DREXEL UNIVERSITY | DSCI591

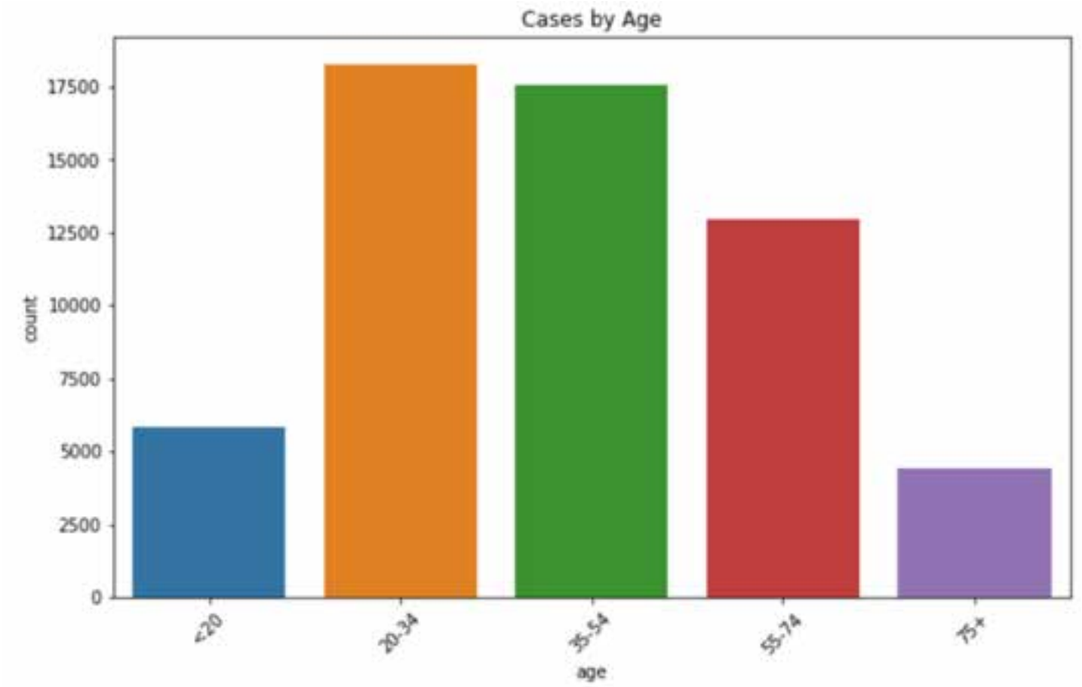
Real Estate Trends & Investigating Relationships With Covid-19

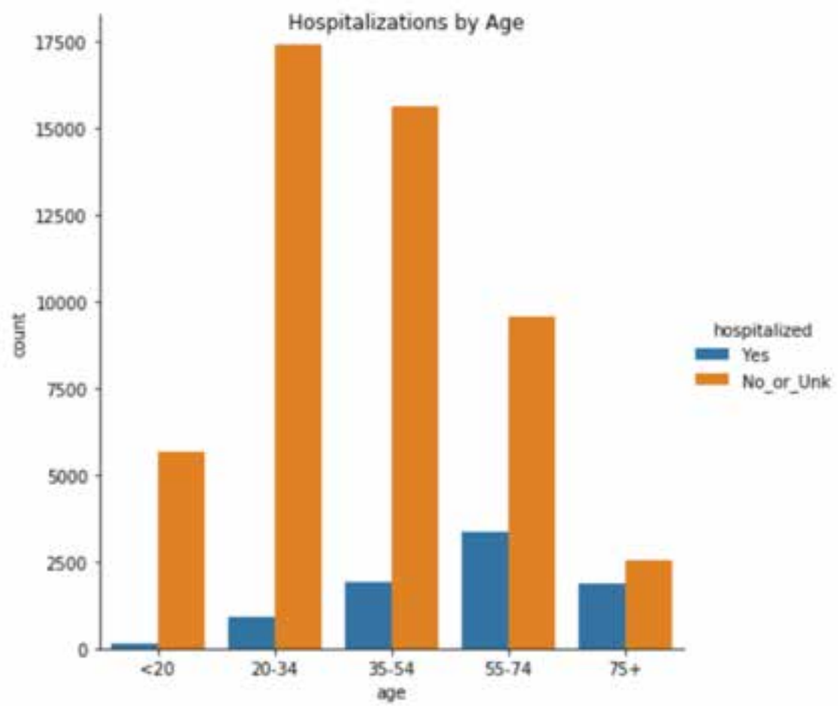
Yan Li, Frank Zhao, Lawrence Love, Gustavo Ferreira

OUTLINE

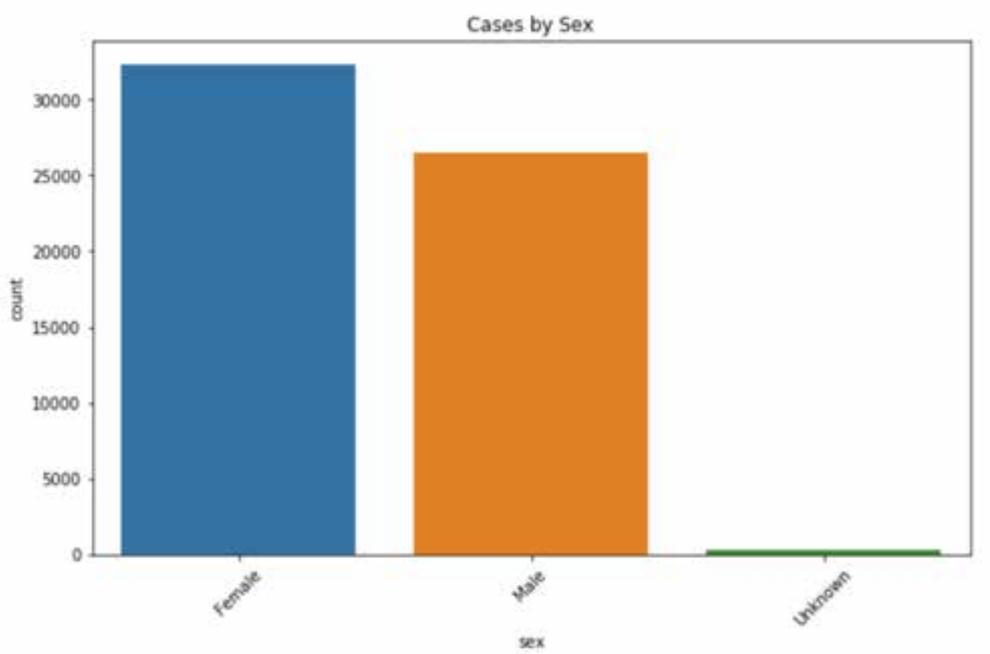
- Covid-19 Cases/Deaths Analysis
- Sold Properties Analysis
- Property for Sale/Rent Analysis
- Statistical Analysis
- Findings

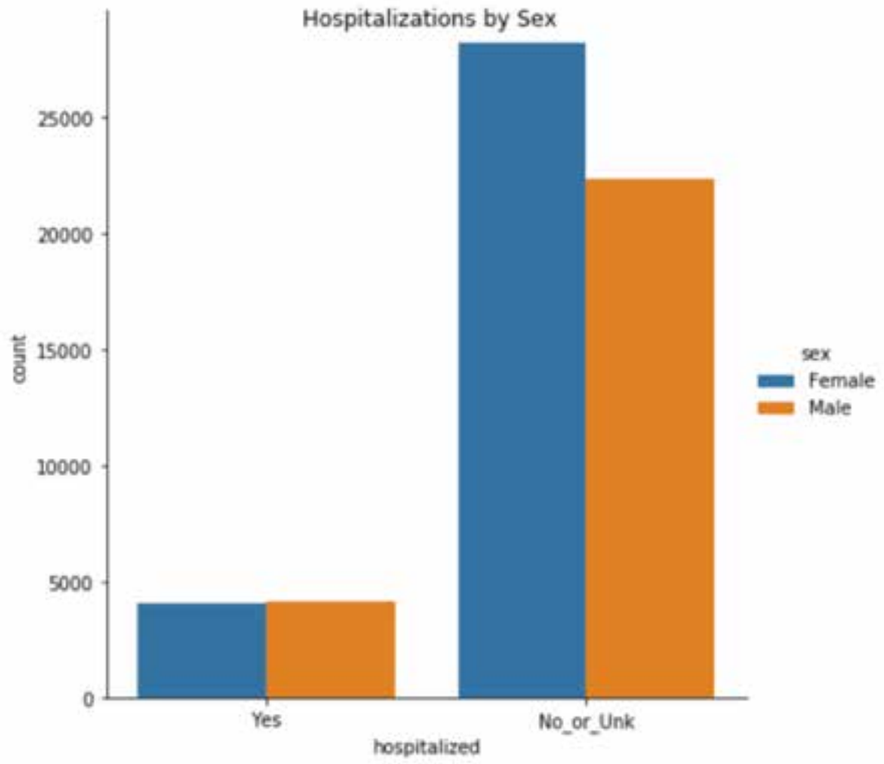
COVID-19 | CASES & HOSPITALIZATIONS - AGE



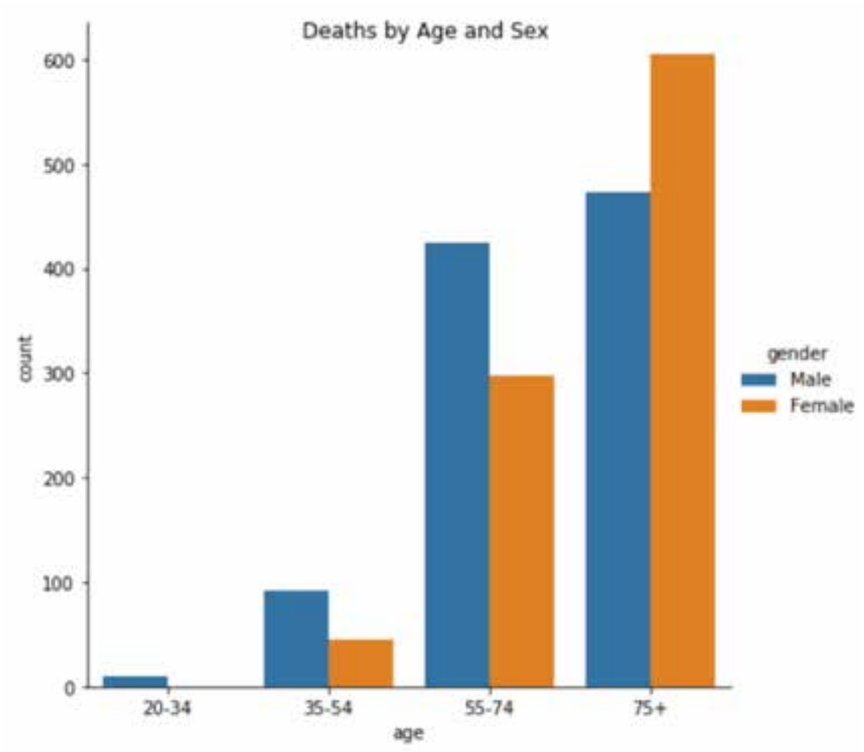


COVID-19 | CASES & HOSPITALIZATIONS - SEX

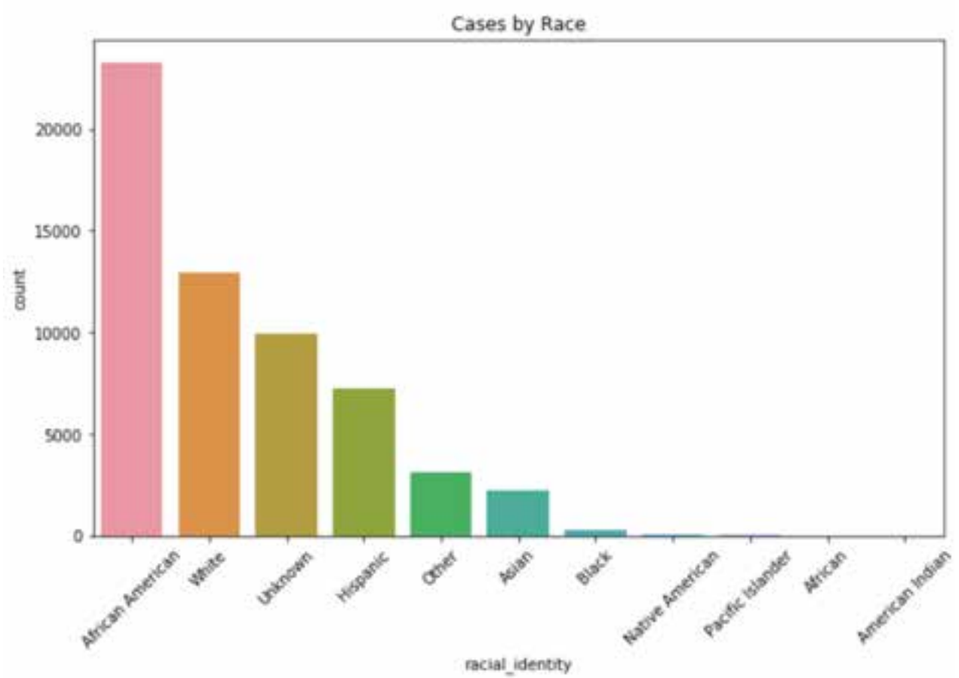


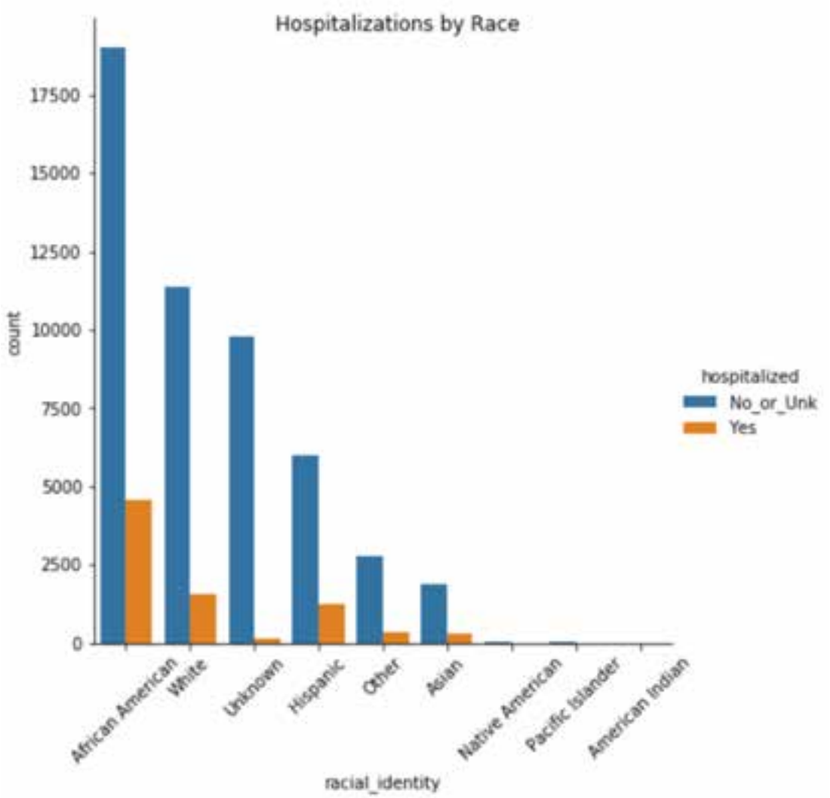


COVID-19 | DEATHS - AGE & SEX

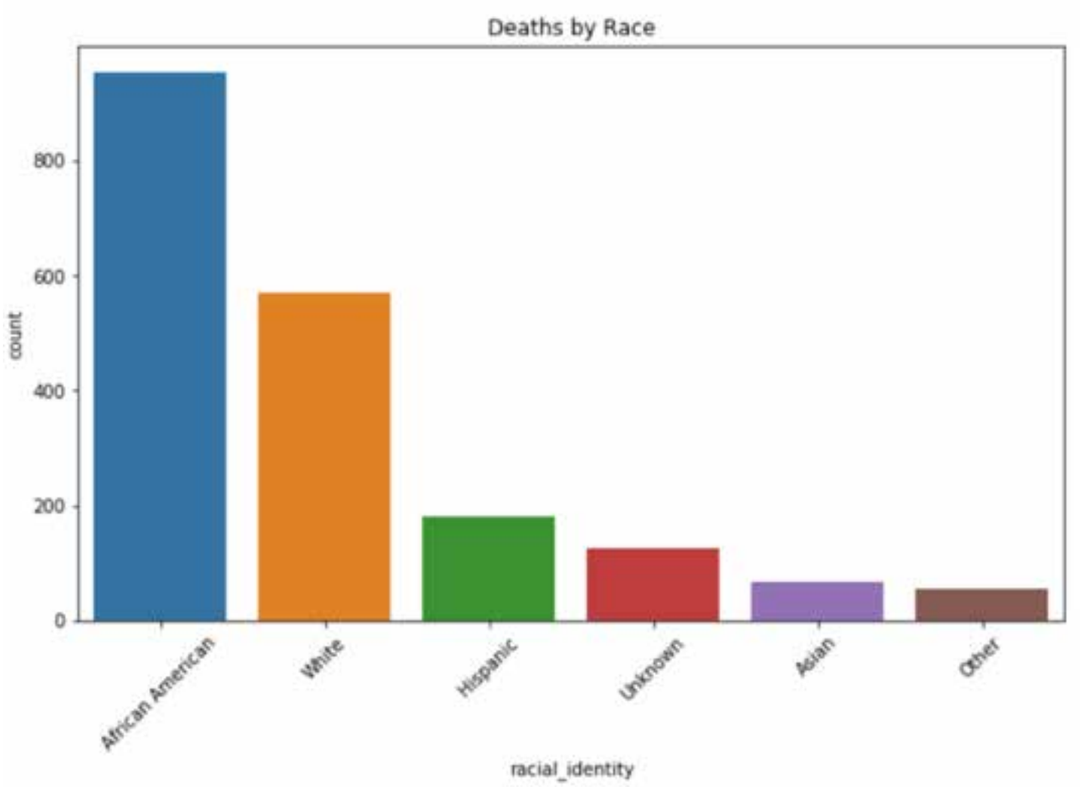


COVID-19 | CASES & HOSPITALIZATIONS - RACE





COVID-19 | DEATHS - RACE



THE DATASET | SOLD PROPERTIES

- *Realtor* API only goes back to April/20
- New dataset very disorganized
- 942,959 instances, 78 columns

DATA CLEANING

- Dropped useless columns
- Timestamp data created
- Transformed instances to months data
- New, organized data, based on the original numbers

DATA CLEANING

```
#The idea was to work with the data divided by month.
#As data are for weeks, some pre-processing was required to convert the numbers to the desired time period.

philadelphia_clean = pd.DataFrame(columns= ['period'] + list(philadelphia.columns)[3:])

for year in ['2019', '2020']:
    for month in ['01', '02', '03', '04', '05', '06', '07', '08', '09', '10', '11', '12']:

        if (year == '2020' and month == '11'):
            break

        else:
            temp_df = philadelphia[philadelphia['period_begin'].str.startswith(year + '-' + month)]
            month_title = year + '-' + month
            total_homes_sold = np.sum(temp_df['total_homes_sold'])
            median_sale_price = np.sum(temp_df['median_sale_price'])
            median_days_to_close = np.average(temp_df['median_days_to_close'])
            price_drops = np.average(temp_df['price_drops'])
            percent_active_listings_with_price_drops = np.average(temp_df['percent_active_listings_with_price_drops
            pending_sales = np.average(temp_df['pending_sales'])
            total_new_listings = np.sum(temp_df['total_new_listings'])
            average_new_listings = np.average(temp_df['average_new_listings'])
            median_new_listing_price = np.average(temp_df['median_new_listing_price'])
            inventory = np.average(temp_df['inventory'])
            total_active_listings = np.average(temp_df['total_active_listings'])
            age_of_inventory = np.average(temp_df['age_of_inventory'])
            homes_delisted = np.sum(temp_df['homes_delisted'])
            median_active_list_price = np.average(temp_df['median_active_list_price'])
            avg_offer_to_list = np.average(temp_df['avg_offer_to_list'])
            median_days_on_market = np.average(temp_df['median_days_on_market'])
            months_of_supply = np.average(temp_df['months_of_supply'])
            percent_total_price_drops_of_inventory = np.average(temp_df['percent_total_price_drops_of_inventory'])

            row_data = [month_title, total_homes_sold, median_sale_price, median_days_to_close, price_drops,
                        percent_active_listings_with_price_drops, pending_sales, total_new_listings,
                        average_new_listings, median_new_listing_price, inventory, total_active_listings,
                        age_of_inventory, homes_delisted, median_active_list_price, avg_offer_to_list,
                        median_days_on_market, months_of_supply, percent_total_price_drops_of_inventory]

            philadelphia_clean.loc[len(philadelphia_clean)] = row_data
```

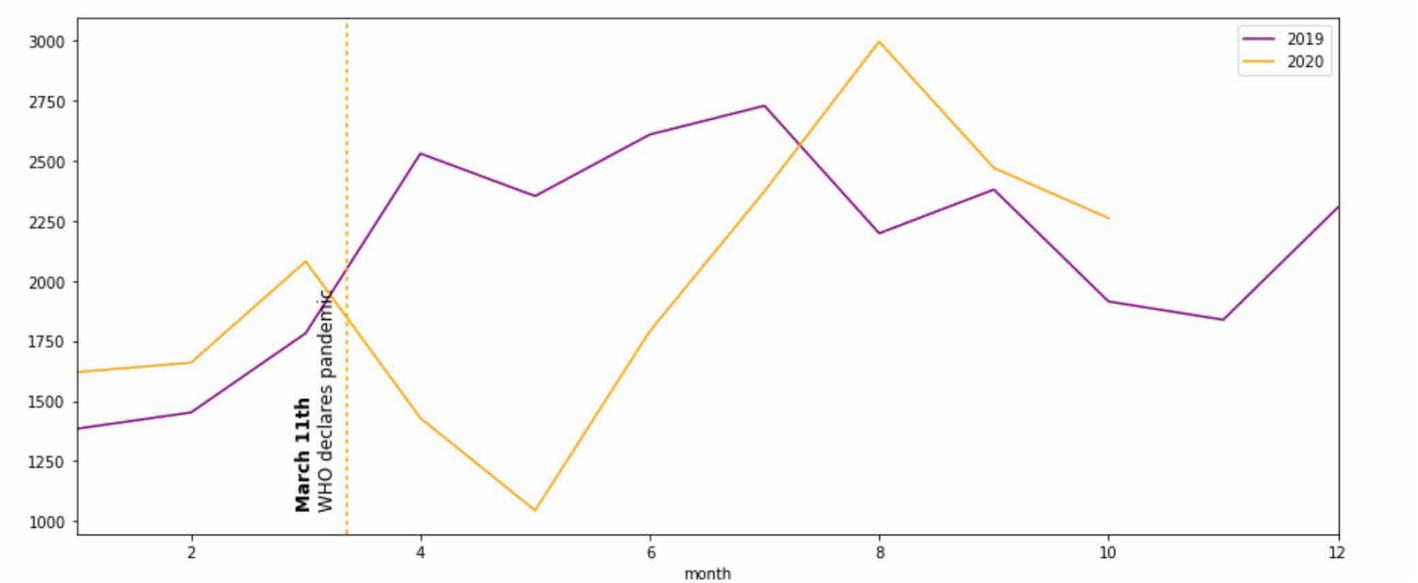
INFLATION

	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC
2020	2.5	2.3	1.5	0.3	0.1	0.6	1.0	1.3	1.4	1.2	NAN	NAN
2019	1.6	1.5	1.9	2.0	1.8	1.6	1.8	1.7	1.7	1.8	2.1	2.3

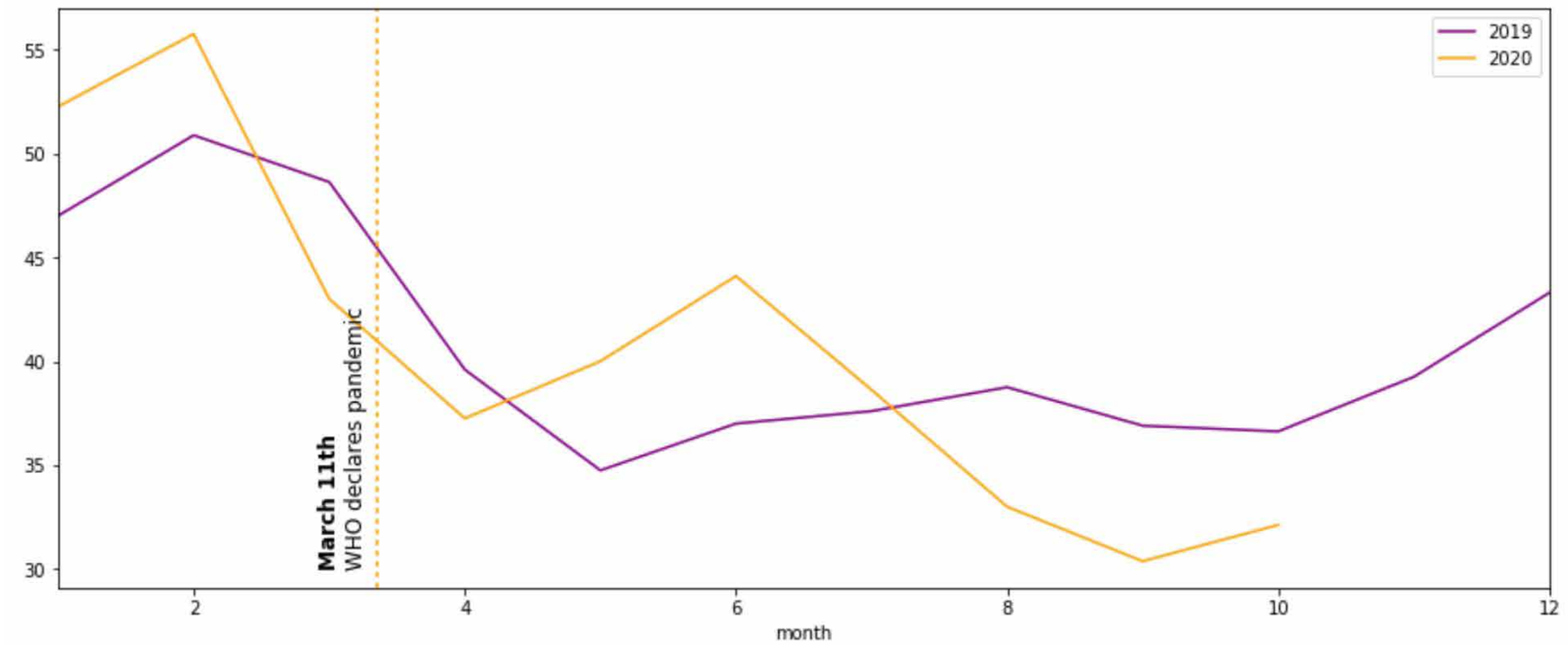
- Updated 2019 financial data
- Better visualization when comparing the graphics

TOTAL OF PROPERTIES SOLD

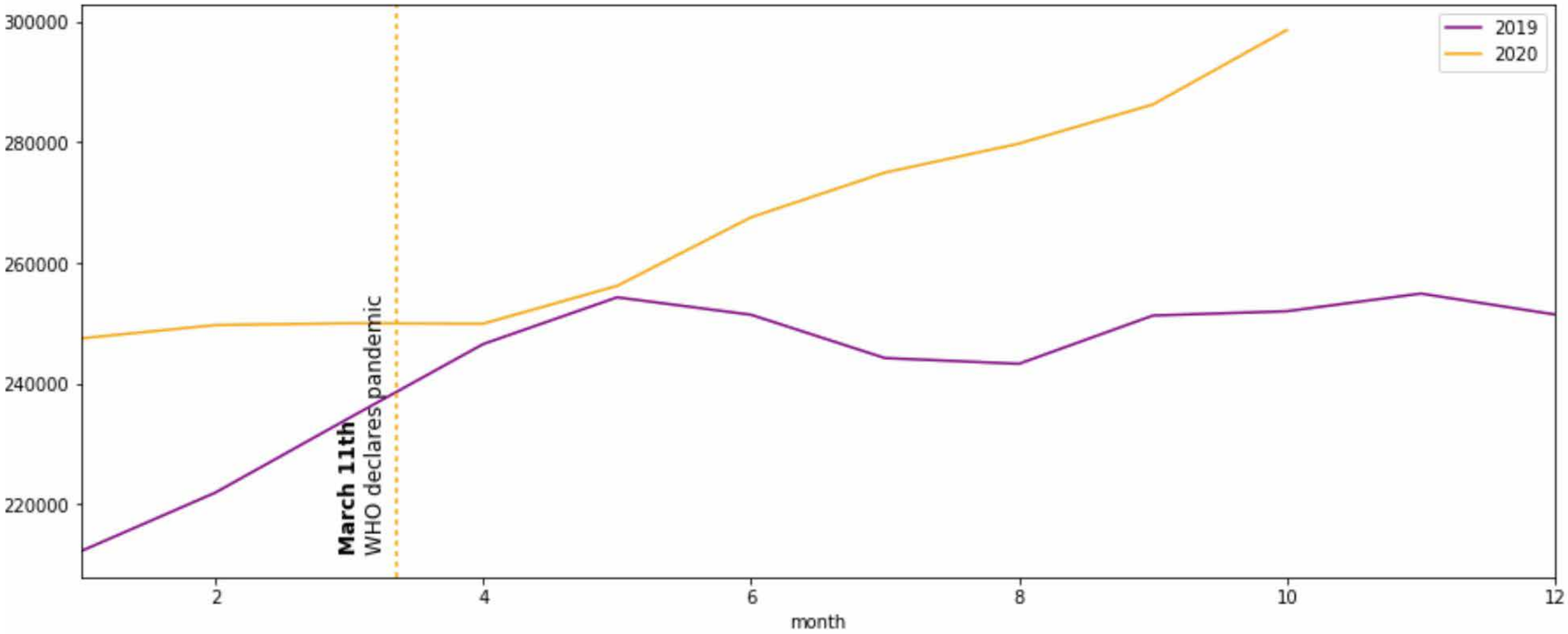
```
months_19 = [dt.datetime.strptime(row, '%Y-%m').month for row in philadelphia19['period']]
philadelphia19['month'] = months_19

months_20 = [dt.datetime.strptime(row, '%Y-%m').month for row in philadelphia20['period']]
philadelphia20['month'] = months_20

ax = philadelphia19.plot(x='month', y='total_homes_sold', label = '2019', figsize=(15,6.3), c = 'Purple')
ax = philadelphia20.plot(x='month', y='total_homes_sold', label = '2020', ax=ax, c = 'Orange')
plt.axvline(3.354, c = 'Orange', dashes = (2, 2))
plt.text(3.2,1065,'WHO declares pandemic',rotation=90, ha='center', size = 'large')
plt.text(3,1065,'March 11th',rotation=90, ha='center', weight = 'bold', size = 'large');
```



DAYS ON THE MARKET



MEDIAN ACTIVE SALES PRICE



SOLD PROPERTIES

Month with the highest median sale price:
2020-06

Month with the highest number of days to close the sales:
2020-05

Month with the highest inventory:
2019-05

Month with the highest number median days on market:
2020-02

PROCEDURES

- DATA ACQUISITION
  API Access | Property for Sale | Property for Rent

- DATA PRE-PROCESSING
  Remove Duplicates & NAs | Extract Info from List

- VISUALIZATION
  Unit Price

DATA ACQUISITION | API ACESS

```
## functions

def realtor_api_request(property_id):
    url = "https://realtor.p.rapidapi.com/properties/v2/detail"

    querystring = {"property_id":property_id}

    headers = {
        'x-rapidapi-host': "realtor.p.rapidapi.com",
        'x-rapidapi-key': AFKii13
        }

    response = requests.request("GET", url, headers=headers, params=querystring)

    return response.json()
```

DATA ACQUISITION
| API ACCESS

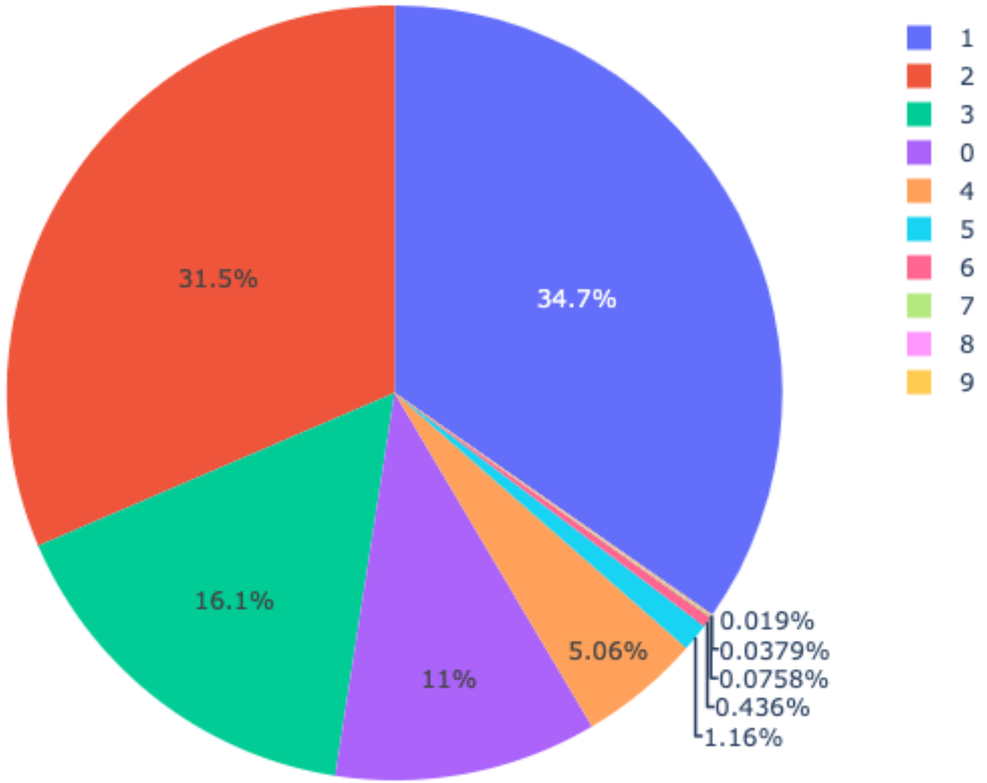

```
{'meta': {'build': '3.23.121',
  'schema': 'core.3',
  'tracking': 'type|meta|data|resource_type|property_detail|query|client_id|rdc_mobile_native,1
3.3.0.53|schema|core.3|tag_version|v2^^^$0|1|2|$3|4|5|$6|7|8|9|A|B]]]',
  'returned_rows': 1,
  'matching_rows': 1,
  'tracking_params': {'ldpPropertyStatus': 'ldp:for_sale',
   'pageType': 'ldp',
   'leadDelivery': 'co_broke',
   'leadEnhancements': 'classic,rcm',
   'listingActivity': 'pending',
   'productType': 'core.agent,core.broker',
   'propertyStatus': 'for_sale',
   'listingId': '2920870158',
   'rentalDataSource': 'unknown',
   'advertiserIdAgent': '1888436',
   'advertiserIdBroker': '2072377',
   'advertiserIdOffice': '2215397',
   'communityId': 'unknown',
   'mprId': '4282337167',
   'listingMls': 'PHPA',
   'planId': 'unknown',
   'subId': 'unknown',
   'city': 'Philadelphia',
   'neighborhood': 'South Philadelphia',
   'state': 'PA',
   'zip': '19148',
   'listingBaths': '2',
   'listingBeds': '3',
   'listingSqFt': '1135',
   'listingEnhancements': 'broker-photo-top,broker-photo-btm,broker-phone-btm',
   'listingPrice': '279500',
   'photoCount': '30',
   'propertyType': 'ldp:condo',
   'version': '1.0'}},
 'properties': [{'property_id': 'M4282337167',
   'prop_status': 'for_sale',
   'listing_id': '2920870158',
   'prop_type': 'condo',
   'list_date': '2020-09-03T04:50:59Z',
   'hoa_fee': 0,
   'hoa_historic_fee': None,
   'last_update': '2020-10-25T17:09:59Z',
   'virtual_tour': {'href': 'https://mls.homejab.com/property/view/114-fitzgerald-st-philadelphi
a-pa-19148-usa'},
   'broker': {'advertiser_id': 2072377,
    'name': 'Houwzer Salaried Realtors',
    'phone1': {'number': '2674635995', 'type': 'broker'}},
   'year_built': 1920,
   'listing_status': 'Pending',
   'beds': 3,
```

DATA ACQUISITION | PROP. FOR SALE

2	M4036371277	condo	townhomes	'Philadelphia', 'line': '5703 N 13th ...	{'list_item': {'name': 'Pre...	for_sale	215000	1.0	2	3.0	{'size': 1360, 'units': 'sqft'}	True, 'photo': None, 'name': 'Kev...	2020-10-13T17:24:20Z	35	1	3
3	M3553029343	single_family	NaN	{'city': 'Philadelphia', 'line': '1009 Rhawn S...	{'listing_office': {'list_item': {'name': 'Re/...	for_sale	394800	1.0	2	3.0	{'size': 1856, 'units': 'sqft'}	[{'primary': True, 'advertiser_id': '4759', 'i...	2020-10-13T17:11:54Z	123	1	4
4	M3649199107	condo	townhomes	{'city': 'Philadelphia', 'line': '3850 N Bouvi...	{'listing_office': {'list_item': {'name': 'Re/...	for_sale	130000	1.0	2	3.0	{'size': 1180, 'units': 'sqft'}	[{'primary': True, 'advertiser_id': '391546', ...	2020-10-13T17:02:13Z	33	1	5
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
9557	M3400474681	condo	townhomes	{'city': 'Philadelphia', 'line': '2077 Bridge ...	{'listing_office': {'list_item': {'name': 'Re/...	for_sale	94900	1.0	1	4.0	{'size': 1296, 'units': 'sqft'}	[{'primary': True, 'photo': None, 'name': ''}]	2020-10-05T12:27:53Z	18	42	33
9558	M3654238233	condo	townhomes	{'city': 'Philadelphia', 'line': '1452 N 57th ...	{'listing_office': {'list_item': {'name': 'Kel...	for_sale	149900	1.0	2	3.0	{'size': 1026, 'units': 'sqft'}	[{'primary': True, 'advertiser_id': '950515', ...	2020-10-01T06:57:05Z	14	42	34
9559	M9041602444	condo	duplex_triplex	{'city': 'Philadelphia', 'line': '7354	{'listing_office': {'list_item': {'name':	for_sale	432990	3.0	4	3.0	{'size': 1614, 'units': 'sqft'}	[{'primary': True, 'advertiser_id':	2020-10-06T09:19:02Z	2	42	35

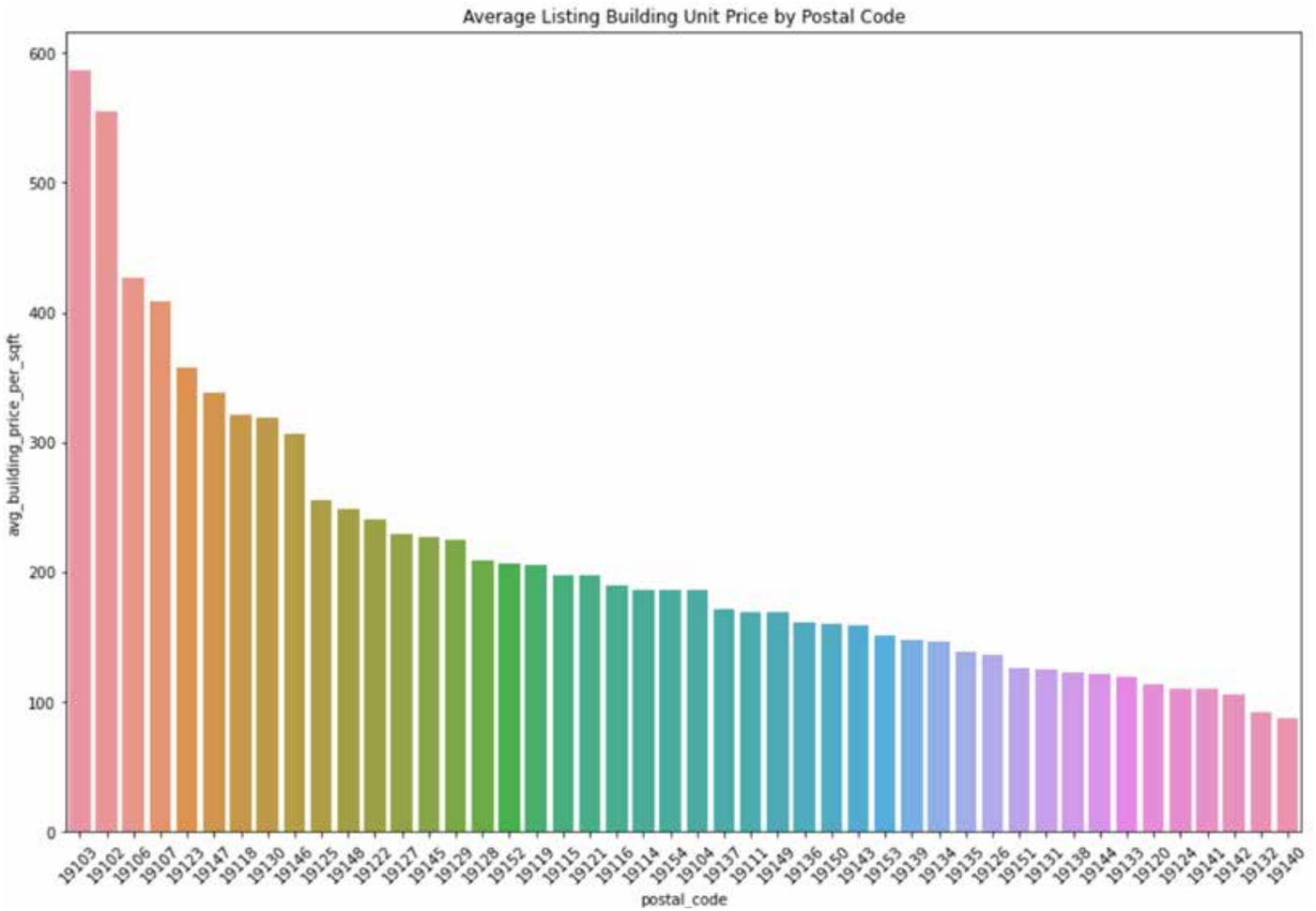
HEATMAP | PROPERTIES FOR RENT

PROPERTIES FOR RENT | BY BEDROOM

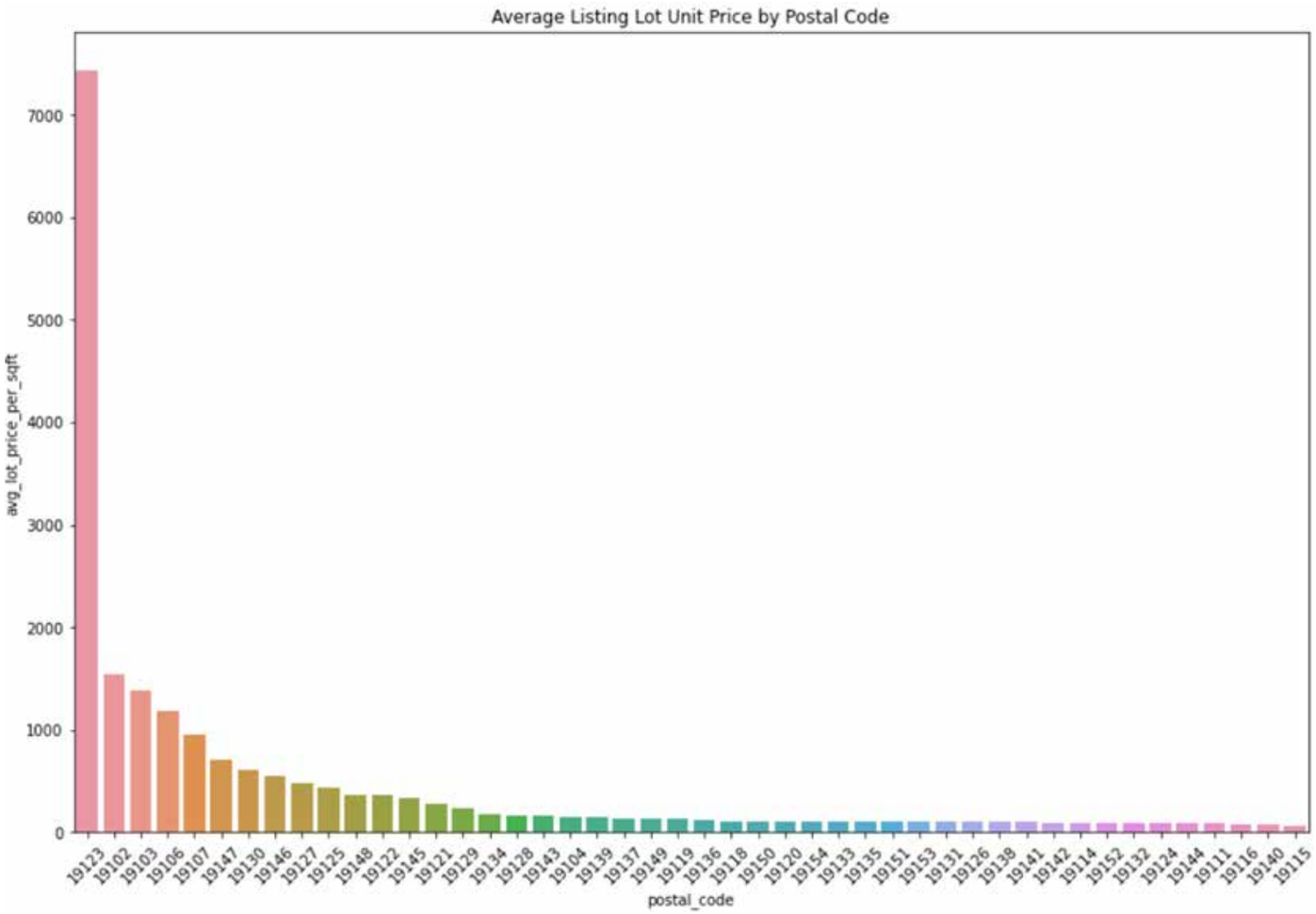


Studios and properties with **1, 2, and 3 bedrooms** are dominant in the house rental market.

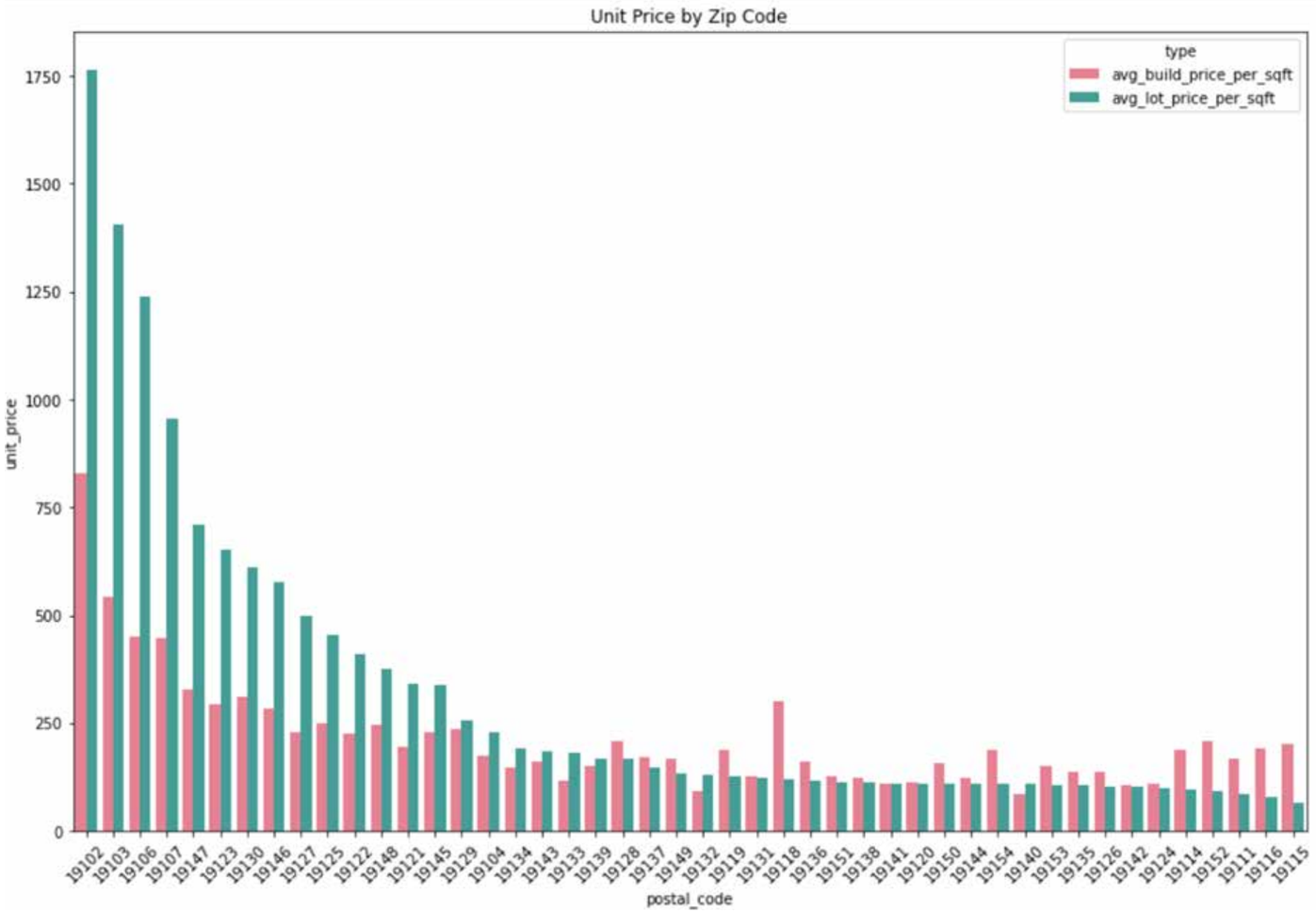
UNIT BUILDING PRICE BY POSTAL CODE



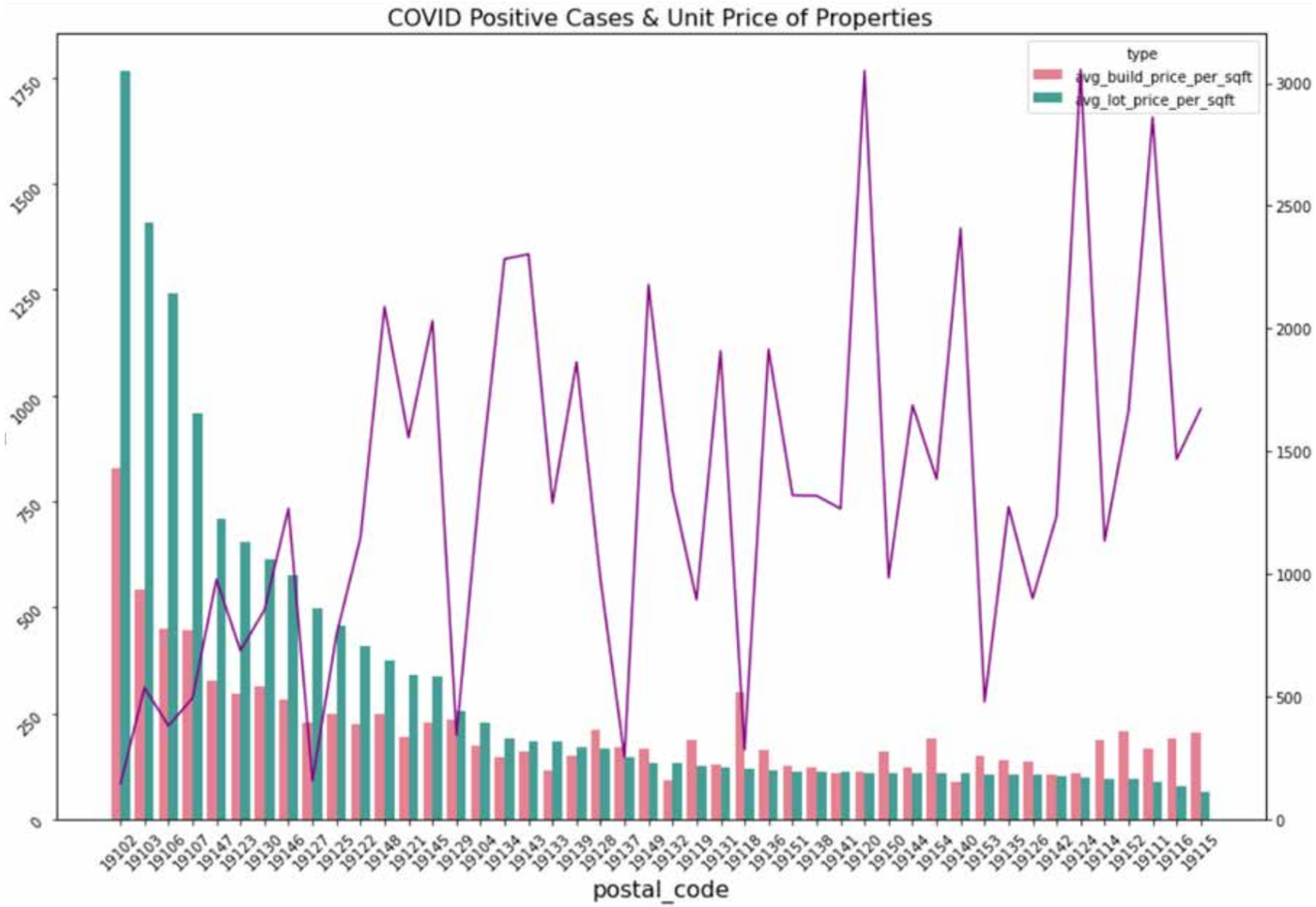
UNIT LOT PRICE BY POSTAL CODE



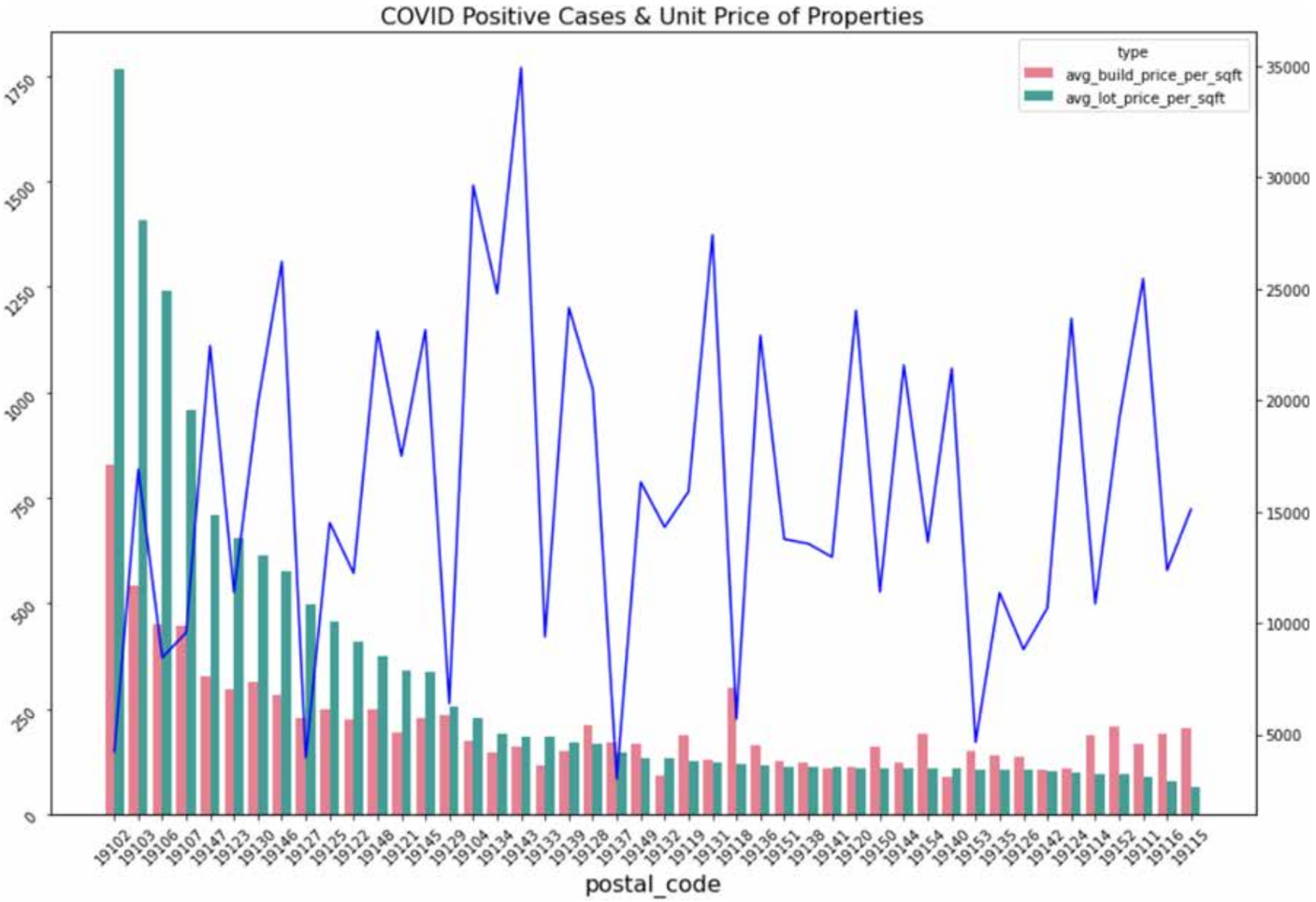
UNIT PRICE BY POSTAL CODE



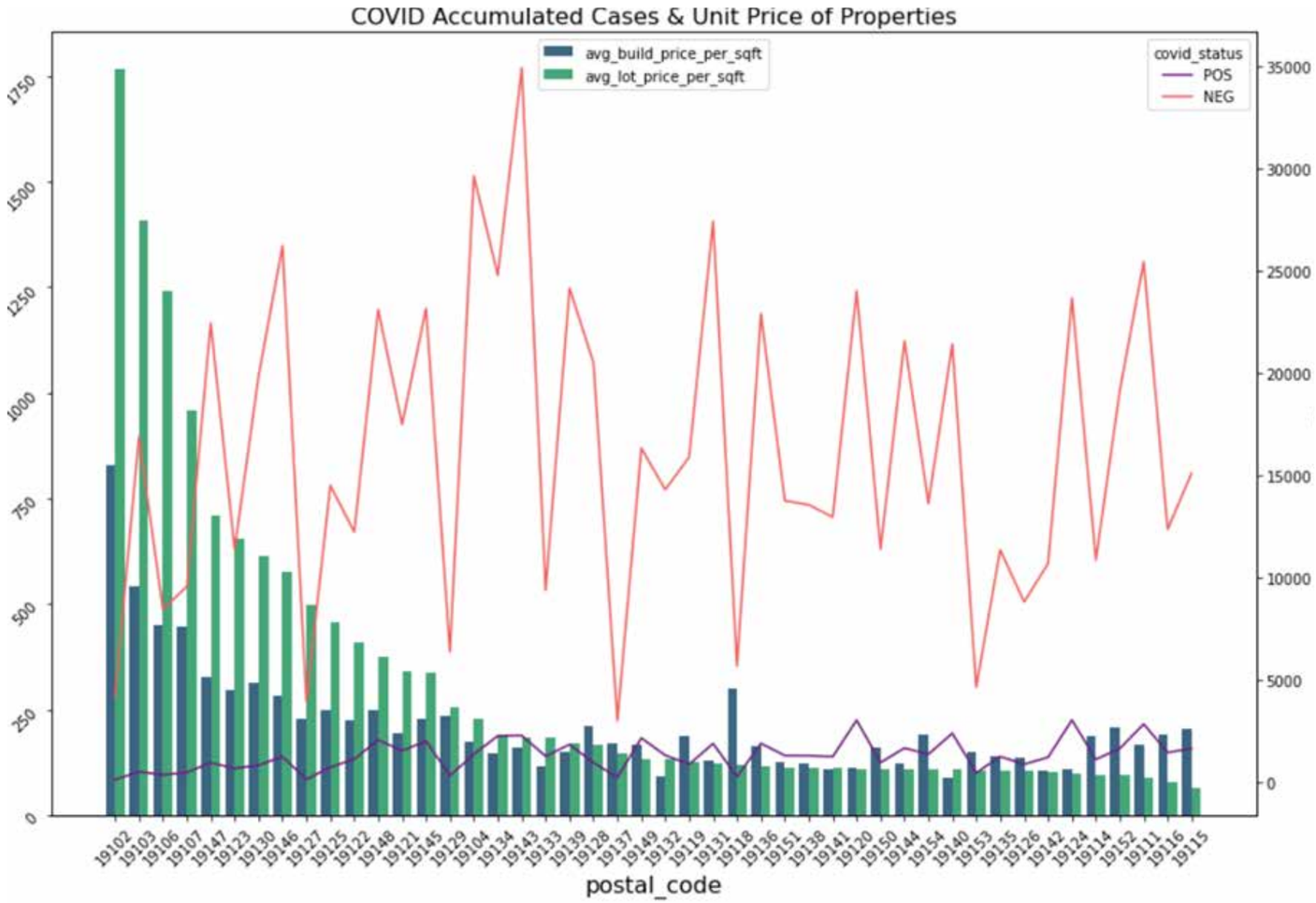
POSITIVE CASES & UNIT PRICE BY POSTAL CODE



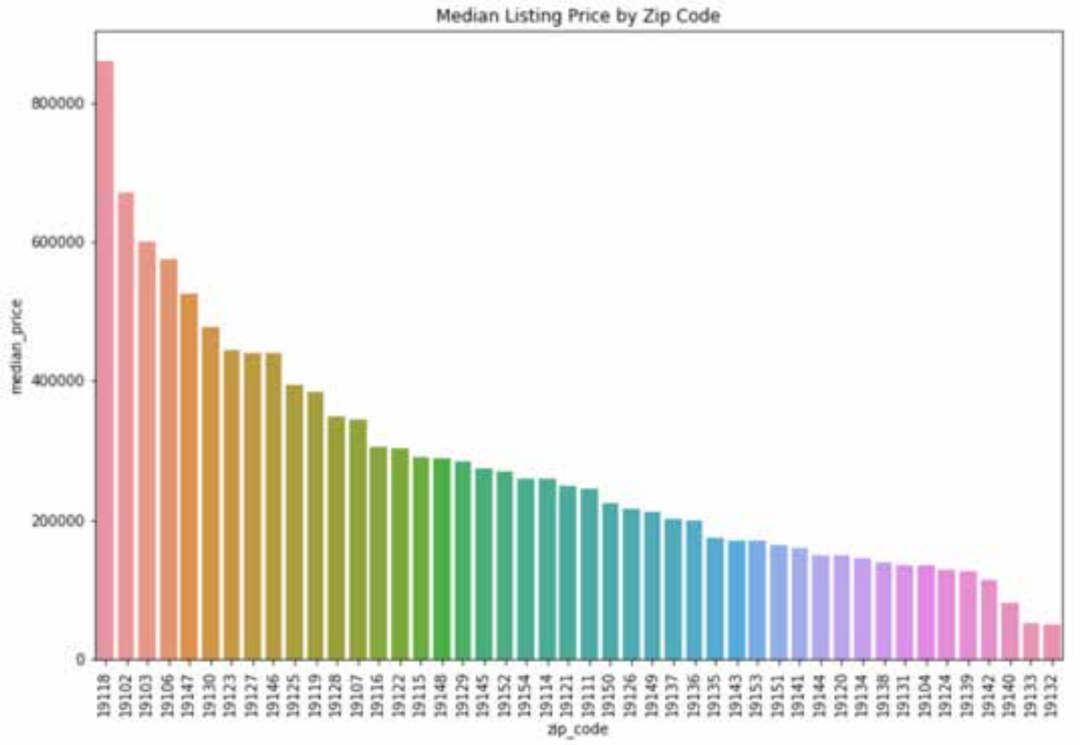
NEGATIVE CASES & UNIT PRICE BY POSTAL CODE



ALL CASES AND UNIT PRICE BY POSTAL CODE



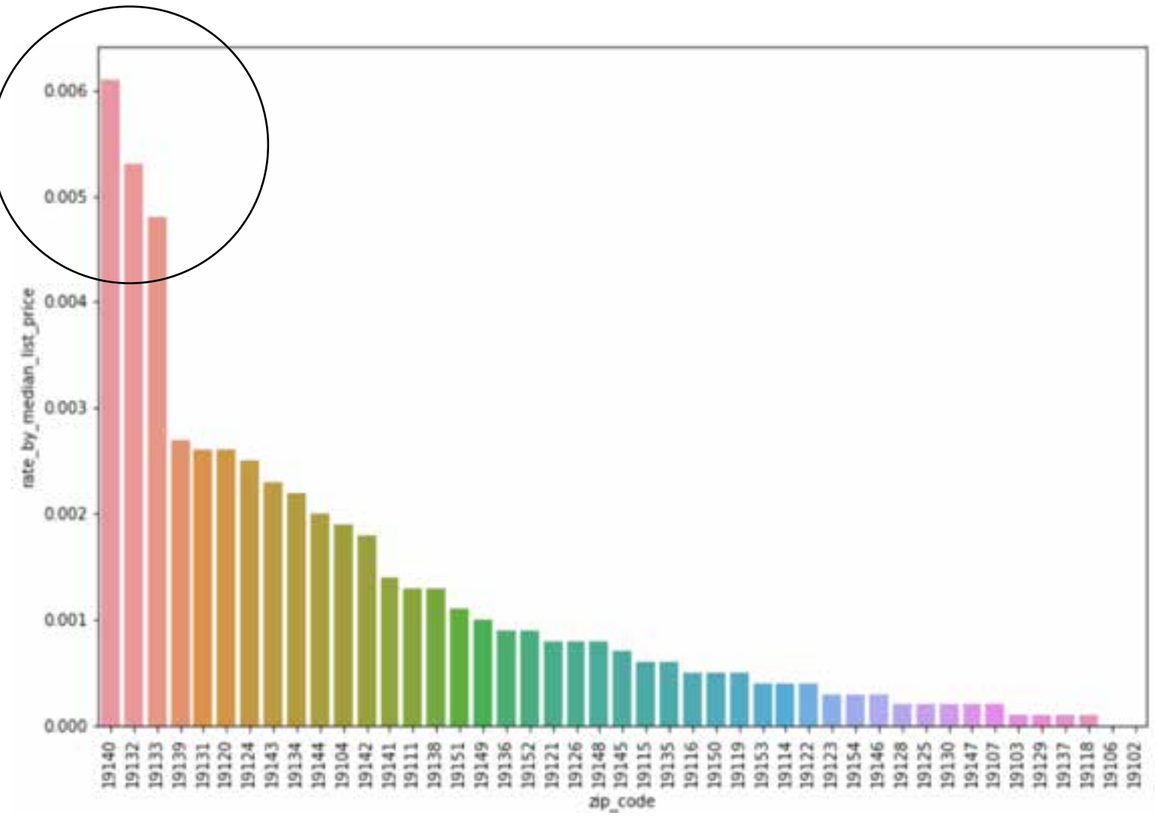
MEDIAN LISTING PRICE BY POSTAL CODE



- **Top 3:**
  19118, 19102, 19103
- **Bottom 3:**
  19140, 19133, 19132

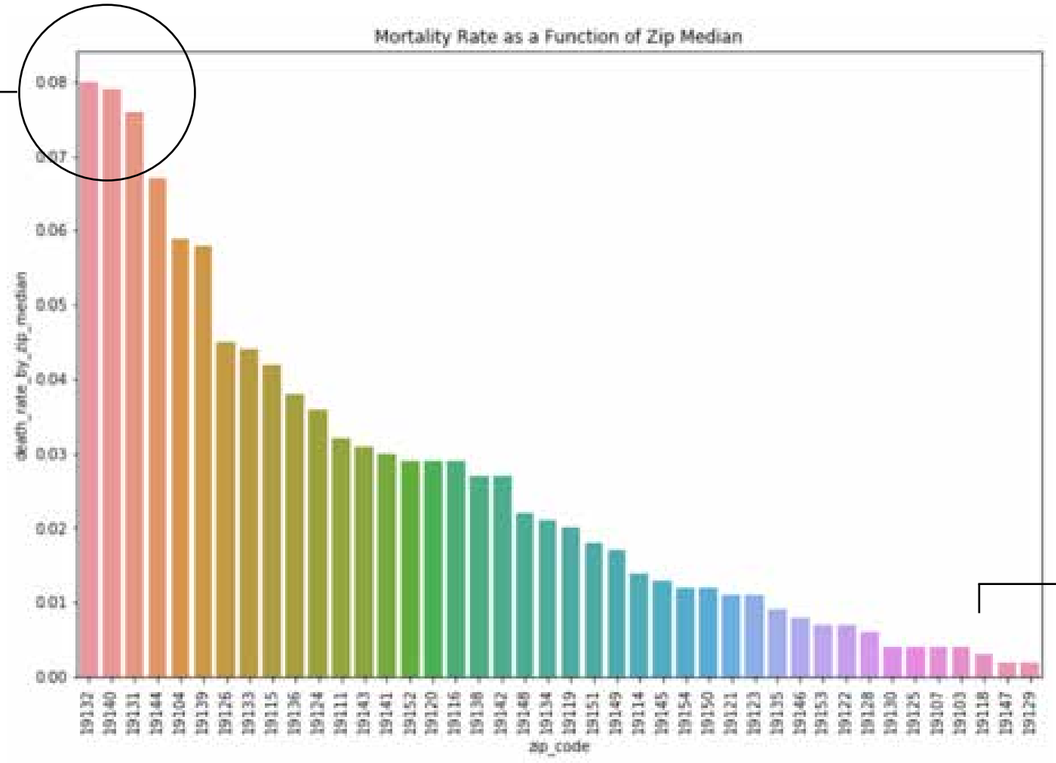
HOSPITALIZATIONS PER MEDIAN LISTING PRICE

- Top 3 zips 19140, 19132, 19133 were **bottom 3 in median listing price**



RATE OF MORTALITY PER MEDIAN LISTING PRICE

Top 2 zips were in **bottom 3 in median listing price**



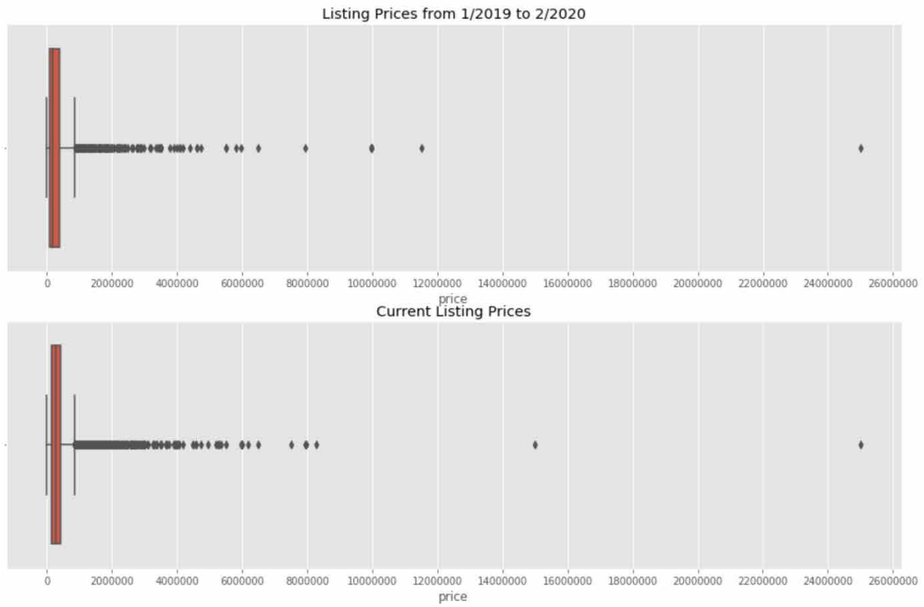
Highest median listing price

STATISTICAL ANALYSIS

- Means of the current and historical real estate data?
- Is there a difference in the top & bottom 10 postal codes effected by COVID-19?
- Differences stay consistent when focusing on the mean price per square foot?

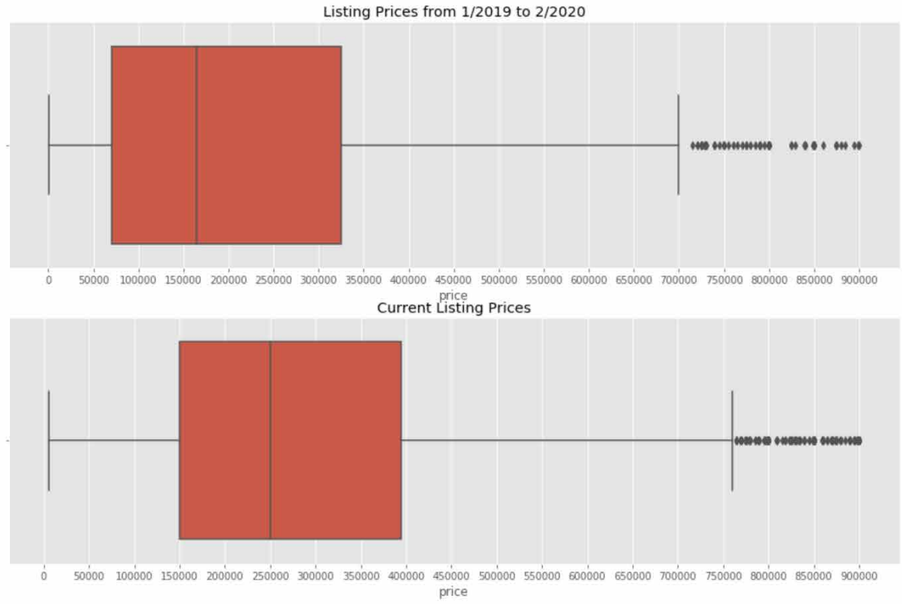
VISUALIZING DISTRIBUTION

FULL DATASETS



UNEQUAL VARIANCE

SOME OUTLIERS REMOVED



EQUAL VARIANCE

2-SAMPLE T-TEST | MEAN LISTING PRICE

MEAN OF HISTORICAL DATA: **$224,190**
MEAN OF CURRENT DATA: **$288,463**

```
tstat, pval = stats.ttest_ind(curr_df2['price'], final_hist2['price'], equal_var=True)
ci = sms.CompareMeans(sms.DescrStatsW(curr_df2['price']), sms.DescrStatsW(final_hist2['price']))
print('p-value =', pval)
print('95% Confidence Interval -', ci.tconfint_diff(usevar='pooled'))
```

```
p-value = 2.1435395293392626e-47
95% Confidence Interval - (55601.26424426666, 72944.90935554293)
```

With 95% of confidence, the difference in the mean listing price between historical data increased between $55,601 and 72,945.

2-SAMPLE T-TEST

MEAN LISTING PRICE TOP 10 ZIPS EFFECTED

MEAN OF HISTORICAL DATA: **$152,856**
MEAN OF CURRENT DATA: **$208,375**

```
tstat, pval = stats.ttest_ind(top10_curr2['price'], top10_hist2['price'], equal_var=True)
ci = sms.CompareMeans(sms.DescrStatsW(top10_curr2['price']), sms.DescrStatsW(top10_hist2['price']))
print('p-value =', pval)
print('95% Confidence Interval -', ci.tconfint_diff(usevar='pooled'))
```

```
p-value = 1.3994086678358032e-27
95% Confidence Interval - (45608.744306212364, 65429.2242495698)
```

With 95% of confidence, the difference in the mean listing price between historical data increased between $45,608 and 65,429.

FINDINGS

- The difference in mean listing price between the historical and current data overall is statistically significant
- The mean listing price overall increased by ~$64,000
- Lower income postal codes saw the highest number of positive cases, and had significant increase in mean listing price. Higher income postal saw a decrease, however, this decrease not to be considered significant
- The mean building price per sqft saw a statistically significant increase by about $17.86 during the pandemic

FINDINGS

- The number of properties sold fell sharply after the start of the pandemic. However, the market recovered in a few months.
- The median active sales price remained stable for about a year but started to increase after the pandemic.
- The average property unit prices are higher in center city than in surrounding areas.
- The covid cases do not seem very significantly impact the unit price of properties.