

Your presentation must include:

Build on Project Proposal

Build on your project proposal (from Milestone 1) that described the client or dataset you chose, the approach you were going to take, your initial hypotheses, and your initial approach. Include descriptive stats and any visualizations from your data exploration. You want to highlight key learnings from your data exploration and any aha's or changes to your plan as a results of your findings:

- Include Client/Hypotheses/Approach
- Include artifacts from previous modules
- Include results (good and bad paths); Correlations / regressions
- Graphics / Visualizations

In this project I looked at the relationships between country data (population, GDP per capita, and urbanization) and olympic success (measured by medals won in olympic games). This analysis could be used by national olympic committees in trying to find routes to olympic success, or by the casual sports analyst or fan.

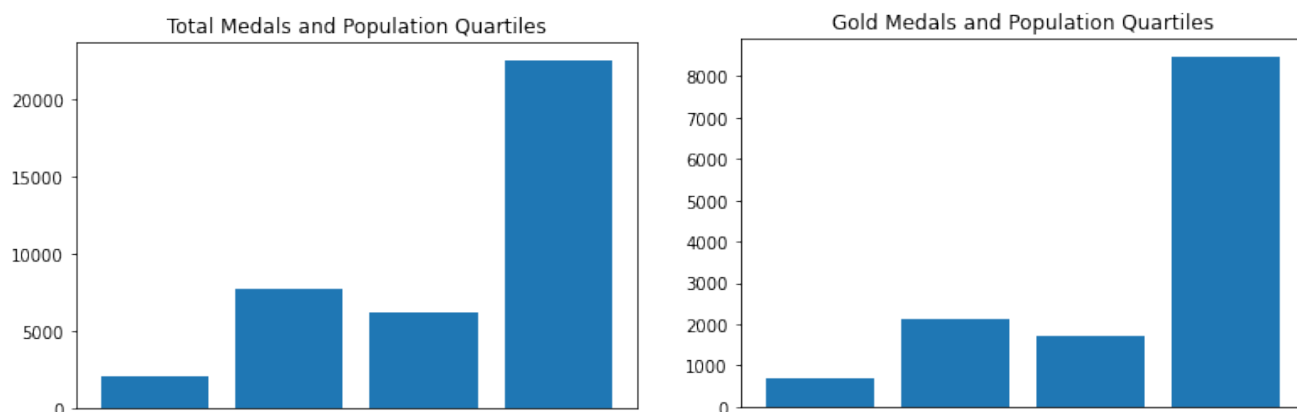
Before looking at the data, I put forward three basic hypotheses:

- 1) More populous countries would have more olympic success.
- 2) More urbanized countries would be more successful.
- 3) Richer countries (defined by having a greater per capita gross domestic product [GDP]) would be more successful.

Using total medals won and total gold medals won as metrics for success (stay tuned) here is how these hypotheses fared:

1) More populous countries would have more olympic success.

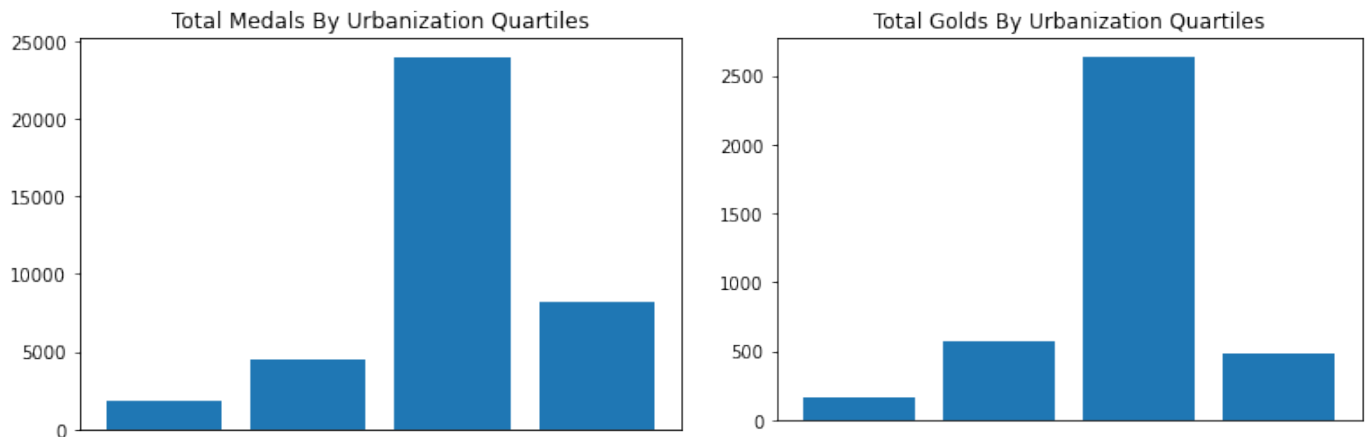
This one fared reasonably well.



The countries in the highest population quartile dominated the gold and total medal counts while the least populous countries won the least. Interestingly the third quartile performed better than the second though...

2) More urbanized countries would be more successful.

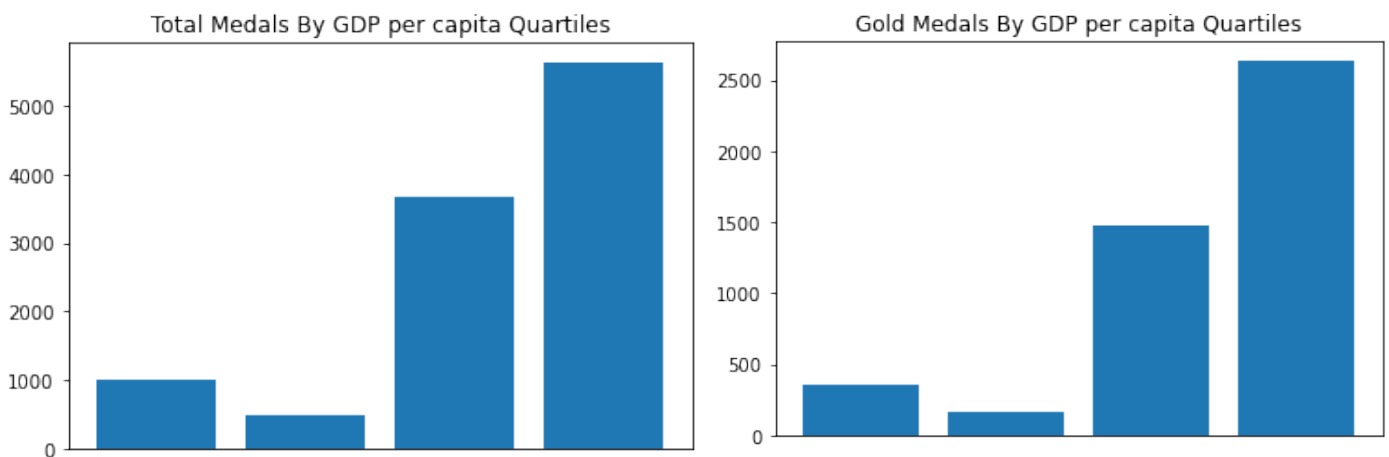
This hypothesis gloriously failed.



The countries in the second quartile of urbanization (defined as what percentage of the population lives in urban areas) dominated the gold and total medal count. With the first and third quartiles comparable to each other.

3) Richer countries (defined by having a greater per capita gross domestic product [GDP]) would be more successful.

This one held up pretty well, with a twist.



As you can see, the countries in the top two quartiles of Gross Domestic Product per capita cleaned up. Though interestingly, the poorest countries have done a little better than those a bit richer than them.

Discuss Insights Discovered

Discuss insights discovered (results from your diving deeper / going broader analysis). This is where you put your spin on what you've discovered

- Discuss your hypotheses and any direct outcomes from whether you were right or wrong. Did you change your hypotheses? Or create new ones?
- Discuss any metrics you created and why?
- Discuss discoveries about relationships in the data / themes discovered.

Before evaluating the results in terms of whether or not they are instructional, I decided to look at “success” from some different angles. I decided to explore two new metrics, “Medals per capita” and “Golds per capita,” to examine olympic success accounting for country size. The results were interesting.

	Medals_per_cap*1000000	Country
0	235.929431	Liechtenstein
1	190.288682	Norway
2	162.389302	Finland
3	151.925282	Sweden
4	117.543703	Hungary
...
84	0.113067	Ecuador
85	0.097842	Jordan
86	0.063697	Mozambique
87	0.041030	Vietnam
88	0.034217	Nepal

89 rows x 2 columns

	Golds_per_cap*1000000	Country
0	69.631289	Norway
1	52.428762	Liechtenstein
2	47.377741	Sweden
3	44.739101	Hungary
4	35.725647	Finland
...
84	0.037755	Cote d'Ivoire
85	0.034217	Nepal
86	0.031848	Mozambique
87	0.030257	Peru
88	0.010257	Vietnam

89 rows x 2 columns

So at the top we have some Scandinavian countries, Liechtenstein and Hungary. And at the bottom some reasonably populous, but poorer countries.

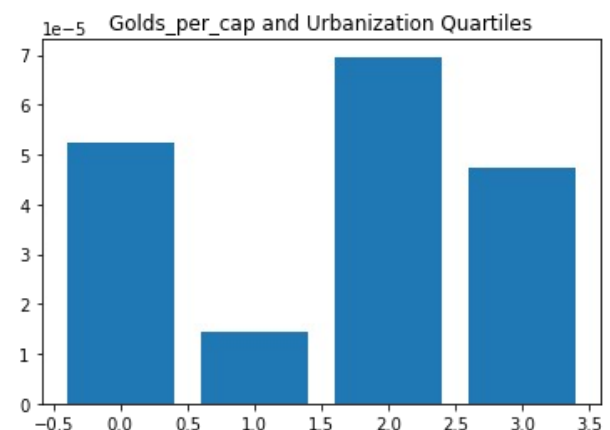
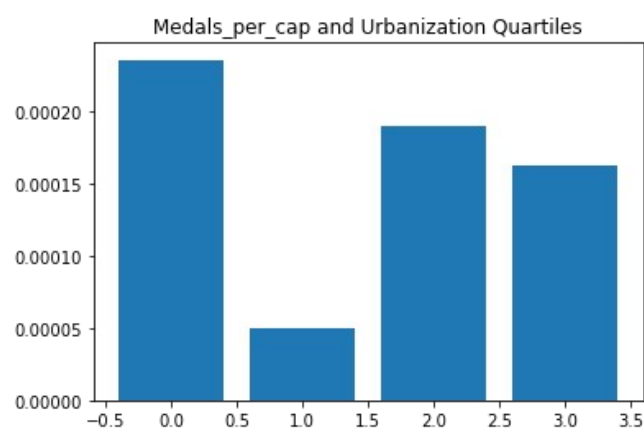
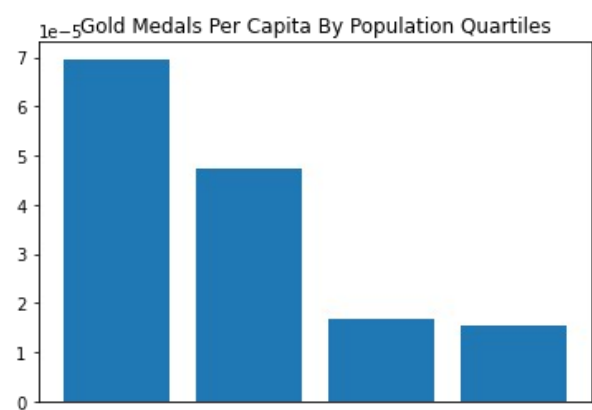
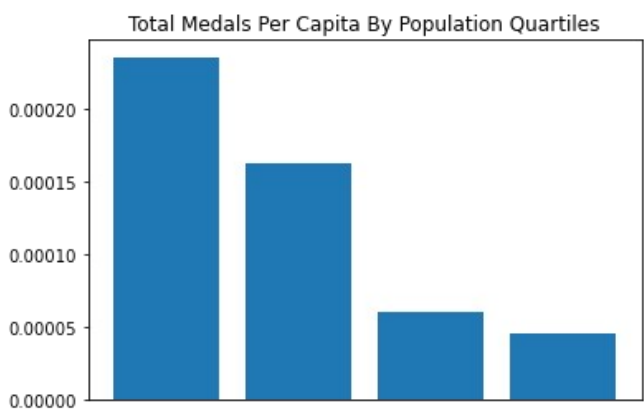
Before looking at how the main country metrics do when success is redefined, I decided to look at one more derived metric, olympic athletes per capita. I was curious what percentage of a country's population became olympic athletes. With this step the Liechtenstein story developed...

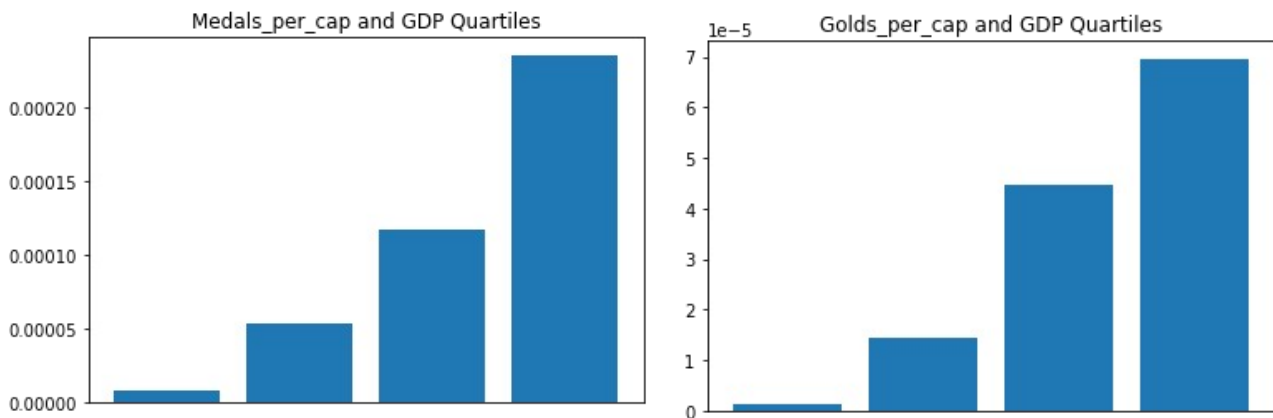
	Country	gold_per_100_athletes	medals_per_100_athletes	Athletes_per_cap
0	United States	13	29	0.000057
1	Russia	13	33	0.000074
2	India	9	13	0.000001
3	Germany	8	23	0.000188
4	Pakistan	7	21	0.000003
...
84	Slovenia	0	4	0.000534
85	Latvia	0	3	0.000505
86	Bahrain	0	2	0.000074
87	Luxembourg	0	0	0.001587
88	Liechtenstein	0	2	0.009673

89 rows x 4 columns

Liechtenstein sends almost one out of every thousand of its citizens to the olympics. Per athlete, they do quite poorly, but if they send enough competitors, they will do well occasionally. On the other side of the spectrum are India and Pakistan, very populous countries who send relatively small olympic contingents, but the athletes they do send bring home medals.

Let's go back to population, urbanization and GDP per capita, and see which metrics correlate with the new metrics of success.





Using golds and total medals *per capita* as our measuring stick for olympic success, the results come out quite differently. Total population comes out INVERSELY correlated to doing well, meaning small countries generally do better per capita. Urbanization... well... not seeing anything to speak of there. Finally, GDP per capita shows a clear correlation. The richer countries are more successful per capita than poorer ones, which is almost exactly what we found when looking at gross medal counts (except going per capita exposed some more populous poorer countries who pushed quartile 4 over quartile 3).

Recommendations and Actions

Summarize the insights you found and make recommendations on what your client should do. What is the next steps or the action that should be taken as a result of your analysis?

The metrics played out in interesting ways. More populous countries won the most medals overall, but the least per capita. The second tier of urbanized countries were far more successful than other countries in terms of gross medals, though didn't fare as well per capita. The GDP per capita analysis showed that richer countries were successful over all and per capita, not a huge surprise, but nice to have it confirmed.

Really, there is nothing that national olympic committees can do about their population, urbanization or GDP. What they can do is look to successful countries who are in their same boat in terms of metrics, and try to emulate them.

The one interesting result which deserves more time is the number of medals earned by the countries in the second tier of urbanization. There is some advantage to this level of urbanization. I recommend further research on the countries in this tier who were also successful per capita. The whole list of the second tier urbanized countries (minus UAE, number 20, and Jordan, number 21) is below.

(if you want to know even more about this project, check out riceanddata.com)

	Country	Gold per Million Pop	Medals per Million Pop
0	Sweden	47.377741	151.925282
1	Finland	35.725647	162.389302
2	Denmark	30.884540	103.005980
3	New Zealand	18.637318	47.214539
4	Netherlands	16.742949	60.671314
5	Australia	13.619988	51.662024
6	Uruguay	8.918702	18.125105
7	Belgium	8.449404	40.350216
8	Luxembourg	6.372430	12.744861
9	Greece	5.953561	24.486419
10	Argentina	2.010275	6.052915
11	Japan	1.953999	7.222677
12	Bahrain	0.584434	1.753302
13	Brazil	0.512165	2.231912
14	Puerto Rico	0.351264	3.161378
15	Dominican Republic	0.276074	0.644172
16	Mexico	0.232259	0.851617
17	Singapore	0.170697	1.536276
18	Chile	0.156702	1.671489
19	Israel	0.115227	1.037044