# G.A. Data Science

Final Project
Part 1: Lightning Talk

Michael Wong

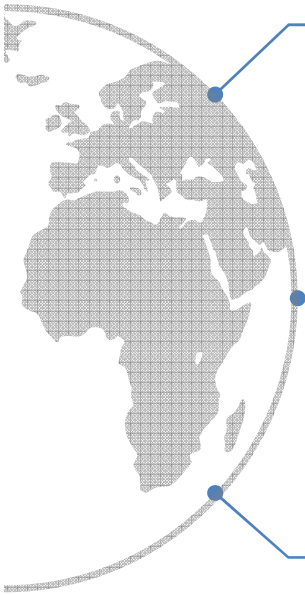
# Timeline & Topics

| 1: Lightning Talk | 2: Research Design | 3: Exploratory Analysis | 4: Model Development | 5: Presentation |
|---|---|---|---|---|
| 8/16 | 9/11 | 9/18 | 9/25 | 10/2 |
| 3 topics x 1 minute | Research & Write-up: problem statement, scope, hypothesis, approach, data sets, risks/issues, resources | Python Notebook: imported and validated data, basic statistics and visualization, correlations | Python Notebook: code and markup, visualizations of results, analysis and conclusions, improvements | 10 minutes |

# 1: Real Estate Markets

| Hypothetical Correlations: | Property Prices | Rents |
|---|---|---|
| Number of Housing Units | Negative | Negative |
| Vacancy rates | Negative | Negative |
| Population | Positive | Positive |
| Employment | Positive | Positive |
| Household income | Positive | Positive |
| Education levels | Positive | Positive |

# 1: Real Estate Markets



Research

Kaggle

Data Sources

# 1: Real Estate Markets

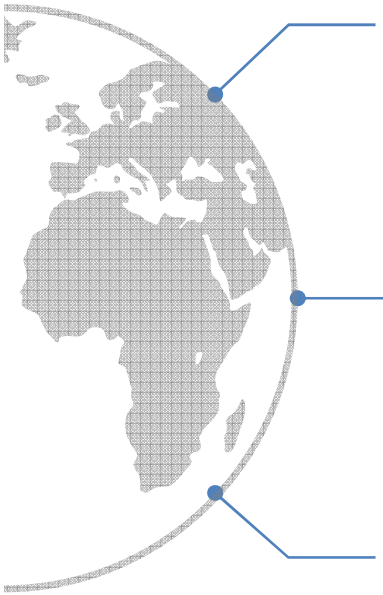
| Annual Census Data 2010-15 | |
|---|---|
| By ZCTA | |
| Housing | Population |
| • Total housing units<br>• Units occupied/vacant<br>• Vacancy rate – rental<br>• Vacancy rate – owner<br>• Housing tenure<br>• Home value<br>• Gross rent<br>• Gross rent / household income | • Population<br>• Sex<br>• Age<br>• Race |
| | Social Characteristics |
| | • Educational attainment by age and race<br>• Household type: married, non-family<br>• Household/family size<br>• Relationships in household<br>• Marital status<br>• Fertility<br>• Grandparents<br>• School enrollment<br>• Residence 1 year ago<br>• Place of birth (foreign/US, region/state)<br>• Citizenship status, year of entry<br>• Computer/internet use |
| Economic Characteristics | |
| • Employment status<br>• Children in household<br>• Commute method/mean travel time<br>• Occupation/industry<br>• Household income<br>• Healthcare coverage (private/public) | |

| Zillow |
|---|
| Monthly Data from 2010 |
| • Zip Code<br>• City<br>• State<br>• Home Type<br>• Inventory for sale<br>• Median listing price<br>• Median listing price/sf<br>• Median sold price<br>• Median sold price/sf<br>• Median rental price<br>• Median rental price/sf<br>• Price cut<br>• Price to rent ratio<br>• Foreclosures out of 10k<br>• Turnover (12mo)<br>• % homes decreasing in value<br>• % homes increasing in value |
| Zillow Metrics (pt in time) |
| • Buyer/Seller Index<br>• Zillow Rent Index<br>• Zillow Home Value Index |

# 2: Spot Election

| **Hypothetical Correlations:** | **Incumbent Party Candidate Victory** |
|---|---|
| GDP | Positive |
| GDP Growth | Positive |
| Inflation | Negative |
| Incumbent Candidate | Positive |
| Duration of Party in Power | Negative |
| War and Terrorism | Neutralizes negative economic conditions |

# 2: Spot Election



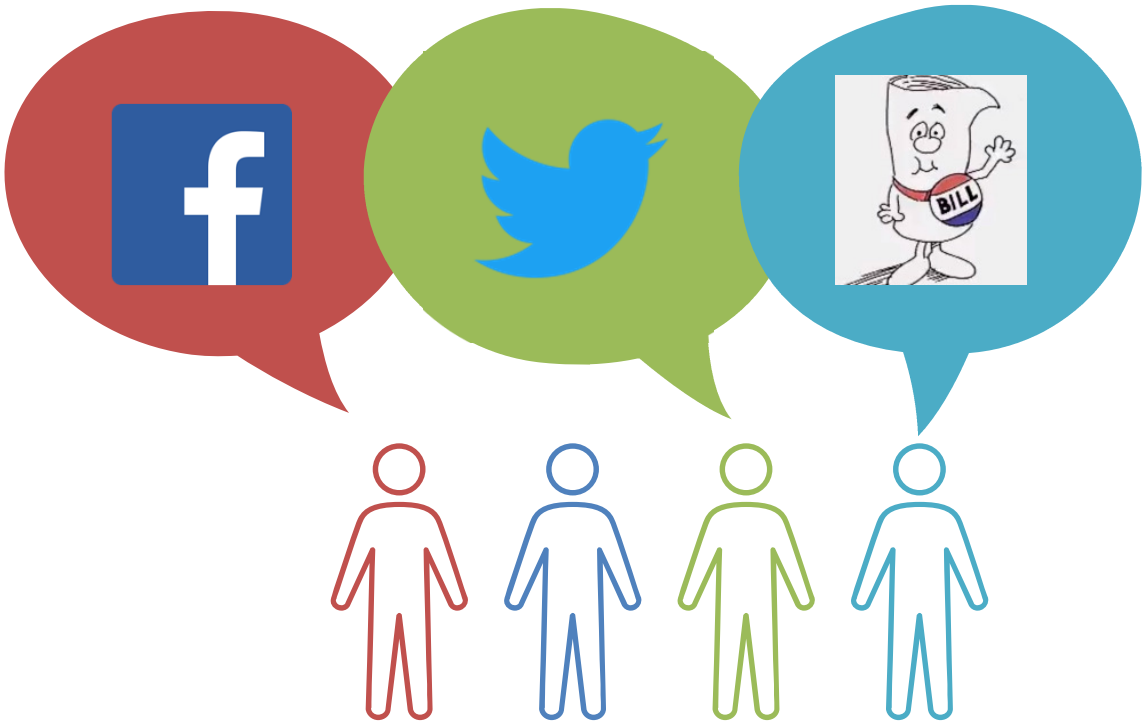
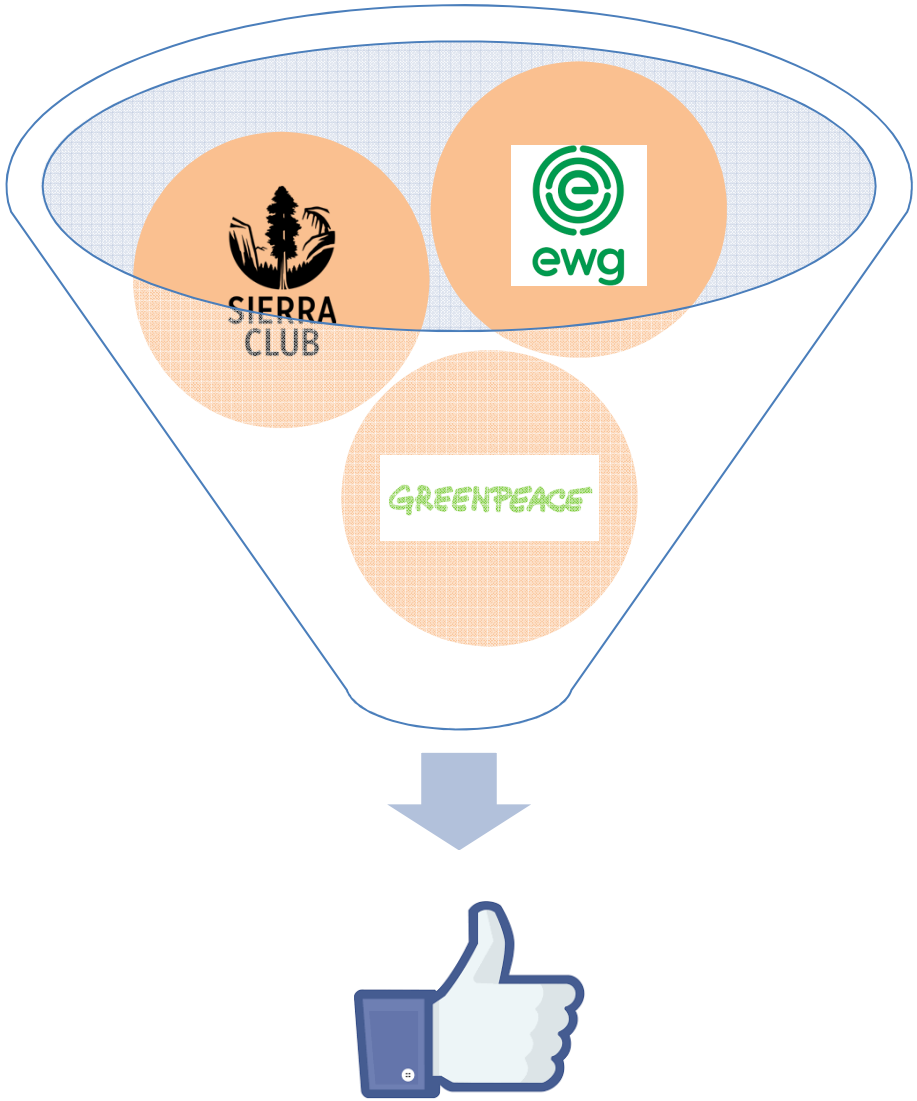
Research

Kaggle

Data Sources

bea
Bureau of Economic Analysis
U.S. DEPARTMENT OF COMMERCE

# 3: Product Sentiment



# 3: Product Sentiment