Toronto Through Data

Neighbourhood Data Exploration

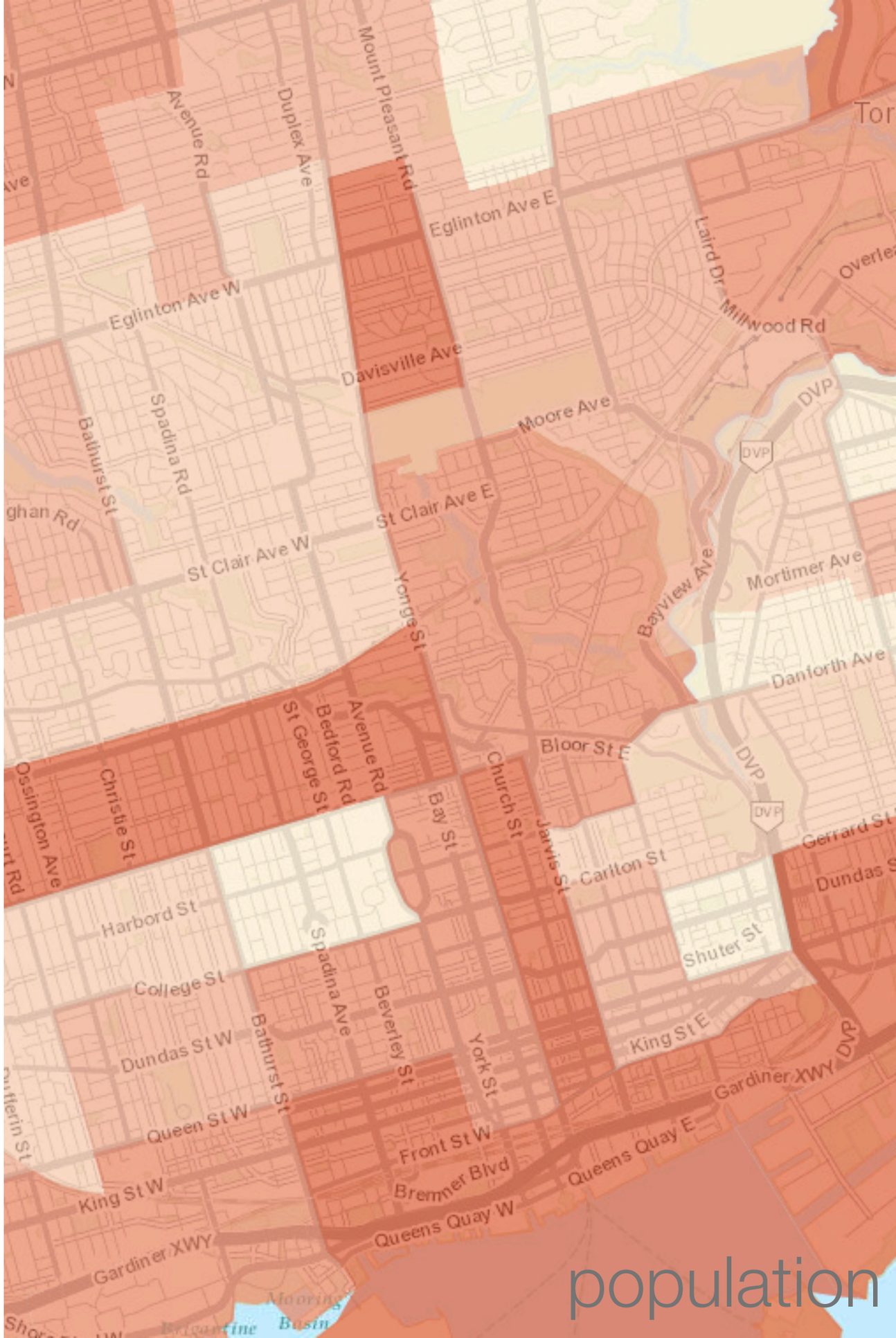
Derek Pyne

Wellbeing Toronto

- Wellbeing Toronto is a dataset collected by the city of Toronto. It contains metrics aggregated by the 140 official neighbourhood.

- Provided to help users make data driven decisions from city planning, to community leaders, to new business owners

- Using the 2013 dataset (the only complete year available), can we find any correlation between home prices and the provided metrics? Can this help us determine what people find most valuable in their neighbourhoods?



population

Let's Explore the Data

civics

culture

environment

housing

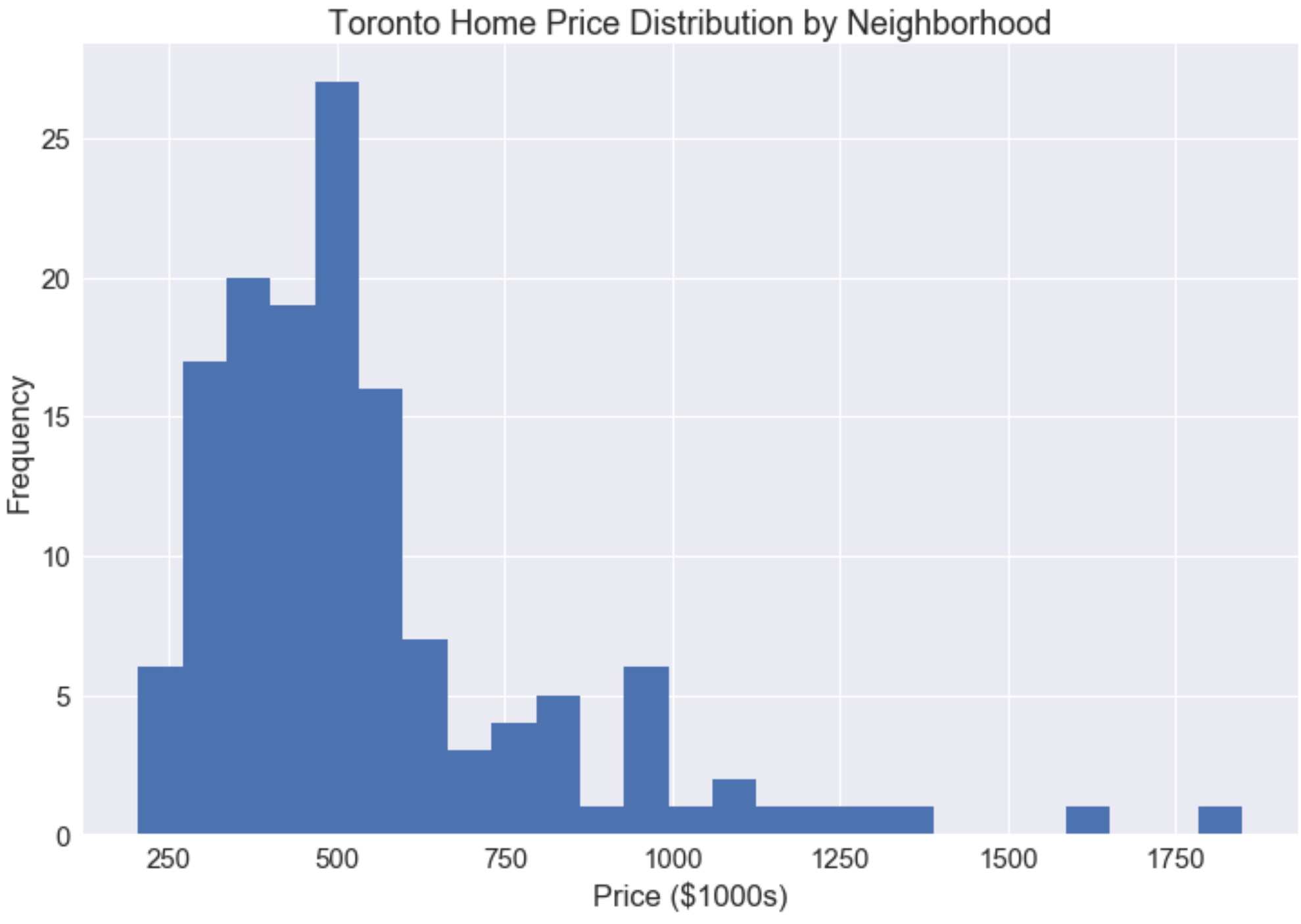
health

recreation

safety

transportation

Home Price - A proxy for demand

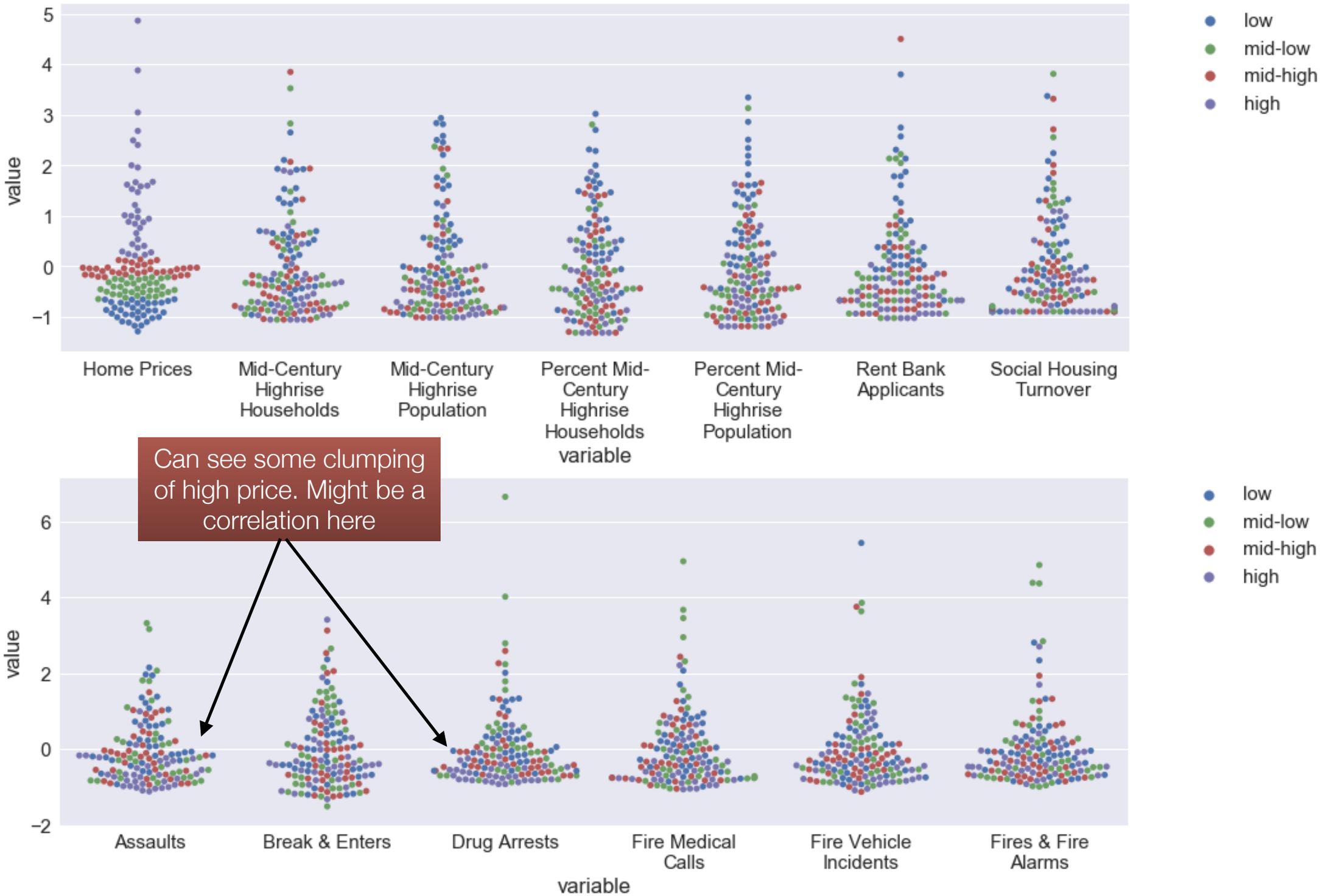


What if we categorize price?

- To help us visualize our data, let's categorize home price by quartiles (high, mid-high, mid-low, low)

- We can then visualize our feature columns with the data coloured by price category

- The following swarm plots restrict data points from overlapping which helps us visualize any clustering.


value
5
4
3
2
1
0
−1
Home Prices
Mid-Century Highrise Households
Mid-Century Highrise Population
Percent Mid-Century Highrise
Percent Mid-Century Highrise
Rent Bank Applicants
Social Housing Turnover
variable
low
mid-low
mid-high
high
Can see some clumping of high price. Might be a correlation here
value
6
4
2
0
−2
Assaults
Break & Enters
Drug Arrests
Fire Medical Calls
Fire Vehicle Incidents
Fires & Fire Alarms
variable
low
mid-low
mid-high
high

Maybe ranking is more important?

- Let’s test the following hypothesis:

  ★ When people are shopping for a home, it isn't as much the absolute value that matters but instead whether or not a neighbourhood is better or worse then other neighbourhoods.

- We can slice the data by using a ranking for each feature column instead of the raw values.

- The following charts have the neighbourhoods ordered by price on the y-axis (highest at the top) and rank on the x axis (highest rank on the right)

Rent Bank Applicants

Social Housing Turnover

Social Housing Units

Social Housing Waiting List

City Grants Funding $

Neighbourhood

Bridle Path-Sunnybrook-York Mills
Forest Hill South
St.Andrew-Windfields
Rosedale-Moore Park
Lawrence Park South
Bedford Park-Nortown
Casa Loma
Leaside-Bennington
Yonge-St.Clair
Annex
Lawrence Park North
Yonge-Eglinton
Princess-Rosethorn
Forest Hill North
Kingsway South
Edenbridge-Humber Valley
Playter Estates-Danforth
North Riverdale
Mount Pleasant East
Lambton Baby Point
Humewood-Cedarvale
Bayview Woods-Steeles
Palmerston-Little Italy
University
The Beaches
Trinity-Bellwoods
Stonegate-Queensway
Runnymede-Bloor West Village
Wychwood
Lansing-Westgate
Newtonbrook East
High Park-Swansea
Banbury-Don Mills
High Park North
Clanton Park
Danforth
Roncesvalles
Old East York
Englemount-Lawrence
Blake-Jones
Bathurst Manor
Newtonbrook West
Parkwoods-Donalda
Cliffcrest
Willowdale West
Willowdale East
Maple Leaf
O'Connor-Parkview
South Riverdale
Cabbagetown-South St.James Town
East End-Danforth
Bayview Village
Junction Area
Highland Creek
Birchcliffe-Cliffside
Woodbine Corridor
Mount Pleasant West
Danforth-East York
Broadview North
Little Portugal
Dovercourt-Wallace Emerson-Junctn
Willowridge-Martingrove-Richview
Markland Wood
Alderwood
Corso Italia-Davenport
Dufferin Grove
South Parkdale
Rustic
Islington-City Centre West
Humber Heights-Westmount
Centennial Scarborough
Greenwood-Coxwell
North St.James Town
Hillcrest Village
Regent Park
Oakwood Village
Don Valley Village
Kensington-Chinatown
Pleasant View
New Toronto
Long Branch
Bay Street Corridor
Moss Park
Guildwood
Woodbine-Lumsden
Steeles
Beechborough-Greenbrook
Mimico
Rouge
Eringate-Centennial-West Deane
Wexford/Maryvale
Yorkdale-Glen Park
Waterfront Communities-The Island
Thistletown-Beaumond Heights
Pelmo Park-Humberlea
Tam O'Shanter-Sullivan
Church-Yonge Corridor
Caledonia-Fairbank
Downsview-Roding-CFB
Niagara
Rexdale-Kipling
Clairlea-Birchmount
Milliken
Weston-Pellam Park
Agincourt North
Briar Hill-Belgravia
Keelesdale-Eglinton West
Westminster-Branson
Henry Farm
Victoria Village
York University Heights
Scarborough Village
L'Amoreaux
Humber Summit
Rockcliffe-Smythe
Brookhaven-Amesbury
Oakridge
Agincourt South-Malvern West
Ionview
Bendale
West Humber-Clairville
Woburn
Humbermede
Thorncliffe Park
West Hill
Glenfield-Jane Heights
Etobicoke West Mall
Kennedy Park
Kingsview Village-The Westway
Morningside
Weston
Dorset Park
Eglinton East
Mount Dennis
Malvern
Taylor-Massey
Mount Olive-Silverstone-Jamestown
Black Creek
Elms-Old Rexdale
Flemingdon Park

No correlation for mid prices but a clear clumping at high and low prices

Neighbourhood Equity Score

Salvation Army Donors

Walk Score

Watermain Breaks

Linguistic Diversity Index

Neighbourhood

Bridle Path-Sunnybrook-York Mills
Forest Hill South
St.Andrew-Windfields
Rosedale-Moore Park
Lawrence Park South
Bedford Park-Nortown
Casa Loma
Leaside-Bennington
Yonge-St.Clair
Annex
Lawrence Park North
Yonge-Eglinton
Princess-Rosethorn
Forest Hill North
Kingsway South
Edenbridge-Humber Valley
Playter Estates-Danforth
North Riverdale
Mount Pleasant East
Lambton Baby Point
Humewood-Cedarvale
Bayview Woods-Steeles
Palmerston-Little Italy
University
The Beaches
Trinity-Bellwoods
Stonegate-Queensway
Runnymede-Bloor West Village
Wychwood
Lansing-Westgate
Newtonbrook East
High Park-Swansea
Banbury-Don Mills
High Park North
Clanton Park
Danforth
Roncesvalles
Old East York
Englemount-Lawrence
Blake-Jones
Bathurst Manor
Newtonbrook West
Parkwoods-Donalda
Cliffcrest
Willowdale West
Willowdale East
Maple Leaf
O'Connor-Parkview
South Riverdale
Cabbagetown-South St.James Town
East End-Danforth
Bayview Village
Junction Area
Highland Creek
Birchcliffe-Cliffside
Woodbine Corridor
Mount Pleasant West
Danforth-East York
Broadview North
Little Portugal
Dovercourt-Wallace Emerson-Juncti
Willowridge-Martingrove-Richview
Markland Wood
Alderwood
Corso Italia-Davenport
Dufferin Grove
South Parkdale
Rustic
Islington-City Centre West
Humber Heights-Westmount
Centennial Scarborough
Greenwood-Coxwell
North St.James Town
Hillcrest Village
Regent Park
Oakwood Village
Don Valley Village
Kensington-Chinatown
Pleasant View
New Toronto
Long Branch
Bay Street Corridor
Moss Park
Guildwood
Woodbine-Lumsden
Steeles
Beechborough-Greenbrook
Mimico
Rouge
Eringate-Centennial-West Deane
Wexford/Maryvale
Yorkdale-Glen Park
Waterfront Communities-The Island
Thistletown-Beaumond Heights
Pelmo Park-Humberlea
Tam O'Shanter-Sullivan
Church-Yonge Corridor
Caledonia-Fairbank
Downsview-Roding-CFB
Niagara
Rexdale-Kipling
Clairlea-Birchmount
Milliken
Weston-Pellam Park
Agincourt North
Briar Hill-Belgravia
Keelesdale-Eglinton West
Westminster-Branson
Henry Farm
Victoria Village
York University Heights
Scarborough Village
L'Amoreaux
Humber Summit
Rockcliffe-Smythe
Brookhaven-Amesbury
Oakridge
Agincourt South-Malvern West
Jonview
Bendale
West Humber-Clairville
Woburn
Humbermede
Thorncliffe Park
West Hill
Glenfield-Jane Heights
Etobicoke West Mall
Kennedy Park
Kingsview Village-The Westway
Morningside
Weston
Dorset Park
Eglinton East
Mount Dennis
Malvern
Taylor-Massey
Mount Olive-Silverstone-Jamestown
Black Creek
Elms-Old Rexdale
Flemingdon Park

Building an interpretable model

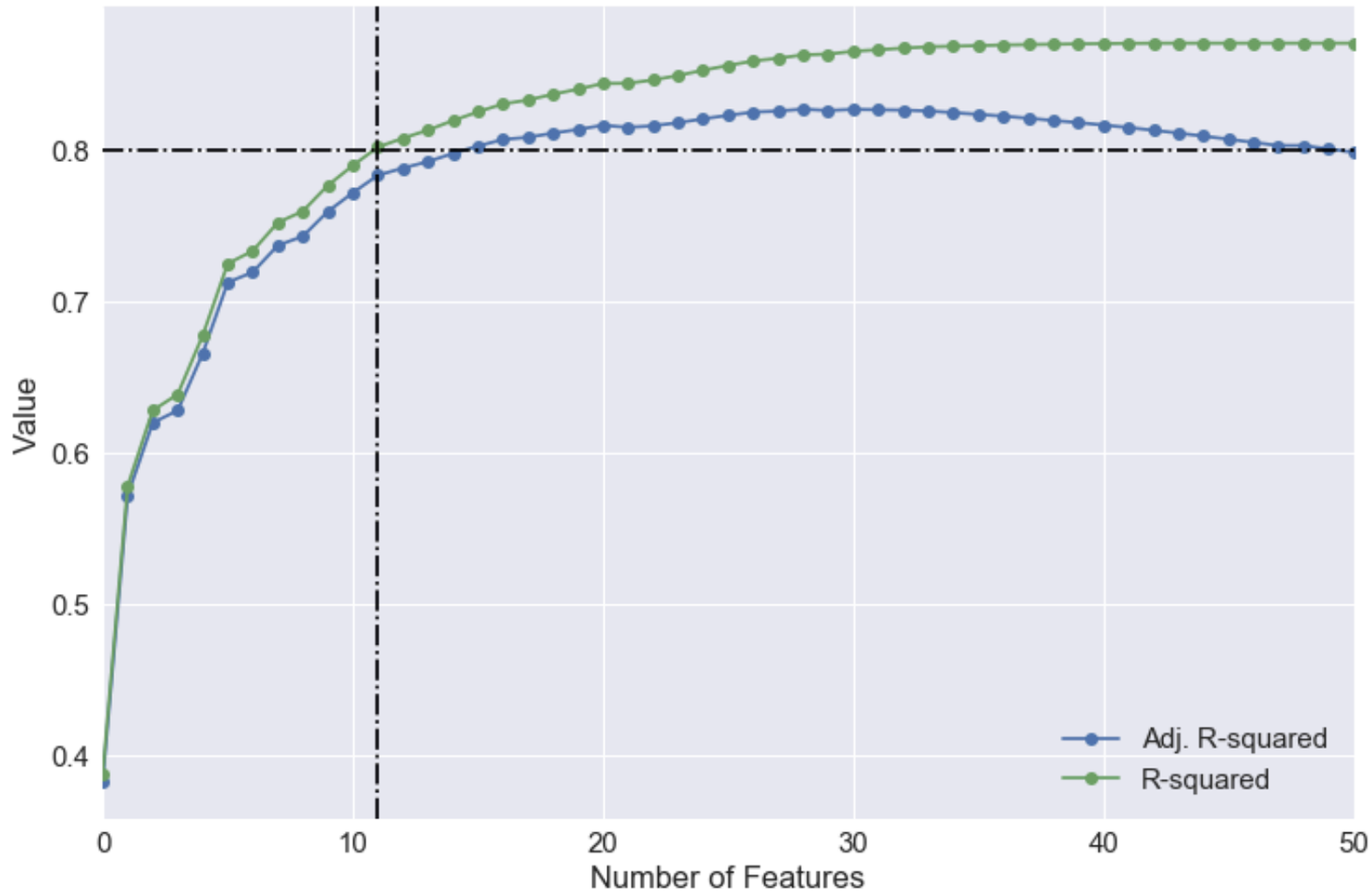
- We want the highest performing model while also keeping it understandable

- We can do this by fitting a linear model, removing the lowest performing feature, and repeating the process until the performance crosses a threshold. We will set this threshold at an R squared of 0.8

- This leaves us with a small set of features that best describe our dataset.

Final Model

Selected Features
Salvation Army Donors
Walk Score
Linguistic Diversity Index
Green Rebate Programs
Green Spaces
Tree Cover
Breast Cancer Screenings
Premature Mortality
Break & Enters
Fires & Fire Alarms
Total Major Crime Incidents



Next Steps?

- Proximity analysis
    - How do the outcomes of bordering neighbourhoods affect each other?
- Time series analysis
    - Can we predict future outcomes with the limited dataset?