San Francisco Housing Cost Analysis

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

```python
In [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 numpy as np
        import os
        from pathlib import Path
        from dotenv import load_dotenv

        import warnings
        warnings.filterwarnings('ignore')
```

```python
In [2]: # Read the Mapbox API key
        load_dotenv()
        map_box_api = os.getenv("mapbox")
```

Load Data

```python
In [3]: # Read the census data into a Pandas DataFrame
        file_path = Path("Data/sfo_neighborhoods_census_data.csv")
        sfo_data = pd.read_csv(file_path, index_col="year")
        sfo_data.head()
```

Out[3]:

year	neighborhood	sale_price_sqr_foot	housing_units	gross_rent
2010	Alamo Square	291.182945	372560	1239
2010	Anza Vista	267.932583	372560	1239
2010	Bayview	170.098665	372560	1239
2010	Buena Vista Park	347.394919	372560	1239
2010	Central Richmond	319.027623	372560	1239

Housing Units Per Year

In this section, you will calculate the number of housing units per year and visualize the results as a bar chart using the Pandas plot function.

Hint: Use the Pandas `groupby` function.

Optional challenge: Use the min, max, and std to scale the y limits of the chart.
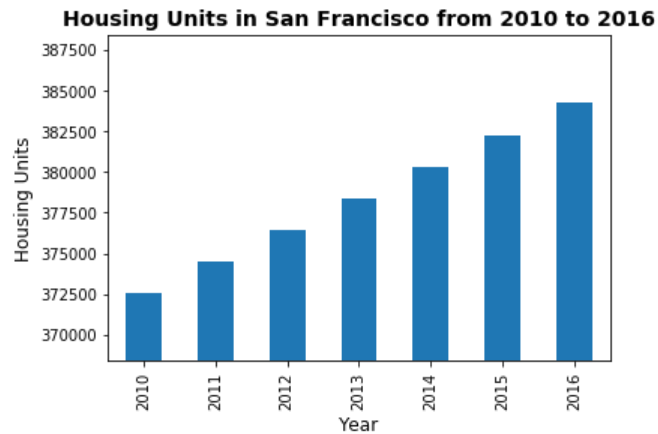
```python
In [4]: # Calculate the mean number of housing units per year (hint: use groupby)
        # YOUR CODE HERE!
```

```
Out[4]: year
        2010    372560
        2011    374507
        2012    376454
        2013    378401
        2014    380348
        2015    382295
        2016    384242
        Name: housing_units, dtype: int64
```

Housing Units in San Francisco from 2010 to 2016

Average Housing Costs in San Francisco Per Year

In this section, you will calculate the average monthly rent and the average price per square foot for each year. An investor may wish to better understand 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. Plot the results as two line charts.

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

Out[7]:

year	sale_price_sqr_foot	gross_rent
2010	369.344353	1239
2011	341.903429	1530
2012	399.389968	2324
2013	483.600304	2971
2014	556.277273	3528
2015	632.540352	3739
2016	697.643709	4390

```
In [8]:  # Create two line charts, one to plot the average sale price per square foot and another for average montly rent

         # Line chart for average sale price per square foot
         # YOUR CODE HERE!

         # Line chart for average montly rent
         # YOUR CODE HERE!
```

Average Price per SqFt by Year

Average Gross Rent by Year
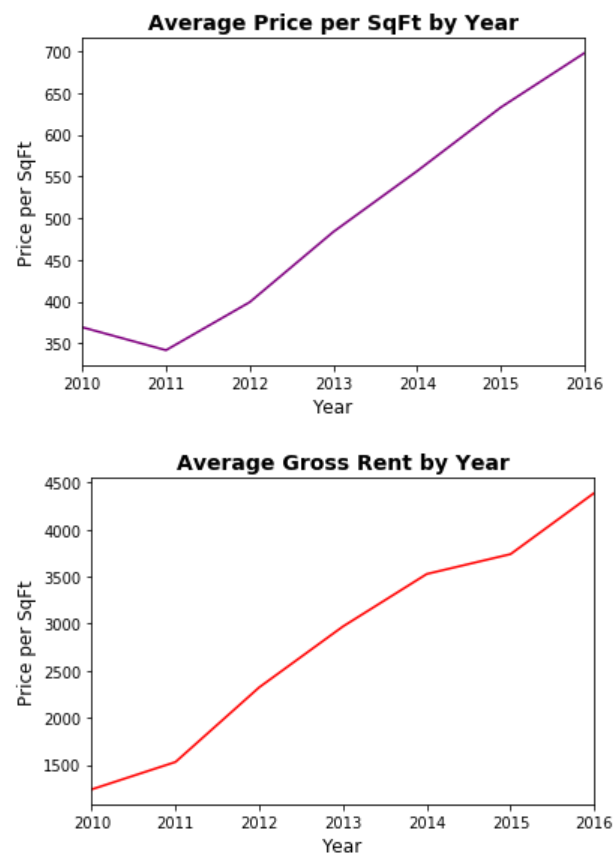
Average Prices by Neighborhood

In this section, you will use hvplot to create two interactive visulizations of average prices with a dropdown selector for the neighborhood. The first visualization will be a line plot showing the trend of average price per square foot over time for each neighborhood. The second will be a line plot showing the trend of average montly rent over time for each neighborhood.

Hint: It will be easier to create a new DataFrame from grouping the data and calculating the mean prices for each year and neighborhood

```
In [9]:  # Group by year and neighborhood and then create a new dataframe of the mean values
         # YOUR CODE HERE!
```

Out[9]:

|	year	neighborhood	sale_price_sqr_foot	housing_units	gross_rent
0	2010	Alamo Square	291.182945	372560	1239
1	2010	Anza Vista	267.932583	372560	1239
2	2010	Bayview	170.098665	372560	1239
3	2010	Buena Vista Park	347.394919	372560	1239
4	2010	Central Richmond	319.027623	372560	1239
5	2010	Central Sunset	418.172493	372560	1239
6	2010	Corona Heights	369.359338	372560	1239
7	2010	Cow Hollow	569.379968	372560	1239
8	2010	Croker Amazon	165.645730	372560	1239
9	2010	Diamond Heights	456.930822	372560	1239

```
In [10]:  # Use hvplot to create an interactive line chart of the average price per sq ft.
          # The plot should have a dropdown selector for the neighborhood
          # YOUR CODE HERE!
```

Out[10]:

neighborhood: Alamo Square
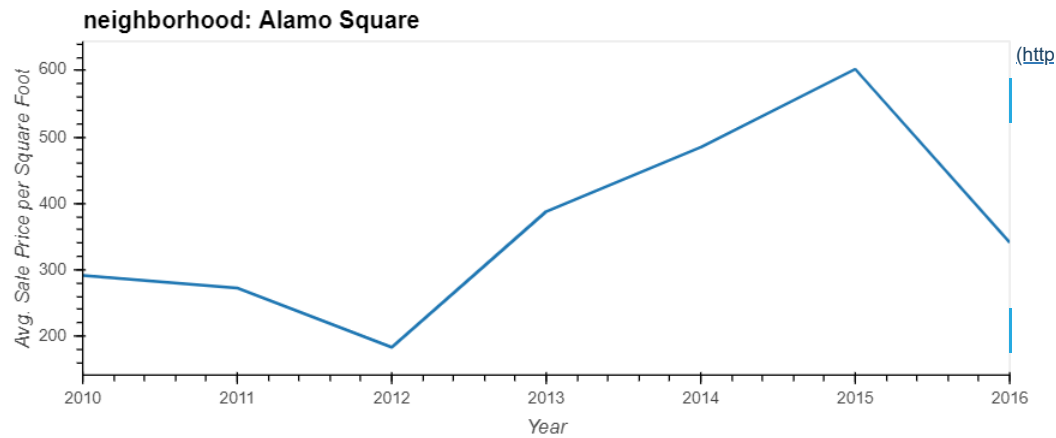

```
In [11]:  # Use hvplot to create an interactive line chart of the average monthly rent.
          # The plot should have a dropdown selector for the neighborhood
          # YOUR CODE HERE!
```

Out[11]:

neighborhood: Alamo Square
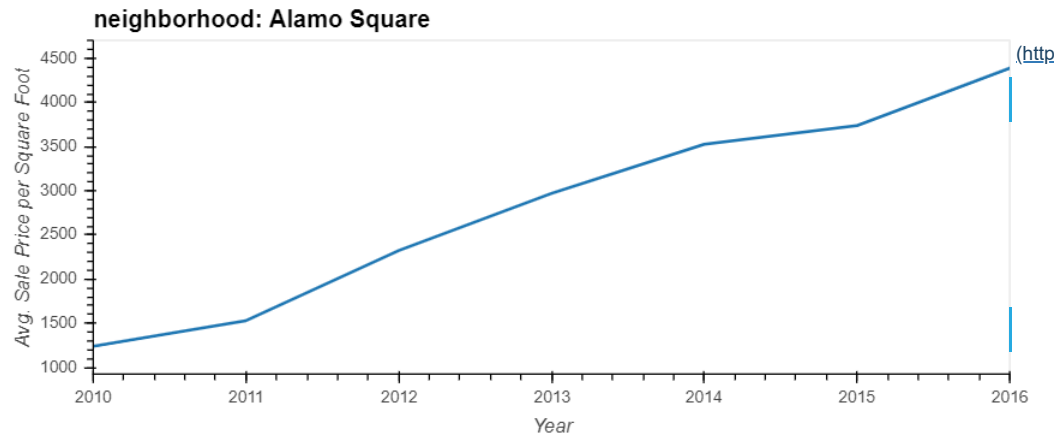
The Top 10 Most Expensive Neighborhoods

In this section, you will need to calculate the mean sale price per square foot for each neighborhood and then sort the values to obtain the top 10 most expensive neighborhoods on average. Plot the results as a bar chart.

```
In [12]:  # Getting the data from the top 10 expensive neighborhoods to own
          # YOUR CODE HERE!
```

Out[12]:

	neighborhood	sale_price_sqr_foot	housing_units	gross_rent
0	Union Square District	903.993258	377427.50	2555.166667
1	Merced Heights	788.844818	380348.00	3414.000000
2	Miraloma Park	779.810842	375967.25	2155.250000
3	Pacific Heights	689.555817	378401.00	2817.285714
4	Westwood Park	687.087575	382295.00	3959.000000
5	Telegraph Hill	676.506578	378401.00	2817.285714
6	Presidio Heights	675.350212	378401.00	2817.285714
7	Cow Hollow	665.964042	378401.00	2817.285714
8	Potrero Hill	662.013613	378401.00	2817.285714
9	South Beach	650.124479	375805.00	2099.000000

```
In [13]:  # Plotting the data from the top 10 expensive neighborhoods
          # YOUR CODE HERE!
```

Out[13]:

Top 10 Expensive Neighborhoods in SFO
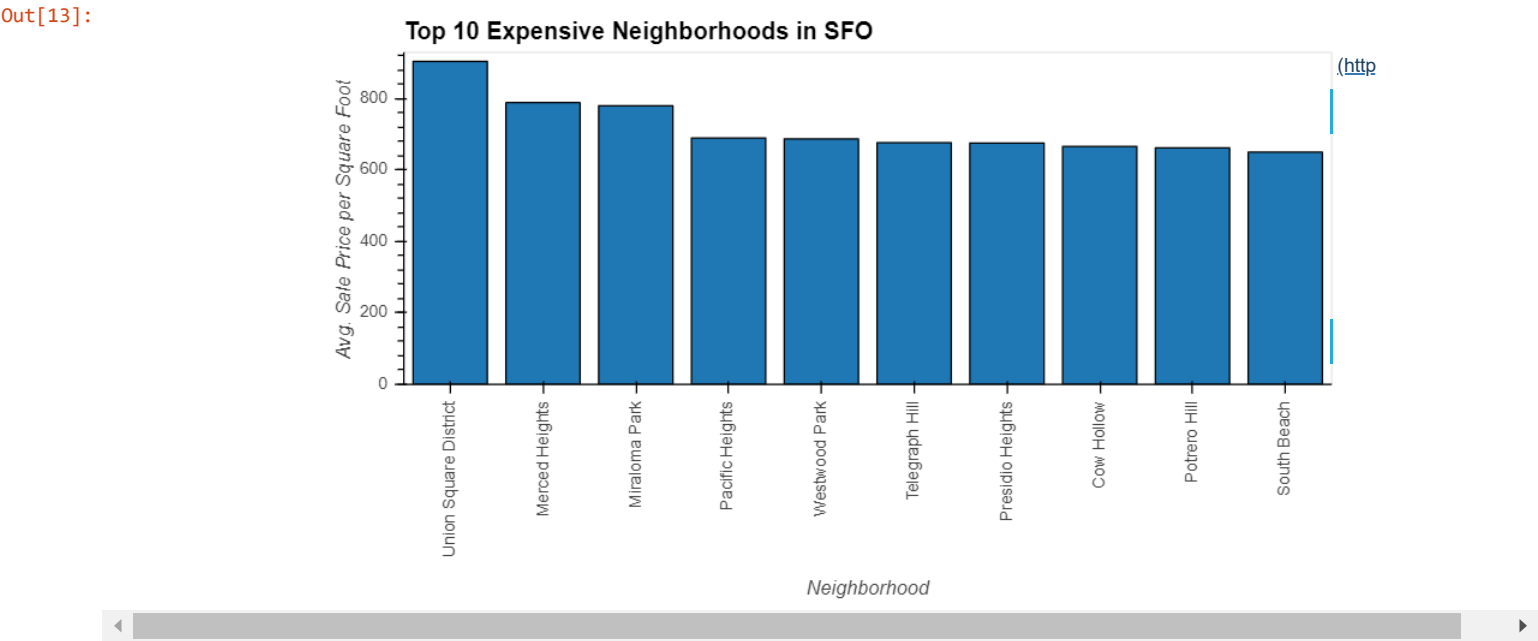
Comparing cost to purchase versus rental income

In this section, you will use `hvplot` to create an interactive visualization with a dropdown selector for the neighborhood. This visualization will feature a side-by-side comparison of average price per square foot versus average montly rent by year.

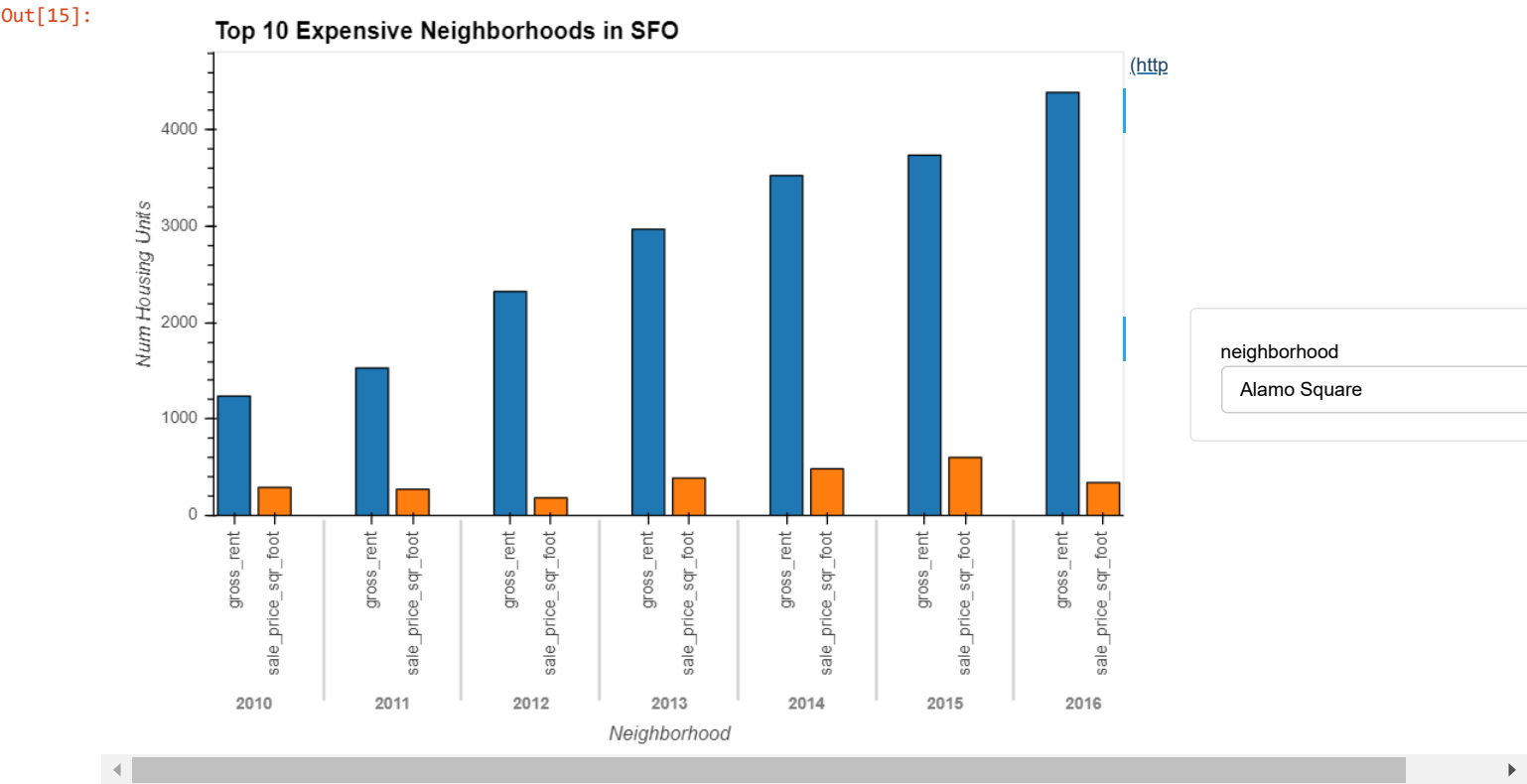**Hint:** Use the `hvplot` parameter, `groupby`, to create a dropdown selector for the neighborhood.

```
In [14]:  # Fetch the previously generated DataFrame that was grouped by year and neighborhood
          # YOUR CODE HERE!
```

Out[14]:

	year	neighborhood	sale_price_sqr_foot	housing_units	gross_rent
0	2010	Alamo Square	291.182945	372560	1239
1	2010	Anza Vista	267.932583	372560	1239
2	2010	Bayview	170.098665	372560	1239
3	2010	Buena Vista Park	347.394919	372560	1239
4	2010	Central Richmond	319.027623	372560	1239
5	2010	Central Sunset	418.172493	372560	1239
6	2010	Corona Heights	369.359338	372560	1239
7	2010	Cow Hollow	569.379968	372560	1239
8	2010	Croker Amazon	165.645730	372560	1239
9	2010	Diamond Heights	456.930822	372560	1239

```
# Plotting the data from the top 10 expensive neighborhoods
# YOUR CODE HERE!
```

Out[15]:

Neighborhood Map

In this section, you will read in neighborhoods location data and build an interactive map with the average house value per neighborhood. Use a `scatter_mapbox` from Plotly express to create the visualization. Remember, you will need your Mapbox API key for this.

Load Location Data

In [16]:
```
# Load neighborhoods coordinates data
# YOUR CODE HERE!
```

Out[16]:

	Neighborhood	Lat	Lon
0	Alamo Square	37.791012	-122.402100
1	Anza Vista	37.779598	-122.443451
2	Bayview	37.734670	-122.401060
3	Bayview Heights	37.728740	-122.410980
4	Bernal Heights	37.728630	-122.443050

Data Preparation

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

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

```
In [17]:  # Calculate the mean values for each neighborhood
          # YOUR CODE HERE!
```

Out[17]:

	Neighborhood	sale_price_sqr_foot	housing_units	gross_rent
0	Alamo Square	366.020712	378401.0	2817.285714
1	Anza Vista	373.382198	379050.0	3031.833333
2	Bayview	204.588623	376454.0	2318.400000
3	Bayview Heights	590.792839	382295.0	3739.000000
4	Bernal Heights	576.746488	379374.5	3080.333333

```
In [18]:  # Join the average values with the neighborhood locations
          # YOUR CODE HERE!
```

Out[18]:

|	Neighborhood	Lat	Lon	sale_price_sqr_foot	housing_units	gross_rent
0	Alamo Square	37.791012	-122.402100	366.020712	378401.0	2817.285714
1	Anza Vista	37.779598	-122.443451	373.382198	379050.0	3031.833333
2	Bayview	37.734670	-122.401060	204.588623	376454.0	2318.400000
3	Bayview Heights	37.728740	-122.410980	590.792839	382295.0	3739.000000
4	Buena Vista Park	37.768160	-122.439330	452.680591	378076.5	2698.833333

Mapbox Visualization

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

```
In [19]:  # Set the mapbox access token
          # YOUR CODE HERE!

          # Create a scatter mapbox to analyze neighborhood info
          # YOUR CODE HERE!
```
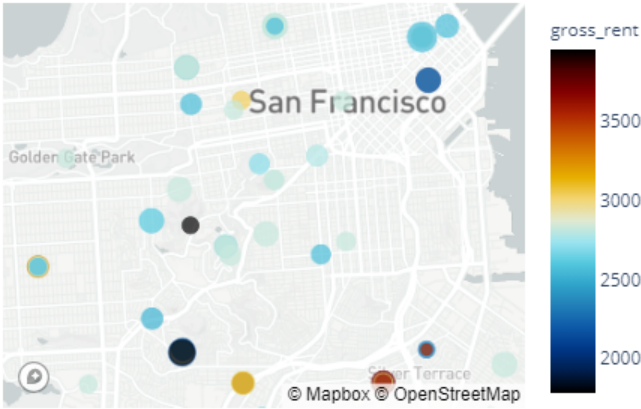

Cost Analysis - Optional Challenge

In this section, you will use Plotly express to create visualizations that investors can use to interactively filter and explore various factors related to the house value of the San Francisco's neighborhoods.

Create a DataFrame showing the most expensive neighborhoods in San Francisco by year

```
# Fetch the data from all expensive neighborhoods per year.
df_expensive_neighborhoods_per_year = df_costs[df_costs["neighborhood"].isin(df_expensive_neighborhoods["neighborhood"])]
df_expensive_neighborhoods_per_year.head()
```

Out[20]:

	year	neighborhood	sale_price_sqr_foot	housing_units	gross_rent
7	2010	Cow Hollow	569.379968	372560	1239
31	2010	Miraloma Park	680.608729	372560	1239
41	2010	Pacific Heights	496.516014	372560	1239
46	2010	Potrero Hill	491.450004	372560	1239
47	2010	Presidio Heights	549.417931	372560	1239

Create a parallel coordinates plot and parallel categories plot of most expensive neighborhoods in San Francisco per year
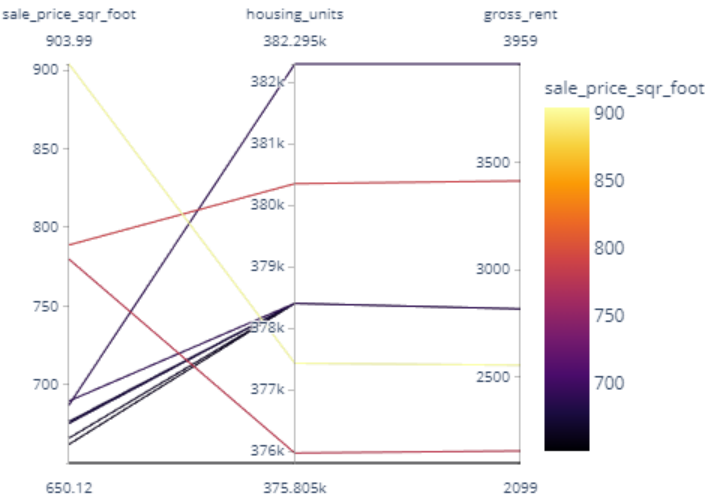
In [21]:

```
# Parallel Categories Plot
# YOUR CODE HERE!
```


Parallel Categories Analysis of Most Expensive San Francisco Ne
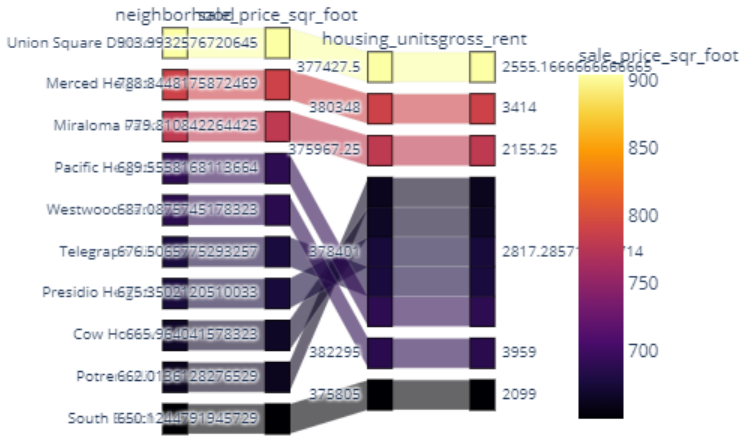
In [22]:

```
# Parallel Coordinates Plot
# YOUR CODE HERE!
```
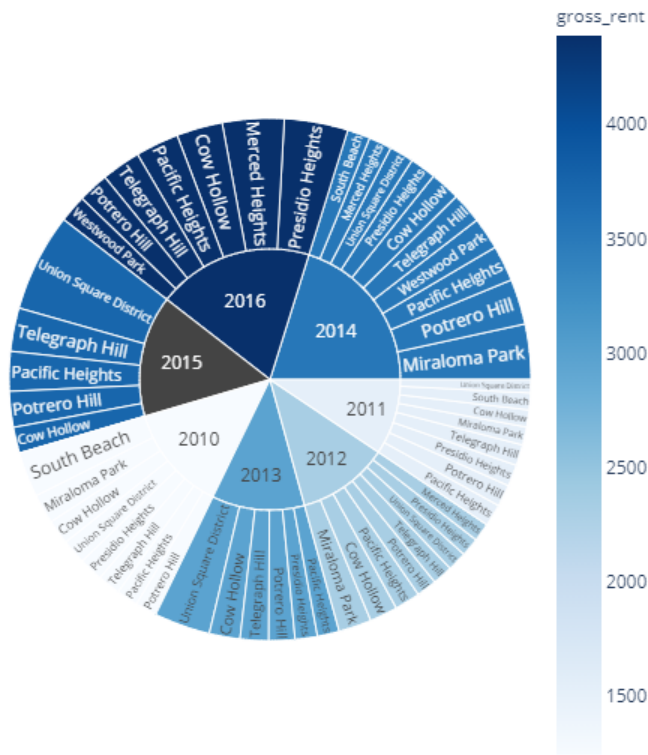

Parallel Coordinates Analysis of Most Expensive San Francisco N

Create a sunburst chart to conduct a costs analysis of most expensive neighborhoods in San Francisco per year

```
# Sunburst Plot
# YOUR CODE HERE!
```


Costs Analysis of Most Expensive neighborhoods in San Francis

In [ ]: