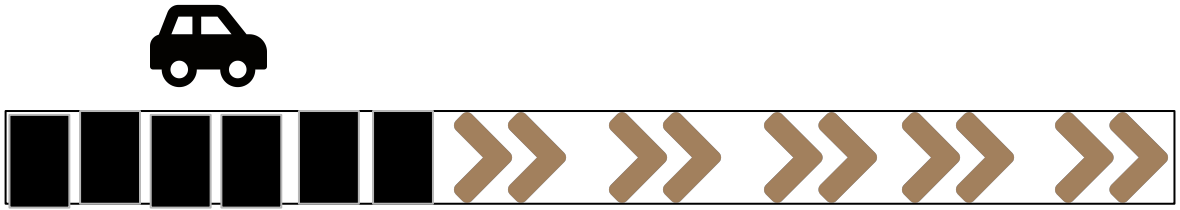
Incorporating Public Transit into Measures of Accessibility

UBCO 2021 Capstone Project

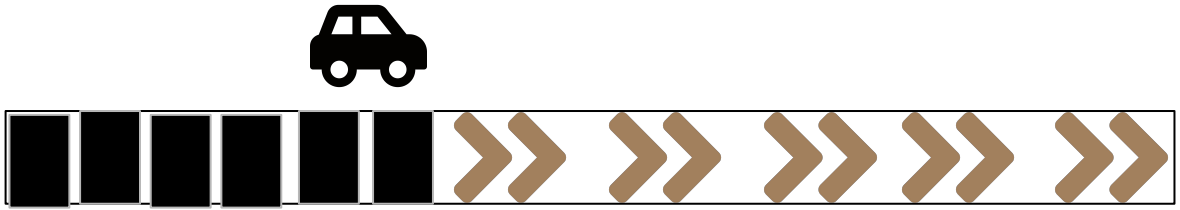
Overall progress

Week 3



Week 3

Overall Progress and Completed Tasks



Last Week’s Tasks

Week 3

Last Week’s Tasks

Week 3

- Created outline of Dashboard in Dash

Last Week’s Tasks

Week 3

- Created outline of Dashboard in Dash
- Developed visualizations of accessibility scores

Last Week’s Tasks

Week 3

- Created outline of Dashboard in Dash
- Developed visualizations of accessibility scores
- Developed weight index

Last Week’s Tasks

Week 3

- Created outline of Dashboard in Dash
- Developed visualizations of accessibility scores
- Developed weight index
- Developed scores incorporating only nearest amenities

Last Week’s Tasks

Week 3

- Created outline of Dashboard in Dash
- Developed visualizations of accessibility scores
- Developed weight index
- Developed scores incorporating only nearest amenities
- Developed individual scores for each amenity type

Last Week's Tasks

Week 3

- Created outline of Dashboard in Dash
- Developed visualizations of accessibility scores
- Developed weight index
- Developed scores incorporating only nearest amenities
- Developed individual scores for each amenity type
- Created model to determine accessibility for each amenity

Summary of team and individual tasks

	Luka	Graham	Yuxuan	Rain	All
R Viz					
Amenity index					
Local Kepler.gl					
Dashboard outline					
Scoring Sets					

Preliminary Results

- Transit Accessibility Score Sets
- Visualizations
  - R (leaflet)
  - Python (plotly)
  - Local Kepler.gl using Python

Recall from last week

- Used R5R to compute Vancouver's travel time matrix
  - **All** dissemination blocks to **All** cultural amenities (15,197 x 432) x 36 > 230,000,000
  - Averaging over departures on Fri/Sat/Sun every hour from 7am - 7pm (= 36)
  - Took roughly ~ 2-3 hours

O	D	T	σ_T
fromId <chr>	toId <chr>	avg_unique_time <dbl>	sd_unique_time <dbl>
59150004004	10	99.76316	5.364721
59150004004	15	72.48718	3.401794
59150004004	157	96.69231	3.001349
59150004004	1759	106.82051	4.388213
59150004004	1760	46.58974	2.642944
59150004004	1822	76.64103	3.990035
59150004004	1839	76.15385	2.680715
59150004004	1840	75.15385	2.680715
59150004004	1916	99.07692	3.571706
59150004004	1930	86.97436	4.923024

o_i — d_i — t_{o_i,d_i} — $\sigma_{t_{o_i,d_i}}$

Recall from last week

- Google Review Data

poi_name <chr>	open_days <dbl>	Total_hours <dbl>	Rating <dbl>	Total_Review <dbl>
Science World At Telus World of Science	7	49.00	4.5	8833
Van Dusen Botanical Garden	7	49.00	4.6	6937
Bloedel Conservatory	7	47.25	4.6	4466
Lynn Canyon Ecology Centre	*NA*	*NA*	4.7	4053
Dr. Sun Yat-Sen Classical Chinese Garden	2	10.00	4.2	3885
Museum of Anthropology	6	42.00	4.7	3406
UBC Museum of Anthropology	6	42.00	4.7	3406
Burnaby Village Museum	*NA*	*NA*	4.6	2167
UBC Botanical Garden	5	35.00	4.5	936
Beaty Biodiversity Museum	6	42.00	4.5	662

Recall from last week

- There are missing values !

poi_name <chr>	open_days <dbl>	Total_hours <dbl>	Rating <dbl>	Total_Review <dbl>
Science World At Telus World of Science	7	49.00	4.5	8833
Van Dusen Botanical Garden	7	49.00	4.6	6937
Bloedel Conservatory	7	47.25	4.6	4466
Lynn Canyon Ecology Centre	*NA*	*NA*	4.7	4053
Dr. Sun Yat-Sen Classical Chinese Garden	2	10.00	4.2	3885
Museum of Anthropology	6	42.00	4.7	3406
UBC Museum of Anthropology	6	42.00	4.7	3406
Burnaby Village Museum	*NA*	*NA*	4.6	2167
UBC Botanical Garden	5	35.00	4.5	936
Beaty Biodiversity Museum	6	42.00	4.5	662

Preliminary Results

- Replaced NAs with column mean

poi_name <chr>	open_days <dbl>	Total_hours <dbl>	Rating <dbl>	Total_Review <dbl>
Science World At Telus World of Science	7	49.00000	4.500000	8833.0000
Van Dusen Botanical Garden	7	49.00000	4.600000	6937.0000
Bloedel Conservatory	7	47.25000	4.600000	4466.0000
Lynn Canyon Ecology Centre	4	35.88077	4.700000	4053.0000
Dr. Sun Yat-Sen Classical Chinese Garden	2	10.00000	4.200000	3885.0000
Museum of Anthropology	6	42.00000	4.700000	3406.0000
UBC Museum of Anthropology	6	42.00000	4.700000	3406.0000
Burnaby Village Museum	4	35.88077	4.600000	2167.0000
UBC Botanical Garden	5	35.00000	4.500000	936.0000
Beaty Biodiversity Museum	6	42.00000	4.500000	662.0000

Preliminary Results

- Normalized data for each Feature

$$Normalize.Values = \frac{Values - Mean}{Max - Min}$$

open_days <dbl>	Total_hours <dbl>	Rating <dbl>	Total_Review <dbl>
1.0000000	0.39423077	0.6875000	1.0000000000
1.0000000	0.39423077	0.7500000	0.7853260870
1.0000000	0.37740385	0.7500000	0.5055480072
0.5000000	0.26808432	0.8125000	0.4587862319
0.1666667	0.01923077	0.5000000	0.4397644928
0.8333333	0.32692308	0.8125000	0.3855298913
0.8333333	0.32692308	0.8125000	0.3855298913
0.5000000	0.26808432	0.7500000	0.2452445652
0.6666667	0.25961538	0.6875000	0.1058650362
0.8333333	0.32692308	0.6875000	0.0748414855

Accomplishments

- Derived the Accessibility Index for Museum

$Nor(Hours)$ $Nor(Rating)$ $Nor(Reviews)$

$Nor(Days)$

open_days <dbl>	Total_hours <dbl>	Rating <dbl>	Total_Review <dbl>	Index <dbl>
1.0000000	0.39423077	0.6875000	1.0000000000	0.77043269
1.0000000	0.39423077	0.7500000	0.7853260870	0.73238921
1.0000000	0.37740385	0.7500000	0.5055480072	0.65823796
0.5000000	0.26808432	0.8125000	0.4587862319	0.50984264
0.1666667	0.01923077	0.5000000	0.4397644928	0.28141548
0.8333333	0.32692308	0.8125000	0.3855298913	0.58957158
0.8333333	0.32692308	0.8125000	0.3855298913	0.58957158
0.5000000	0.26808432	0.7500000	0.2452445652	0.44083222
0.6666667	0.25961538	0.6875000	0.1058650362	0.42991177
0.8333333	0.32692308	0.6875000	0.0748414855	0.48064947

$$Naive\_Score = \frac{Nor(Hour) + Nor(Days) + Nor(Rating) + Nor(Reviews)}{N_i}$$

Interactions with the client

Pitfall (save for discussion)

- Different amenity types should be weighed differently or kept separate (ie. museums, libraries, and theatres are very different)

- Accessibility to the nearest 1-3 amenities gives a more realistic score than when considering all amenities

Can be addressed with modifications in aggregation.

Accomplishments

Visualization Methods:

1. Python
2. R
3. Kepler.gl
    a. All amenities
    b. Different types
    c. Shortest travel time

Accomplishments

- Shiny Dashboard

Roadblocks

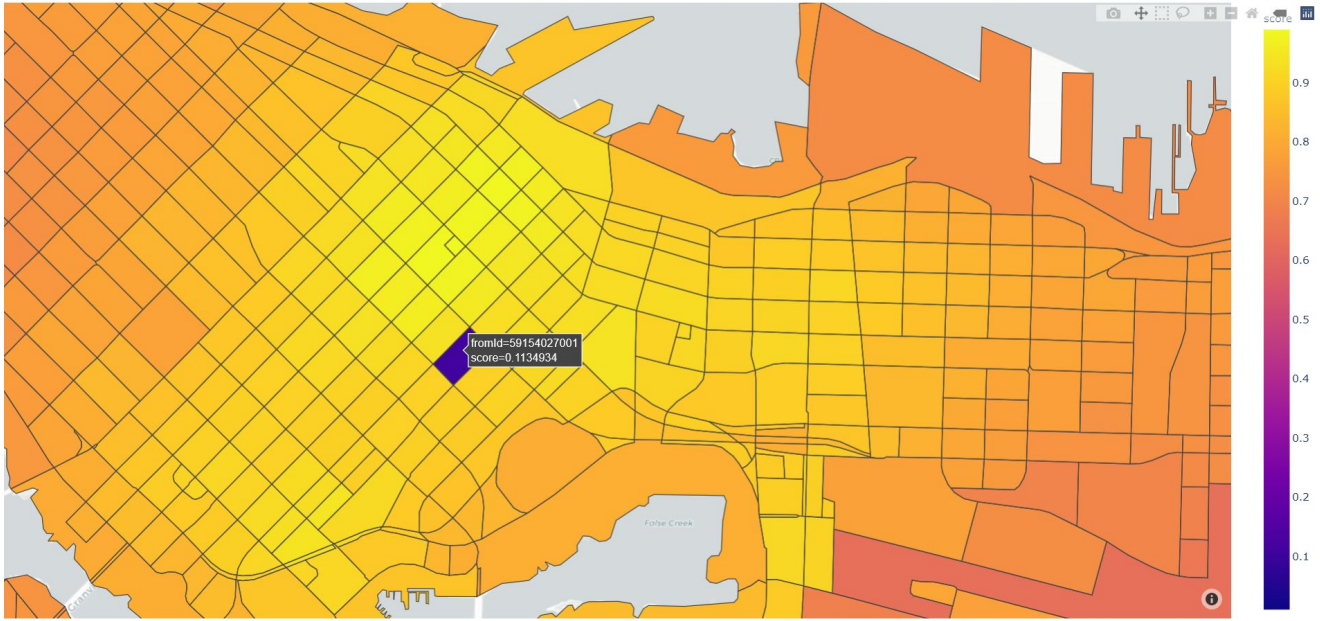
- Selecting the prefered Dashboard Language (R vs Python)
  - Incorporate all amenities
  - Dashboard style

Roadblocks

- Selecting the prefered Dashboard Language (R vs Python)
  - Incorporate all amenities
  - Dashboard style
- Libraries Information
  - Lack of information for each individual library
  - Datasets are mainly for different library systems

Roadblocks

- Strange low access outlier in city centre (to be investigated)



Week 4 Plan

- Construct dashboard in shiny
- Embed kepler.gl html into R dashboard
- Start methodology report writing
- Amenity accessibility feature engineering
- Create score set for transit efficiency

Week 3 Timeline

Weekly Tasks *This week's theme:* Dashboards and Scoring Models	Mon. 5/17/2021	Tue. 5/18/2021	Wed. 5/19/2021	Thu. 5/20/2021	Fri. 5/21/2021	Sat. 5/22/2021
Sketch our dashboard outline	Graham ☑					
First Dashboard build with plotly (R or Python?)	Graham ☑	Graham ☑	Graham ☑	Graham ☑		
Interactive plotly visualization of accessibility scores			Graham ☑	Graham ☑		
Using Urban Access + Pandana for Time computation (a second methodology for benchmarking r5r)	Rain ☑	Rain ☑				
Using Urban Access + Pandana for visualization	Rain ☑	Rain ☑	Rain ☑			
Add weight factor for different categories (libraries)			Rain ☑	Rain ☑		
Developed museum weight index	Yuxuan ☑	Yuxuan ☑				
Keplr.gl in local dash file (deck.gl)				Rain ☑	Rain ☑	
R top down network visualization	Luka ☑	Luka ☑	Luka ☑			
Develop multiple sets of scores to incorporate Joseph's "nearest amenity type" feedback		Luka + Anyone ☑	Luka + Anyone ☑	Luka + Anyone ☑		
Document all models for score computation		Luka + Anyone ☑	Luka + Anyone ☑	Luka + Anyone ☑		
Presentation Slides	All ☑					
Client meeting and feedback	All ☑					
Develop the amenity popuarity weight index using Yuxuan's API data on Google reviews		All ☑	All ☑			
Visualize new score models (to be done on a local Keplr instance and ggplot)				All ☑	All ☑	
Visualizng Transit efficiency by integrating DB populations (not a priority until dashboard and concensus scoring models are build						
Facebook, Twitter API for popularity weights (not a priority if using nearest amenity score model)			Yuxuan ☑	Yuxuan ☑	Yuxuan ☑	

Week 4 - Timeline

Weekly Tasks *This week's theme:* **Dashboards and Scoring Models**	Mon. 5/24/2021	Tue. 5/25/2021	Wed. 5/26/2021	Thu. 5/27/2021	Fri. 5/28/2021	Sat. 5/29/2021
Embed HTML Kepler visualization in R		All ☐	All ☐	All ☐	All ☐	
Determine prefered language for dashboard --> R	All ☑	All ☐	All ☐	All ☐	All ☐	
Create outline for dahsboard in shiny	Graham ☐	Graham ☐	Graham ☐	Graham ☐	Graham ☐	
Create weekly presentation	All ☑					
Report writing - Methodologies - Weight Index	Yuxuan ☐	Yuxuan ☐	Yuxuan ☐	Yuxuan ☐	Yuxuan ☐	
Report writing - Methodologies - Score Index	Luka ☐	Luka ☐	Luka ☐	Luka ☐	Luka ☐	
Report writing - Methodologies - Visualization + algorithm comparison	Rain ☐	Rain ☐	Rain ☐	Rain ☐	Rain ☐	
Report writing - Intro, Background						
Feature engineering for amenity accessibility (combining Google API features)	Yuxuan ☐	Yuxuan ☐	Yuxuan ☐	Yuxuan ☐	Yuxuan ☐	
Create Score set for transit efficiency (using population in conjuction to access)			Luka ☐	Luka ☐	Luka ☐	
Incorporate library capacity into the weight of the scores			Rain ☐	Rain ☐	Rain ☐	

Closing Remarks

Questions?