**Identifying Areas for Light Rail Expansion in the Phoenix Metro Area**

**Aaron Jones**

**Index**

**- Introduction and Purpose 1**

**- Method for Analysis 2**

**- Data Sources 3**

**- Analysis 4-6**

**- Conclusion and Further Analysis 7**

**- Appendix 8-12**

**Introduction and Purpose**

The Phoenix metro area is rapidly growing as people from neighboring states, particularly California, move in for the beautiful weather and relatively low cost of living.

With this growth comes vehicular traffic. The area was developed mostly after car culture took hold of the US and is therefore well planned for cars, but even great foresight has its limits. To alleviate traffic and increase mobility for non-car owners, local governments subsidize Valley Metro, an organization that builds, maintains, and runs public transit in the area. Part of that service is the light rail – an above-ground light train with stops across the city.

The light rail is an efficient, clean, and comfortable mode of transportation, but currently serves a very limited area. In this study, I'll define areas that would most benefit from light rail service and would provide the density and desirable attractions necessary for high ridership.

**Method for Analysis**

To identify these areas, we first need to define them. An area that is a good candidate for light rail service has the following characteristics.

1. Relatively High Occupied Housing Density

For any public transit project to be a success, it needs riders en masse. As the light rail in the area is in its infancy, we want to set these new stations/lines up for success by focusing on areas with high housing density. Based on occupied units per square kilometer, I’ll determine the top 25% most dense census tracts (based on 2019 data) and use those as the base candidates for service.

2. Neighborhood Venues

For a destination to be desirable, it needs to have places people want to go within walking distance; restaurants and bars as well as entertainment (movie theaters, parks, sports stadiums etc.). I’ll create a choropleth map based on venues per square kilometer and use it to determine the best locations for new light rail lines and stations.

**Data Sources**

**US Census Data**

[**https://data.census.gov**](https://data.census.gov)

The US Census provides housing unit estimates at a census tract level. On a broad scale, these tracts can be large and might not show the detail desired. However, in a dense metro area like the one studied here, census tracts cover a small geographic area and give a high level of detail – this is why this data type was chosen. This housing data has many features that can help us learn about a given area. Additionally, the US Census provides shape files (used for geographic applications) that define the census tracts themselves. This will be essential when determining housing density for a given area.

**Foursquare API**

[**https://foursquare.com**](https://foursquare.com)

Foursquare provides information on venues and searches can be centered on whatever points we want (centroids for census tracts). "Venue" is a pretty broad term in this context; the best way to define how Foursquare defines "venue" is "a place where people eat, drink, shop, workout, are entertained, or visit the doctor."

**Valley Metro**

[**https://en.wikipedia.org/wiki/List\_of\_Valley\_Metro\_Rail\_stations**](https://en.wikipedia.org/wiki/List_of_Valley_Metro_Rail_stations)

Using Valley Metro's light rail station list (via Wikipedia), I'll plot the stations and use them as a guide for any new lines/stations. Additionally, there are several expansion projects in the works. These will be taken into consideration even if it's not yet possible to plot stations. This data comes in the form of a .kml file that will need to be processed with the kml2geojson library before it can be plotted and/or analyzed.

**Analysis**

The first step in the process is to clean and analyze the study area – Maricopa County. I used geopandas, a pandas extension for geographic analysis, to import the shape file data from the US Census. The resultant dataframe (Figure 1 in Appendix) has 1526 rows and 13 columns. The key columns are

* COUNTYFP – denotes the county code for the given tract
  + Maricopa is 013
* NAME – the census tract number
* INTPTLAT/INTPTLON – the latitude and longitude of the center point of the tract
* geometry – the polygonal shape of the tract
  + used for plotting on a map

These data were then used to create a map of all census tracts and their center points in the state of Arizona (Figure 2). Census tracts don’t have a set size but are instead determined by population with the typical goal of having each tract cover about 4000 people. Given this information, it’s easy to see where the population of the state is clustered based on the center points of each tract.

From here, the dataframe was trimmed to show only census tracts in Maricopa County – some 916 rows with the same 13 columns. This greatly increased the processing speed for the dataset, as it was about 40% smaller. A final preparation step for this dataset was to determine the surface area of each census tract. This was done with geopandas method “.area” and resulted in a square meters measure of area. By dividing by 10 ^ 6, square kilometers were determined.

US Census data on occupied housing units by census tract were then merged with the Maricopa tract dataframe. To find the occupied housing unit density for each census tract, the number of occupied housing units was divided by the area in square kilometers. Based on the occupied housing density, a choropleth map centered on the Phoenix area was created as a base for determining how the existing light rail line serves the area.

The final step in the initial area analysis was to import and plot existing light rail stations. Valley Metro provides a list of stations, but it is formatted for the web and cannot be easily downloaded. Wikipedia, however, keeps an up-to-date list of existing stations that can be downloaded as a .kml file (similar to a shape file but includes three-dimensional data in a lat/lon/Z format). The python library kml2geojson has simple methods - .x and .y – to get the lat/lon data from this 3d format. From here, the existing light rail stations were plotted on top of the Phoenix area occupied housing density choropleth (Figure 3).

Based on this map, we can confidently say a couple of things about the light rail service. The line seems to do a good job of serving some densely populated areas, but doesn’t cover a lot of neighborhoods – the service is currently too limited to seriously impact how people choose to get around. Second, it’s built as a hub and spoke system – the lines run from suburban areas into the downtown hub (can be seen where the line makes an L-shaped turn in the center of the map). This is further confirmed by the future projects on Valley Metro’s website – a line will run due south from downtown and another due west. While this is great for people who want to commute to downtown jobs, those citizens who want to go neighborhood to neighborhood to socialize or conduct business have no real way to get around without going through the central hub which can mean much longer trip times – especially when compared to the alternative of driving/taking the bus.

The dataframe was then pared down to only those tracts in the top 25% of occupied housing density – a total of 229 tracts. The center points of these dense tracts were then used as search points for venues via the Foursquare API.

Only venues within 1000m of these center points were chosen – basically, 1000m is walkable. We have to remember the riders of the light rail probably won’t have bikes and certainly won’t have cars; our new stations need to put them close to where they want to be.

Of these venues, we have some that shouldn’t be listed as potential destinations for riders – namely convenience stores, gas stations, and intersections. These types of venues were removed to give a more realistic look at which census tracts have the highest density of walkable venues. To determine venue density, the total number of walkable venues was divided by the area of the tract. Resultant dataframe in Figure 4.

Once again, existing light rail stations were plotted on top of a choropleth map of these venue-dense census tracts. The difference is that this time, two new lines were laid out and plotted alongside the existing lines (Figure 5 – new lines in purple). These lines run from some of the most densely populated and venue-dense areas while passing through similar areas – a must for making sure we have riders from all over the city instead of just a couple of endpoints. Additionally, both lines cross the existing lines and connect with each other in the Scottsdale area.

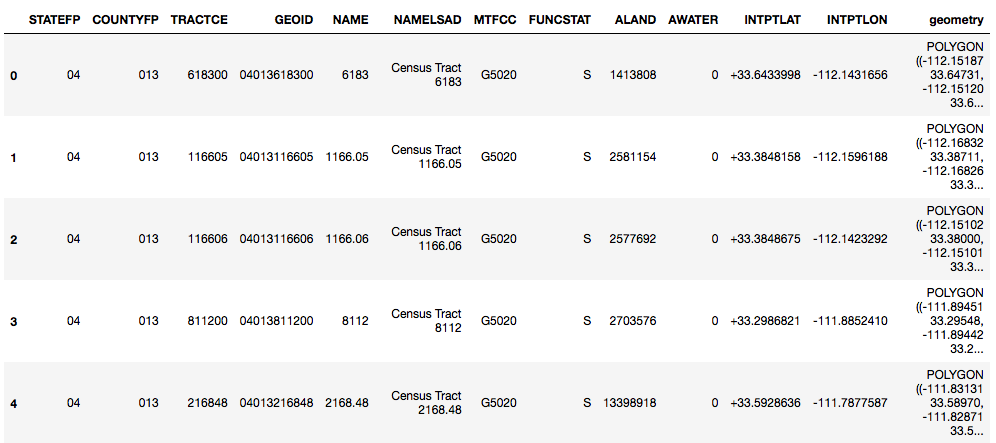
**Conclusion and Further Analysis**

The goal of this project was to determine where light rail service in the Phoenix metro area could be expanded to best serve the people and businesses in the area. I believe we've succeeded in that goal - we have two new lines running opposite to the existing lines, through densely populated areas, and into places with plenty of bars, restaurants, and venues for entertainment. If you know the Phoenix area, you know we've hit Old Town Scottsdale (a walkable location with shops, bars, and restaurants), Tempe (home to Arizona State University), and Westgate (a place with plenty of great restaurants, bars, and sports stadiums for NFL and NHL teams).

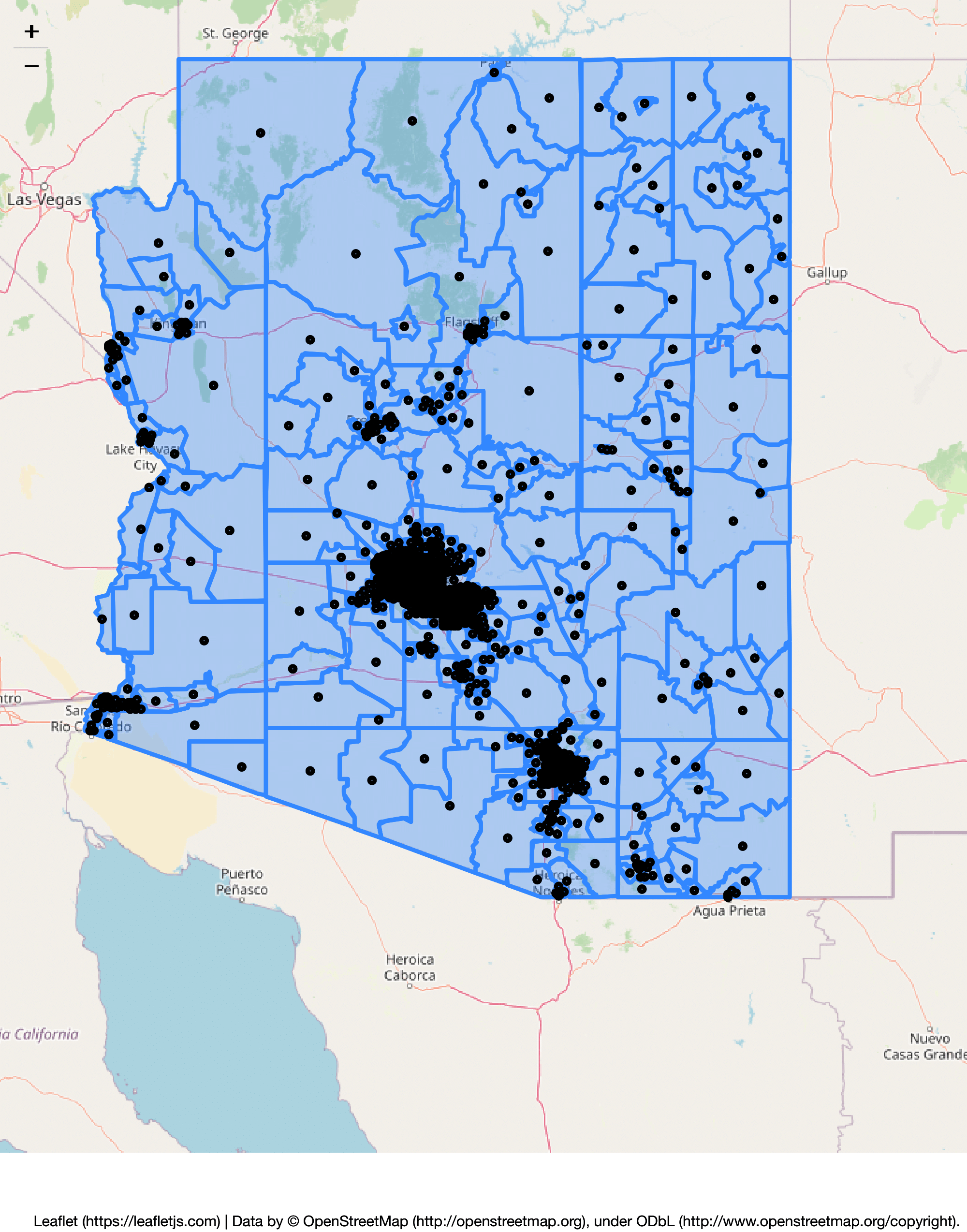
This kind of service would have benefits for residents, businesses, and the city as a whole - imagine how incidents of DUI would drop if this many nightlife locations were accessible without a car? I believe this project can be deemed a success, though it's certainly far from perfect - there are a lot of more factors to be considered when expanding public transit and this is just a surface-level view of how the system is planned and how it serves the community.

From here, I think it would be prudent to expand the project ideas by determining how venues in each of these "destination areas" are rated; we want to route people to places they want to go, not just to places where there are restaurants, bars, and stadiums. I'd also be interested in looking at traffic studies for the area to see how building new lines would impact the streets they run on as well as the number of cars on the road.

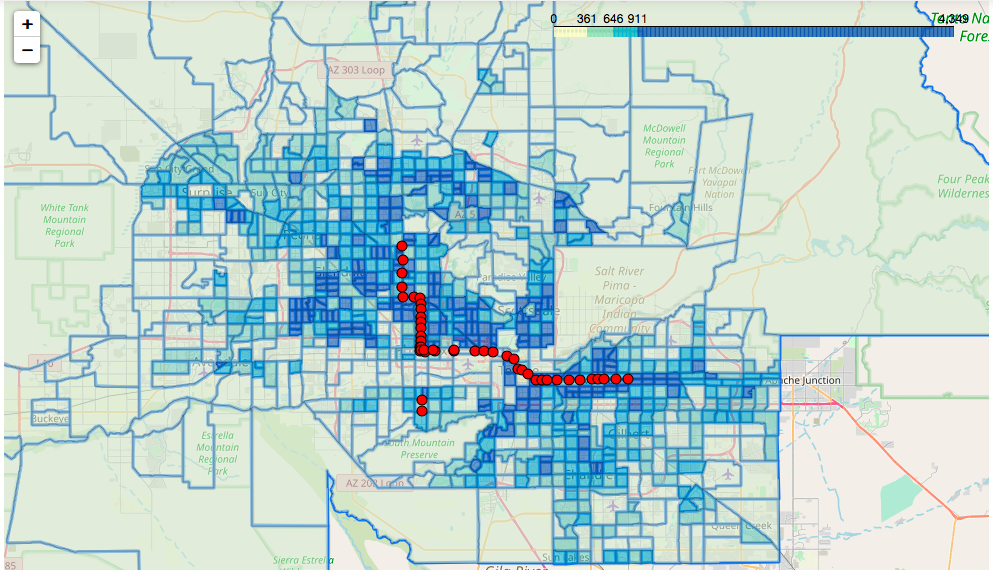
**Figure 1**

****

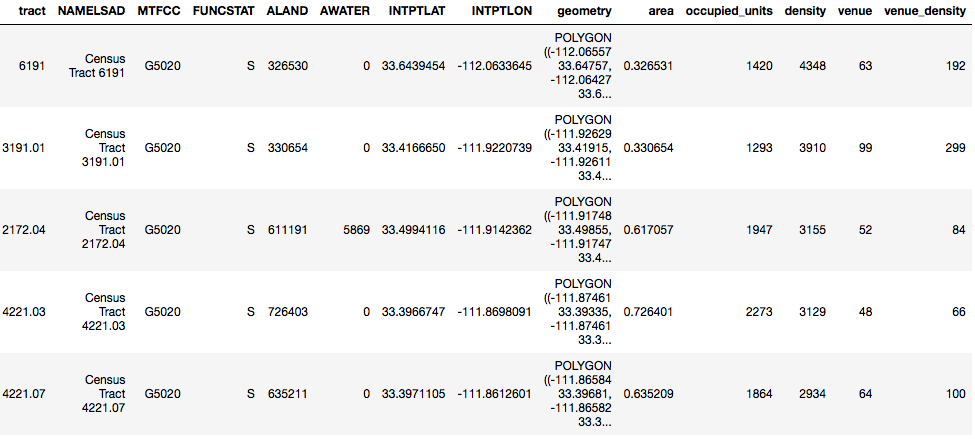
**Figure 2**

****

**Figure 3**

****

**Figure 4**

****

**Figure 5**

