

IBM Applied Data Science Capstone Final Project: **“France in the U.S.A”**

By: Kregg Davenport

02/2021



Table of Contents:

1. Project Problem and Background Information
2. Find a County with graphs
3. Find the Top 4 Employers in WA State including Geo Coordinates plus map
4. Use FourSquare to get venues surrounding each of the 4 sites identified while also using one-hot encoding to analyze those venues
5. Results and Summary
6. Discussion and Conclusion
7. Sources and Acknowledgements

1. Buisness Problem and Background

Buisness Problem:

“Where is a good place for a U.S. expansion of a French style restaurant?”



Cont..

Initial Conversation/background into Buisness Problem:

While living in France a friend there approached regarding an expansion of their French restaurant into the U.S. market. Knowing that I am from the U.S. they requested my assistance in determining a suitable location, within the continental U.S., for this expansion. I gladly accepted. After our initial interview it was determined that we would focus on the State of Washington in the U.S. due to its proximity to the Pacific Ocean (fresh seafood daily), strong economy (Microsoft, Amazon, Google etc) and a large international community (According to the American Immigration Council 1 in 7 residents in Washington State is an immigrant and of those 38% have a college degree or higher (1)). Now the question is where within the State of Washington should my friend startup a new French Restaurant? This is where the Data Science comes in. To choose a location we first must decide which county to focus on then we will narrow the focus down to 2-3 cities within the county. After identifying the potential cities we will make an API call to foursquare enabling us to explore each location and, along with one-hot encoding, we will examine the venues for each location giving us a better idea of where a new French style restaurant may flourish.

Cont..

Steps for Buisness Problem resolution including data:

- Folium map of Washington State
- Download, clean and analyze tables of data to decide on which County to focus on
- Create a graph, or chart, to help visually identify qualifying sites
- Create a Folium map of deciding county within Wash. St.
- Analyze a table of criteria to identify 4 sites of interest to explore on foursquare
- Generate Foursquare API calls for each site identified and examine local venues using one-hot encoding
- Rank the cities based on One-hot encoding and foursquare
- Discuss the results with the friend/customer and examine next steps
- Conclusion and Discussion of results

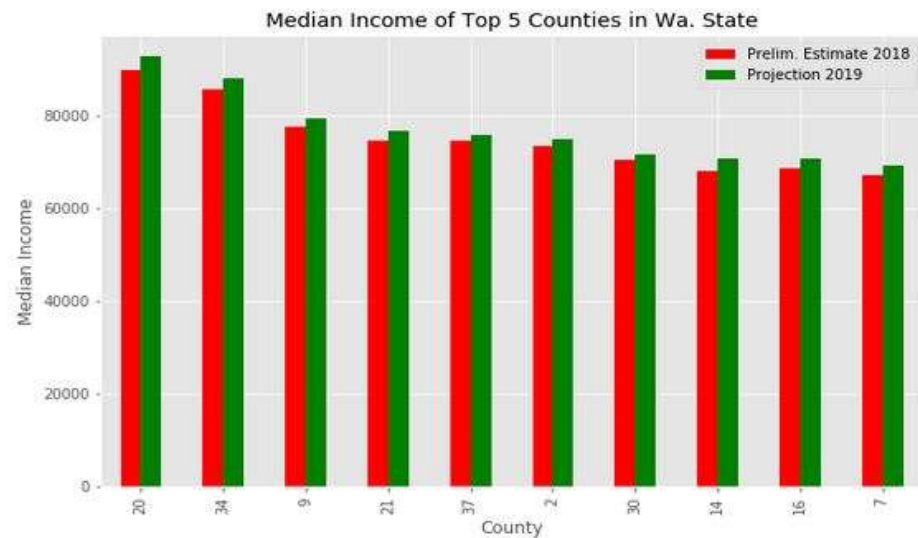
Map of U.S with WA State highlighted



2. Find a County in WA. State

After gathering and cleaning data we have King County identified as the having the highest Median Income at nearly \$93,000 as projected in 2019

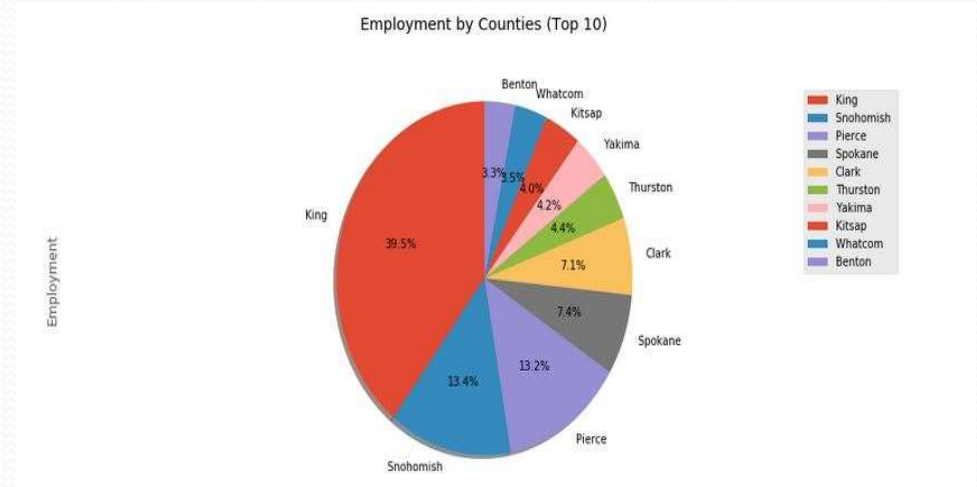
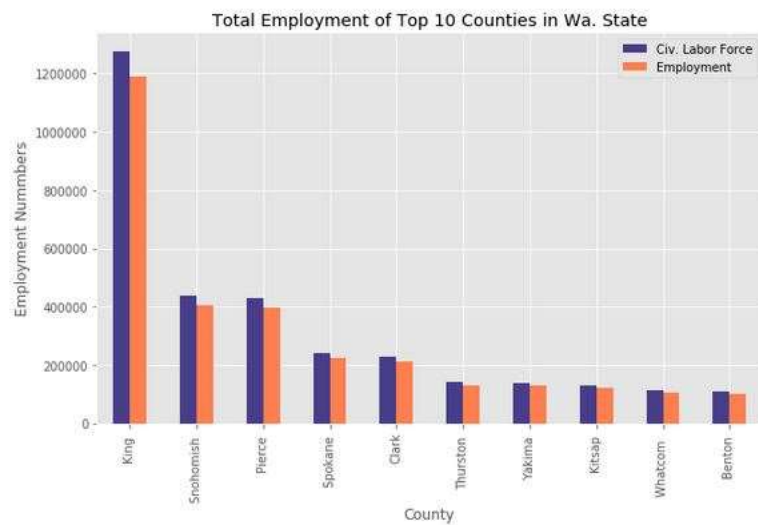
	County	Prelim. Estimate 2018	Projection 2019
20	King	89881.39	92693.33
34	Snohomish	85757.76	88117.87
9	Clark	77458.09	79399.47
21	Kitsap	74728.82	76753.47
37	Thurston	74717.53	75923.70
2	Washington	73294.46	74992.22
30	Pierce	70320.76	71629.19
14	Franklin	68179.08	70638.03
16	Grant	68731.39	70622.45
7	Chelan	67058.48	69250.81



Cont....

The same procedure with the same results (King County #1) was done for Total Labor Workforce

County	Civ. Labor Force	Employment	Unemployment	Unemployment rate
King	1276143	1189555	86588	6.80%
Snohomish	437537	403429	34108	7.80%
Pierce	431281	398510	32771	7.60%
Spokane	240915	223152	17763	7.40%
Clark	229137	213092	16045	7.00%



Cont..

As the charts show, King County is the clear leader in both employment and Civil Labor Workforce. We will focus our search on King County (see image below) . For the next steps we will look into the top employers in King County to narrow the search down to 4 possible locations for expansion



3. Finding Top 4 Employers in WA State

To determine areas to investigate we will utilize the top 4 largest employers in the state which all happen to be in King County. To get this data the following website (<https://www.zippia.com/advice/largest-companies-in-washington/>) was scraped giving a table of the top 100 employers in the state of washington. The results were limited to the top15 of which the top 4 were used for exploration

```
url1 = 'https://www.zippia.com/advice/largest-companies-in-washington/'  
  
Jobs_df = pd.read_html(url1, header=0)[0]  
Jobs_df.head(15)
```

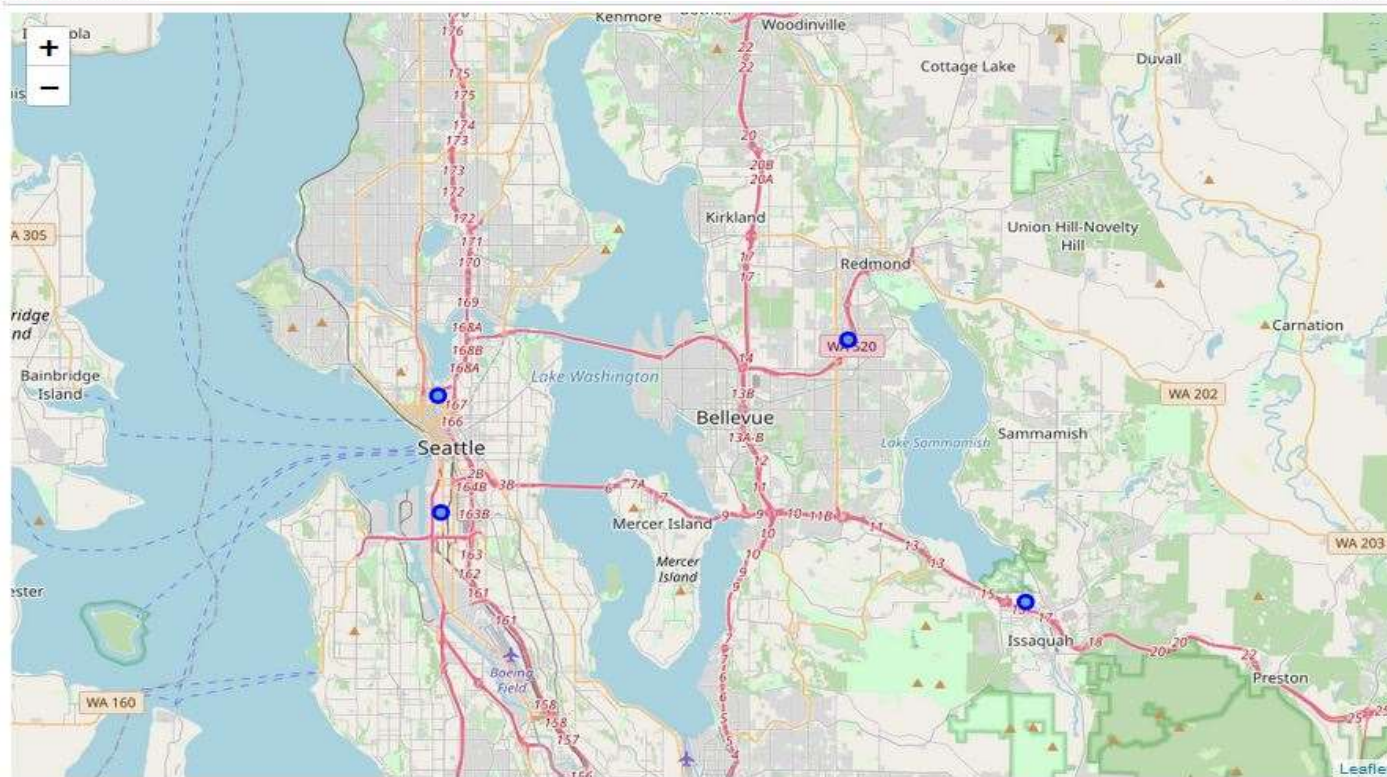
	Rank	Company	City	Employees
0	1	Amazon	Seattle	600000
1	2	Starbucks	Seattle	291000
2	3	Costco Wholesale	Issaquah	245000
3	4	Microsoft	Redmond	134944
4	5	Barrett Business Services	Vancouver	115746
5	6	Nordstrom	Seattle	74000
6	7	T-Mobile	Bellevue	51000
7	8	The Hotel Group	Edmonds	35330
8	9	University of Washington	Seattle	26110
9	10	Fortive	Everett	26000
10	11	Paccar	Bellevue	25000
11	12	Providence Health & Services	Renton	20640
12	13	Expedia Group	Bellevue	20075
13	14	Savers	Bellevue	20000
14	15	Alaska Air Group	Seattle	19214

	Rank	Company	City	Employees
0	1	Amazon	Seattle	600000
1	2	Starbucks	Seattle	291000
2	3	Costco Wholesale	Issaquah	245000
3	4	Microsoft	Redmond	134944

Cont...

Now that the top 4 employers have been identified the next step was to convert the HeadQuarters street addresses to Geo coordinates and to load those into a searchble dataframe while also identifying each location on a map of the area

	Company	Street Address	city	Zip Code	Lat	Long
0	Amazon	410 Terry Ave North	Seattle	98109	47.622150	-122.336810
1	Starbucks	2401 Utah Ave South	Seattle	98134	47.580769	-122.335243
2	Costco	999 Lake Drive	Issaquah	98027	47.548740	-122.050810
3	MicroSoft	15010 N.E. 36th St.	Redmond	98052	47.642440	-122.137030



4. Using FourSquare to explore 4 potential expansion sites

Exploring the area around The Amazon HQ



Cont..

Here we can see the top 10 closest venues to the Amazon Head Quarters. There are 58 unique venues with Coffee Shops being the most abundant. Also listed are the most commonly visited venues

	name	categories	lat	lng
0	Herkimer Coffee	Coffee Shop	47.621510	-122.339621
1	The Bar Method	Gym / Fitness Center	47.619577	-122.338256
2	Cherry Street Coffee	Coffee Shop	47.623291	-122.334234
3	Serious Pie	Pizza Place	47.622131	-122.338576
4	Veggie Grill	Vegetarian / Vegan Restaurant	47.623078	-122.336920
5	Glazer's Camera	Camera Store	47.623186	-122.340542
6	Cascade Coffee Works	Coffee Shop	47.620797	-122.334713
7	Elm Coffee Roasters	Coffee Shop	47.620570	-122.339559
8	Gold Bar	Bar	47.620559	-122.339892
9	Evoke Coffee Co	Coffee Shop	47.620780	-122.339828

```
print('There are {} unique categories.'.format(len(nearby_venues['categories'].unique())))
```

There are 58 unique categories.

```
amazon_nv_df = nearby_venues['categories'].value_counts()  
amazon_nv_df.head()
```

```
Coffee Shop      13  
Hotel             7  
Museum           4  
Gym / Fitness Center  3  
Yoga Studio       3  
Name: categories, dtype: int64
```

categories	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
	Bar	Yoga Studio	Diner	Liquor Store	Jazz Club

Cont...

Here is the venue information for the Costco site. 20 unique categories with a tie between Sandwich place and Mexican Restaurant for most abundant. Mostly food venues in this location

Exploring the area around The Costco HQ



categories	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
	Vietnamese Restaurant	Thai Restaurant	Bakery	Breakfast Spot	Burger Joint

	name	name	name	lat	lng
3	Cafe Rio Mexican Grill	grubhub	/delivery_provider_grubhub_20180129.png	47.544628	-122.049124
7	12th Ave Cafe	grubhub	/delivery_provider_grubhub_20180129.png	47.543426	-122.050426
8	Madras Dosa Corner	grubhub	/delivery_provider_grubhub_20180129.png	47.544033	-122.055337
11	Five Guys	grubhub	/delivery_provider_grubhub_20180129.png	47.543476	-122.050023
16	Noodle Boat Thai Restaurant	grubhub	/delivery_provider_grubhub_20180129.png	47.543654	-122.047050
23	Coho Cafe	grubhub	/delivery_provider_grubhub_20180129.png	47.548459	-122.041449
24	Qdoba Mexican Grill	grubhub	/delivery_provider_grubhub_20180129.png	47.548065	-122.040641
25	Top Pot Doughnuts	grubhub	/delivery_provider_grubhub_20180129.png	47.543313	-122.058046
29	Orenji Sushi & Noodles	grubhub	/delivery_provider_grubhub_20180129.png	47.550897	-122.045622
32	Garlic Crush	grubhub	/delivery_provider_grubhub_20180129.png	47.546011	-122.061997

```
print('There are {} uniques categories.'.format(len(nearby_venues1['categories'].unique())))
```

There are 20 uniques categories.

```
costco_nv_df = nearby_venues1['categories'].value_counts().head(10)
costco_nv_df
```

Sandwich Place	2
Mexican Restaurant	2
Burger Joint	2
Coffee Shop	1
Donut Shop	1
Greek Restaurant	1
American Restaurant	1
Thai Restaurant	1

Cont..

Starbucks local has 69 unique venues with Coffee shops as the most abundant. This is the only site, besides Costco that has even 1 restaurant in the top 5 most visited venue

Exploring the area around The Starbucks HQ



	name	categories	lat	lng
0	Living Computer Museum	Museum	47.582506	-122.334363
1	Starbucks Reserve SoDo	Coffee Shop	47.580899	-122.335145
2	Starbucks HQ	Coffee Shop	47.580584	-122.335378
3	Cafe Con Leche	Cuban Restaurant	47.577601	-122.334401
4	Pacific Galleries Antique Mall & Auction House	Antique Shop	47.579647	-122.330769
5	Macrina Bakery & Cafe	Bakery	47.584557	-122.334302
6	Monkey Loft	Other Nightlife	47.577207	-122.334355
7	Esquin Wine Merchants	Wine Shop	47.579706	-122.328802
8	Derby	American Restaurant	47.582718	-122.326619
9	CC Filson Co.	Accessories Store	47.587099	-122.334543

```
print('There are {} unique categories.'.format(len(nearby_venues2['categories'].unique())))
```

There are 69 unique categories.

```
STRnv_df = nearby_venues2['categories'].value_counts().head(10)
STRnv_df
```

```
Coffee Shop      6
Brewery          4
Food Truck       4
Sporting Goods Shop 4
Furniture / Home Store 3
Marijuana Dispensary 3
Fast Food Restaurant 3
Business Service  2
Grocery Store    2
Taco Place       2
Name: categories, dtype: int64
```

categories	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Winery	Cuban Restaurant	Business Service	Caribbean Restaurant	Clothing Store
--------	------------------	------------------	----------------------	----------------

Cont..

Microsoft has 64 unique venues with hotels and cafes the most abundant. This location had the 2nd most food places of the 4 sites

Exploring the area around The Microsoft HQ!



	name	categories	lat	lng
0	Microsoft Company Store	Electronics Store	47.642362	-122.137166
1	MOD Pizza	Pizza Place	47.644552	-122.136737
2	Pro Sports Club	Gym	47.651755	-122.145046
3	Aloft Seattle Redmond	Hotel	47.635439	-122.137189
4	ProSports Club Spa	Spa	47.651269	-122.144759
5	Bai Tong Thai Restaurant	Thai Restaurant	47.631761	-122.141686
6	Post Alley Café	Café	47.644501	-122.137633
7	Regent Bakery & Cafe	Bakery	47.630899	-122.138304
8	FLEE Ultimate Escape Game	Escape Room	47.630577	-122.137181
9	Microsoft Field @ Studios	Soccer Field	47.644282	-122.138396

```
: print('There are {} uniques categories.'.format(len(nearby_venues4['categories'].unique())))
```

```
There are 64 uniques categories.
```

```
: MSnv_df = nearby_venues4['categories'].value_counts().head(10)
MSnv_df
```

```
: Hotel 6
Café 6
Coffee Shop 4
Burger Joint 3
Food Truck 3
Mobile Phone Shop 3
Japanese Restaurant 3
Bus Stop 3
Chaat Place 2
Chinese Restaurant 2
Name: categories, dtype: int64
```

categories

1st Most
Common
Venue

2nd Most
Common
Venue

3rd Most
Common
Venue

4th Most
Common
Venue

5th Most
Common
Venue

Vietnamese
Restaurant

Dessert Shop

Gym /
Fitness
Center

Gym

Grocery
Store

5. Results and Summary

Results/Summary:

Based on the results illustrated in the table listed 3 of the sites (Amazon, Starbucks, and Microsoft) are quite similar and 1 (Costco) stands out on its own. Amazon, Starbucks and Microsoft all have a high number of unique venues with no restaurants listed among their top 5 most abundant venues. In contrast, the top 5 most abundant venues for Costco all involve dine-in/sit down food related venues. In fact for Costco the most commonly visited venues also all include food establishments with the 2nd most being Starbucks at 2. Below is a table of the most abundant venues and the most commonly visited venues for each of the sites explored. The next slides show the sites that I have ranked from 4 to 1 with 4 being the least recommended site and 1 being the most recommended site.

Column 1	1st MAV	2nd MAV	3rd MAV	4th MAV	5th MAV	1st MCV	2nd MCV	3rd MCV	4th MCV	5th MCV	Mean Venues
Amazon	coffee shop	hotel	museum	bakery	yoga	yoga	dessert shop	hotel	gym/fitness center	grocery store	59
Costco	mexican restaurant	burger joint	sandwich shop	japanese restaurant	greek restaurant	Vietnamese restaurant	juice bar	bakery	breakfast spot	burger joint	18
Starbucks	coffee shop	food truck	brewery	sporting goods store	furniture home store	winery	cuban restaurant	business service	caribbean restaurant	clothing store	70
Microsoft	hotel	cafe	food truck	coffee shop	bus stop	Vietnamese restaurant	discount store	gym/fitness center	grocery store	go-cart track	60
MAV = Most Abundant Venue											
MCV = Most Commonly Visited Venue											



Ranking the 4 sites from 4 to 1 with 4 being the least recommended and 1 being the most recommended



#4 – Starbucks located in the South Downtown (SODO) district of Seattle, WA



#3 – Amazon located in the South Lake Union district of Seattle, WA



#2 – Microsoft located in Redmond, WA

#1 Recommended site is.....

Based on the results of this process I am, with the utmost confidence, recommending that the site for a new French restaurant be located near the Costco Headquarters in Issaquah, Washington.

#1 Winning location is Costco in Issaquah, Wa



#6 Conclusion:

Target Audience:

This particular project is intended for anybody looking to move to a different city, state, even country and who would be looking into starting up a business there. The project demonstrated the process in identifying a "landing spot" by evaluating different areas through Data Science and API calls.

Conclusion:

This was a fictional problem rooted in reality. I am from Washington State and I now live in French speaking Switzerland. I have often wondered how an American style Brew pub would go over here in Switzerland and is where I got the idea for a French Restaurant in Washington. Only I obviously reversed the concept. After completion of this project, I will probably tackle my real-life question now that I have a much better background into Data Science. This is where my project has real world applications. Anybody looking into relocating and starting up a business can benefit from a service such as this. In fact my other topic idea was about starting up a construction type company on the Big Island of Hawaii. Building permits are easily accessible from the Hawaii.gov site which shows what kind of work is most often being done. The business uses for this are almost endless and could provide a much needed service.

#6 Discussion:

Discussion:

The task was given to help a friend, here in France, pick a new location in the United States for an expansion of their French restaurant. Knowing that the restaurant has an abundance of seafood on its menu, I suggested the state of Washington, in the Pacific Northwest region of the U.S., for its high quality of fresh seafood daily from local fishermen. However, choosing a particular location required the use of Data Science. The first step was to break Washington down into counties. To do this I downloaded, and cleaned, excel spreadsheets for the Median Income and the Civilian Labor force by county. The results were then filtered largest to smallest and graphed. The graphs both showed that King County, with a median income of \$92,693.33 a work force of 1,276,143 and a modest unemployment rate of 6.8%, was the clear winner. For the next step I scraped the website www.zipppia.com to obtain a list of the top 100 companies in the state of Washington taking the top 4 companies, all of which happened to be headquartered in King County, and using the headquarters as the starting points for our searches. The 4 companies are Amazon, Costco, Starbucks and Microsoft which involved 3 different cities: Issaquah, Redmond and Seattle. Using an API call to foursquare, and one-hot encoding, I was able to analyze the different venues (limit set at 100) in a 1500m radius around each companies headquarters. The results were quite telling and left 1 company as the clear favorite for a French restaurant expansion. The area surrounding the Costco headquarters, located in Issaquah, Wa, showed a large amount of not just restaurants, but ethnic restaurants leading towards a more diverse culture in the area. This is a good sign for non-domestic oriented cuisine. The final step is to discuss the results with the friend and discuss the next steps of the process. I would suggest a phase 2 project taking a similar approach to phase 1 but zeroing in on specific neighborhoods in Issaquah, searching for open building permits (great indicator of growth), tax rates (business, personal, property etc), housing prices and so on

#7 Sources and Acknowledgements:

Resources/Acknowledgments

<https://www.zippia.com/advice/largest-companies-in-washington/>
<https://www.arcgis.com/apps/Embed/index.html?webmap=d7b85866ae4b411083632e41aa94654b&extent=-122.7201,47.036,-120.9472,47.8044&zoom=true&scale=true&search=true&searchextent=true&legend=true&theme=light>
<https://www.kingcounty.gov/services/gis/GISData/metadata.aspx>
https://www.americanimmigrationcouncil.org/research/immigrants-in-washington?__cf_chl_jschl_tk__=92bf28816455576db75bce805816af5ed7190850-1612704203-0-AXUEPE2bcfoq9KH5KMX5Gv_-zWPRJLW2nFRMhT1zb60pGPHY_2iAzfypUYsxxdL_AkCP_fhwDIUsF3Uyj1fUfVSqNgDp6CA2CEz8eu2qXl1mTPPzJ6yn60X4kHkft58Y46_kx5EgbrUTdyauCqwrhFtgMEJOS9wf-9wYH7TRvyWagBAGAKZT_FGYW003LeukxH4A6TdQLD_6wY5jNG8jTtxtS-HSBm2EwMLW07krjlmE31S1KtLMQddYAWD7UoXa2w6_IEqsUCNVDQ2r05UrV3dlAAuivwf80zn5rCAIMUqUyvxF-uAttWaL4J19EsZFZSK8ICs193fUo4nEo1dn_AmQCJh0dFh6s02Ta77_CGx6BttsU1AjoTGvN2luyqo0r-frzQf0fTuOuiYGj_PTbfP-2Yl5mXLrwr3ZMWIZUh
<https://esd.wa.gov/labormarketinfo/labor-force>
<https://kuow-prod.imgix.net/store/aa24c79416e67b62e97f40f09a908409.jpg?ixlib=rails-2.1.4&auto=format&crop=faces&fit=crop&h=634&w=924>
https://uploads.visitseattle.org/2015/02/LakeUnion_NickHallPhotography.jpg
<https://www.issaquahwa.gov/ImageRepository/Document?documentID=6378>
<https://kuow-prod.imgix.net/store/1f566eb1ee357e56bfedc3f75864c7b9.jpg?ixlib=rails-2.1.4&auto=format&crop=faces&fit=crop&h=634&w=924>
<https://urbonner.files.wordpress.com/2013/04/crepes.jpg>
<https://www.ofm.wa.gov/washington-data-research/economy-and-labor-force/median-household-income-estimates>