
1.4 Average House Value per Year

In this section, you want to determine the average house value per year. An investor may want to understand better the sales price of the rental property over time. For example, a customer will want to know if they should expect an increase or decrease in the property value over time so they can determine how long to hold the rental property. You will visualize the `average_house_value` per year as a bar chart.

```
[13]: # Calculate the average house value per year
average_house_value = to_data["average_house_value"].groupby("year").mean()
average_house_value
```

```
[13]: year
2001    289882.885714
2006    424059.664286
2011    530424.721429
2016    664068.328571
Name: average_house_value, dtype: float64
```

```
[14]: # Plot the average house value per year as a line chart
create_line_chart(data=average_house_value, title="Average House Value in Toronto ",
                  xlabel="Year", ylabel="Average House Value", color="red")
```



1.5 Average House Value by Neighbourhood

In this section, you will use `hvplot` to create an interactive visualization of the average house value with a dropdown selector for the neighbourhood.

Hint: It will be easier to create a new DataFrame from grouping the data and calculating the mean house values for each year and neighbourhood.

```
[15]: # Create a new DataFrame with the mean house values by neighbourhood per year
avg_house_value_by_neighbourhood = □
    ↪to_data[["neighbourhood", "average_house_value"]].reset_index()
avg_house_value_by_neighbourhood.head(10)
```

```
[15]:   year      neighbourhood  average_house_value
0  2001      Agincourt North          200388
1  2001  Agincourt South-Malvern West          203047
2  2001             Alderwood          259998
3  2001              Annex          453850
4  2001  Banbury-Don Mills          371864
5  2001      Bathurst Manor          304749
6  2001  Bay Street Corridor          257404
7  2001      Bayview Village          327644
8  2001  Bayview Woods-Steeles          343535
```

```
[16]: # Use hvplot to create an interactive line chart of the average house value per
      ↪neighbourhood
      # The plot should have a dropdown selector for the neighbourhood
      avg_house_value_by_neighbourhood.hvplot.line(x="year", y="average_house_value",
      ↪xlabel="Year", ylabel="Average House Value", groupby="neighbourhood")
```

```
[16]: :DynamicMap    [neighbourhood]
      :Curve      [year]    (average_house_value)
```

1.6 Number of Dwelling Types per Year

In this section, you will use hvplot to create an interactive visualization of the average number of dwelling types per year with a dropdown selector for the neighbourhood.

```
[17]: # Fetch the data of all dwelling types per year
      dwelling_types_per_year = to_data.drop(columns=["average_house_value",
      ↪"shelter_costs_owned", "shelter_costs_rented" ])
      dwelling_types_per_year.head(10)
```

```
[17]:
```

	neighbourhood	single_detached_house	\
year			
2001	Agincourt North	3715	
2001	Agincourt South-Malvern West	3250	
2001	Alderwood	3175	
2001	Annex	1060	
2001	Banbury-Don Mills	3615	
2001	Bathurst Manor	2405	
2001	Bay Street Corridor	10	
2001	Bayview Village	2170	
2001	Bayview Woods-Steeles	1650	
2001	Bedford Park-Nortown	4985	

	apartment_five_storeys_plus	movable_dwelling	semi_detached_house	\
year				
2001	1480	0	1055	
2001	1835	0	545	
2001	315	0	470	
2001	6090	5	1980	
2001	4465	0	240	
2001	1550	0	130	
2001	7575	0	0	
2001	630	0	170	
2001	1715	0	925	
2001	2080	0	45	

	row_house	duplex	apartment_five_storeys_less	other_house
year				
2001	1295	195	185	5
2001	455	105	425	0
2001	50	185	370	0
2001	605	275	3710	165
2001	380	15	1360	0
2001	130	375	745	0
2001	15	0	240	0
2001	765	15	640	0
2001	105	10	170	5
2001	40	210	1235	15

```
[18]: # Use hvplot to create an interactive bar chart of the number of dwelling types
      ↪ per neighbourhood
      # The plot should have a dropdown selector for the neighbourhood
      dwelling_types_per_year.hvplot.bar(groupby="neighbourhood",rot=90,
      ↪ ylabel="Dwelling Type Units", xlabel="Year", height=500)
```

```
[18]: :DynamicMap    [neighbourhood]
      :Bars    [year,Variable]    (value)
```

1.7 The Top 10 Most Expensive Neighbourhoods

In this section, you will need to calculate the house value for each neighbourhood and then sort the values to obtain the top 10 most expensive neighbourhoods on average. Plot the results as a bar chart.

```
[19]: # Getting the data from the top 10 expensive neighbourhoods
      ten_most_expensive_neighbourhoods = to_data.groupby("neighbourhood").mean().
      ↪ sort_values(by="average_house_value", ascending=False).reset_index()
      ten_most_expensive_neighbourhoods.head(10)
```

```
[19]:
```

	neighbourhood	single_detached_house \
0	Bridle Path-Sunnybrook-York Mills	2260.00
1	Forest Hill South	1742.50
2	Lawrence Park South	3472.50
3	Rosedale-Moore Park	2498.75
4	St.Andrew-Windfields	3225.00
5	Casa Loma	916.25
6	Bedford Park-Nortown	4865.00
7	Forest Hill North	1488.75
8	Kingsway South	2326.25
9	Yonge-St.Clair	565.00

	apartment_five_storeys_plus	movable_dwelling	semi_detached_house \
--	-----------------------------	------------------	-----------------------

0	331.25	0.00	36.25
1	2031.25	1.25	61.25
2	773.75	0.00	126.25
3	4641.25	0.00	486.25
4	1670.00	0.00	185.00
5	2310.00	0.00	288.75
6	1981.25	0.00	43.75
7	3392.50	0.00	12.50
8	576.25	0.00	66.25
9	3948.75	0.00	425.00

	row_house	duplex	apartment_five_storeys_less	other_house \
0	90.00	25.0	40.00	0.00
1	45.00	75.0	1027.50	3.75
2	38.75	225.0	966.25	16.25
3	245.00	327.5	1618.75	2.50
4	552.50	97.5	586.25	5.00
5	201.25	162.5	1192.50	2.50
6	57.50	287.5	1275.00	88.75
7	16.25	82.5	402.50	1.25
8	48.75	20.0	336.25	2.50
9	212.50	172.5	1308.75	6.25

	average_house_value	shelter_costs_owned	shelter_costs_rented
0	1526485.75	2360.75	2321.75
1	1195992.50	1781.00	1313.75
2	1094027.75	1954.00	1372.75
3	1093640.00	1909.75	1537.25
4	999107.00	1880.25	1384.50
5	981064.25	1873.75	1547.75
6	930415.25	1786.75	1255.00
7	851680.50	1722.75	1245.50
8	843234.25	1736.75	1622.00
9	813220.25	1680.75	1369.00

```
[20]: # Plotting the data from the top 10 expensive neighbourhoods
ten_most_expensive_neighbourhoods.head(10).hvplot.bar(rot=90, ylabel="Average_
↪House Value", xlabel="Neighbourhood", y="average_house_value",
↪x="neighbourhood",height=500)
```

```
[20]: :Bars [neighbourhood] (average_house_value)
```

1.8 Neighbourhood Map

In this section, you will read in neighbourhoods location data and build an interactive map with the average house value per neighbourhood. Use a `scatter_mapbox` from Plotly express to create

the visualization. Remember, you will need your Mapbox API key for this.

1.8.1 Load Location Data

```
[21]: # Load neighbourhoods coordinates data
file_path = Path("Data/toronto_neighbourhoods_coordinates.csv")
df_neighbourhood_locations = pd.read_csv(file_path)
df_neighbourhood_locations.head()
```

```
[21]:
```

	neighbourhood	lat	lon
0	Agincourt North	43.805441	-79.266712
1	Agincourt South-Malvern West	43.788658	-79.265612
2	Alderwood	43.604937	-79.541611
3	Annex	43.671585	-79.404001
4	Banbury-Don Mills	43.737657	-79.349718

1.8.2 Data Preparation

You will need to join the location data with the mean values per neighbourhood.

1. Calculate the mean values for each neighbourhood.
2. Join the average values with the neighbourhood locations.

```
[22]: # Calculate the mean values for each neighborhood
mean_data_neighbourhoods = to_data.groupby("neighbourhood").mean().reset_index()
mean_data_neighbourhoods.head()
```

```
[22]:
```

	neighbourhood	single_detached_house	\
0	Agincourt North	3435.00	
1	Agincourt South-Malvern West	2897.50	
2	Alderwood	2903.75	
3	Annex	751.25	
4	Banbury-Don Mills	3572.50	

	apartment_five_storeys_plus	movable_dwelling	semi_detached_house	\
0	1947.50	2.50	863.75	
1	2180.00	1.25	375.00	
2	302.50	1.25	503.75	
3	7235.00	1.25	1375.00	
4	5388.75	1.25	273.75	

	row_house	duplex	apartment_five_storeys_less	other_house	\
0	1406.25	512.50	547.50	10.00	
1	456.25	523.75	628.75	32.50	
2	76.25	302.50	502.50	1.25	
3	613.75	355.00	4605.00	83.75	
4	626.25	32.50	1340.00	0.00	

	average_house_value	shelter_costs_owned	shelter_costs_rented
0	329811.5	1109.00	983.50
1	334189.0	1131.25	985.00
2	427922.5	1166.75	1003.25
3	746977.0	1692.75	1315.25
4	612039.0	1463.50	1242.75

```
[23]: # Join the average values with the neighbourhood locations
mean_data_neighbourhoods.set_index(keys="neighbourhood",inplace=True)
df_neighbourhood_locations.set_index(keys="neighbourhood",inplace=True)
neighbourhood_with_location = pd.concat([mean_data_neighbourhoods,
↳df_neighbourhood_locations ], join="inner", axis="columns")
neighbourhood_with_location
```

```
[23]: single_detached_house \
```

neighbourhood	
Agincourt North	3435.00
Agincourt South-Malvern West	2897.50
Alderwood	2903.75
Annex	751.25
Banbury-Don Mills	3572.50
...	...
Wychwood	1056.25
Yonge-Eglinton	1468.75
Yonge-St.Clair	565.00
York University Heights	1355.00
Yorkdale-Glen Park	2286.25

```
apartment_five_storeys_plus movable_dwelling \
```

neighbourhood		
Agincourt North	1947.50	2.50
Agincourt South-Malvern West	2180.00	1.25
Alderwood	302.50	1.25
Annex	7235.00	1.25
Banbury-Don Mills	5388.75	1.25
...
Wychwood	1236.25	0.00
Yonge-Eglinton	1638.75	0.00
Yonge-St.Clair	3948.75	0.00
York University Heights	5165.00	1.25
Yorkdale-Glen Park	1347.50	0.00

```
semi_detached_house row_house duplex \
```

neighbourhood			
Agincourt North	863.75	1406.25	512.50
Agincourt South-Malvern West	375.00	456.25	523.75
Alderwood	503.75	76.25	302.50

Annex	1375.00	613.75	355.00
Banbury-Don Mills	273.75	626.25	32.50
...
Wychwood	992.50	298.75	325.00
Yonge-Eglinton	470.00	33.75	328.75
Yonge-St.Clair	425.00	212.50	172.50
York University Heights	1316.25	662.50	188.75
Yorkdale-Glen Park	73.75	450.00	377.50

	apartment_five_storeys_less	other_house	\
neighbourhood			
Agincourt North	547.50	10.00	
Agincourt South-Malvern West	628.75	32.50	
Alderwood	502.50	1.25	
Annex	4605.00	83.75	
Banbury-Don Mills	1340.00	0.00	
...	
Wychwood	1878.75	17.50	
Yonge-Eglinton	1385.00	6.25	
Yonge-St.Clair	1308.75	6.25	
York University Heights	1085.00	33.75	
Yorkdale-Glen Park	722.50	7.50	

	average_house_value	shelter_costs_owned	\
neighbourhood			
Agincourt North	329811.50	1109.00	
Agincourt South-Malvern West	334189.00	1131.25	
Alderwood	427922.50	1166.75	
Annex	746977.00	1692.75	
Banbury-Don Mills	612039.00	1463.50	
...	
Wychwood	565976.50	1390.75	
Yonge-Eglinton	809745.75	1799.50	
Yonge-St.Clair	813220.25	1680.75	
York University Heights	305899.50	1116.75	
Yorkdale-Glen Park	430861.25	1122.50	

	shelter_costs_rented	lat	lon
neighbourhood			
Agincourt North	983.50	43.805441	-79.266712
Agincourt South-Malvern West	985.00	43.788658	-79.265612
Alderwood	1003.25	43.604937	-79.541611
Annex	1315.25	43.671585	-79.404001
Banbury-Don Mills	1242.75	43.737657	-79.349718
...
Wychwood	1017.25	43.676919	-79.425515
Yonge-Eglinton	1347.75	43.704689	-79.403590

Yonge-St.Clair	1369.00	43.687859	-79.397871
York University Heights	937.50	43.765736	-79.488883
Yorkdale-Glen Park	942.50	43.714672	-79.457108

[140 rows x 13 columns]

[]:

1.8.3 Mapbox Visualization

Plot the average values per neighbourhood using a Plotly express `scatter_mapbox` visualization.

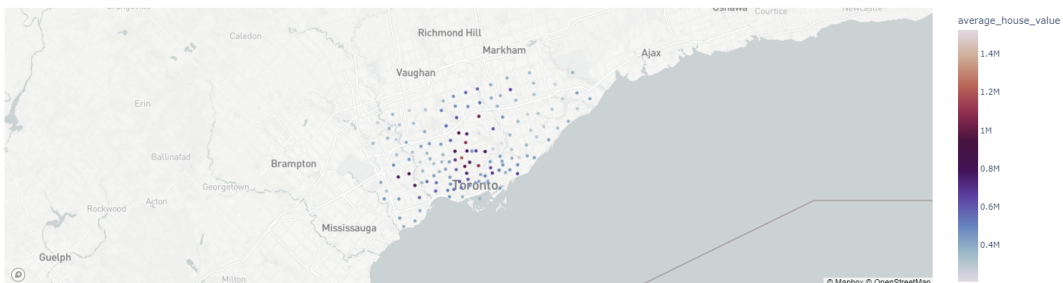
```
[24]: # Create a scatter mapbox to analyze neighbourhood info
average_house_value_map= px.scatter_mapbox(

    neighbourhood_with_location,
    lat="lat",
    lon="lon",
    color="average_house_value",
    color_continuous_scale=px.colors.cyclical.Twilight,
    title="Average House Value in Toronto",
    zoom=9,
    width=700,
    height=600,

)

average_house_value_map.show()
```

Average House Value in Toronto



1.9 Cost Analysis - Optional Challenge

In this section, you will use Plotly express to a couple of plots that investors can interactively filter and explore various factors related to the house value of the Toronto's neighbourhoods.

1.9.1 Create a bar chart row facet to plot the average house values for all Toronto's neighbourhoods per year

```
[25]: mean_data_neighbourhoods.reset_index(inplace=True)
      to_data
```

```
[25]:
```

	neighbourhood	single_detached_house \
year		
2001	Agincourt North	3715
2001	Agincourt South-Malvern West	3250
2001	Alderwood	3175
2001	Annex	1060
2001	Banbury-Don Mills	3615
...
2016	Wychwood	920
2016	Yonge-Eglinton	1400
2016	Yonge-St.Clair	520
2016	York University Heights	1235
2016	Yorkdale-Glen Park	2165

	apartment_five_storeys_plus	movable_dwelling	semi_detached_house \
year			
2001	1480	0	1055
2001	1835	0	545
2001	315	0	470
2001	6090	5	1980
2001	4465	0	240
...
2016	1295	0	880
2016	1995	0	465
2016	4315	0	450
2016	5505	0	1360
2016	1185	0	80

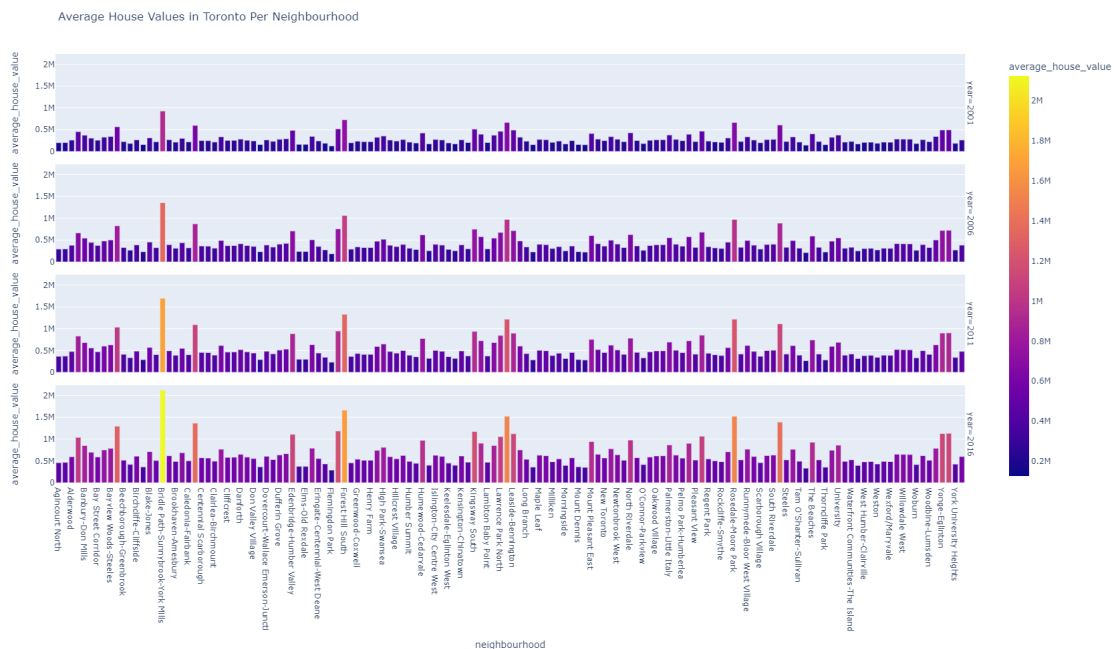
	row_house	duplex	apartment_five_storeys_less	other_house \
year				
2001	1295	195	185	5
2001	455	105	425	0
2001	50	185	370	0
2001	605	275	3710	165
2001	380	15	1360	0
...
2016	290	395	2080	35

2016	60	310	1445	0
2016	220	130	1370	0
2016	775	280	995	0
2016	600	465	830	5

	average_house_value	shelter_costs_owned	shelter_costs_rented
year			
2001	200388	810	870
2001	203047	806	892
2001	259998	817	924
2001	453850	1027	1378
2001	371864	1007	1163
...
2016	787760	1864	1146
2016	1127052	2398	1535
2016	1131888	2192	1619
2016	425769	1444	1122
2016	599698	1451	1128

[560 rows x 12 columns]

```
[26]: px.bar(to_data, x="neighbourhood",
            y="average_house_value", color="average_house_value", facet_row=to_data.
            index,height=1000, title="Average House Values in Toronto Per Neighbourhood")
```



1.9.2 Create a sunburst chart to conduct a costs analysis of most expensive neighbourhoods in Toronto per year

```
[27]: # Fetch the data from all expensive neighbourhoods per year.
ten_most_expensive_neighbourhoods
```

```
[27]:
```

	neighbourhood	single_detached_house \
0	Bridle Path-Sunnybrook-York Mills	2260.00
1	Forest Hill South	1742.50
2	Lawrence Park South	3472.50
3	Rosedale-Moore Park	2498.75
4	St.Andrew-Windfields	3225.00
..
135	Malvern	3693.75
136	Thorncliffe Park	3.75
137	Mount Olive-Silverstone-Jamestown	1730.00
138	Taylor-Massey	622.50
139	Flemingdon Park	5.00

	apartment_five_storeys_plus	movable_dwelling	semi_detached_house \
0	331.25	0.00	36.25
1	2031.25	1.25	61.25
2	773.75	0.00	126.25
3	4641.25	0.00	486.25
4	1670.00	0.00	185.00
..
135	2982.50	0.00	1317.50
136	6318.75	0.00	0.00
137	5916.25	1.25	176.25
138	4795.00	0.00	203.75
139	6368.75	0.00	182.50

	row_house	duplex	apartment_five_storeys_less	other_house \
0	90.00	25.00	40.00	0.00
1	45.00	75.00	1027.50	3.75
2	38.75	225.00	966.25	16.25
3	245.00	327.50	1618.75	2.50
4	552.50	97.50	586.25	5.00
..
135	3410.00	878.75	552.50	55.00
136	141.25	0.00	187.50	5.00
137	1033.75	603.75	127.50	1.25
138	216.25	150.00	218.75	15.00
139	630.00	3.75	591.25	0.00

	average_house_value	shelter_costs_owned	shelter_costs_rented
0	1526485.75	2360.75	2321.75

1	1195992.50	1781.00	1313.75
2	1094027.75	1954.00	1372.75
3	1093640.00	1909.75	1537.25
4	999107.00	1880.25	1384.50
..
135	256880.25	1228.00	998.75
136	255264.75	1304.25	1088.50
137	253182.25	1159.50	955.25
138	239414.00	1129.25	919.25
139	207799.00	1199.75	973.00

[140 rows x 12 columns]

```
[28]: # Create the sunburst chart
px.sunburst(
    ten_most_expensive_neighbourhoods,
    path=[ten_most_expensive_neighbourhoods.index, 'neighbourhood'],
    values='average_house_value', color='average_house_value',
    color_continuous_scale='blues',
    height=500
)
```



[]: