

Traffic Congestion Analysis for Rialto

DS-Path Summer Fellowship 2022

Hosted by

University of California, Riverside

Team- Traffic Congestion Analysis for Rialto

- Mentors:



Dr. Hou, Yunfei; Professor at CSUSB



Dr. Raffi, Der Wartanian, Professor at CSUSB

- Teammates:



Baldonado, Eugene
BA Economics, UCR



Fukushima, Eiko
MS Mathematics, CSUSB



Menon, Vishal
BS Computer Science, UCR



Shinde, Sonal
MS Information Systems and
Technology, CSUSB



Tong, Spencer
BS Computer Science, UCR

Agenda

- Introduction (Slide 4)
- Methods (Slides 6-7)
- Data Analysis (Slide 9-11)
- Results (Slides 13-17)
- Challenges Faced (Slide 19)
- Conclusion & Future Scope (Slide 21)

Introduction



Objective: Determine points of traffic congestion by performing trip analysis between Rialto itself, Rialto and San Bernardino county, and Rialto and surrounding counties



Field of study: Rialto, CA



Data Sources: ESRI Living Atlas, OSM (Open – Source Mapping), StreetLight Data InSight, SCAG, LEHD, US Census Bureau.

Methods

Second
Lines

Tangent
Line

$x+h$

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

$$f'(x) = \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h}$$

$$= \lim_{h \rightarrow 0} \frac{x^2 + 2xh + h^2 - x^2}{h}$$

$$= \lim_{h \rightarrow 0} \frac{2xh + h^2}{h}$$

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

Methods

- For our project, we decided to focus on 3 different type of trips: Rialto to Rialto, Rialto to another city within San Bernardino County, and Rialto to the other 3 counties (Los Angeles, Riverside, and Orange)
- We obtained our data using Streetlight Data, a big data transportation analytics software. Streetlight was chosen as our data vendor because they support the Trip to/from Pre-set Geography analysis necessary for our project
 - Data is collected through pings of mobile devices, pinging stopped = end of trip
- Analysis resulted in 4,406,200 datapoints
- Software used: Streetlight, ArcGIS, Python, and Excel
 - For obtaining the data, Streetlight Data and ArcGIS were mainly used
 - In order to obtain the SB County city block groups in the form of CSV files, Excel was used
 - For the actual Data Analysis, Python and ArcGIS were used



Methods (Cont'd)

Data Analysis

- Data was obtained by first uploading a shapefile of Rialto, CA to Streetlight Data and then performing a Trip to/from Pre-set Geography analysis
 - Shapefile of Rialto, CA was created using ArcGIS
 - A "Select by Location" query was performed where the target layers were the USA census BGs and the source layer was the Rialto, CA boundary
 - Spatial selection method was chosen as "have their centroid within source"
 - Creates a "lookup" table that corresponds census BG GEOIDs to our boundary
 - Zipped shapefile was uploaded as zone set
 - Streetlight Analysis was then performed; key parts of the analysis were the basic info, time periods, zones, and add-ons

The screenshot shows the Streetlight Data analysis interface with several panels:

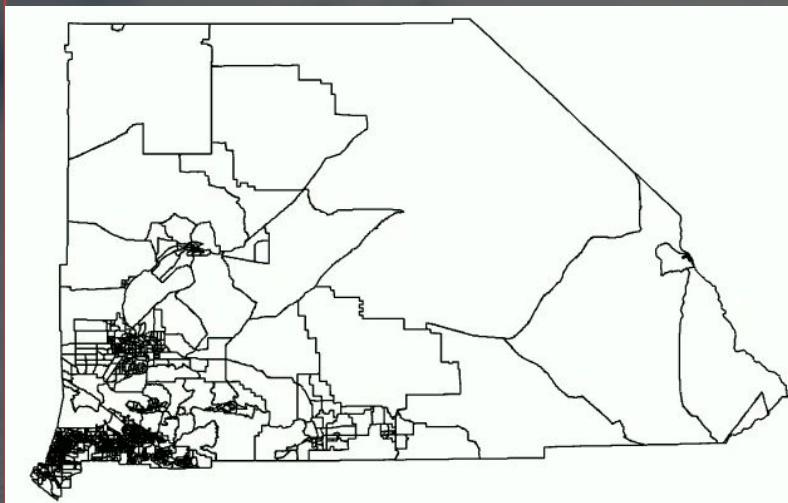
- Unit of Measurement:** Miles
- Day Types:** Segmentation for the days of data in your analysis.

Name	Start	End
All Days	Monday	Sunday
Weekday	Monday	Thursday
M-F	Monday	Friday
Tuesday	Tuesday	Tuesday
Workday	Tuesday	Thursday
Fri-Sun	Friday	Sunday
Weekend Day	Saturday	Sunday
- Day Parts:** Segmentation of parts of the day (in hours) for your analysis.

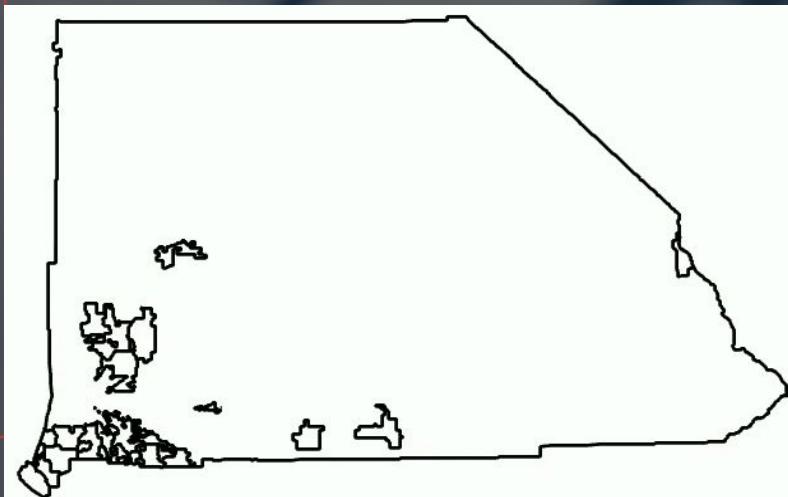
Name	Start	End
All Day	12am	12am
Early AM	12am	6am
Peak AM	6am	10am
Mid-Day	10am	3pm
Peak PM	3pm	7pm
Late PM	7pm	12am
- Pre-Set Geography:** 2020 US Census Block Groups
- Zones:** FinalRialtoBGB 52 ZONE(S)
- Output Type:** StreetLight Index (Device Trips)

Data Analysis

Block Group

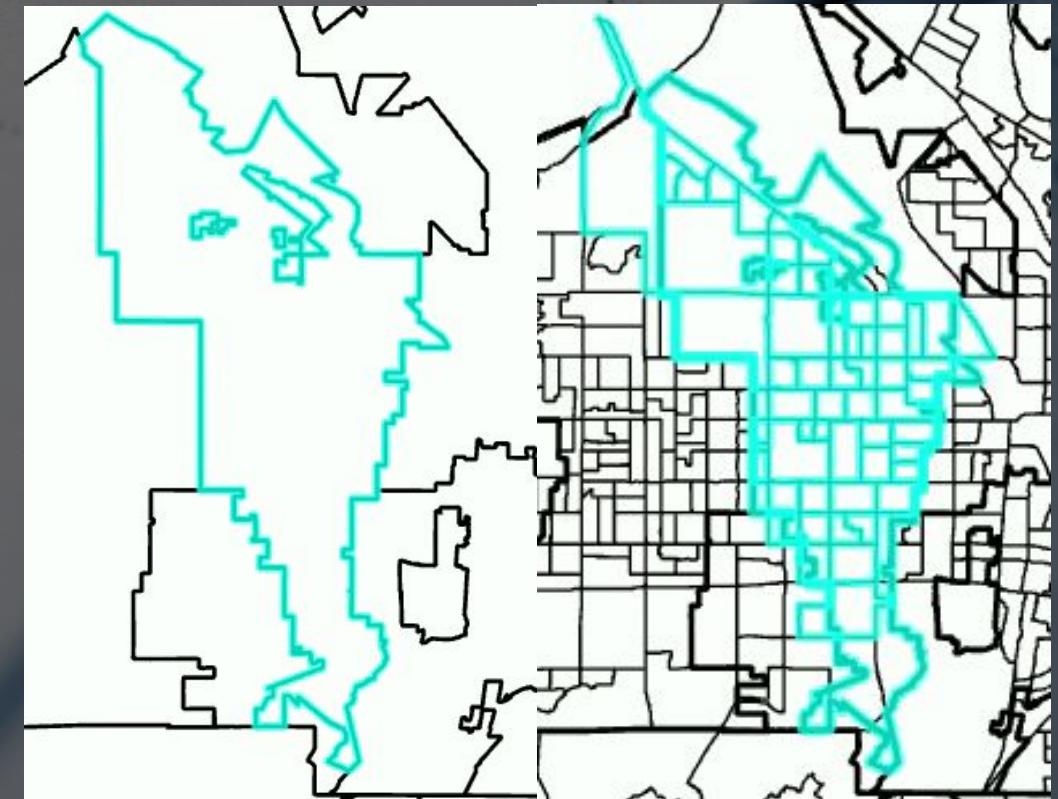


San Bernardino County



Data Analysis

- Create Selection in ArcGIS showing which block groups are associated with Rialto



- Repeat process for each city in San Bernardino County (And for other counties), tagging each for which city they're associated to in a .csv file

Data Cleaning

- Using Python, read the tables created, and then merge them with StreetLight data for Traffic Percentages by the block group's FIPS code, creating a dummy variable for the City they're associated with
 - Remove unnecessary columns, and drop NA values (No percentages in any Traffic Metric)
 - Aggregate each of the Traffic percentages from each City's block-groups (And Counties)

```
cityList2 = { "isAdelantoBG", "isAppleValleyBG", "isBarstowBG", "isBigBearLakeBG", "isChinoBG", "isChinoHillsBG", "isColtonBG",  
  
for city in cityList2:  
    globals()[f"ADF{city}"] = caliBG.loc[caliBG[f"{city}"] == True, 'totalTrafficAllDayFrom'].sum()  
    globals()[f"PMT{city}"] = caliBG.loc[caliBG[f"{city}"] == True, 'totalTrafficPmTo'].sum()  
    globals()[f"PMF{city}"] = caliBG.loc[caliBG[f"{city}"] == True, 'totalTrafficPmFrom'].sum()
```

- Create new dataframe with aggregated data
- Convert dataframe to .csv format and add them to a geodatabase in ArcGIS Pro to join tables and shapefiles attribute tables

```
for city in cityList:  
    matches = set(globals()[f"{city}"]['FIPS'])  
    caliBG[f"is{city}"] = caliBG['FIPS'].isin(matches).astype(int)
```

Join : ScagCitiesJoin/TripPercentsJoins

Cardinality	One to many
Primary key	location
Foreign key	CITY
Forward join	No
Join type	Left outer join

Source table : ScagCitiesJoin

Data Type	File Geodatabase Feature Class
Database	\lapporto.com\dfs\UCR\Users\ebald009_ucr\Di
Feature Class	ScagCitiesJoin
Alias	ScagCitiesJoin
Feature Type	Simple
Geometry Type	Polygon
Coordinates have Z value	No
Coordinates have M value	No
Attachments	No Attachments

Data Analysis

Steps to cleaning a 4,406,200 row Dataset

- Highlight "All day" time slot
 - Drop any NA (No trip volume block groups)
- Match Trip Purpose Block Group's Table with the Table associating each Block Group with its City
- Find the mean for each of the trip purpose to know what portion each City's trips were for each purpose
- Multiply Trip Purpose proportions and Trip Percentages for the Percentage of trips for each Trip Purpose

```
matches = {"0: All Day (12am-12am)"}  
tripPurpose = tripPurpose[tripPurpose['Day Part'].isin(matches)]
```

```
rialtoTrafficPercentages = {  
    'location' : "Rialto",  
    'Home to Work %' : tripPurpose.loc[tripPurpose['isRialtoBG'] == True, 'Home to Work'].mean(),  
    'Home to Other %' : tripPurpose.loc[tripPurpose['isRialtoBG'] == True, 'Home to Other'].mean(),  
    'Non-Home Based Trip' : tripPurpose.loc[tripPurpose['isRialtoBG'] == True, 'Non-Home Based Trip'].mean()  
}
```

	location	Home to Work %	Home to Other %	Non-Home Based Trip
0	Rialto	0.090244	0.570095	0.339661
1	Fontana	0.111304	0.619260	0.269435
2	San Bernardino	0.116110	0.592565	0.291325
3	Unincorporated	0.126627	0.599617	0.273757
4	Colton	0.117379	0.502237	0.380384

```
htw = tripPurposePercentages['Home to Work %'] * tripPurposePercentages['totalTrafficAllDayFrom']  
hto = tripPurposePercentages['Home to Other %'] * tripPurposePercentages['totalTrafficAllDayFrom']  
nhbt = tripPurposePercentages['Non-Home Based Trip'] * tripPurposePercentages['totalTrafficAllDayFrom']  
tripPurposePercentages['Home to Work'] = htw  
tripPurposePercentages['Home to Other'] = hto  
tripPurposePercentages['Non-Home Based Trip'] = nhbt
```

	location	Home to Work	Home to Other	Non-Home Based Trip
0	Rialto	0.039951	0.252381	0.150368
1	Fontana	0.015716	0.087440	0.038044
2	San Bernardino	0.013585	0.069330	0.034085
3	Unincorporated	0.010358	0.049049	0.022393
4	Colton	0.005435	0.023254	0.017612

Results



Community Profile & Mode of Share for Rialto in 2019

Population



103,549

Household



26,134

Income



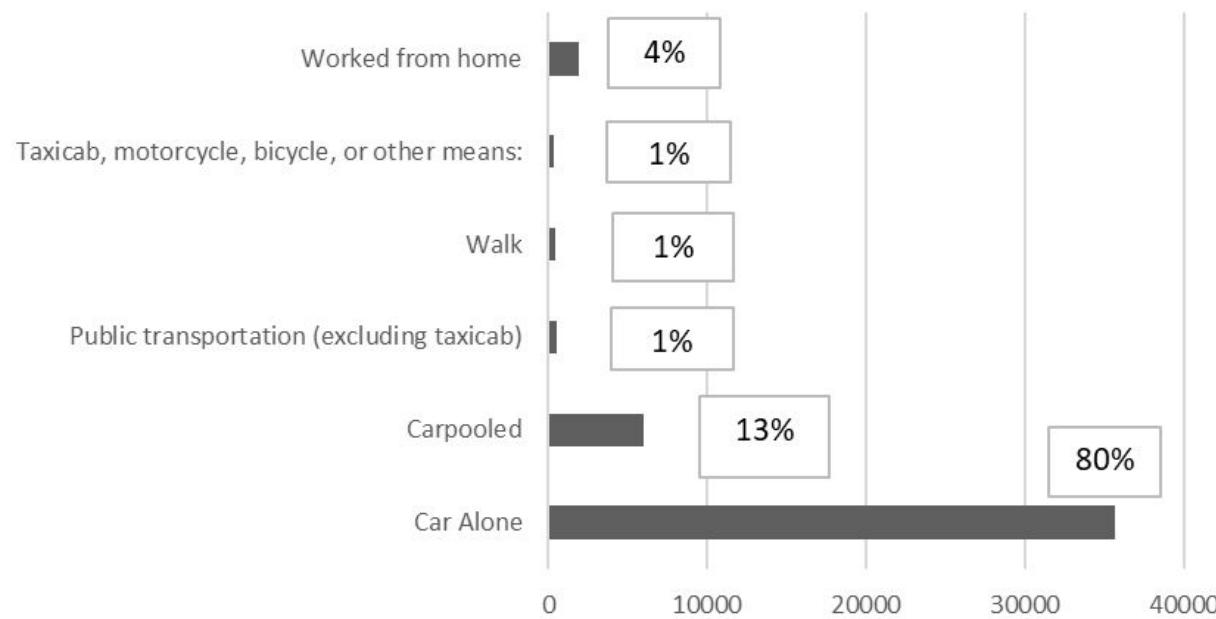
\$ 65,538

Employment

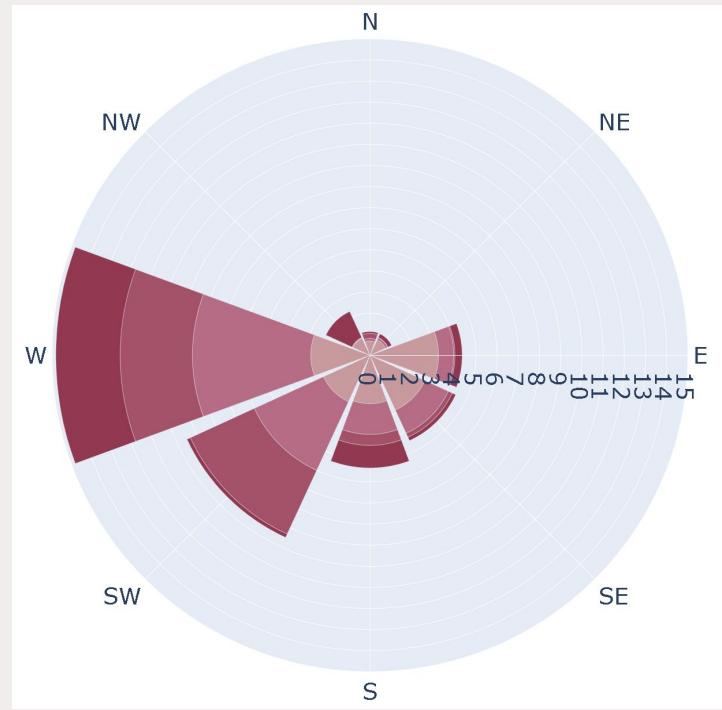


44,920

Mode Share for all Trips

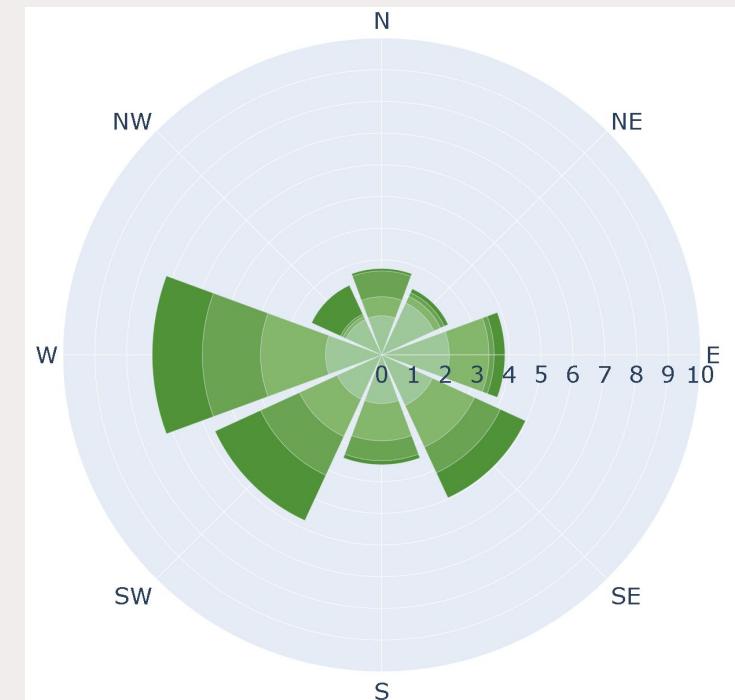


Commute Trip Distance and Direction with Live and Work share



Commute Distance and Direction from

Less Than 10 miles	15,970	37.1%
10-24 miles	12,817	29.8%
25-50 miles	7,800	18.1%
Greater than 50 miles	6,428	14.9%

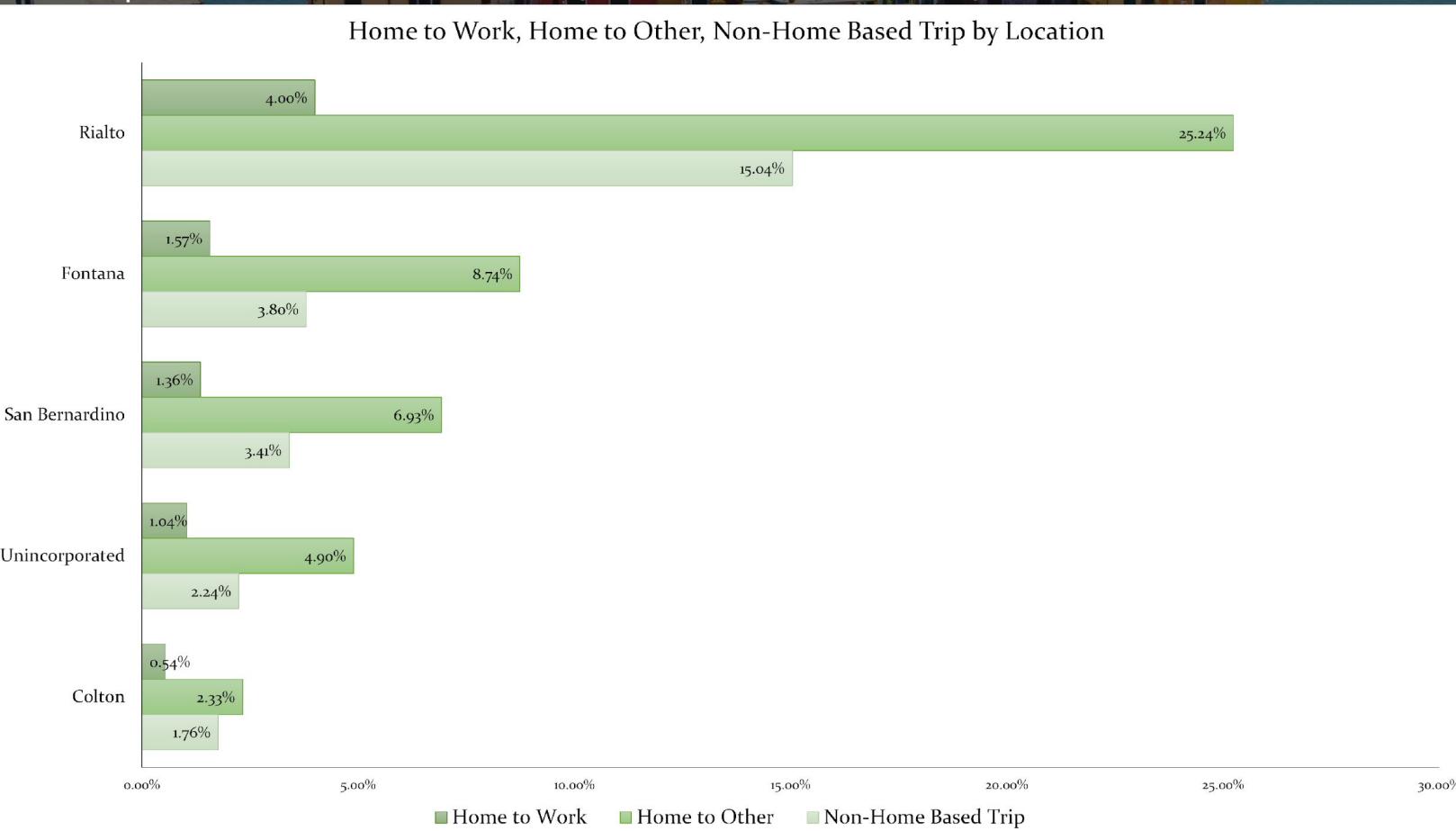


Commute Distance and Direction to Rialto

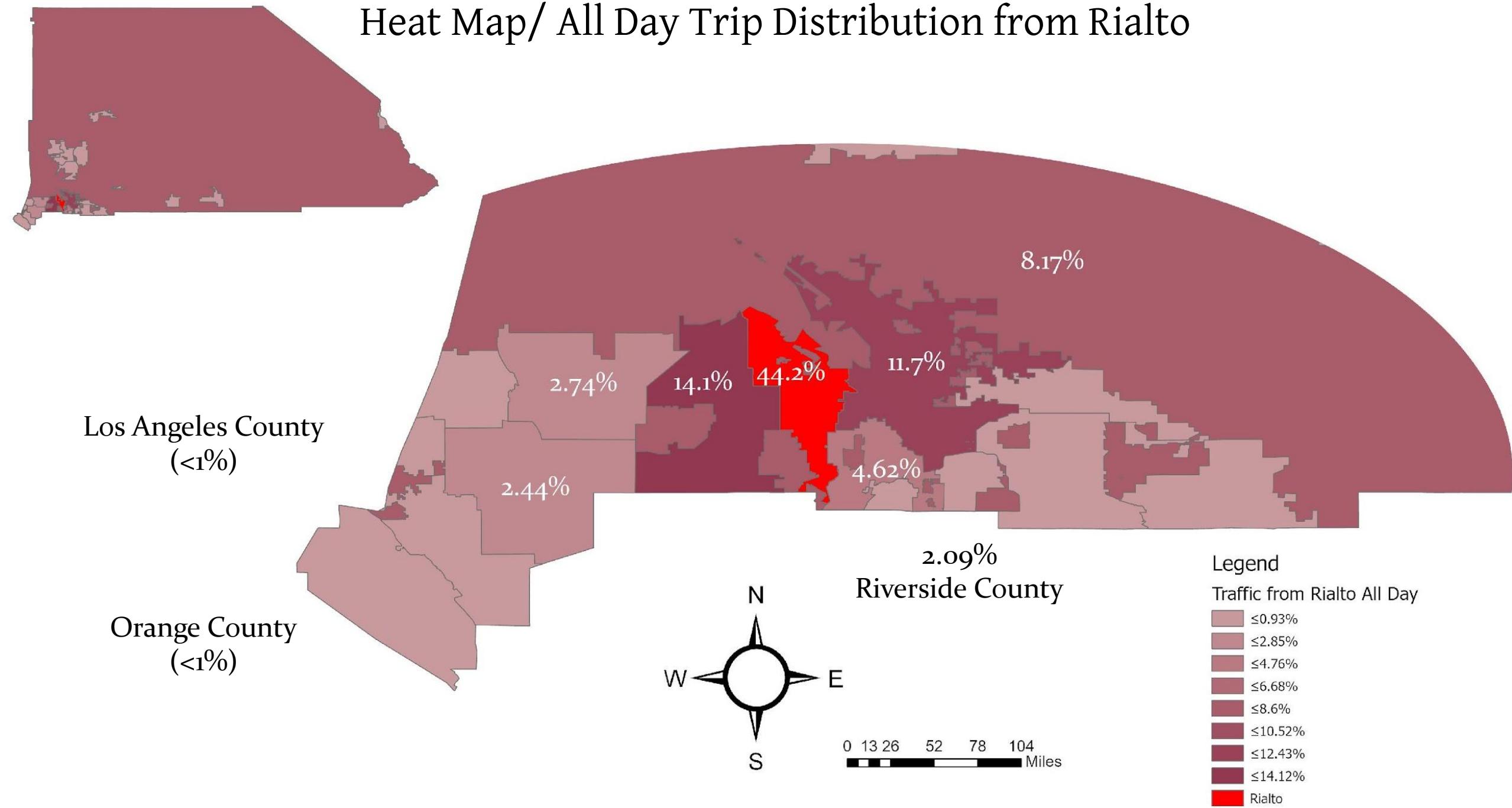
Less Than 10 miles	12,985	41.5%
10-24 miles	8,147	26.1%
25-50 miles	5,878	18.8%
Greater than 50 miles	4,242	13.6%

Top Places Visited from Rialto, CA (All Day avg.)

- The top destinations visited from Rialto, CA were Rialto (44.27%), Fontana (14.12%), San Bernardino (11.7%), the Unincorporated BG (8.18%) and Colton (4.63%)
- The top county visited from Rialto, CA was San Bernardino County (89.59%), indicating that a majority of the trips made from Rialto stayed within the county

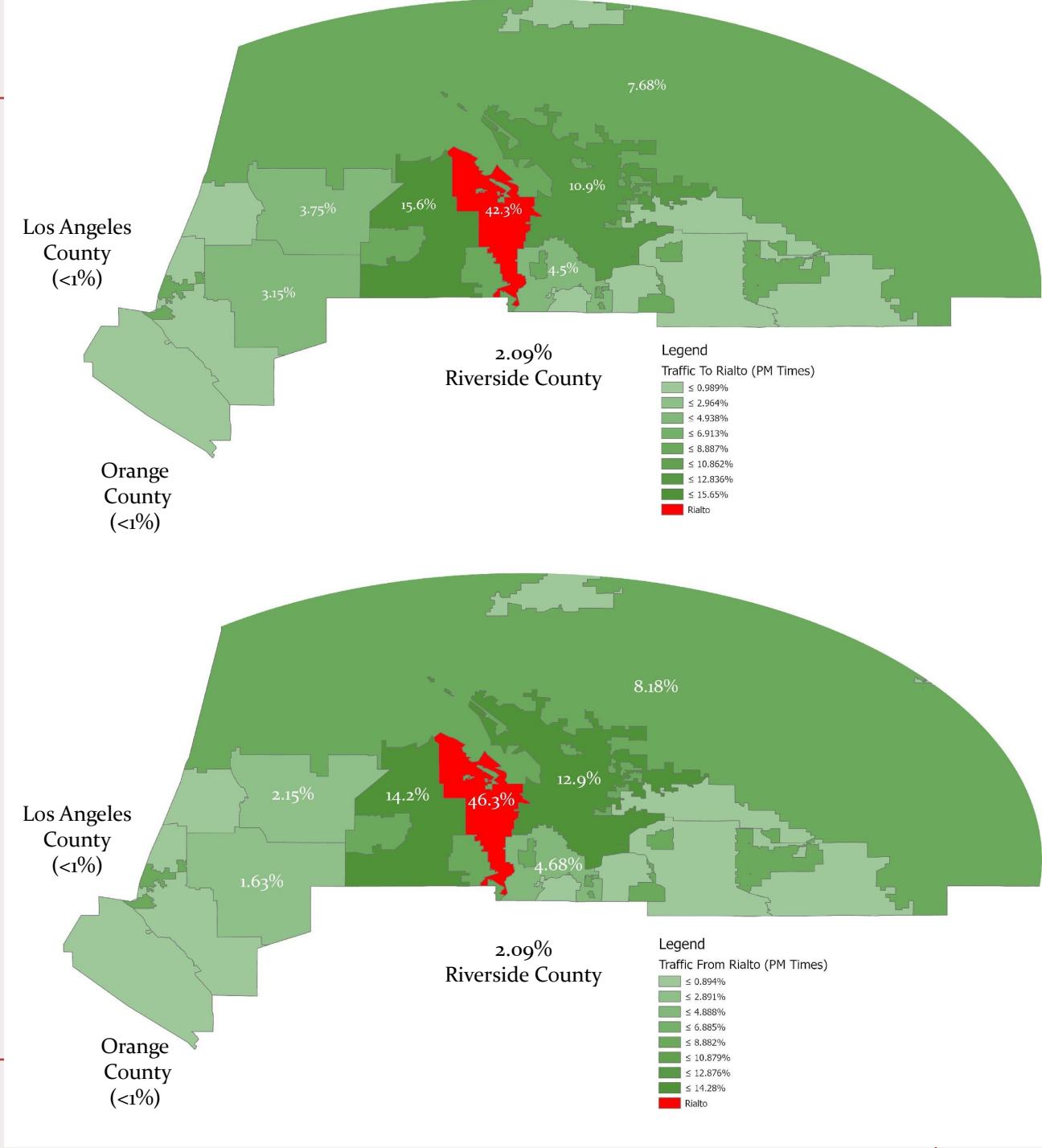


Heat Map/ All Day Trip Distribution from Rialto



PM Trip Distribution To and From Rialto

	totalTrafficAllDayFrom	totalTrafficPmTo	totalTrafficPmFrom	location
0	0.0036	0.0059	0.0004	Los Angeles County
1	0.0007	0.0012	0.0000	Orange County
2	0.0209	0.0233	0.0155	Riverside County
3	0.8959	0.8982	0.9182	San Bernardino County
4	0.4427	0.4238	0.4639	Rialto
5	0.1170	0.1093	0.1297	isSanBernardinoBG
6	0.0000	0.0000	0.0000	isAdelantoBG
7	0.0000	0.0000	0.0000	isBarstowBG
8	0.0024	0.0026	0.0027	isRedlandsBG
9	0.0000	0.0000	0.0002	isHesperiaBG
10	0.0000	0.0000	0.0000	isBigBearLakeBG
11	0.1412	0.1565	0.1428	isFontanaBG
12	0.0002	0.0002	0.0000	isChinoHillsBG
13	0.0818	0.0768	0.0818	isUnincorporatedBG
14	0.0013	0.0020	0.0007	isMontclairBG
15	0.0020	0.0032	0.0005	isChinoBG
16	0.0020	0.0034	0.0006	isUplandBG
17	0.0000	0.0000	0.0006	isVictorvilleBG
18	0.0463	0.0450	0.0469	isColtonBG
19	0.0025	0.0025	0.0019	isLomaLindaBG
20	0.0000	0.0000	0.0000	isYuccaValleyBG
21	0.0274	0.0375	0.0215	isRanchoCucamongaBG
22	0.0244	0.0315	0.0163	isOntarioBG
23	0.0000	0.0000	0.0000	isTwentyninePalmsBG
24	0.0000	0.0000	0.0000	isNeedlesBG
25	0.0017	0.0015	0.0022	isGrandTerraceBG
26	0.0000	0.0000	0.0002	isYucaipaBG
27	0.0030	0.0024	0.0057	isHighlandBG
28	0.0000	0.0000	0.0000	isAppleValleyBG



Challenges



Challenges

- Time constraint of 10 weeks
- Deciding the best way to obtain the data
- Merging data from ArcGIS and Streetlight
- Interpreting and visualizing commute distance and direction data from the LEHD Website
- Not having access to Python Geoprocessing libraries when performing data analysis
- Creating Heat Maps in ArcGIS

Conclusion



Conclusion & Future Scope



Conclusions:



The highest percentage of trips made were within Rialto (44.27%). City officials should think of policies in suburban areas that support efficient mobility within Rialto.



Since the next highest percentage of trips was made to Fontana (14.12%), the city of Rialto should come up with new transportation policies to better connect Rialto and Fontana.



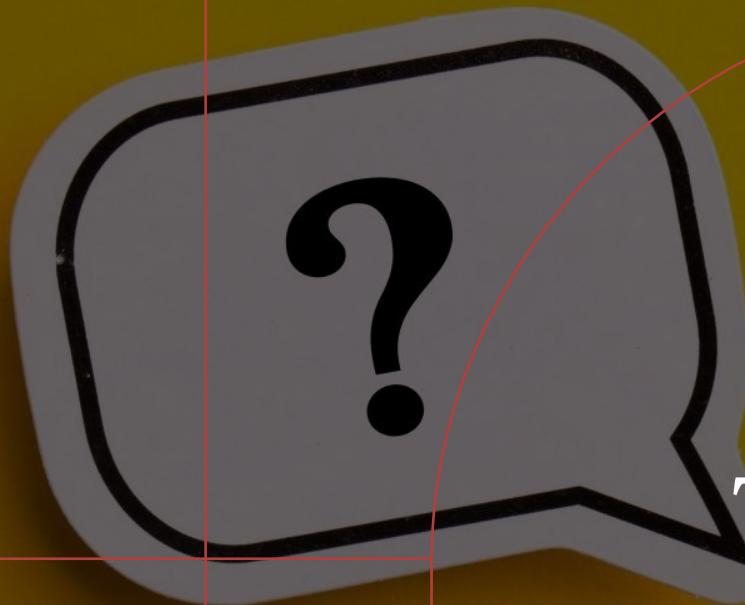
Trips from Rialto to San Bernardino County made up the highest percentage (89.59%). Additional work should be done by district legislatures to build more transportation networks within San Bernardino county.



For trip types in Rialto, home to other and home made up the highest percentage (25.2%). Urban developers and housing development projects must take into consideration how to build transportation networks to enhance the mobility of people within Rialto.



Future scope of study: Traffic patterns during pre-Covid Time times vs. now, Traffic congestion patterns in specific cities of interest in other counties



Thank you

