P

Core Message

Using multiple categories of data we can determine which U.S. cities have a higher quality of life.

Questions and Motivation

- What sort of patterns can we discern comparing individual categories and overall quality of life?
- Does geographical location or population density have any impact on scores?
- Interest in areas such as housing, travel, healthcare, environmental, cost of living, safety, education, income, culture, etc.

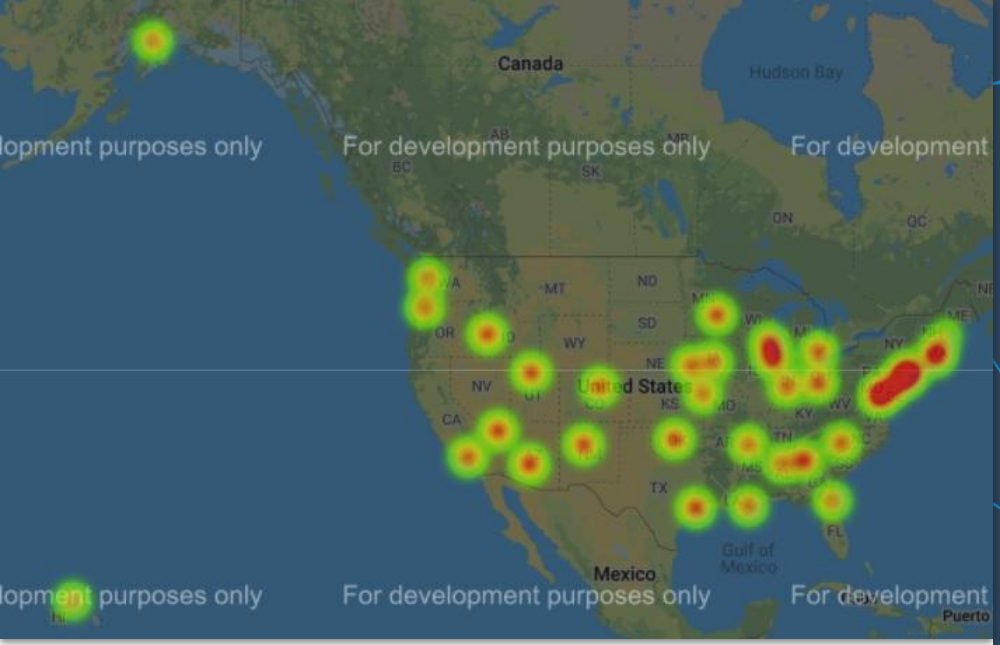
- Compare the locations with highest quality of life ratings to lowest quality of life ratings, and see if any of the above listed variables (i.e. healthcare, income) have any great impact on average or total

KH

P

Summary of Where and How We Found Our Data

- **Teleport API:** Compare cities on quality of life, cost of living, salaries and more.
  Explore where to move based on your personal preferences.
  https://developers.teleport.org/api/
- **US Bureau of Economic Analysis:** Income Data
  https://www.bea.gov/
- **Centers for Disease Control - CDC Wonder API:** Mortality Data by County by Year
  https://wonder.cdc.gov/ucd-icd10.html
- **Centers for Disease Control - Socrata Open Data:** NCHS - Drug Poisoning Mortality
  https://dev.socrata.com/foundry/data.cdc.gov/rpvx-m2md
- **Federal Communications Commision:** County Info (FIPS)
  https://geo.fcc.gov/api/census/

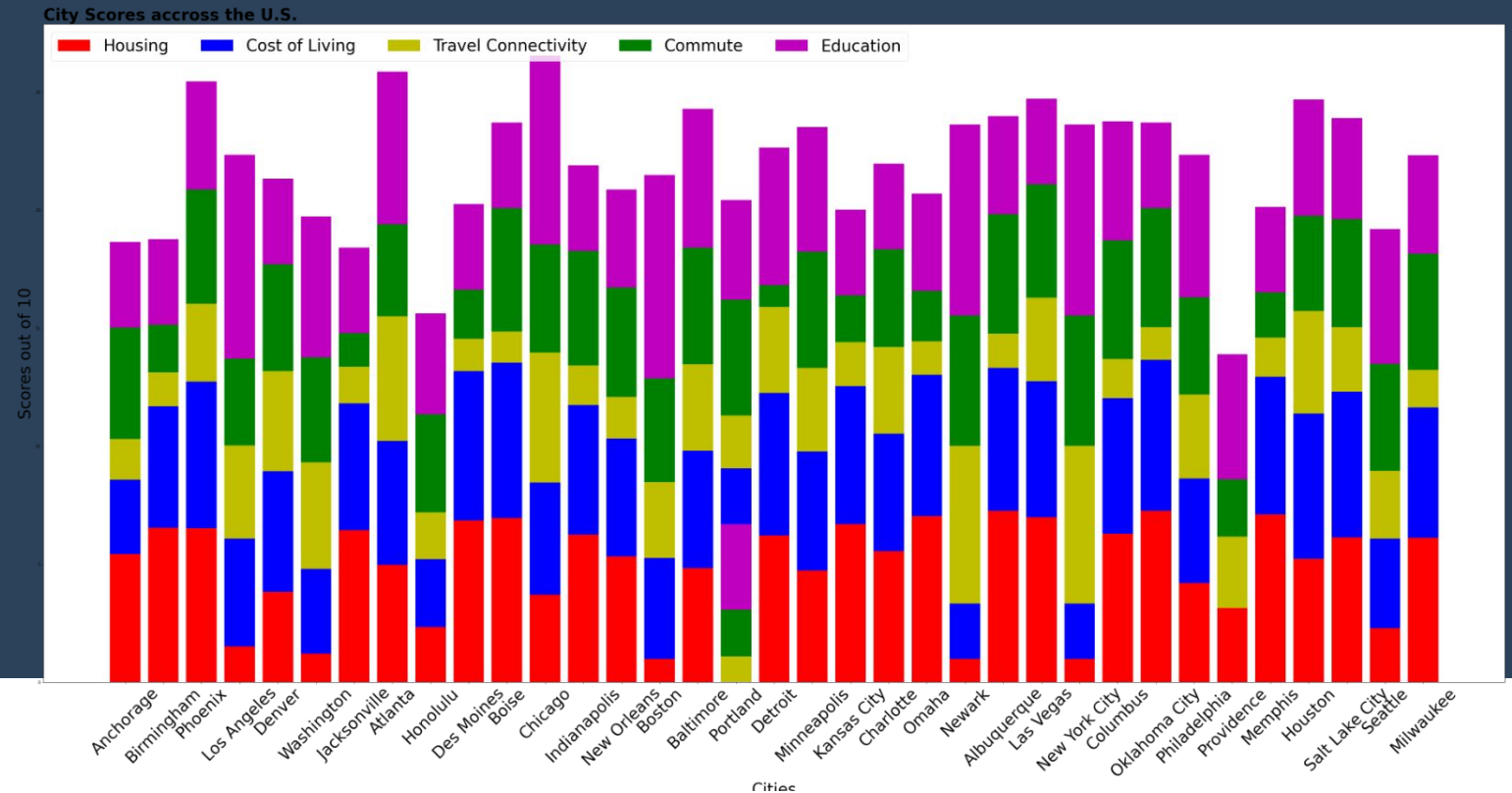


Total Score of 5 Categories (where data is available) Heat Map

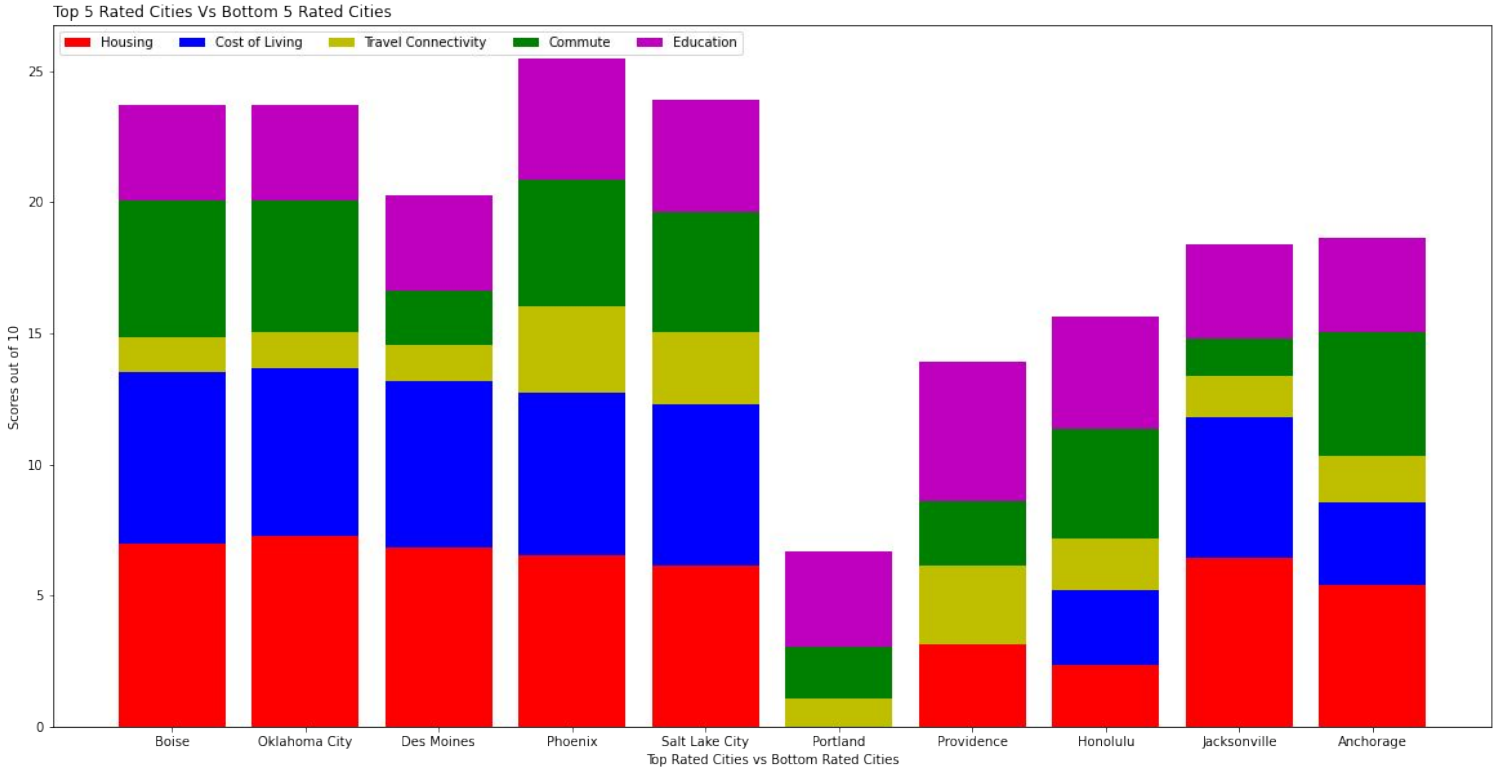
Comparison of quality of life based on scores rated 1 out of 10 for:

Housing, Cost of Living, Travel Connectivity, Commute, and Education across major U.S. Cities

Quality of Life Scores in 5 Categories Across the U.S.



Looking at 5 specific categories that we are interested in, and we can determine which scores are higher or lower within major U.S. cities We can also combine them for a total score.

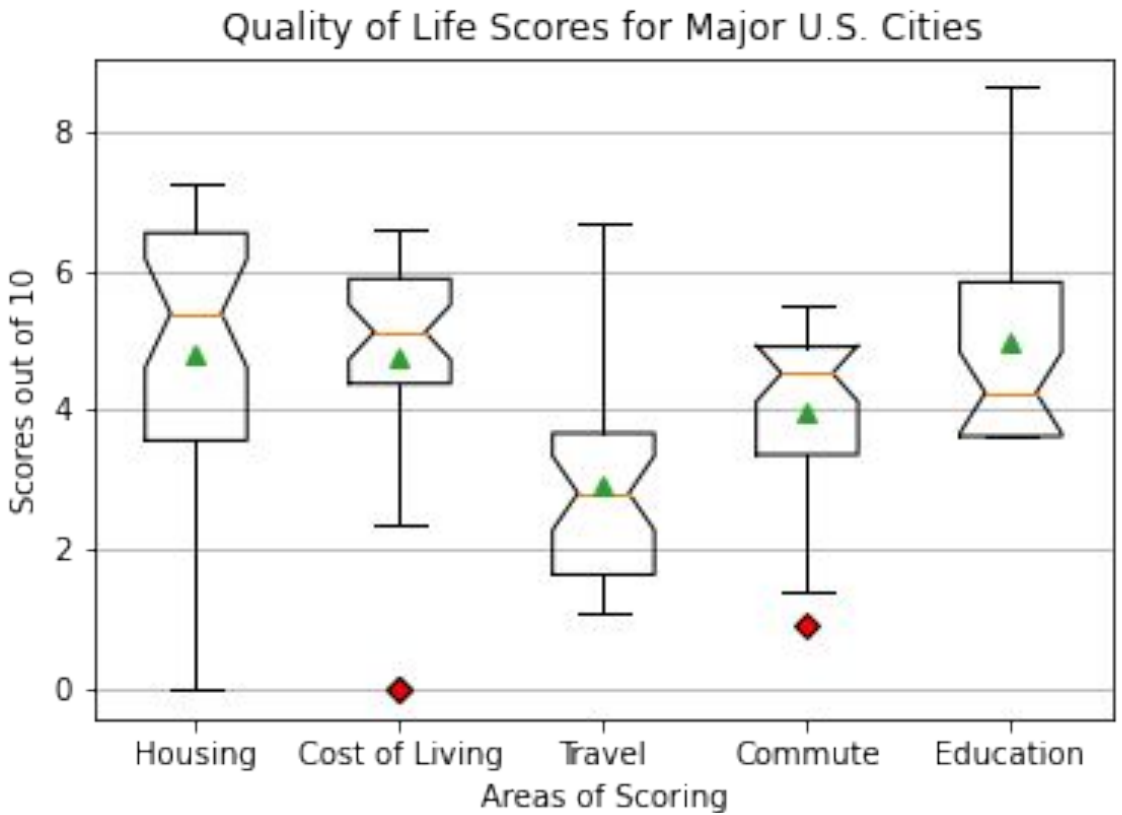


Using our data we can create graphs to compare highest and lowest rated cities based on specific categories of our choosing. We can use this to determine where we would prefer to live or work based on the category that is most important to us.

Statistical Analysis

	Population	Housing	Cost of Living	Travel Connectivity	Commute	Education	Total Score
count	3.600000e+01	36.000000	36.000000	36.000000	36.000000	36.000000	36.000000
mean	9.847224e+05	4.785542	4.776389	2.922931	3.967465	4.970806	21.423132
std	1.458472e+06	2.190898	1.618827	1.527034	1.306818	1.602809	3.722342
min	6.688100e+04	0.000000	0.000000	1.102000	0.933000	3.624500	6.696750
25%	3.851270e+05	3.564750	4.381250	1.643000	3.363250	3.624500	20.114625
50%	6.227980e+05	5.386250	5.117500	2.810750	4.526875	4.248000	22.150750
75%	8.277522e+05	6.538625	5.881500	3.669625	4.909750	5.834000	23.729063
max	8.175133e+06	7.262000	6.578000	6.675000	5.519250	8.624500	26.554250

Here we can view a description of each category to compare the average scores for our categories of interest. We can determine the amount of deviation from each scores average, as well as the lowest and highest scores in each category, along with the quartiles. We can display this information in a box plot as well to get a better visualization of our statistical analysis.



We can do a statistical analysis to compare the differences between the average scores of interest, and determine which scores vary and which tend to stay consistent across cities. For example, we can see here that Housing tends to vary the most, and that Travel tends to get the lowest score consistently. We can also see that Cost of living and Education tend to score higher, while Cost of living and Commute vary the least in their scores.

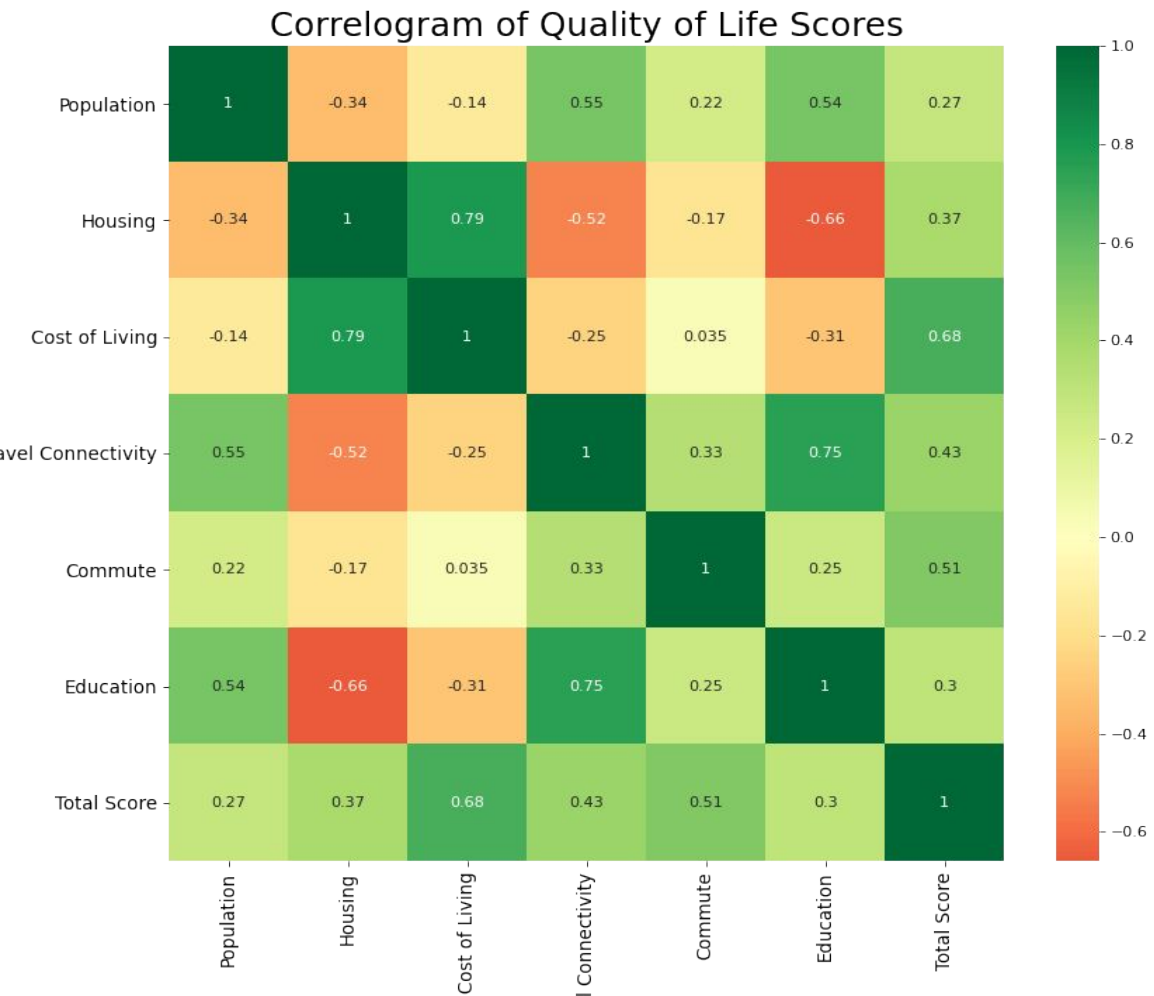
Lower quartile of **Housing** is: **3.56**
Upper quartile of Housing is: **6.54**
Interquartile range of Housing is: **2.98**
Median of Housing is: **5.39**
Values below **-0.91** could be **outliers**
Values above **11.01** could be **outliers**

Lower quartile of **Cost of Living** is: **4.38**
Upper quartile of cost of living is: **5.88**
Interquartile range of cost of living is: **1.5**
Median of cost of living is: **5.39**
Values below **2.13** could be **outliers**
Values above **8.13** could be **outliers**

Lower quartile of **Travel Connectivity** is: **1.64**
Upper quartile of travel connectivity is: **6.54**
Interquartile range of travel connectivity is: **4.9**
Median of travel connectivity is: **2.81**
Values below **-5.71** could be **outliers**
Values above **13.89** could be **outliers**

Lower quartile of **Commute** is: **3.36**
Upper quartile of commute is: **4.91**
Interquartile range of commute is: **1.55**
Median of commute is: **4.53**
Values below **1.03** could be **outliers**
Values above **7.24** could be **outliers**

Lower quartile of **Education** is: **3.62**
Upper quartile of education is: **5.83**
Interquartile range of education is: **2.21**
Median of education is: **4.25**
Values below **0.30** could be **outliers**
Values above **9.15** could be **outliers**

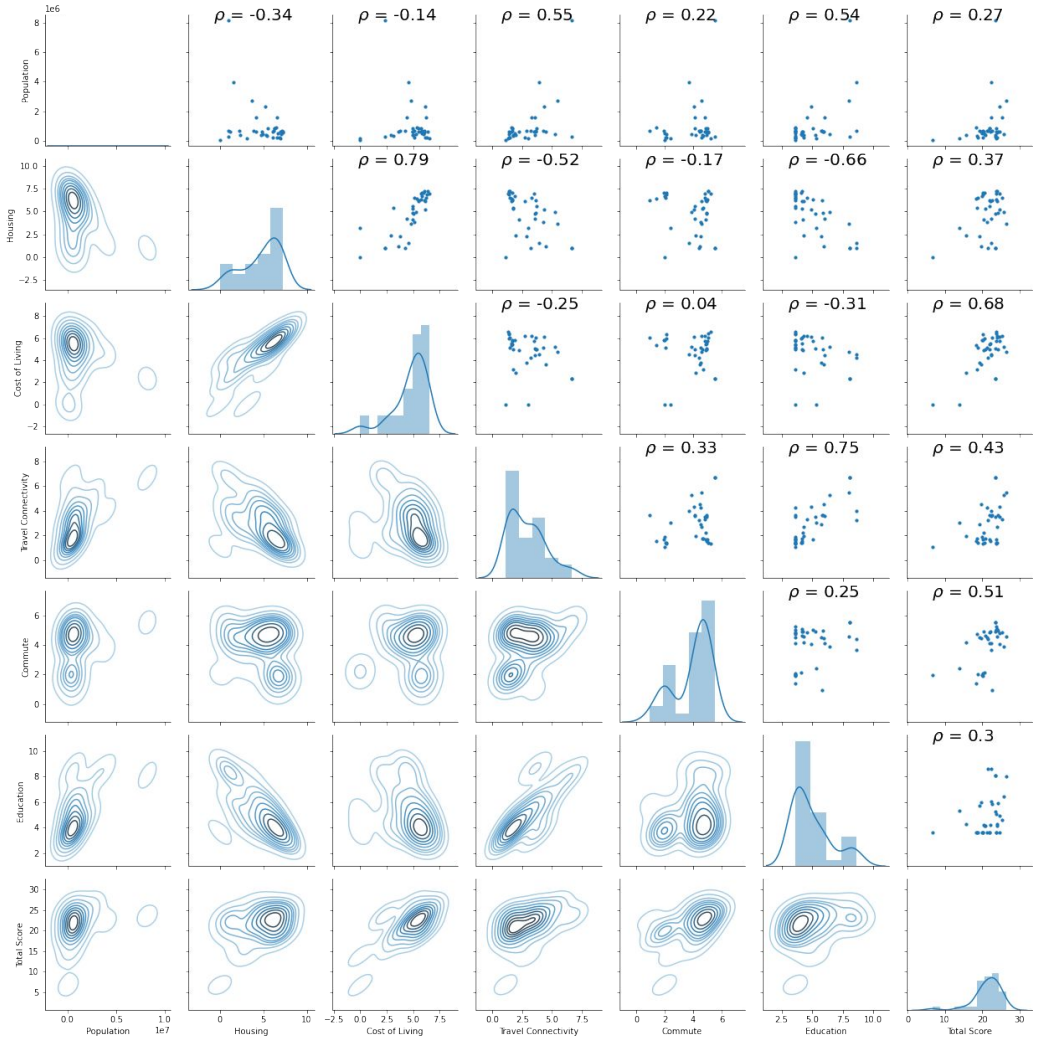


We can also compare the relationships between our scores of our choosing to see which scores tend to rate closely to each other across the major U.S. cities.

Here we can see that cities' Housing and Cost of Living scores tend to correlate, and score similarly across cities, while Education and Housing scores tend not to have any sort of relationship among cities.

There does seem to be a correlation, however, between Education and Travel Connectivity score consistency according to this graph.

Another interesting trend to consider is that Cost of Living tends to be the highest influencer in each city's total score among these categories!



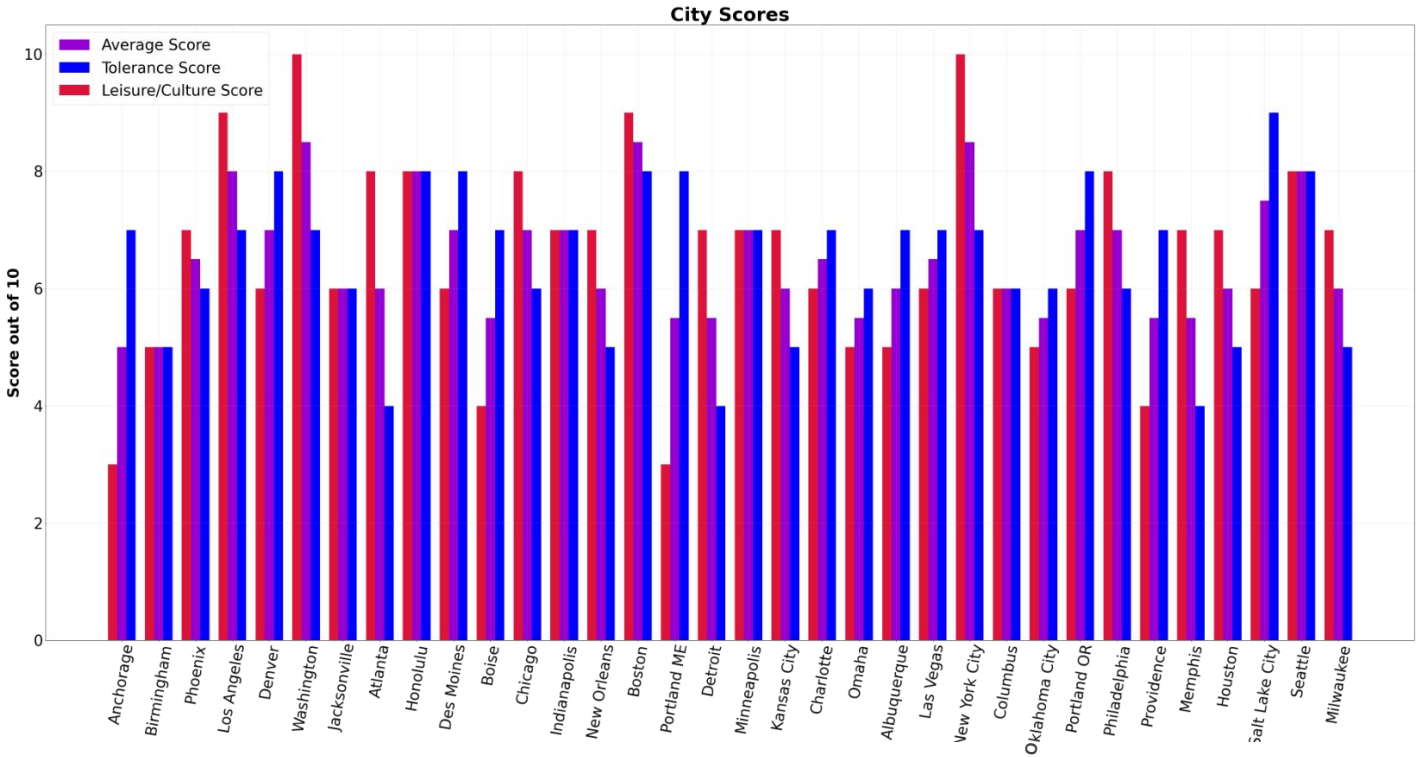
Similar to the previous graph, these show relationships between each score as well as population and total score.

A correlation coefficient has been added here for statistical observation purposes.

This is another way to view the previous findings while viewing each point on a graph for a closer look into the trends.

For example, we can see that Travel Connectivity scores tend to increase along with Education scores, as well as Housing scores with Cost of Living Scores.

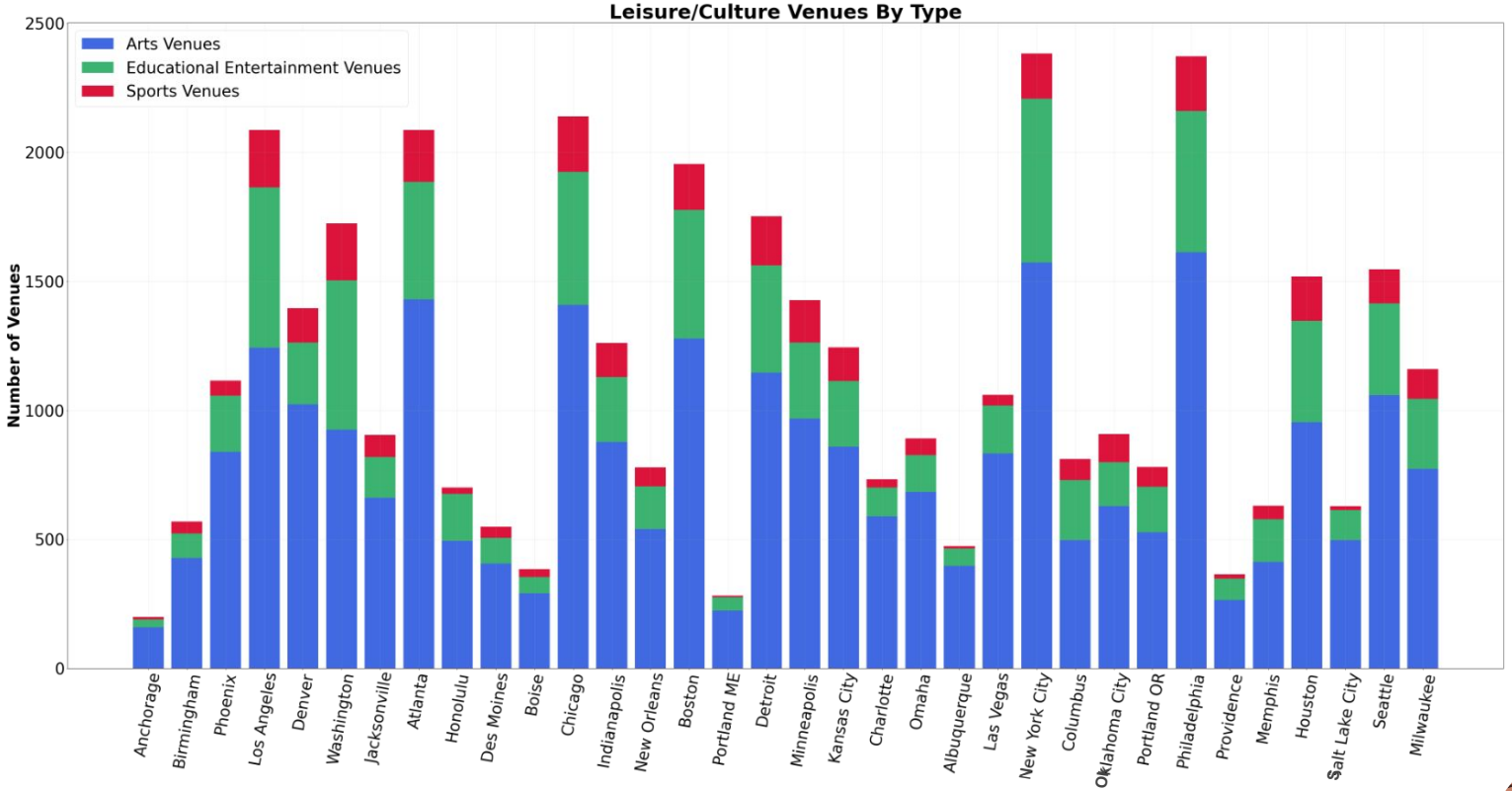
We can also see that Cost of Living scores greatly impact the total score, but also that a decent amount of cities have higher Cost of Living Scores in general.



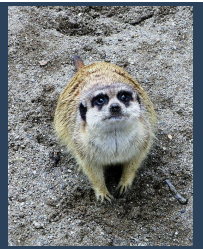
Leisure and Culture-
Based on available sites/venues for leisure/entertainment and culture

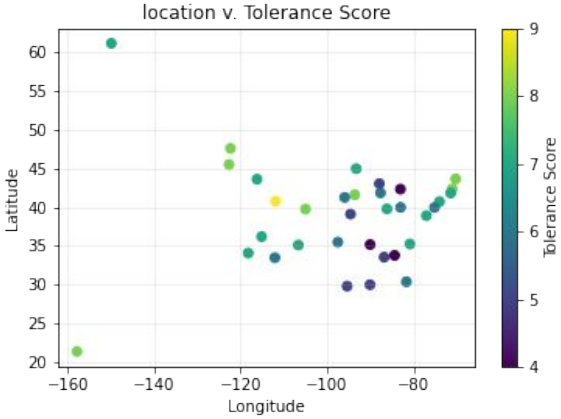
Tolerance-
Based on various LGBT rights and tolerance of minorities

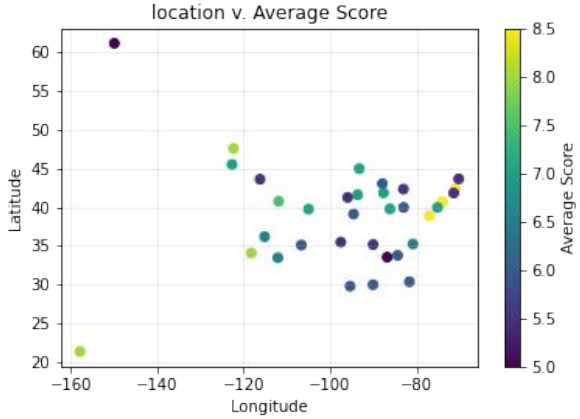
K

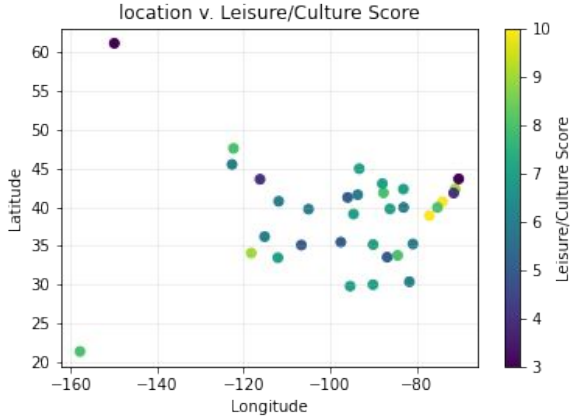


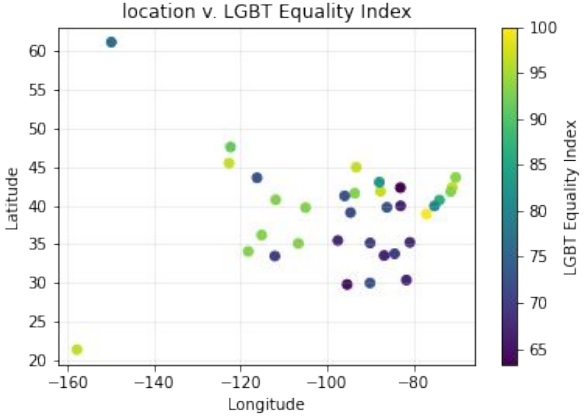
Broke down the amount of venues by type to see how the different categories would compare to each other and across all our cities

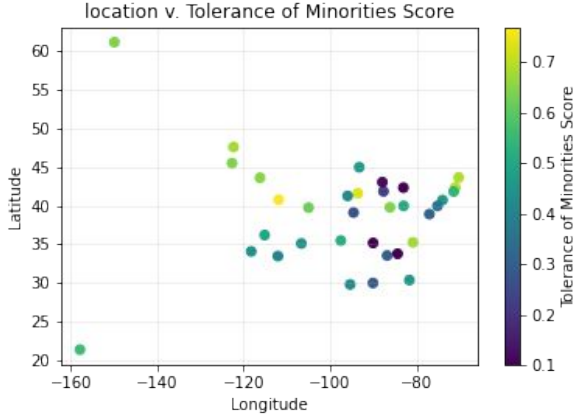


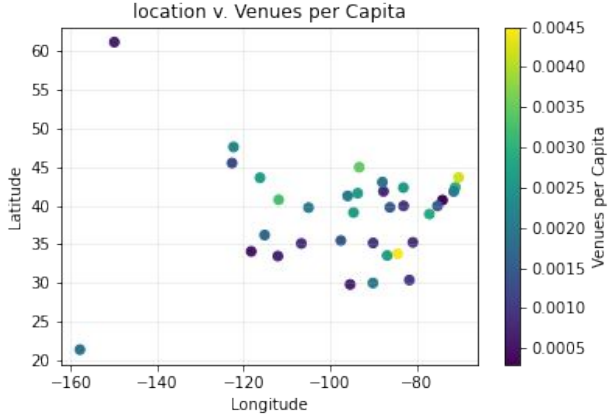

location v. Tolerance Score
Latitude
Longitude
Tolerance Score

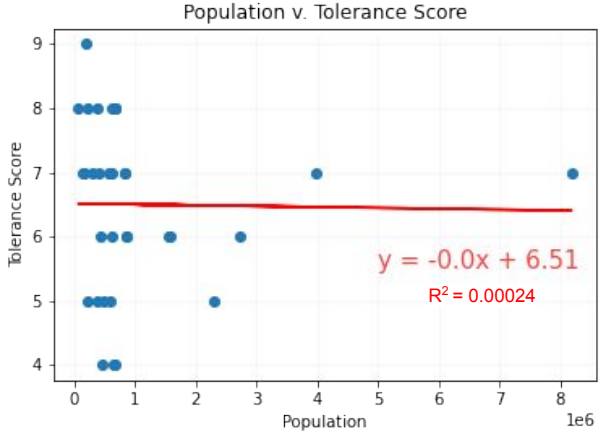

location v. Average Score
Latitude
Longitude
Average Score

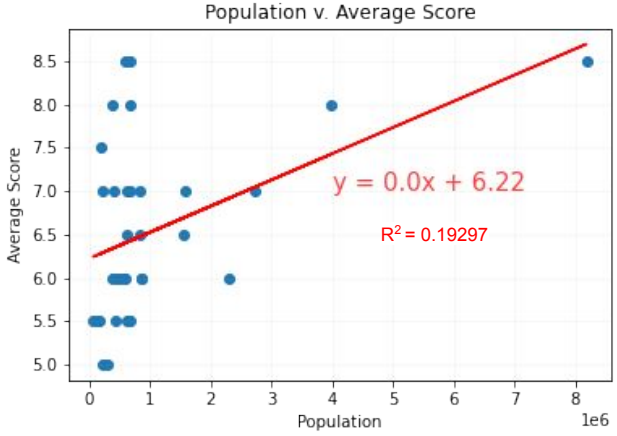

location v. Leisure/Culture Score
Latitude
Longitude
Leisure/Culture Score

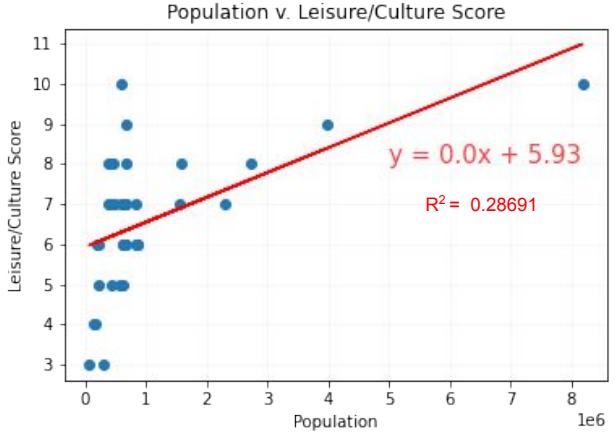

location v. LGBT Equality Index
Latitude
Longitude
LGBT Equality Index

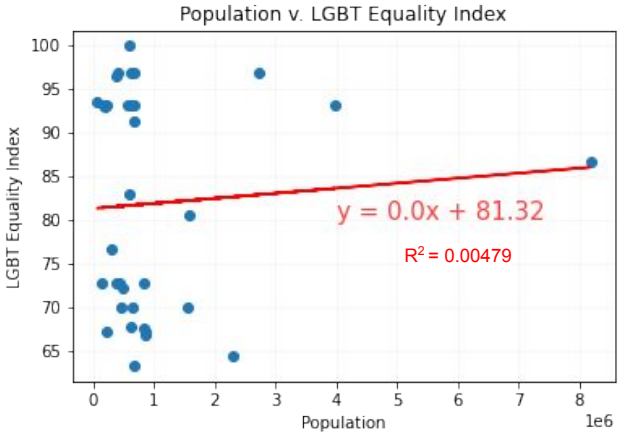

location v. Tolerance of Minorities Score
Latitude
Longitude
Tolerance of Minorities Score

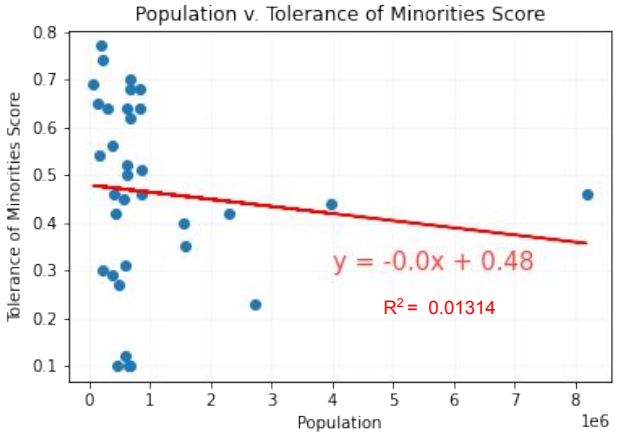

location v. Venues per Capita
Latitude
Longitude
Venues per Capita

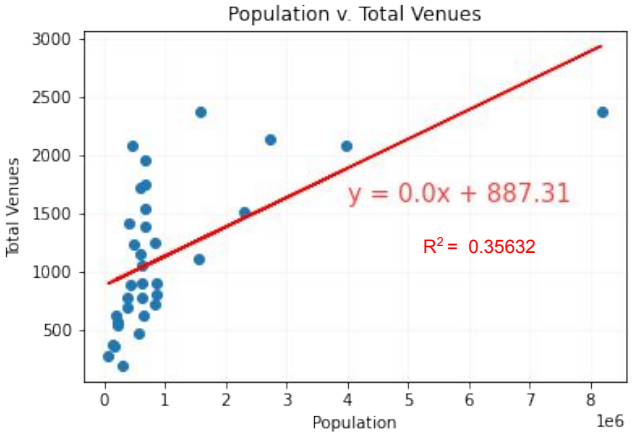

Population v. Tolerance Score
Tolerance Score
Population
1e6
y = -0.0x + 6.51
R² = 0.00024


Population v. Average Score
Average Score
Population
1e6
y = 0.0x + 6.22
R² = 0.19297


Population v. Leisure/Culture Score
Leisure/Culture Score
Population
1e6
y = 0.0x + 5.93
R² = 0.28691


Population v. LGBT Equality Index
LGBT Equality Index
Population
1e6
y = 0.0x + 81.32
R² = 0.00479


Population v. Tolerance of Minorities Score
Tolerance of Minorities Score
Population
1e6
y = -0.0x + 0.48
R² = 0.01314

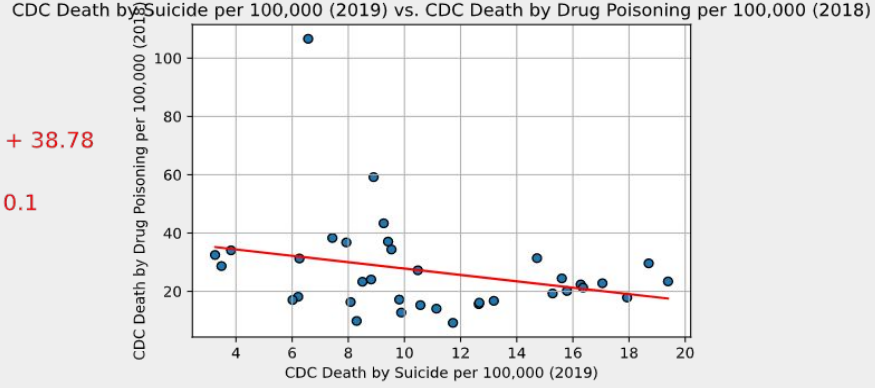

Population v. Total Venues
Total Venues
Population
1e6
y = 0.0x + 887.31
R² = 0.35632

Deaths of Despair and Teleport's Quality of Life Algorithm

Paul

The Problem

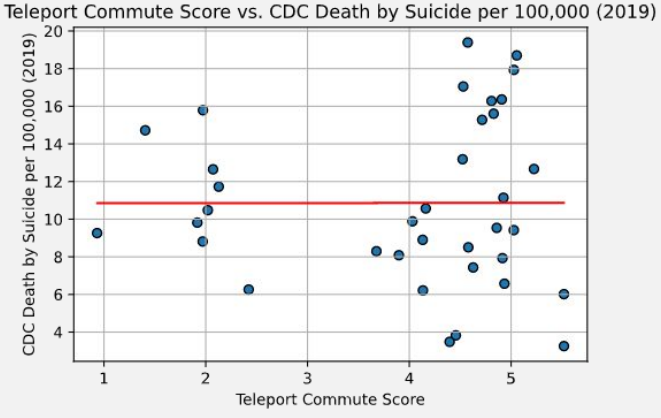
y = -1.1x + 38.78
ρ =-0.28
pvalue = 0.1



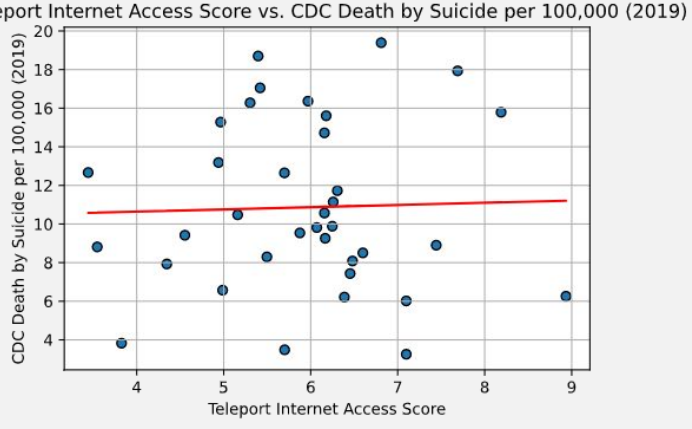
	CDC Death by Suicide per 100,000 (2019)	CDC Death by Drug Poisoning per 100,000 (2018)
count	36.000000	36.000000
mean	10.861933	26.874916
std	4.405619	17.169873
min	3.254169	9.193817
25%	8.044993	16.979066
50%	9.851847	23.030857
75%	14.860890	31.663485
max	19.389247	106.635164

Death by Suicide

y = 0.01x + 10.84
ρ =0.0
pvalue = 0.99

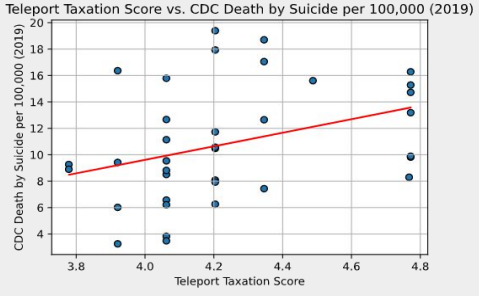


y = 0.12x + 10.18
ρ =0.03
pvalue = 0.86

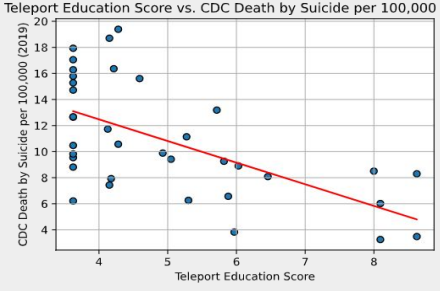


Death by Suicide

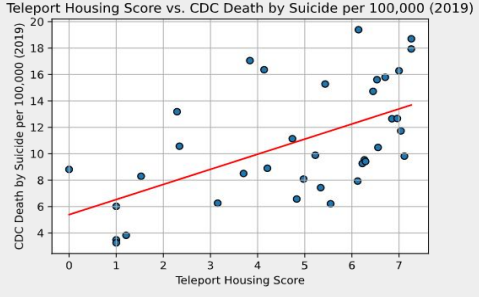
y = 5.13x + -10.91
ρ =0.35
pvalue = 0.03



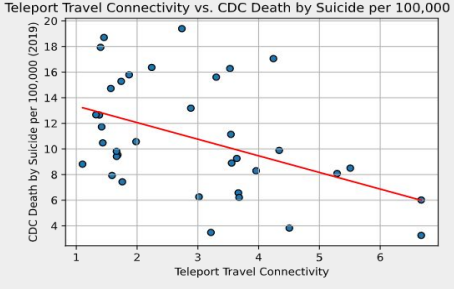
y = -1.66x + 19.12
ρ =-0.6
pvalue = 0.0



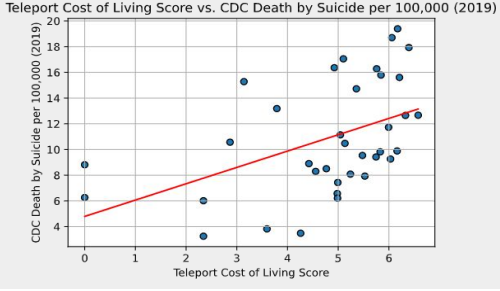
y = 1.14x + 5.39
ρ =0.57
pvalue = 0.0



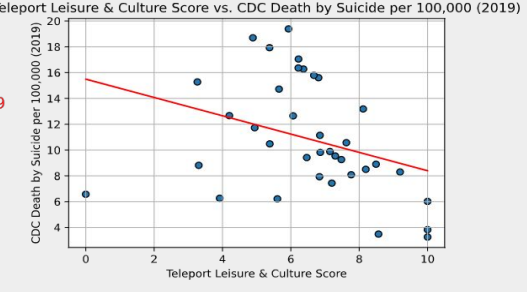
y = -1.3x + 14.65
ρ =-0.45
pvalue = 0.01



y = 1.27x + 4.79
ρ =0.47
pvalue = 0.0

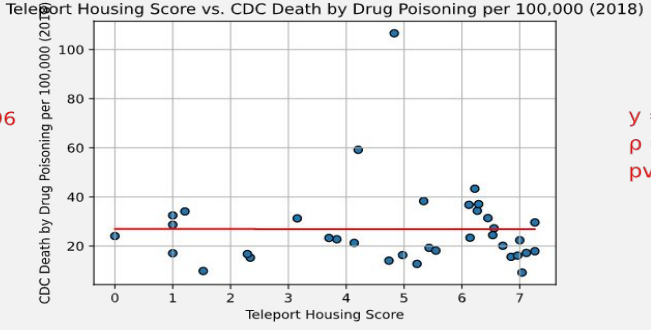


y = -0.71x + 15.49
ρ =-0.33
pvalue = 0.05

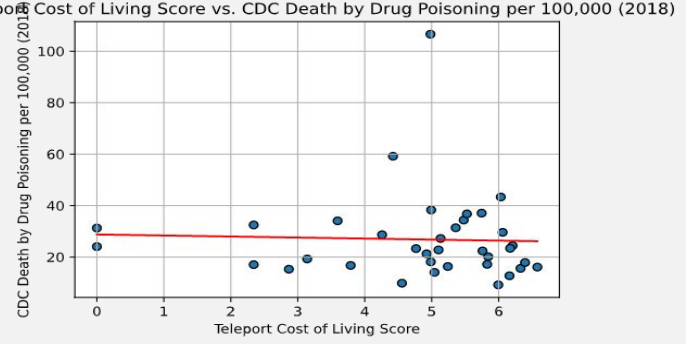


Death by Drug Poisoning

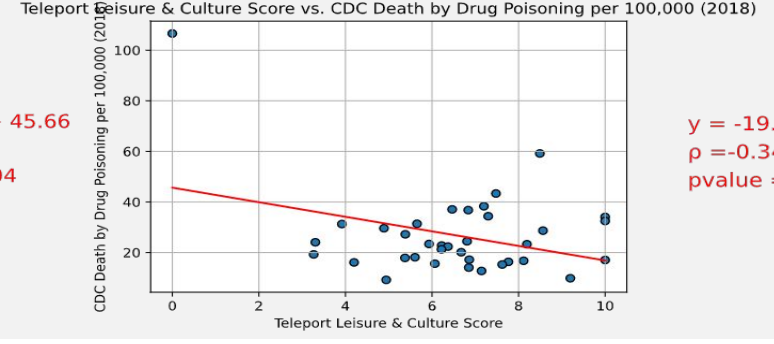
y = -0.02x + 26.96
ρ =-0.0
pvalue = 0.99



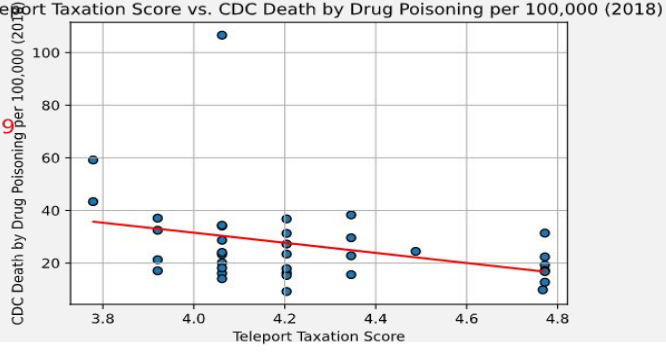
y = -0.4x + 28.78
ρ =-0.04
pvalue = 0.83



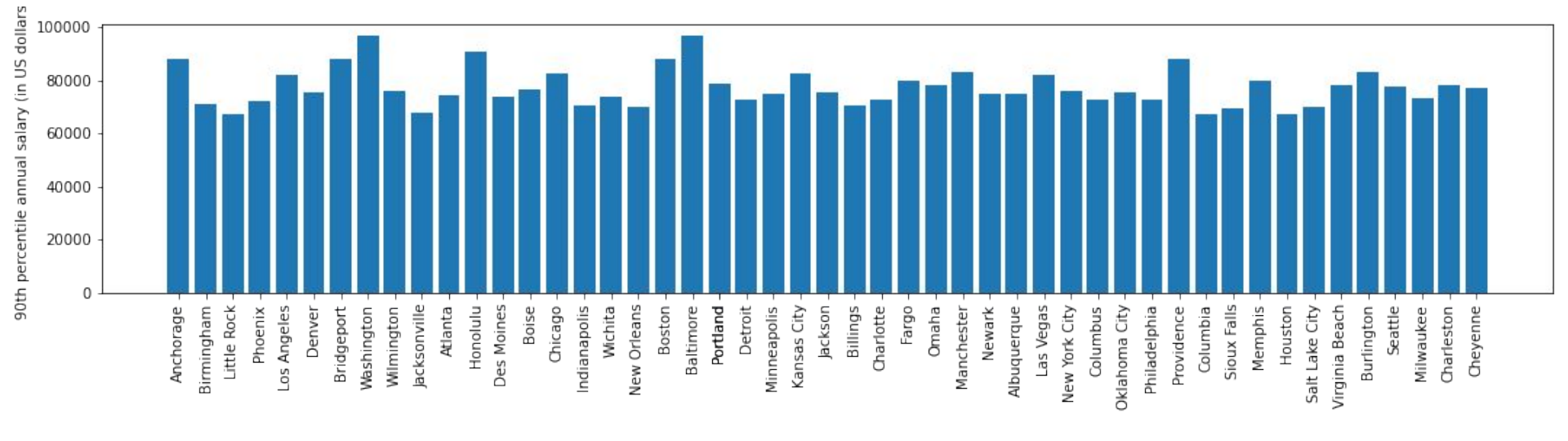
y = -2.88x + 45.66
ρ =-0.34
pvalue = 0.04

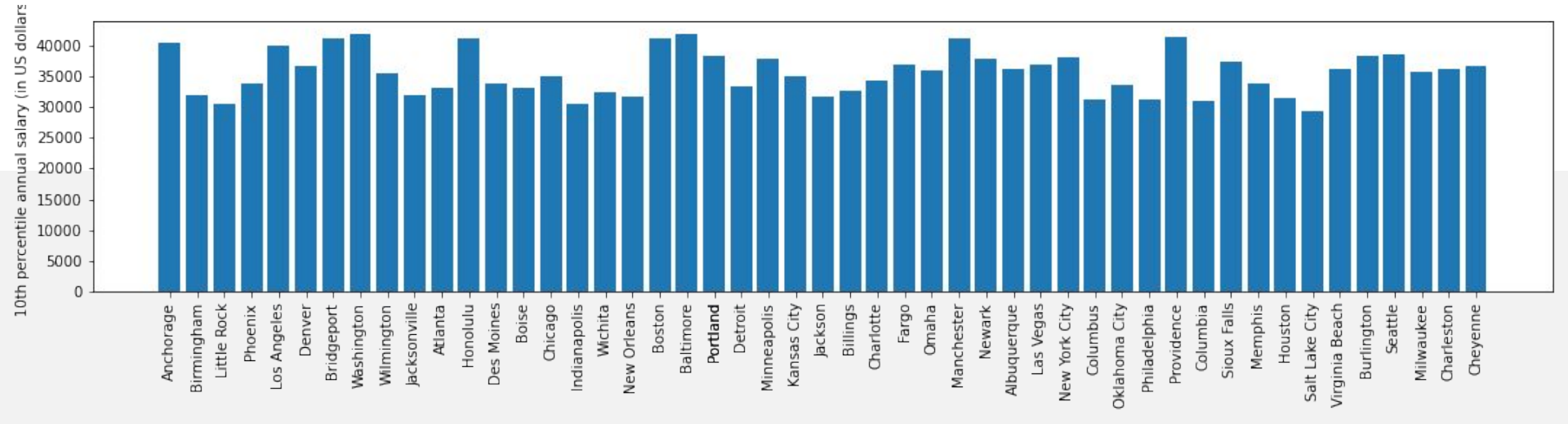


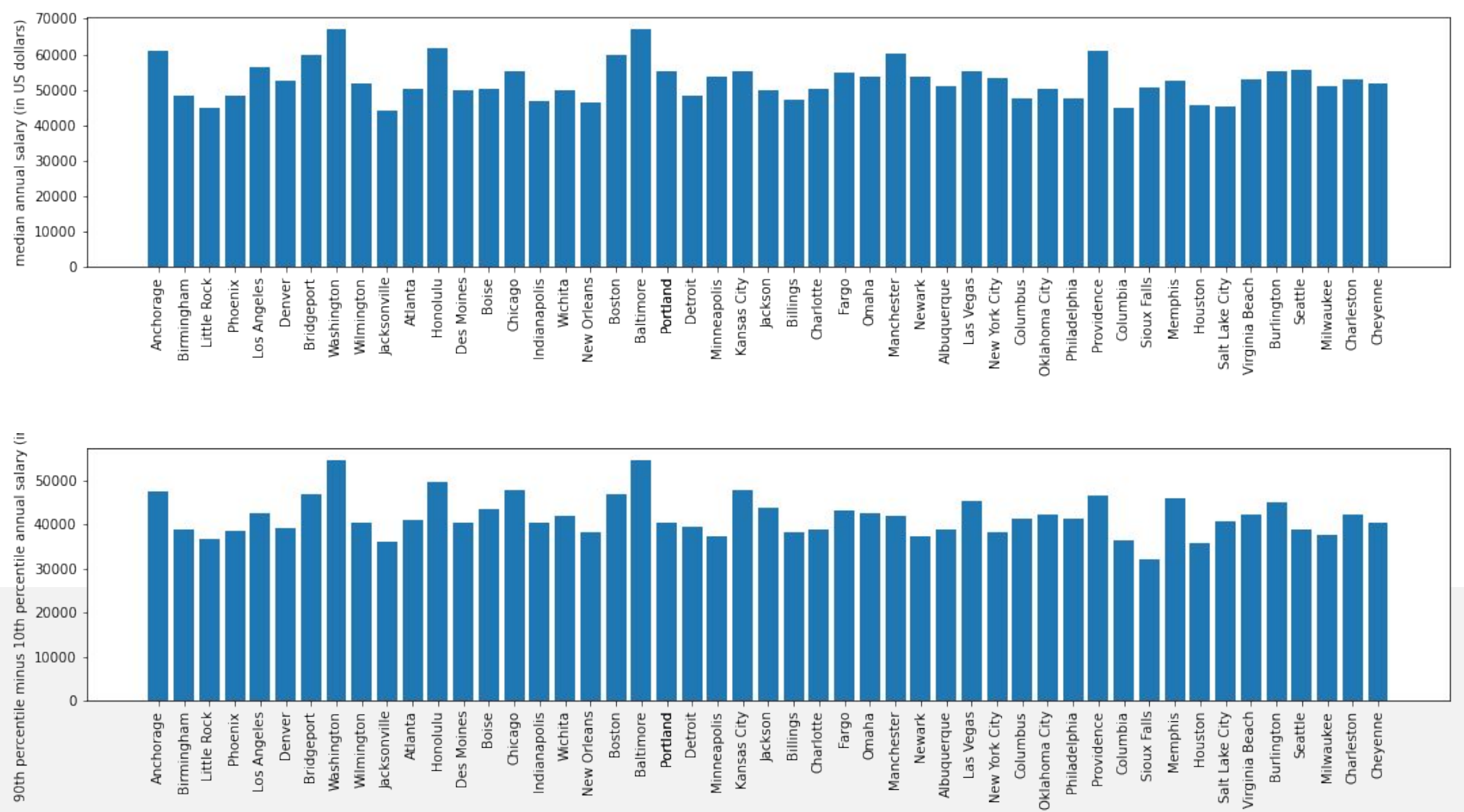
y = -19.16x + 108.19
ρ =-0.34
pvalue = 0.04

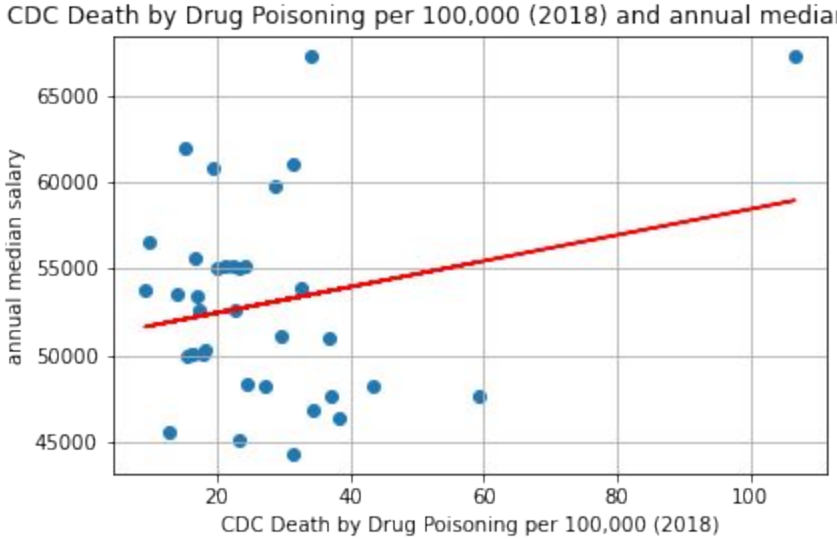


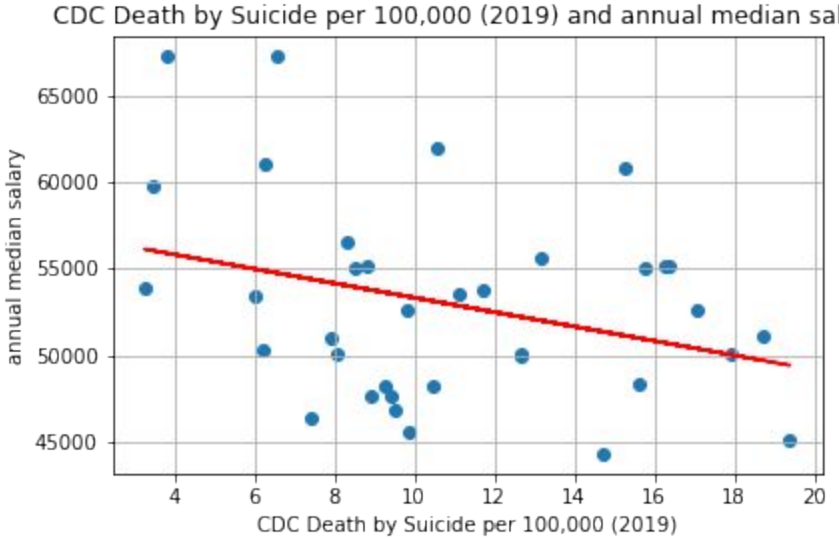
Income


90th percentile annual salary (in US dollars
100000
80000
60000
40000
20000
0
Anchorage
Birmingham
Little Rock
Phoenix
Los Angeles
Denver
Bridgeport
Washington
Wilmington
Jacksonville
Atlanta
Honolulu
Des Moines
Boise
Chicago
Indianapolis
Wichita
New Orleans
Boston
Baltimore
Portland
Detroit
Minneapolis
Kansas City
Jackson
Billings
Charlotte
Fargo
Omaha
Manchester
Newark
Albuquerque
Las Vegas
New York City
Columbus
Oklahoma City
Philadelphia
Providence
Columbia
Sioux Falls
Memphis
Houston
Salt Lake City
Virginia Beach
Burlington
Seattle
Milwaukee
Charleston
Cheyenne


10th percentile annual salary (in US dollars
40000
35000
30000
25000
20000
15000
10000
5000
0
Anchorage
Birmingham
Little Rock
Phoenix
Los Angeles
Denver
Bridgeport
Washington
Wilmington
Jacksonville
Atlanta
Honolulu
Des Moines
Boise
Chicago
Indianapolis
Wichita
New Orleans
Boston
Baltimore
Portland
Detroit
Minneapolis
Kansas City
Jackson
Billings
Charlotte
Fargo
Omaha
Manchester
Newark
Albuquerque
Las Vegas
New York City
Columbus
Oklahoma City
Philadelphia
Providence
Columbia
Sioux Falls
Memphis
Houston
Salt Lake City
Virginia Beach
Burlington
Seattle
Milwaukee
Charleston
Cheyenne


median annual salary (in US dollars)
70000
60000
50000
40000
30000
20000
10000
0
Anchorage
Birmingham
Little Rock
Phoenix
Los Angeles
Denver
Bridgeport
Washington
Wilmington
Jacksonville
Atlanta
Honolulu
Des Moines
Boise
Chicago
Indianapolis
Wichita
New Orleans
Boston
Baltimore
Portland
Detroit
Minneapolis
Kansas City
Jackson
Billings
Charlotte
Fargo
Omaha
Manchester
Newark
Albuquerque
Las Vegas
New York City
Columbus
Oklahoma City
Philadelphia
Providence
Columbia
Sioux Falls
Memphis
Houston
Salt Lake City
Virginia Beach
Burlington
Seattle
Milwaukee
Charleston
Cheyenne
90th percentile minus 10th percentile annual salary (i
50000
40000
30000
20000
10000
0
Anchorage
Birmingham
Little Rock
Phoenix
Los Angeles
Denver
Bridgeport
Washington
Wilmington
Jacksonville
Atlanta
Honolulu
Des Moines
Boise
Chicago
Indianapolis
Wichita
New Orleans
Boston
Baltimore
Portland
Detroit
Minneapolis
Kansas City
Jackson
Billings
Charlotte
Fargo
Omaha
Manchester
Newark
Albuquerque
Las Vegas
New York City
Columbus
Oklahoma City
Philadelphia
Providence
Columbia
Sioux Falls
Memphis
Houston
Salt Lake City
Virginia Beach
Burlington
Seattle
Milwaukee
Charleston
Cheyenne





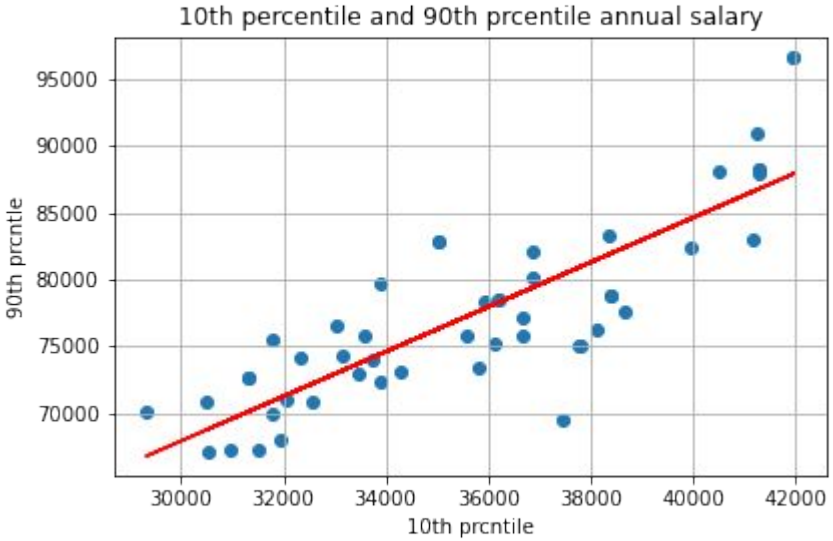
Neither negative nor positive quality of life metrics correlated strongly with the median salary of a city.

This was also true when selecting only the top or bottom 10 percent of earners in each city

Scores from https://developers.teleport.org/api/

(the absolute value of all presented correlation coefficients were less than 0.2)

One Exception



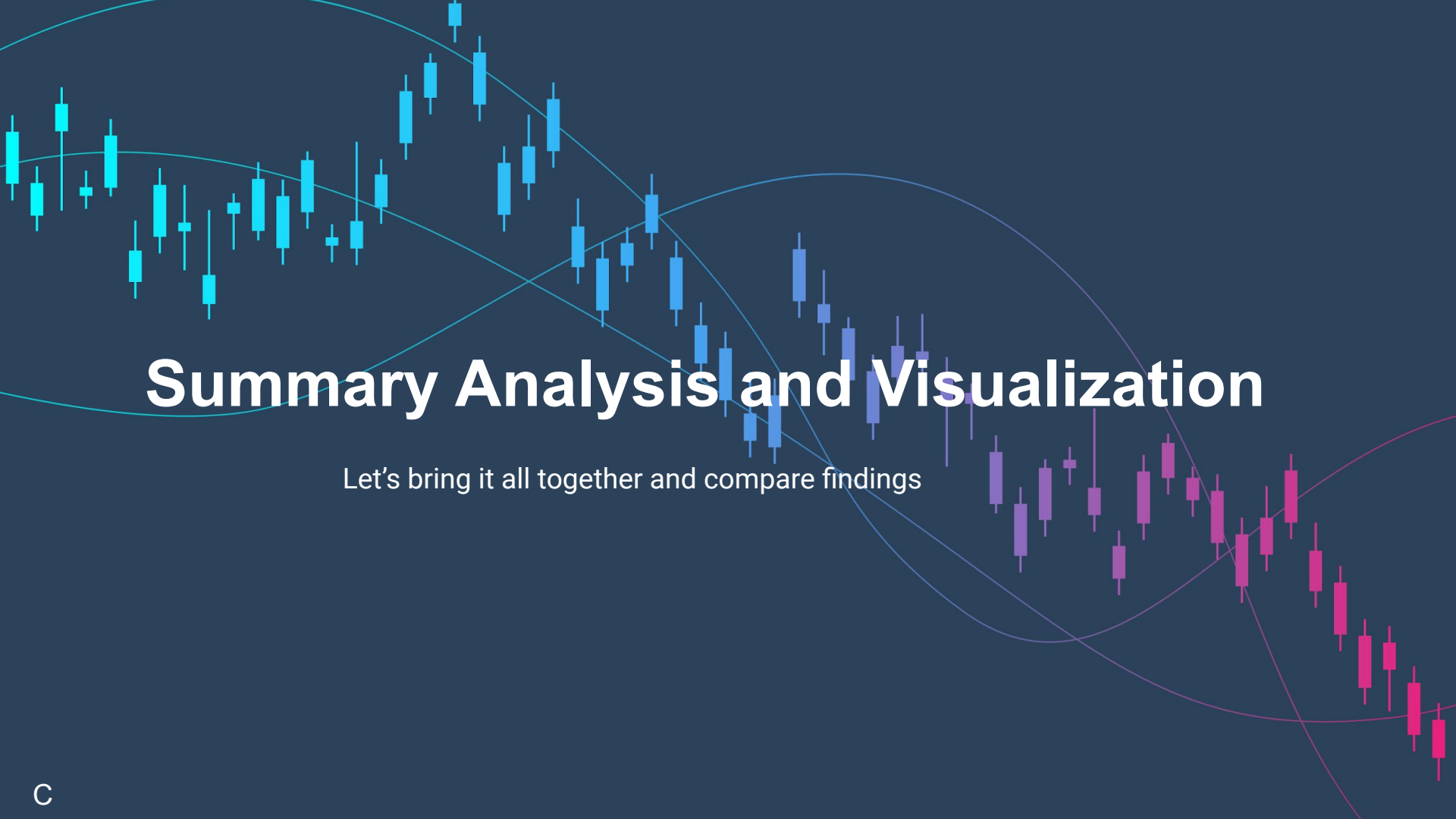
(correlation coefficient of 0.700736998634241)

(when the median salary is paired with either the 90th or the 10th percentile, the correlation is even stronger (0.93 and 0.85) as would be expected)

Income Conclusion:

In the largest U.S. cities in each state, quality of life does not appear to be affected by the annual earnings of its populus.

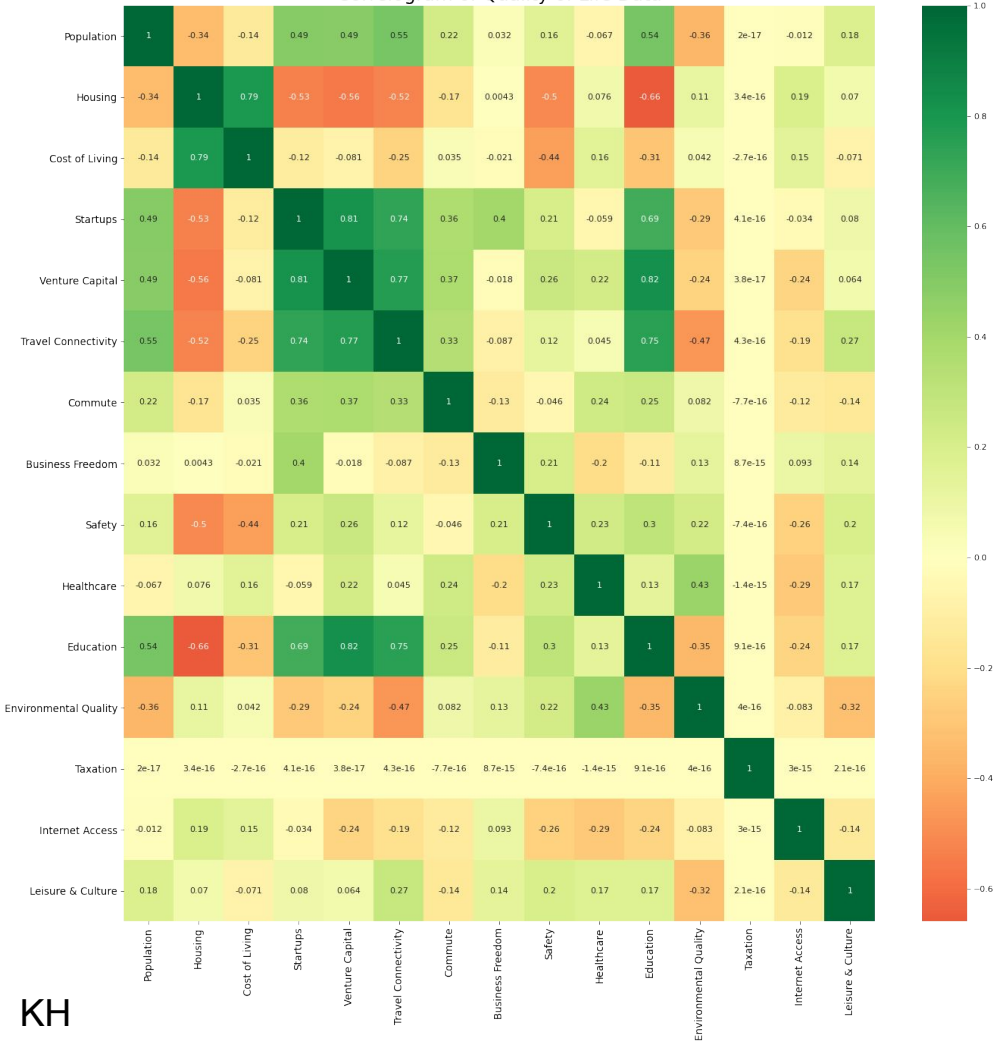
(and the income gap is relatively stable from city to city)



Summary Analysis and Visualization

Let’s bring it all together and compare findings

C

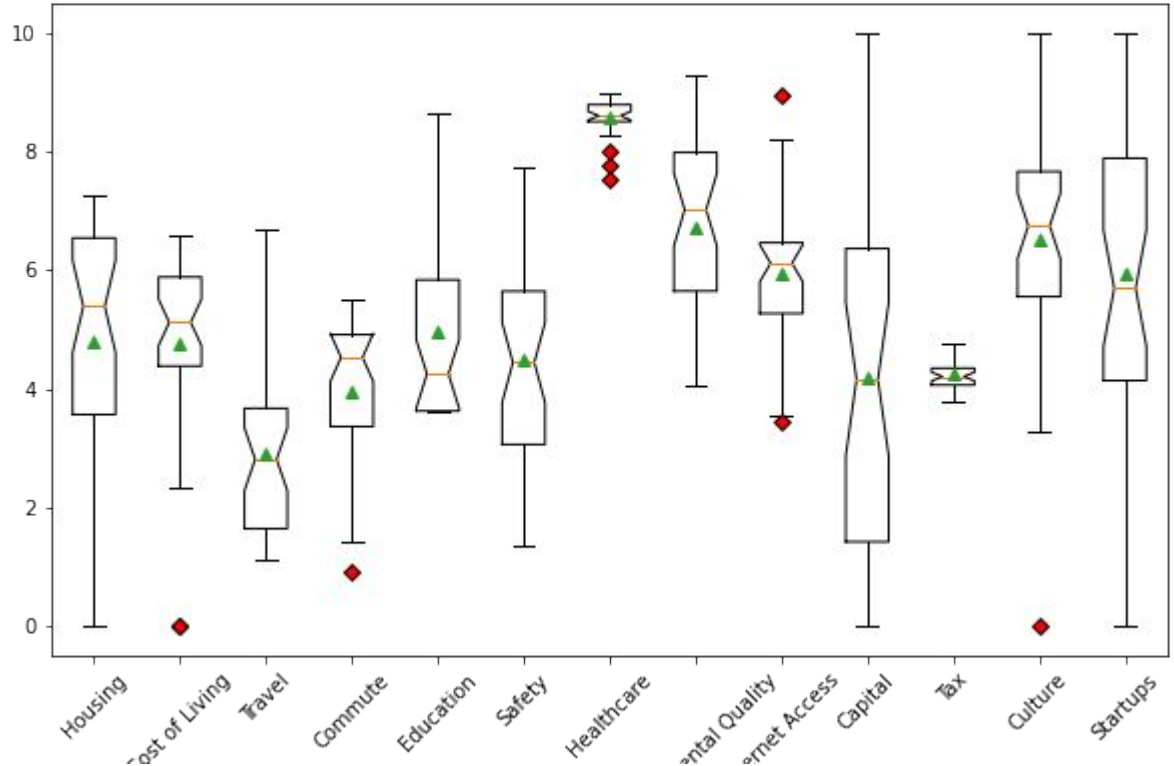


Referring back to some of the analysis we did among our varying categories of interest, we can take a look at the broader range of categories and their quality of life scores to see if we can find any sort of trends among the cities.

Here we can see the the categories that correlate most tend to be: Housing and Cost of Living, and Venture Capital with both Startups and Travel Connectivity as well as Education.

The scores that seems to correlate least among cities are Education and Housing.

KH

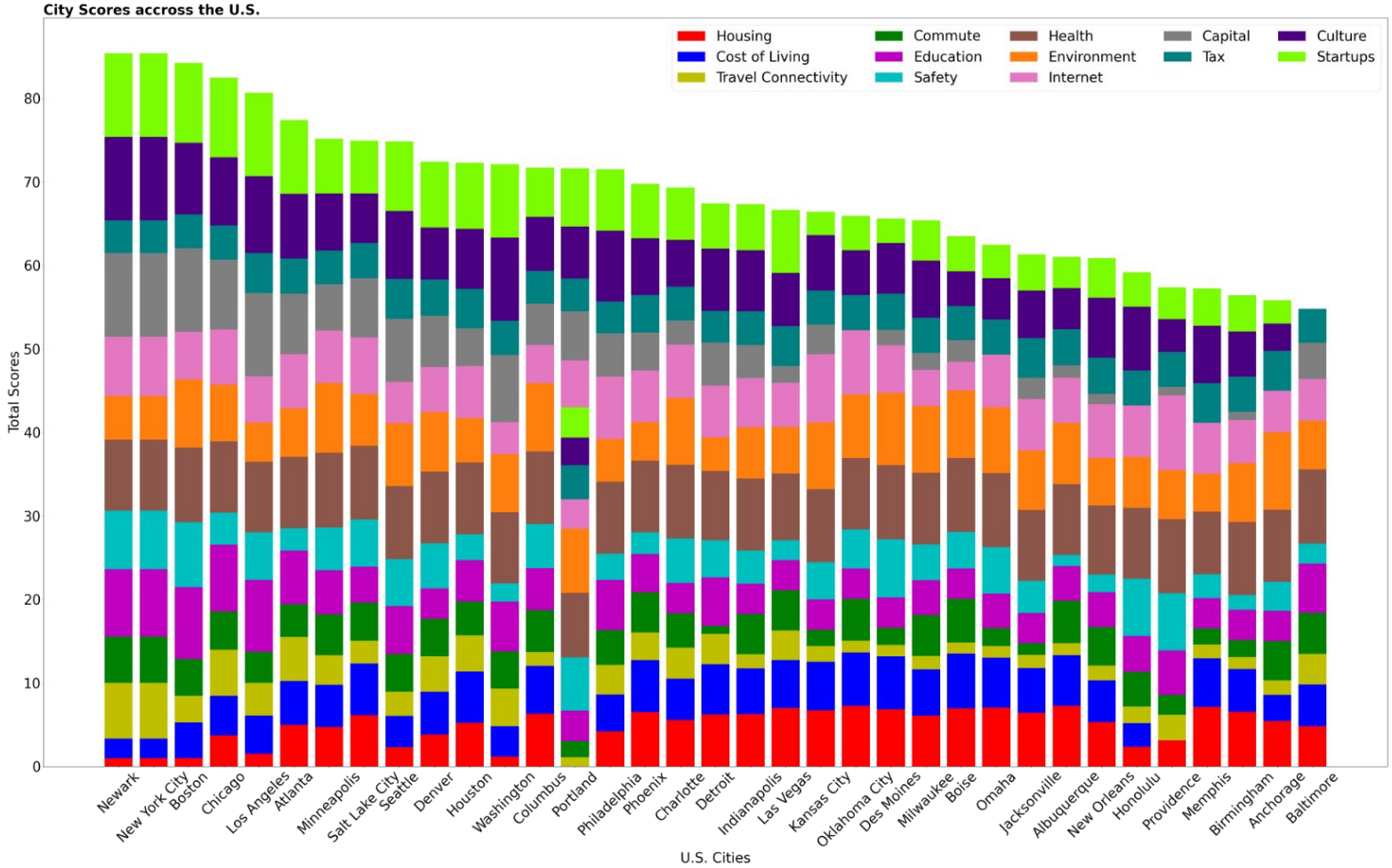


Comparing all of the categories we can find which scores tend to vary most and least. We can see that scores vary greatly among U.S. cities for Capital and Startups, while scores tend to be about the same among major U.S. cities for Healthcare and Taxation.

This is something that could help determine which categories may be the most important to delve deeper into when deciding on what area of quality of life you may be most interested in when searching for a city to move or work in.

Comparing these broad categories also leaves one wondering what specifically may cause differences in variation of scores among categories such as capital and healthcare.

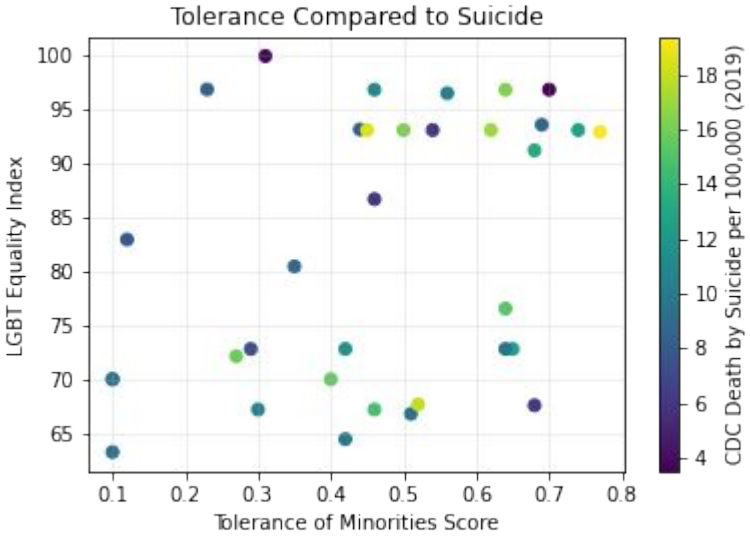
KH

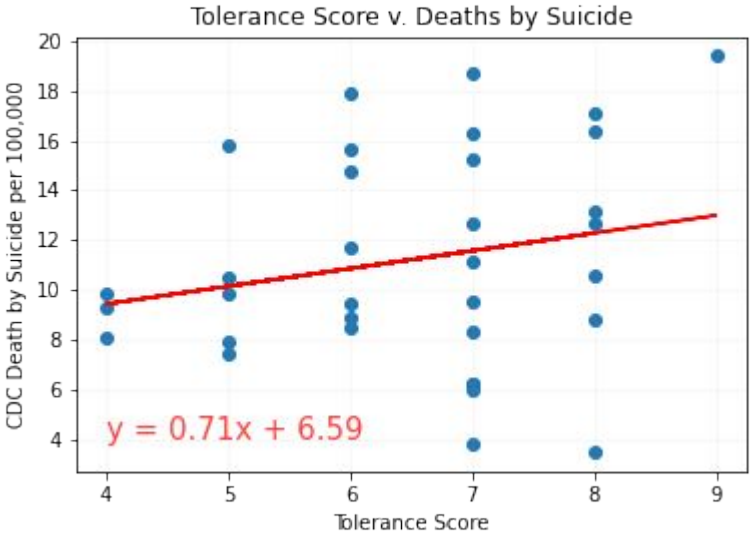


If we do not have any specific categories of interest and want to consider the overall total Quality of Life Scores we can view this bar graph to find the highest vs. lowest rated cities within the city list we provided.

KH

Keep in mind we do not have data listed here on all cities and states within the U.S. This is something we would be interested in adding in the future.





Comparisons of Teleport tolerance data to CDC deaths by suicide (2019)

K

After Thoughts

- Commercial data???
- County vs City
- Modules are amazing.
- Excited to do real work.

P

Implications

- Major U.S. cities tend to be relatively consistent in most quality of life metrics
- Most scores do not seem to change based on geographical location or population, when they are per capita

C

Things to Consider

- Teleport is not clear on how their scores for each category are calculated which can make comparisons slightly unclear
- It’s also important to note that user surveys have an impact on Telegram scores which could easily cause bias in the data
- The data is not all from the same time, they’re close but things can change in those gaps
- City data is only from a few big cities

K


FIN