Joseph Vandenorth

Unit 3 Project: Cities and Rents

Problem
Rents across the U.S. are on the rise while incomes stagnate. It's important for real estate development professionals and government service providers to understand where the biggest need is for housing so they can develop more units and services. With urban areas growing due to service sector jobs, larger cities across the U.S. likely are experiencing the highest rents on average.

Proposed Test
Compare the top 50 cities by U.S. Census data population and compare to Zillow rents in recent quarters, showing cities experiencing the highest rents.

Measurable Outcome
Show how larger cities compare to smaller cities when it comes to average rents. Order by population size.

UNIT 3 PROJECT RESULTS - PYTHON AND JUPYTER

Conclusion
In running exploratory analysis of our comparison of population and rent, we found a modest positive correlation with an R2 of 0.21. When looking at a scatter plot, we found this validated though there are some outliers like smaller cities with high rents (San Francisco).

```
merge = pd.merge(rents, cities, on = "City", how = "left")

rents[['City', 'rent']].describe()
```

Analyses

Exploratory Data Analyses

Descriptive statistics and correlations for the variables of interest are provided below.

```
[30]: rents[['City', 'rent']].describe()
```

[30]:

	rent
count	49.000000
mean	1636.612245
std	725.194107
min	750.000000
25%	1197.000000
50%	1391.000000
75%	1974.000000
max	4382.000000

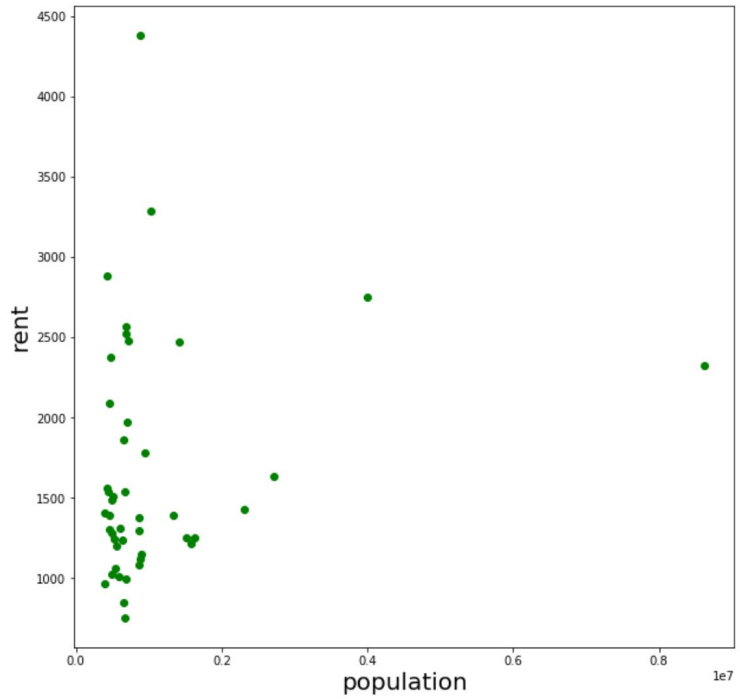
merge.loc[:, ['population', 'rent']].corr()

```
[26]: merge.loc[:, ['population', 'rent']].corr()
```

[26]:

	population	rent
population	1.000000	0.210552
rent	0.210552	1.000000

```
plt.figure(figsize=(10, 10))
ax = plt.axes()
ax.scatter(merge['population'], merge['rent'],color='g')
ax.set_xlabel('population', size = 20)
ax.set_ylabel('rent', size = 20)
plt.show()
```



Joseph Vandenorth

UNIT 2 PROJECT RESULTS - SQL

The analysis found that by grouping all larger than average cities by population (grouped as 1) and comparing to all cities smaller than average (grouped as 0), rents where over $300 higher for larger sized cities on average.

popgroup avg(rent.rent)
0 1586.8857 1 1899.2000

Columns
Rent table: city; rent
Pop table: city; population

CODE

```
SELECT rent.city, rent.rent, pop.population
FROM rent
LEFT JOIN pop ON rent.city = pop.city
ORDER BY pop.population;

SELECT AVG(population), STDDEV(population), MIN(population), MAX(population)
FROM pop;

alter table pop
add column popgroup varchar(30);

UPDATE pop
SET popgroup = case
when population >= 1000000 then 1
when population < 1000000 then 0
End;

SELECT AVG(rent)
GROUP BY popgroup
FROM pop

select pop.popgroup, avg(rent.rent) FROM pop
left join rent on rent.city = pop.city
group by pop.popgroup;
```