Caribbean Economic Analysis
IBM Data Science Capstone

Sergio Smith
February 2021

Objectives

1. Collect data on Caribbean economics and venue prices

2. Build machine learning model using data collected

3. Evaluate model and assess results

Procedure

1. Data collection and cleaning
2. Data analysis
   a. Preliminary investigation
   b. Primary analysis
   c. Follow-up analysis
3. Results evaluation

Software tools

1. Python via Anaconda3
   a. Jupyter Notebook
2. Python packages
   a. Pandas (data management)
   b. Numpy (math features)
   c. Matplotlib (data visualisation)
   d. etc.

Data sources

1. Coordinate data
    a. https://lab.lmnixon.org/4th/worldcapitals.html
2. GDP data
    a. https://en.wikipedia.org/wiki/List_of_Latin_American_and_Caribbean_countries_by_GDP_(PPP)
    b. https://en.wikipedia.org/wiki/List_of_Latin_American_and_Caribbean_countries_by_GDP_(nominal)
3. Venue price data
    a. Foursquare API

Data gathered

Country	GDP (PPP) per capita	GDP (nominal) per capita	Capital	Latitude	Longitude	Food average price tier	Nightlife average price tier
Antigua and Barbuda	27542	14159	Saint John's	17.20	-61.48	1.894737	2.400000
Barbados	18866	16082	Bridgetown	13.05	-59.30	1.526316	1.875000
Belize	8467	3734	Belmopan	17.18	-88.30	1.500000	2.500000
Dominica	9726	7709	Roseau	15.20	-61.24	1.736842	2.357143
Grenada	16033	9824	Saint George's	12.05	-61.75	1.722222	1.947368
Guyana	8524	8649	Georgetown	6.50	-58.12	1.125000	2.200000
Haiti	1940	732	Port-au-Prince	18.40	-72.20	1.600000	2.312500
Jamaica	9726	5221	Kingston	18.00	-76.50	1.277778	2.066667
Saint Kitts and Nevis	29098	15246	Basseterre	17.17	-62.43	1.894737	2.125000
Saint Vincent and the Grenadines	11965	7033	Kingstown	13.10	-61.10	1.533333	2.250000
Suriname	15362	4199	Paramaribo	5.50	-55.10	1.315789	2.333333
The Bahamas	33516	30027	Nassau	25.05	-77.20	1.650000	2.000000
Trinidad and Tobago	33026	16197	Port of Spain	10.67	-61.52	1.157895	1.700000

Caribbean countries with GDP (PPP)

Caribbean capitals with food prices (red) and nightlife prices (blue)

Questions asked

Preliminary: Are nominal and PPP GDP correlated? Are Food venue and Nightlife venue prices correlated?

Primary: Can Food venue and Nightlife venue prices predict GDP?

Follow-up: Does a Nightlife price-only model fit the data better than the multiple regression model?
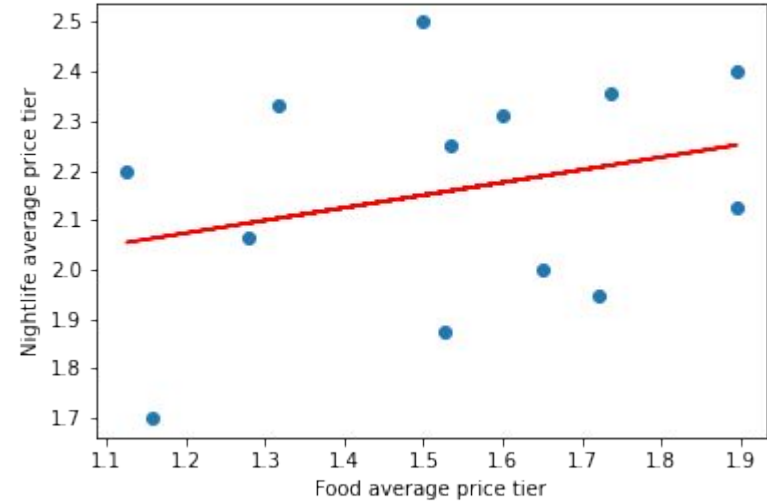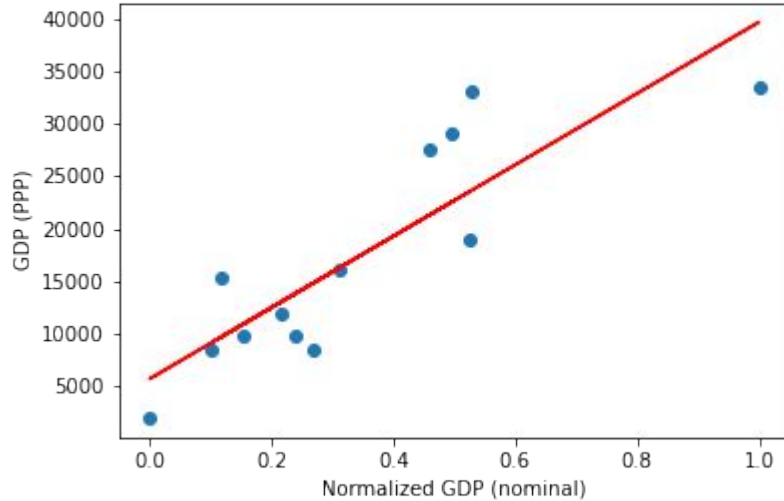
Machine learning tools

Simple linear regression (initial and follow-up investigation)

$$y = \boldsymbol{\beta}_0 + \boldsymbol{\beta}x$$

Multiple linear regression

$$y = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 x_1 + \boldsymbol{\beta}_2 x_2 + \ldots$$

Results: preliminary investigation

Results: preliminary investigation

Property	Value
Intercept	4822
Slope	1.16
Regression Score	0.741

(a) Nominal vs PPP regression

Property	Value
Intercept	1.77
Slope	0.255
Regression Score	0.078

(b) Food vs Nightlife price regression

Table 1: Preliminary analysis results

Results: primary investigation
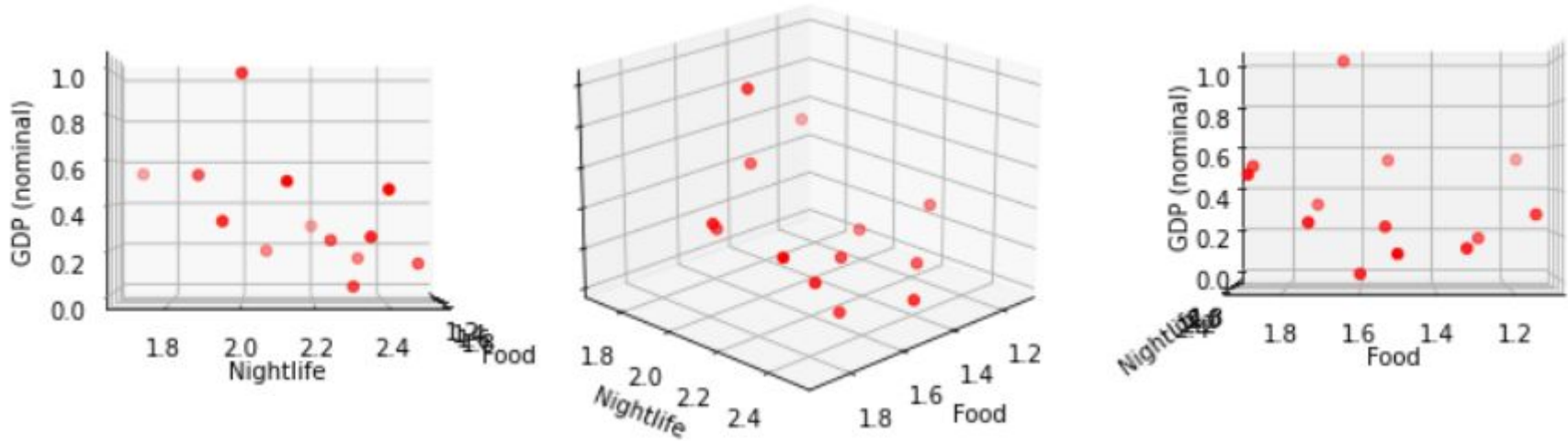
Results: primary investigation

Property	Value
Intercept	0.520
Food Slope	0.326
Nightlife Slope	-0.589
Regression Score	0.464

Property	Value
Intercept	0.702
Food Slope	0.386
Nightlife Slope	-0.684
Regression Score	0.433

(a) Nominal GDP multiple regression results

(b) PPP GDP multiple regression results

Table 2: Multiple regression results

Results: follow-up investigation
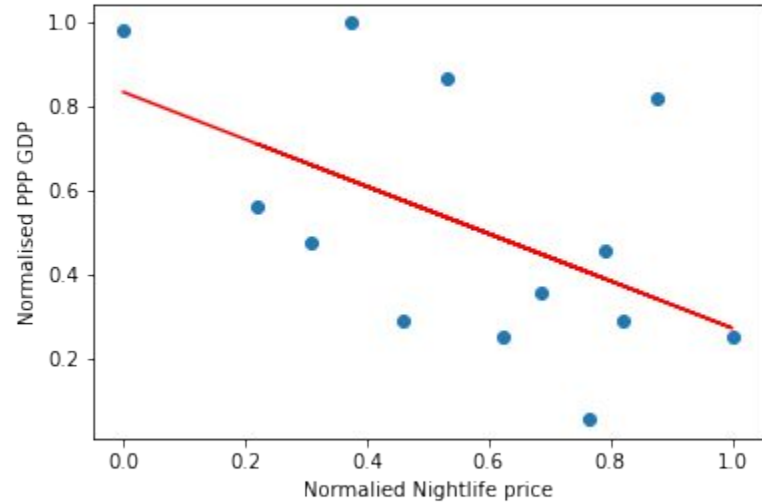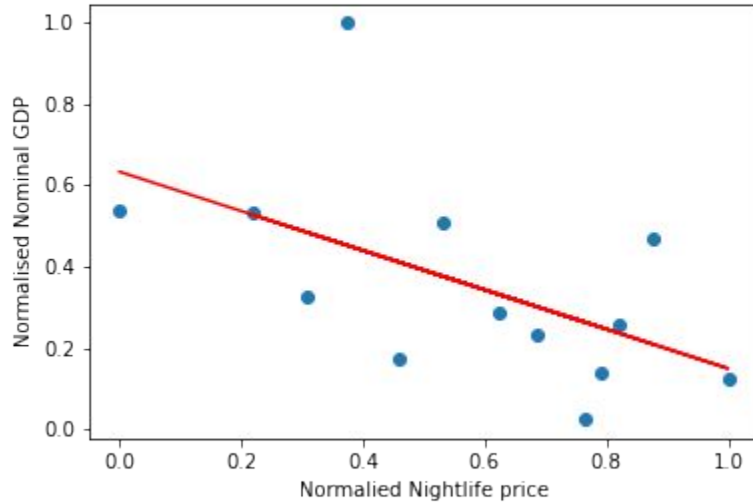
Results: follow-up investigation

Property	Value
Intercept	0.634
Slope	-0.485
Regression Score	0.301

(a) GDP (nominal) vs. Nightlife price regression

Property	Value
Intercept	0.836
Slope	-0.562
Regression Score	0.277

(b) GDP (PPP) vs. Nightlife price regression

Table 3: Follow-up analysis results

Conclusions

1. Nominal and PPP GDP are strongly correlated
2. Food and Nightlife prices are not strongly correlated with each other
3. **Food and Nightlife prices together reasonably well predict GDP**
4. Nightlife price alone is not a strong predictor of GDP

Data on Caribbean venues is very limited.

Follow-up with more data is warranted.