Human Development in Detroit

Jeff Warchall

Detroit Non-profit Day, 2021

Background – Old Detroit

- In 1950 Detroit's population was 1,850,000
- Second only to Chicago in the Great Lakes Region

Image credit: history.com

Modern Detroit

- Today Detroit's population is 639,000

- Fully 2/3rds of its infrastructure is now unused and decaying

Image credit: history.com

Current Renaissance

- Despite this there are signs that Detroit has hit bottom
- Major investment in downtown

Image credit: autonews.com

Still work to do…

- Criticism that the “Renaissance” is not reaching the locals.
- Many different ideas of what is important.

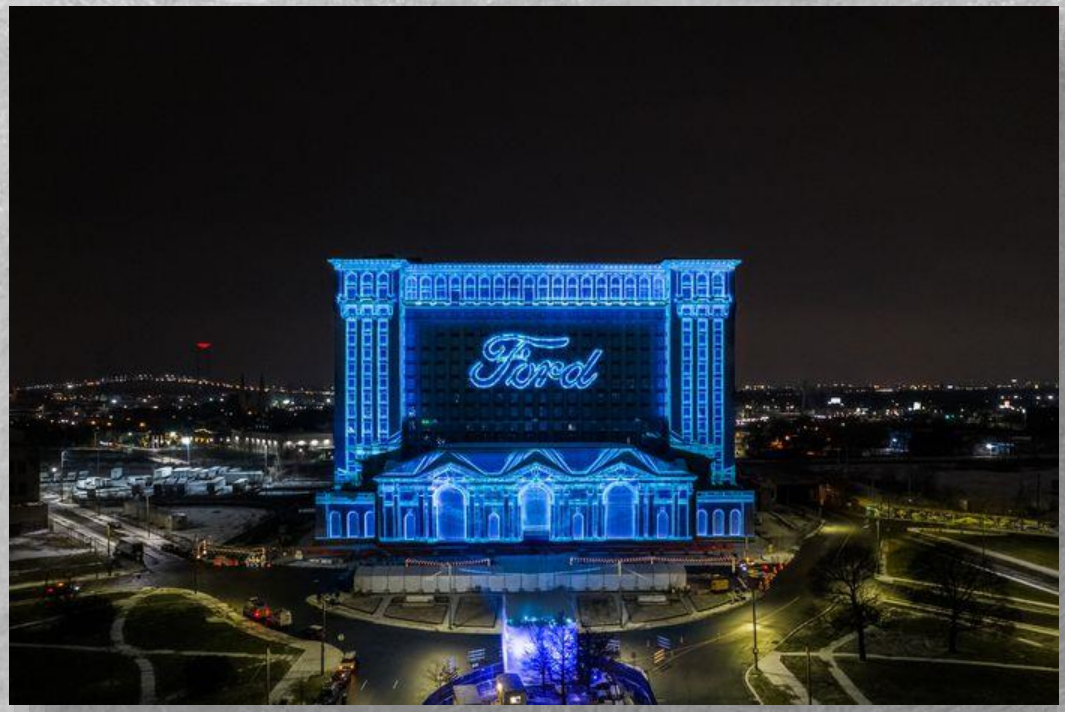


Image credit: detroitmetrotimes.com

Problem Statement

Can we use Data Science to make inferences about where money or time can be best spent to improve the quality of life in Detroit?

The Approach

- Use the Human Development Index (HDI) as a target for improvement.

- Compare Wayne County, Michigan to other counties across the United States to train a model.

Human Development Index

- The HDI was developed by the United Nations to compare the overall quality of life in different countries.

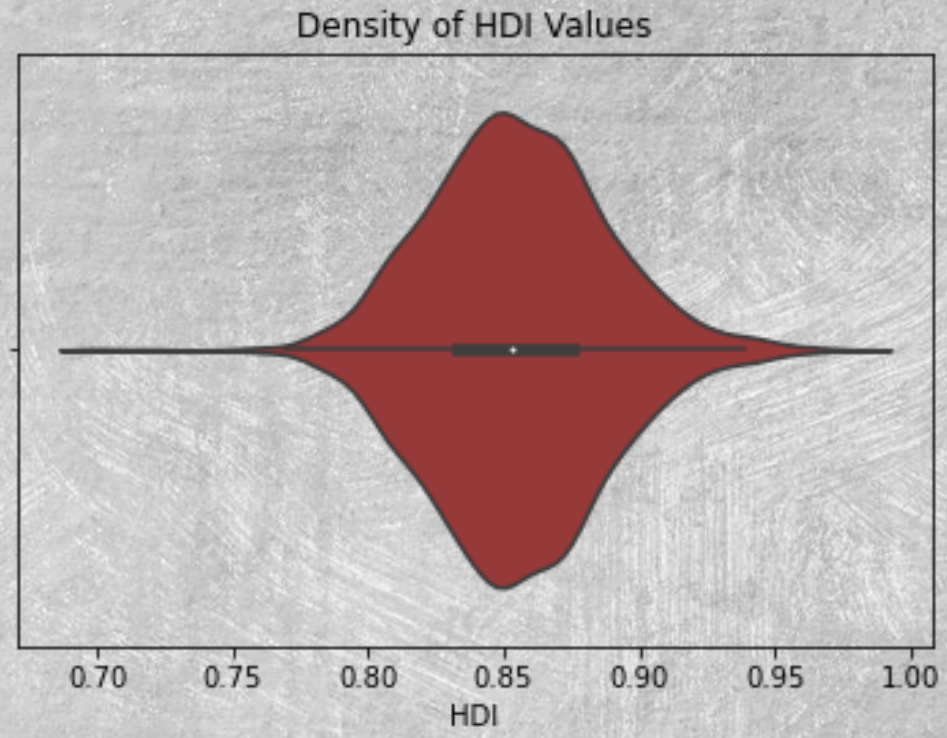
$$HDI = \sqrt[3]{\left(\frac{LE - 20}{65}\right) \times \left(\frac{\frac{MYS}{15} + \frac{EYS}{18}}{2}\right) \times \left(\frac{\ln(GNIpc) - \ln(100)}{\ln(75{,}000) - \ln(100)}\right)}$$

Data Sources

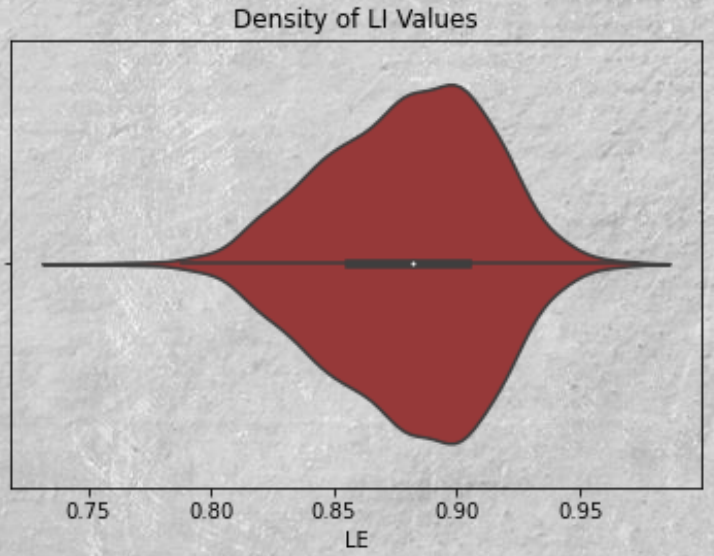
- Economic
  - U.S. Economic Development Administration
- Health and Life Expectancy
  - University of Wisconsin
- Crime
  - University of Michigan
- Education
  - Social Science Research Council

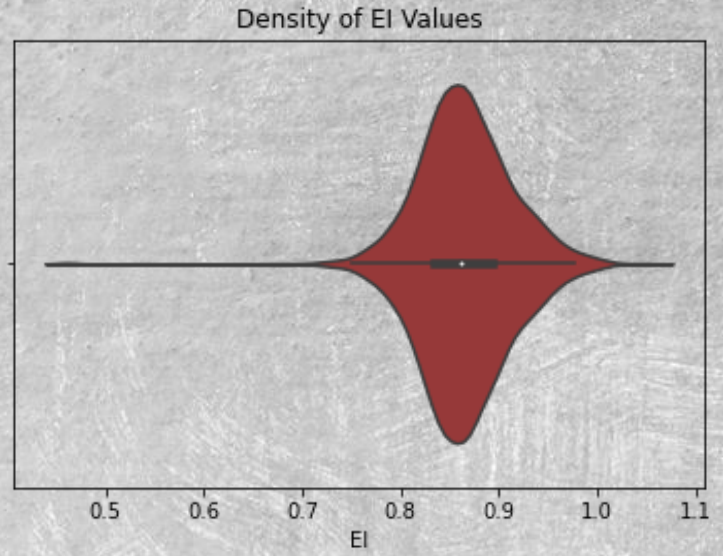
HDI in Context

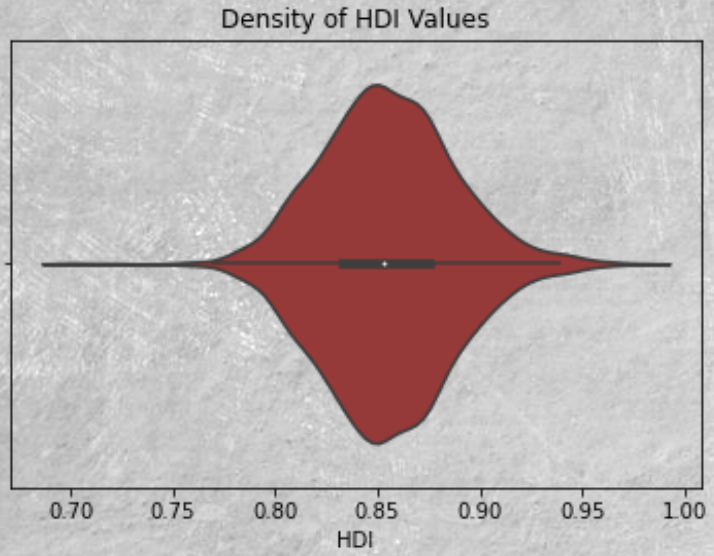
- Lowest HDIs in the world is 0.394
- The Highest is Norway at 0.957
- Three standard deviations in HDI among counties in the United States (99.7%) span the 4th through 89th positions with:
  - Lowest similar to the Dominican Republic or Azerbaijan
  - Highest similar to Hong Kong or Switzerland

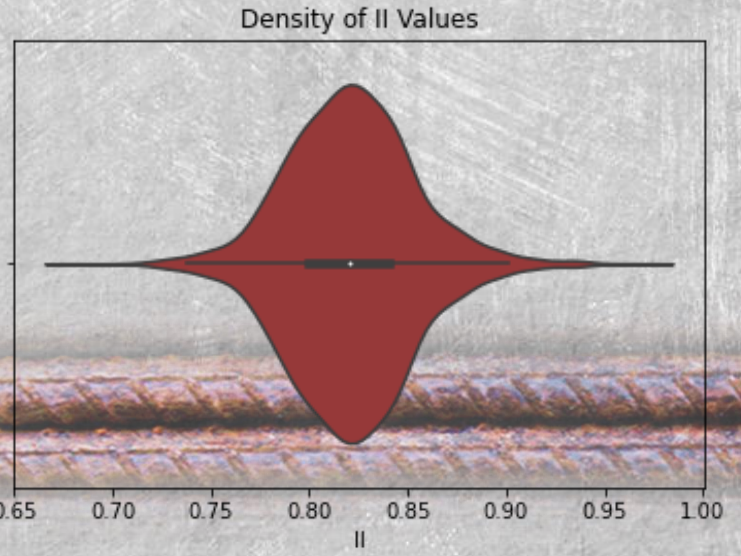


Comparing Indices









High Collinearity Problem

- Data had 186 features
- After collinearity removal there were 117 features

Some Examples:

Percent of returns with retirement plan
Percent of returns with S-Corp income
Percent of returns self-emp health deduction
AGI on itemized returns
Percent of returns with farm income
Percent of returns with taxable Social Security

Modeling Results

- Four models were trained:

Model	R Squared	RMSE
Multivariate Linear	0.744	0.015
K-Nearest Regressor	0.507	0.016
Random Forest	0.792	0.013
Boosted Forest	0.832	0.012

- Tree based models performed best
- But, we proceed with the linear model for interpretability

Importance of Coefficients

	parameter	coef	importance
19	Percent of population that didn't work over th...	-0.008419	0.008419
21	Percent of returns with itemized deductions	0.00585	0.005850
100	% Smokers	-0.003856	0.003856
64	Total civilian population aged 18-64 rate	0.0035	0.003500
23	Percent of returns with salaries and wages	-0.0028	0.002800
32	Population growth, 2010-2016	0.002669	0.002669
17	Percent of civilian population aged 25-34 that...	-0.00264	0.002640
41	Rural-urban continuum code (1-9)	-0.002192	0.002192
13	Non-rent seeking organizations per 10,000 popu...	0.002075	0.002075
43	Share of emp. in all local industries	0.001935	0.001935

Report

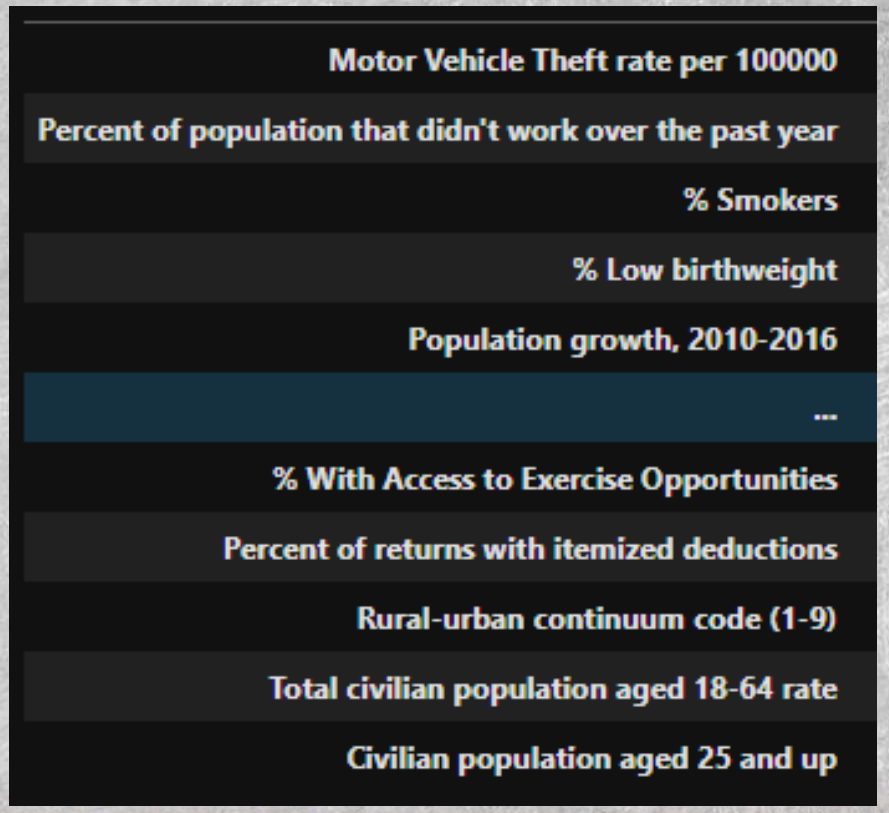
```
Input County Wayne County
Input Two Letter State Code MI
With an HDI of 0.841, Wayne County, MI, compared to the international community, currently has an HDI most similar to  Brunei.
According to our model, if Wayne County, MI successfully remediated the top 5 items it could have an HDI more simalar to  Slovakia.
```

	Mean Value	County Value	County SDs from Mean	Amount of Change	Impact on HDI
Motor Vehicle Theft rate per 100000	106.465	859.079	7.354	752.615	0.006
Percent of population that didn't work over the past year	26.435	32.402	0.641	5.966	0.005
% Smokers	21.345	24.000	0.693	2.655	0.003
% Low birthweight	8.134	11.000	1.580	2.866	0.002
Population growth, 2010-2016	0.343	-2.911	-0.696	-3.253	0.002
...	...	...	...	...	...
% With Access to Exercise Opportunities	64.255	95.000	1.490	30.745	-0.002
Percent of returns with itemized deductions	21.344	24.243	0.348	2.899	-0.002
Rural-urban continuum code (1-9)	4.359	1.000	-1.517	-3.359	-0.003
Total civilian population aged 18-64 rate	0.578	0.617	0.877	0.039	-0.003
Civilian population aged 25 and up	67527.664	1168342.000	5.004	1100814.336	-0.003

Interpretation

- Some of this we already “know”
- This allows us to get more focused on detail

- Also shows where Detroit is doing well



Conclusions

- Politicians could use this to identify key agenda items
- Voters could use this to evaluate political platforms
- Nonprofit organizations could focus effort
- Citizen volunteers could

Generalization

- Motivation for the project was Detroit, but generalization to *any* county was quite simple.

HDI Map

Limitations and Future Work

- This is a first step
- We have discovered that Motor Vehicle Theft is a major problem
- Next research would need to be done to determine how to address these issues specifically

- High correlation to income

- Integration of Python into visualization tool

Sources

Our World in Data
https://ourworldindata.org/grapher/expected-years-of-schooling

Measure of America
https://measureofamerica.org/maps/?county^ed^all_all^HDI^hdiD

Macrotrends
https://www.macrotrends.net/countries/USA/united-states/gni-per-capita

StatsAmerica (US Economic Development Administration)
https://www.statsamerica.org/About.aspx

University of Michigan
https://www.icpsr.umich.edu/web/pages/

Any Questions?