



Technology Review: Real Estate Project

Corey Christopherson
Richard Todd
Maggie Weatherly
Ruian Yang • 05.12.2020



Project Overview

Problem : Buying a home is the largest single purchase that most people will make in their entire lives¹. Additionally, the residential real estate market is one of the largest components of the American economy². Despite this prominence, good data on real estate trends and transactions is almost impossible to find

Solution : create a tool to collect real estate listing and transaction data for Seattle, WA from the King County Assessor's office and real estate brokerage firm Redfin, then allow users to visualize this data

1- See for example "More people are buying a home — the biggest financial decision of their lives — sight unseen", Seattle Times, 20 July 2018

<https://www.seattletimes.com/business/real-estate/more-people-are-buying-a-home-the-biggest-financial-decision-of-their-lives-sight-unseen/>

2 - See for example Congressional Research Service report: Introduction to U.S. Economy: Housing Market (October 2019) <https://fas.org/sgp/crs/misc/IF11327.pdf>



Use Case - Visualizing Real Estate Pricing

User is researching the best neighborhoods in Seattle for making investment.

- User: Rank neighborhood by selling/listing price ratio
 - display zip codes with top selling/listing price ratio
- User: Select multiple zip code for comparison
 - display price trends for selected zip codes



Technology Option - Plot.ly

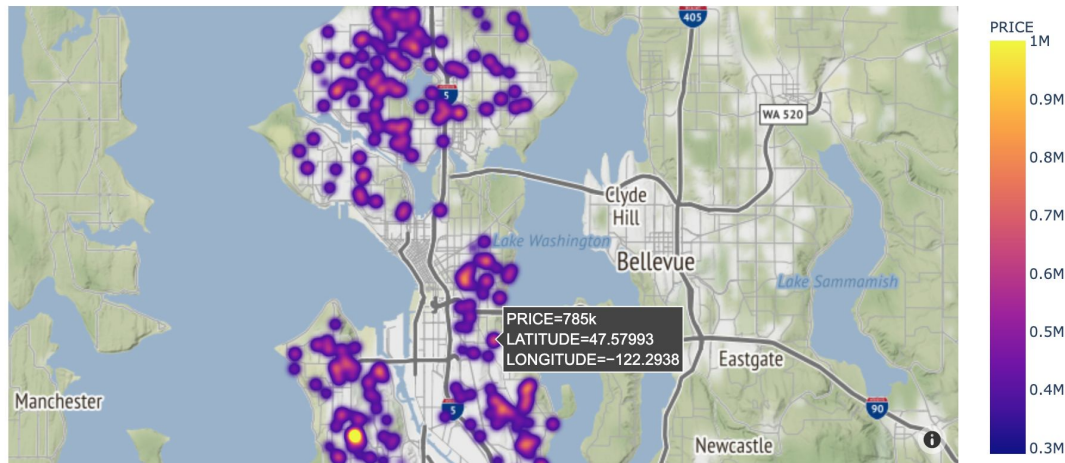
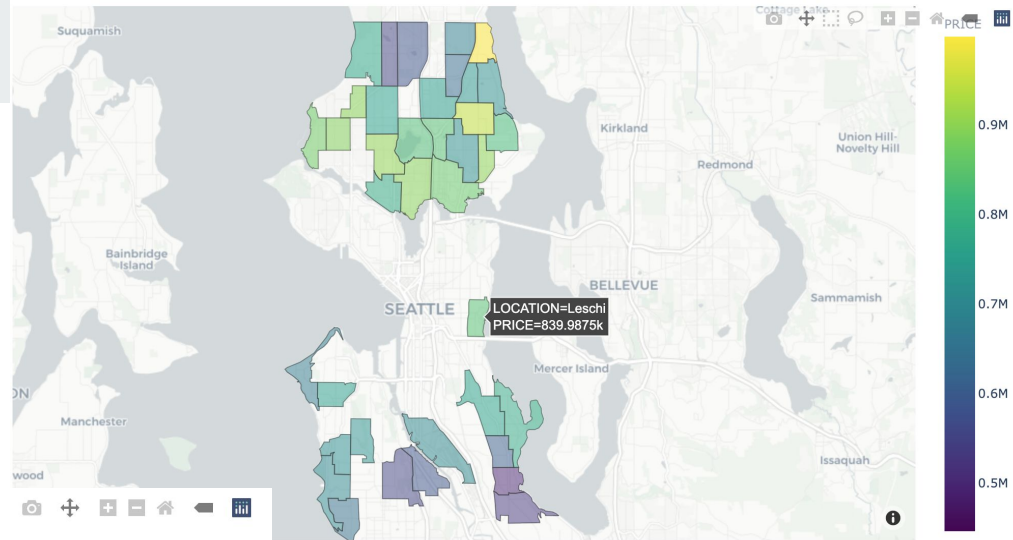
Pros

- Excellent visualizations - very customizable
- Very user friendly

Cons

- Geospatial aggregations are limited - can be done but requires more work
- No hexbin plotting functionality

Plot.ly Example





Technology Option - GeoPandas

Pros

- Excellent at aggregating based on spatial joins
- Easily create plots with simple `.plot()` command
- Can leverage matplotlib plotting functionality for hexbin plots and other plotting functions

Cons

- Native visualization abilities are lacking - uses built in matplotlib libraries
- Visualizations are not natively interactive, takes additional lines of code to add labels, titles, etc

GeoPandas Example

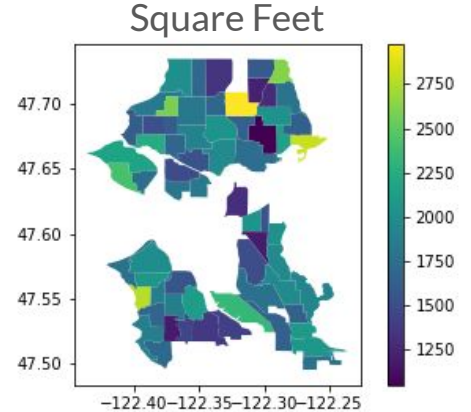
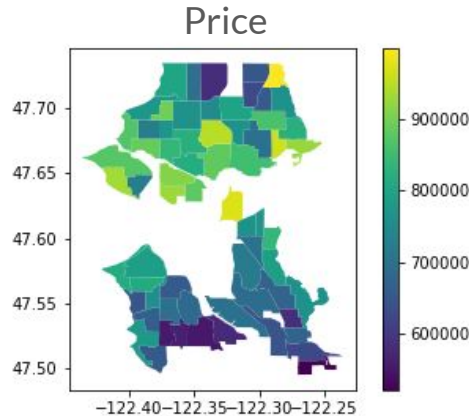
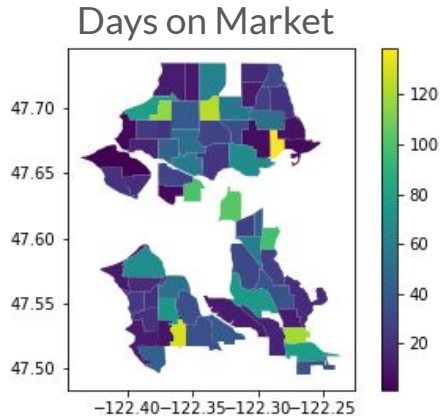
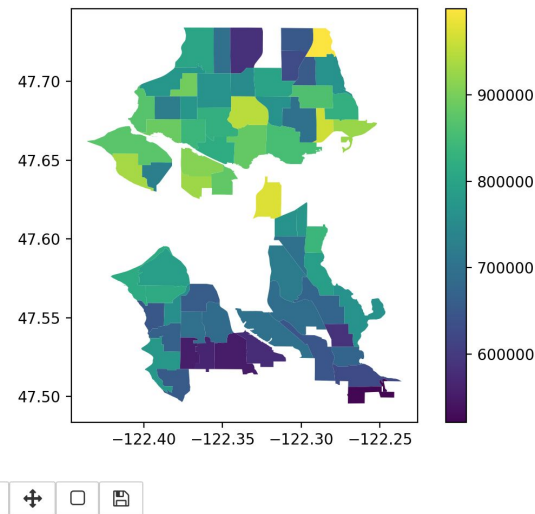
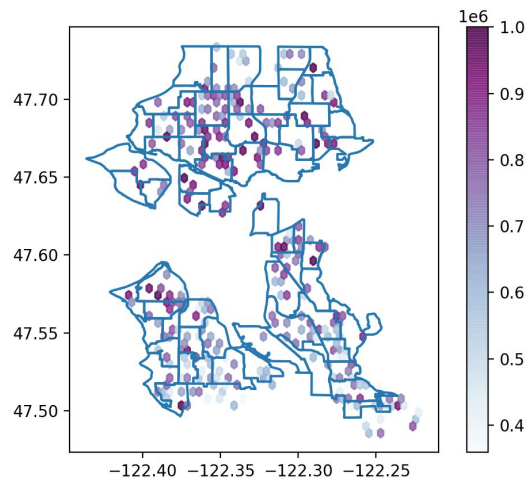


Figure 3

GeoPandas Example

With added line “%matplotlib notebook”
the user gets better interactive experience
such as zoom, pan, and minimal hover experience



Can easily leverage `matplotlib.pyplot.hexbin()`
with a GeoPandas GeoDataFrame to create
hexbin plot

Conclusion

	Plot.ly	GeoPandas	Winner
Ease of adding & visualizing multiple geometry layers	Excellent	Decent	Plot.ly
Able to handle geometric aggregations	Ok	Excellent	GeoPandas
User experience - how interactive	Very Good	Poor	Plot.ly
Documentation	Excellent	Fine	Plot.ly

We choose → Both!

Because of our need do to complex geospatial aggregation and joins, Plot.ly won't be able to do everything. We will be using geopandas for aggregations and Plot.ly for visualizations.