

An Analysis of Minority Population Distribution and Organ Transplantation Centers in Pennsylvania

Jing Li

May 2023

1 Introduction

Organ availability is currently the biggest concern in the organ transplantation system. The allocation of an organ is dependent on proximity to donor hospital, immune system compatibility, and necessary matching criteria for a successful transplant. Although organ donation is matched based on a computerized system, inequitable access to healthcare due to racial disparities by geographic locations can be difficult to calculate. It is evident that areas with more minorities has lower economic status that leads to limited healthcare access. Understanding the patterns of race distribution and areas of low socio-economic status is critical in developing policies for resolving inequities in health.

The aim of this analysis is to provide a bi-variate overview to investigate the distribution of minority groups in comparison with nearby transplantation centers. The study will focus on the state of Pennsylvania using QGIS and ArcGIS Pro to measure each county's average proximity to transplantation centers. To compare with dot density maps for Asian, Black/African American, and Hispanic/Latino groups.

2 Data

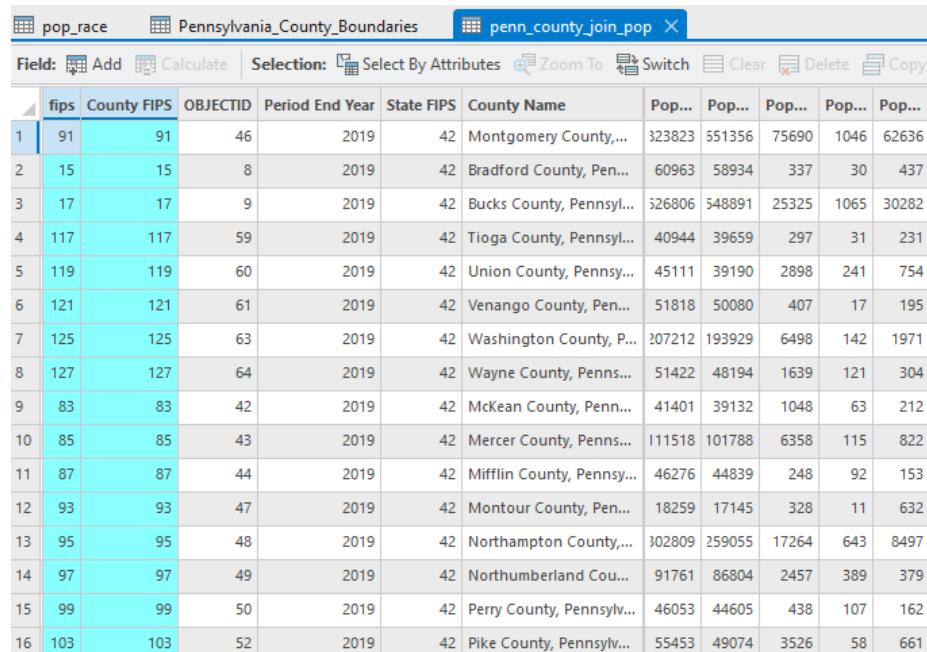
Population data (2019) is extracted from the Pennsylvania government website to be joined with Pennsylvania county boundary. Transplant centers are provided by Gift of Life Donor Program to be geo-coded as data points. The resulting file contains 67 features for 67 counties and 11 points for 11 transplantation centers in Pennsylvania. Methodologies for both software will be explored to determine differences and/or precision of the outcome.

3 ArcGIS Pro

3.1 Methodology

3.1.1 Import Files

Population data was imported using excel to table tool, followed by adding the county boundary data and joined using an unique ID (county FIPS). Projection of all layers: NAD 1983 StatePlane Pennsylvania North FIPS 3701 (Meters).

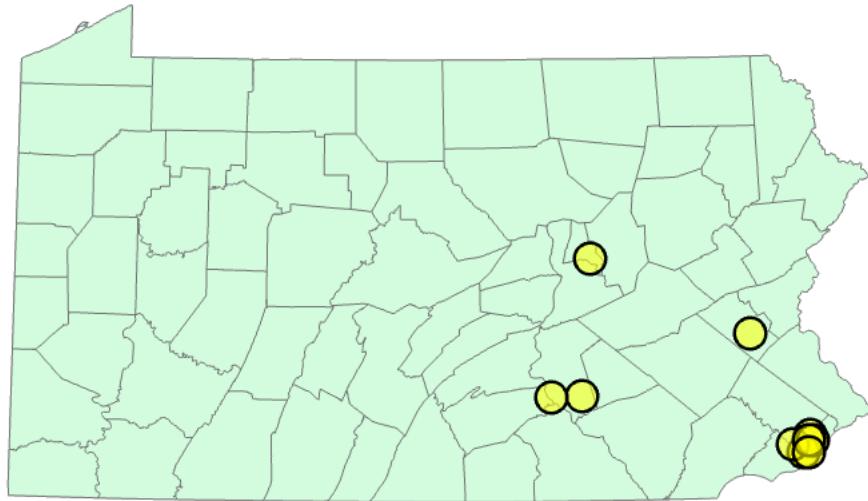


A screenshot of the ArcGIS Pro interface showing a joined layer table. The table has 16 rows and 12 columns. The columns are labeled: fips, County FIPS, OBJECTID, Period End Year, State FIPS, County Name, Pop..., Pop..., Pop..., Pop..., Pop..., and Pop... . The first column (fips) contains values like 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16. The second column (County FIPS) contains values like 91, 15, 17, 117, 119, 121, 125, 127, 83, 85, 87, 93, 95, 97, 99, 103. The third column (OBJECTID) contains values like 46, 8, 9, 59, 60, 61, 63, 64, 42, 43, 44, 47, 48, 49, 50, 52. The fourth column (Period End Year) contains values like 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019, 2019. The fifth column (State FIPS) contains values like 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42, 42. The sixth column (County Name) contains values like Montgomery County, Bradford County, Bucks County, Tioga County, Union County, Venango County, Washington County, Wayne County, McKean County, Mercer County, Mifflin County, Montour County, Northampton County, Northumberland Cou..., Perry County, Pike County, The remaining columns represent population counts for different race categories.

	fips	County FIPS	OBJECTID	Period End Year	State FIPS	County Name	Pop...	Pop...	Pop...	Pop...	Pop...
1	91	91	46	2019	42	Montgomery County, Penns...	323823	551356	75690	1046	62636
2	15	15	8	2019	42	Bradford County, Pennsyl...	60963	58934	337	30	437
3	17	17	9	2019	42	Bucks County, Pennsylv...	526806	548891	25325	1065	30282
4	117	117	59	2019	42	Tioga County, Pennsylv...	40944	39659	297	31	231
5	119	119	60	2019	42	Union County, Pennsylv...	45111	39190	2898	241	754
6	121	121	61	2019	42	Venango County, Pennsyl...	51818	50080	407	17	195
7	125	125	63	2019	42	Washington County, Penn...	207212	193929	6498	142	1971
8	127	127	64	2019	42	Wayne County, Pennsylv...	51422	48194	1639	121	304
9	83	83	42	2019	42	McKean County, Pennsylv...	41401	39132	1048	63	212
10	85	85	43	2019	42	Mercer County, Pennsylv...	111518	101788	6358	115	822
11	87	87	44	2019	42	Mifflin County, Pennsylv...	46276	44839	248	92	153
12	93	93	47	2019	42	Montour County, Pennsyl...	18259	17145	328	11	632
13	95	95	48	2019	42	Northampton County, Penn...	302809	259055	17264	643	8497
14	97	97	49	2019	42	Northumberland County, Penn...	91761	86804	2457	389	379
15	99	99	50	2019	42	Perry County, Pennsylv...	46053	44605	438	107	162
16	103	103	52	2019	42	Pike County, Pennsylv...	55453	49074	3526	58	661

Joined layer is exported as a permanent layer, each country polygon now contains values of population by race.

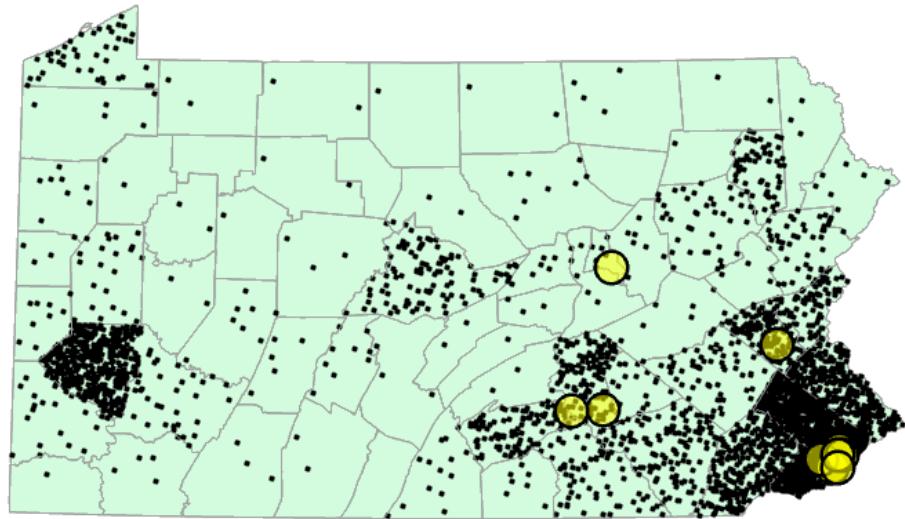
3.1.2 Geocoding



ObjectID *	Loc_name	Shape *	Status	Score	Match_type	Match_addr	LongLabel	ShortLabel
1	World	Point	M	91.24	A	100 N Academy Ave, D...	100 N Academy Ave, D...	100 N Academy Ave
2	World	Point	M	99.57	A	5501 Old York Rd, Ste...	5501 Old York Rd, Ste...	5501 Old York Rd, Ste PA
3	World	Point	M	96.96	A	S 34th St & Civic Cente...	S 34th St & Civic Cente...	S 34th St & Civic Cente...
4	World	Point	M	98	A	3400 Spruce St, Philad...	3400 Spruce St, Philad...	3400 Spruce St
5	World	Point	M	98.48	A	500 University Dr, Hers...	500 University Dr, Hers...	500 University Dr
6	World	Point	M	98.3	A	1200 S Cedar Crest Blv...	1200 S Cedar Crest Blvd	1200 S Cedar Crest Blvd
7	World	Point	M	96.97	A	100 E Lancaster Ave, W...	100 E Lancaster Ave, W...	100 E Lancaster Ave
8	World	Point	M	97.92	A	160 E Erie Ave, Philade...	160 E Erie Ave, Philade...	160 E Erie Ave
9	World	Point	M	96.77	A	111 S Front St, Harrisb...	111 S Front St, Harrisb...	111 S Front St
10	World	Point	M	96.86	A	3401 N Broad St, Phila...	3401 N Broad St, Phila...	3401 N Broad St
11	World	Point	M	96.76	A	111 S 11th St, Philadel...	111 S 11th St, Philadel...	111 S 11th St

Addresses of transplantation centers were geo-coded as a layer of data points over the previously joined county feature layer.

3.1.3 Create Dot Density



Population of Asians by county was selected first to produce a dot density map. With one dot equaling 100 people. Based on the results, the centers seem to be located on the major cluster of Asian population and isolated to one side of Pennsylvania only.

3.1.4 Generate Near Table

Every transplant center (11 points) will be analyzed against every county, to generate a table of distances between each county and the 11 point values. Resulting in 737 features (67×11). IN_FID is the unique ID for each transplant center, with NEAR_FID as the objective ID that is assigned as index to each county. The NEAR_DIST measures the distance in meters from one county to a point value.

transplant_cen_GenerateNearT2 X				
Field: <input type="button" value="Add"/> <input type="button" value="Calculate"/> Selection: <input type="button" value="Select By Attributes"/>				
OBJECTID *	IN_FID	NEAR_FID	NEAR_DIST	NEAR_RANK
1 25	1	67	81997.65979	25
2 77	2	67	90353.67801	10
3 147	3	67	98868.094456	13
4 214	4	67	98714.236505	13
5 300	5	67	118886.068316	32
6 343	6	67	27919.869176	8
7 416	7	67	93442.019132	14
8 480	8	67	94095.73949	11
9 672	9	67	124207.064440	26

3.1.5 Summary Statistics

The average distance of each county from a transplant center was calculated using summary statistics. To give a broad overview of approximately how far a patient has to travel from their county to a designated transplant center.

transplant_cen_Ge_Statistics				
Field: <input type="button" value="Add"/> <input type="button" value="Calculate"/> Selection: <input type="button" value="Select By Attributes"/>				
	OBJECTID *	NEAR_FID	FREQUENCY	MEAN_NEAR_DIST
1	1		11	30685.039703
2	2		11	167430.363148
3	3		11	43362.4376
4	4		11	193169.041468
5	5		11	132604.051028
6	6		11	350736.567006
7	7		11	360728.397241
8	8		11	132481.368734
9	9		11	271406.47803

To clarity, the tool assigned objectID to each county instead of FIPS code. The figure down below reveals the unique objectID that the county was matched with.

transplant_cen_Ge_Statistics								transplant_cen_GenerateNearT2		penn_county_join_pop_Project	
Field: <input type="button" value="Add"/> <input type="button" value="Calculate"/> Selection: <input type="button" value="Select By Attributes"/> <input type="button" value="Zoom To"/> <input type="button" value="Switch"/> <input type="button" value="Clear"/> <input type="button" value="Delete"/>											
	OBJECTID_12 *	fips	Shape	OBJECTID *	MSLINK	COUNTY_NAM	COUNTY_NUM	AREA_SQ_MI			
1	1	91	Polygon	1409	46	MONTGOMERY	46	487.42705			
2	2	15	Polygon	1410	8	BRADFORD	08	1161.337905			
3	3	17	Polygon	1411	9	BUCKS	09	622.083587			
4	4	117	Polygon	1412	58	TIOGA	58	1137.247962			
5	5	119	Polygon	1413	59	UNION	59	319.189297			
6	6	121	Polygon	1414	60	VENANGO	60	683.367616			
7	7	125	Polygon	1415	62	WASHINGTON	62	862.107693			
8	8	127	Polygon	1416	63	WAYNE	63	750.828634			
9	9	83	Polygon	1417	42	MCKEAN	42	985.269974			

Frequency indicates how many transplant centers was calculated into the average (11 for all rows).

Summary stats with MEAN_NEAR_DIST is joined with the previous population county joined layer by objectID.

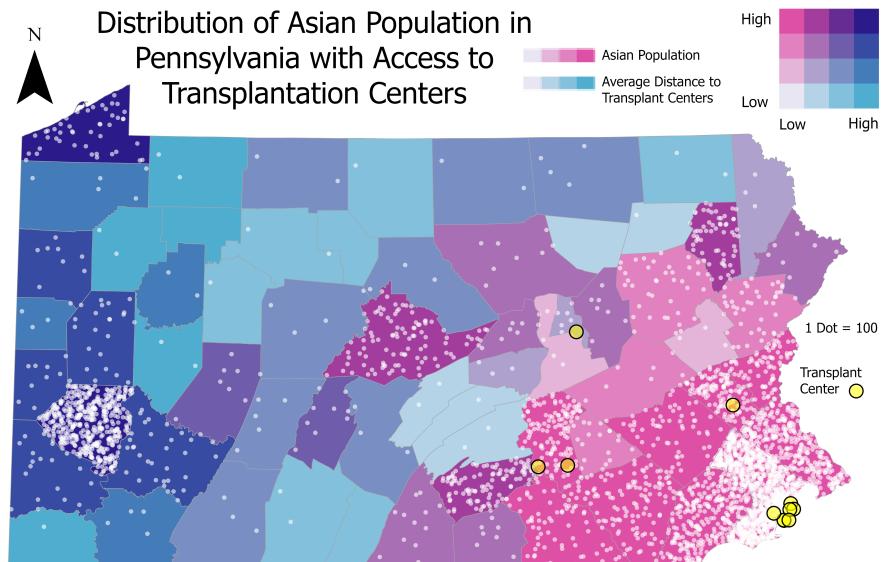
3.1.6 Bivariate Analysis

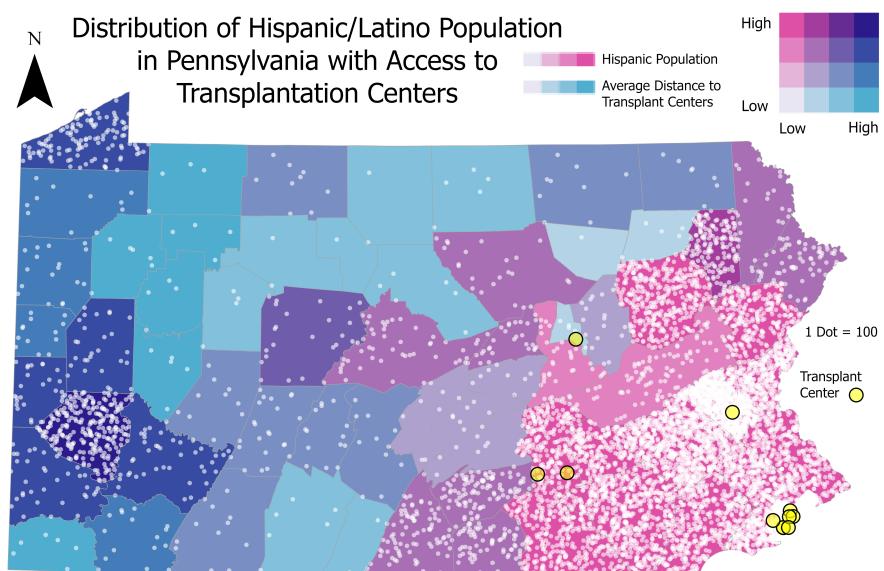
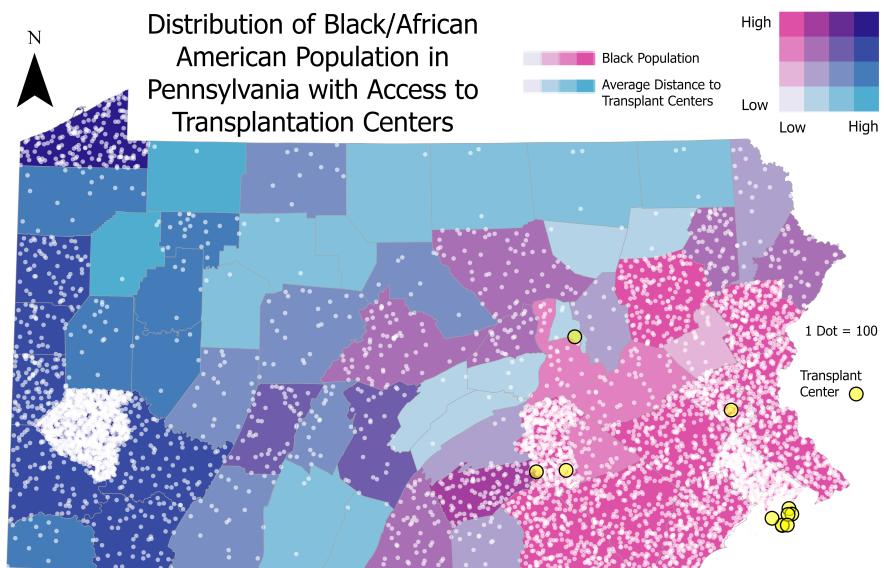
The resulting layer after a join will be used to create a layout containing graduate variables of Asian population and average distance to transplant centers.

Field:		Add	Calculate	Selection:	Select By Attributes	Zoom To	Switch	Clear	Delete	Copy
		OBJECTID	OBJECTID_12_13 *	OBJECTID	County FIPS	County Name	MEAN_NEAR_DIST	Population by Race: Asian Alone (B02001_005E)		
1	1409	1		1	91	Montgomery County, ...	30685.039703	62636		
2	1410	2		2	15	Bradford County, Pen...	167430.363148	437		
3	1411	3		3	17	Bucks County, Pennsyl...	43362.4376	30282		
4	1412	4		4	117	Tioga County, Pennsyl...	193169.041468	231		
5	1413	5		5	119	Union County, Pennsy...	132604.051028	754		
6	1414	6		6	121	Venango County, Pen...	350736.567006	195		
7	1415	7		7	125	Washington County, P...	360728.397241	1971		
8	1416	8		8	127	Wayne County, Penns...	132481.368734	304		
9	1417	9		9	83	McKean County, Penn...	271406.47803	212		

3.2 Layouts

Methodologies were repeated for Black/African American and Hispanic/Latino populations.





4 QGIS

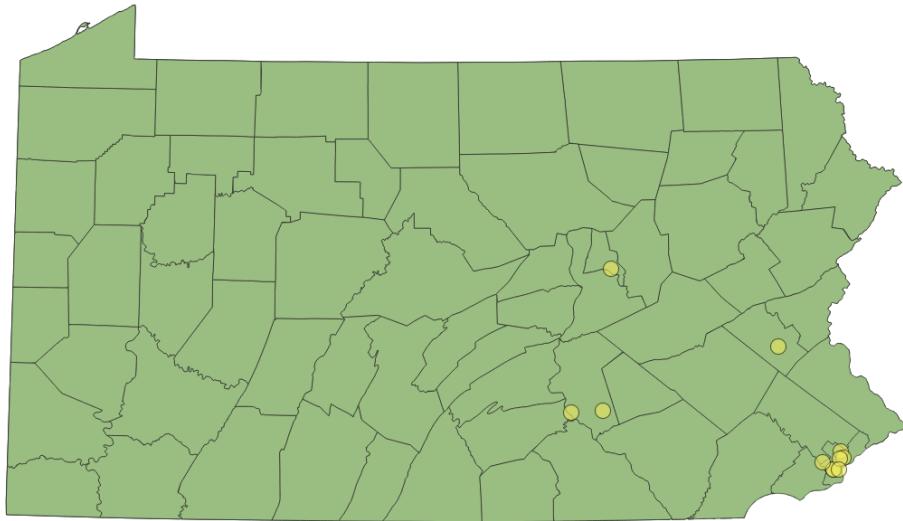
4.1 Methodology

4.1.1 Import Files

Population data was imported into QGIS as a table using the add delimited text layer option, followed by adding the county boundary shapefile and joined using an unique ID (county FIPS). Projected as the same coordinate system as ArcGIS Pro: NAD 1983 StatePlane Pennsylvania North FIPS 3701 (Meters).

Pennsylvania_County_Boundaries — Features Total: 67, Filtered: 67, Selected: 0									
OBJECTID	MSLINK	COUNTY_NAM	COUNTY_NUM	AREA_SQ_Mi	fips	FIPS1	n_By_Race_Perio	cation_By_Race_Stat	on_By_Race_Cou
1409	46	MONTGOMERY	46	487.427050...	91.0000000...	91	2019	42	Montgomery ...
1410	8	BRADFORD	08	1161.337905...	15.0000000...	15	2019	42	Bradford Cou...
1411	9	BUCKS	09	622.083586...	17.0000000...	17	2019	42	Bucks County...
1412	58	TIOGA	58	1137.247961...	117.0000000...	117	2019	42	Tioga County...
1413	59	UNION	59	319.1892970...	119.0000000...	119	2019	42	Union County...
1414	60	VENANGO	60	683.367615...	121.0000000...	121	2019	42	Venango Cou...
1415	62	WASHINGTON	62	862.1076930...	125.0000000...	125	2019	42	Washington ...
1416	63	WAYNE	63	750.828633...	127.0000000...	127	2019	42	Wayne Count...
1417	42	MCKEAN	42	985.269974...	83.0000000...	83	2019	42	McKean Cou...
1418	43	MERCER	43	682.3598120...	85.0000000...	85	2019	42	Mercer Count...
1419	44	MIFFLIN	44	413.7323319...	87.0000000...	87	2019	42	Mifflin County...
1420	47	MONTOUR	47	132.2815229...	93.0000000...	93	2019	42	Montour Cou...
1421	48	NORTHAMPT...	48	377.5675390...	95.0000000...	95	2019	42	Northampton...
1422	49	NORTHUMBE...	49	475.312944...	97.0000000...	97	2019	42	Northumberl...
1423	50	PERRY	50	557.920839...	99.0000000...	99	2019	42	Perry County...
1424	51	PIKE	51	567.9975359...	103.0000000...	103	2019	42	Pike County...

4.1.2 Geocoding



transplant_centers — Features Total: 11, Filtered: 11, Selected: 0

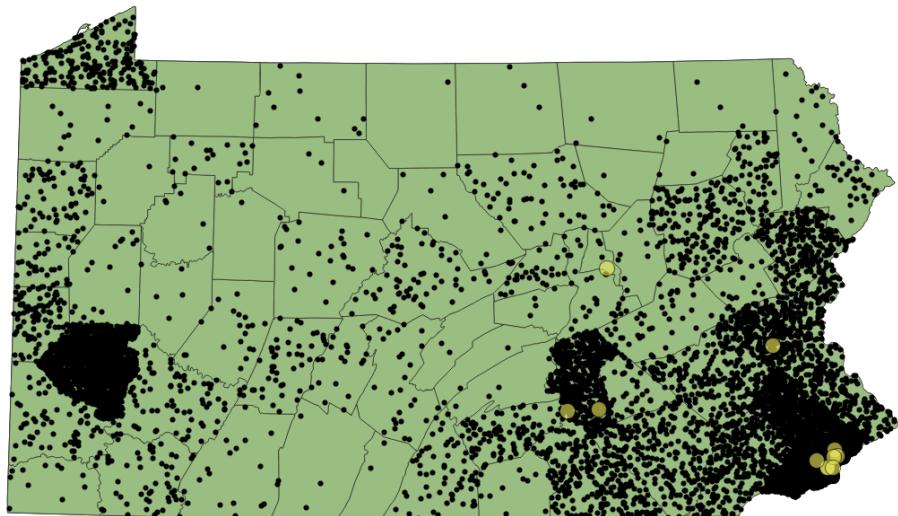
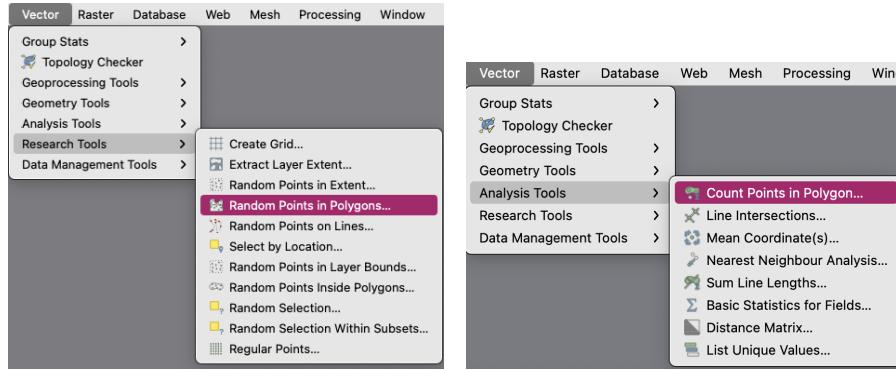
	name	address	city	state	zip	country	field_7	lat	long
1	Children's Ho...	34th Street a...	Philadelphia	PA	19104	United States	NULL	39.9487	-75.1939
2	Einstein Medi...	5501 Old Yor...	Philadelphia	PA	19141	United States	NULL	40.0366	-75.1424
3	Geisinger Me...	100 N.Acade...	Danville	PA	17822	United States	NULL	40.968	-76.606
4	Hospital of th...	3400 Spruce...	Philadelphia	PA	19104	United States	NULL	39.9501	-75.1939
5	Lankenau Ho...	100 East Lan...	Wynnewood	PA	19096	United States	NULL	39.9884	-75.2622
6	Lehigh Valley...	1200 South C...	Allentown	PA		United States	NULL	40.56633	-75.52403
7	Penn State M...	500 Universit...	Hershey	PA	17033	United States	NULL	40.264	-76.6766
8	St.Christoph...	160 East Erie ...	Philadelphia	PA		United States	NULL	40.0063	-75.1257
9	Temple Univ...	3401 North B...	Philadelphia	PA	19140	United States	NULL	40.0049	-75.1511
10	Thomas Jeff...	111 South 11t...	Philadelphia	PA	19107	United States	NULL	39.9492	-75.1575
11	UPMC Harris...	111 South Fro...	Harrisburg	PA	17101	United States	NULL	40.2579	-76.8801

Used longitude and latitude coordinates instead of geocoding addresses.

4.1.3 Create Dot Density

Since there wasn't a dot density option in QGIS because it was joined to a polygon layer, a vector tool was used to create random points in polygons based on the population (Black/African American) field. one dot = 100 people.

Another vector analysis tool was used to count up points within a polygon to get a field of total points.



4.1.4 Distance Matrix

The distance matrix from QGIS is similar to the generate near table in pro. Instead of calculating distance with polygon and points, the matrix calculates the distance all dot density points and transplant center points.

The screenshot shows the QGIS application interface. The top menu bar includes Vector, Raster, Database, Web, Mesh, Processing, and Window. A context menu is open under the Vector menu, listing Group Stats, Topology Checker, Geoprocessing Tools, Geometry Tools, Analysis Tools (which is currently selected), Research Tools, and Data Management Tools. Below this, a submenu for Analysis Tools is displayed, containing Count Points in Polygon..., Line Intersections..., Mean Coordinate(s)..., Nearest Neighbour Analysis..., Sum Line Lengths..., Basic Statistics for Fields..., Distance Matrix..., and List Unique Values... The Distance Matrix... option is highlighted with a pink rectangle.

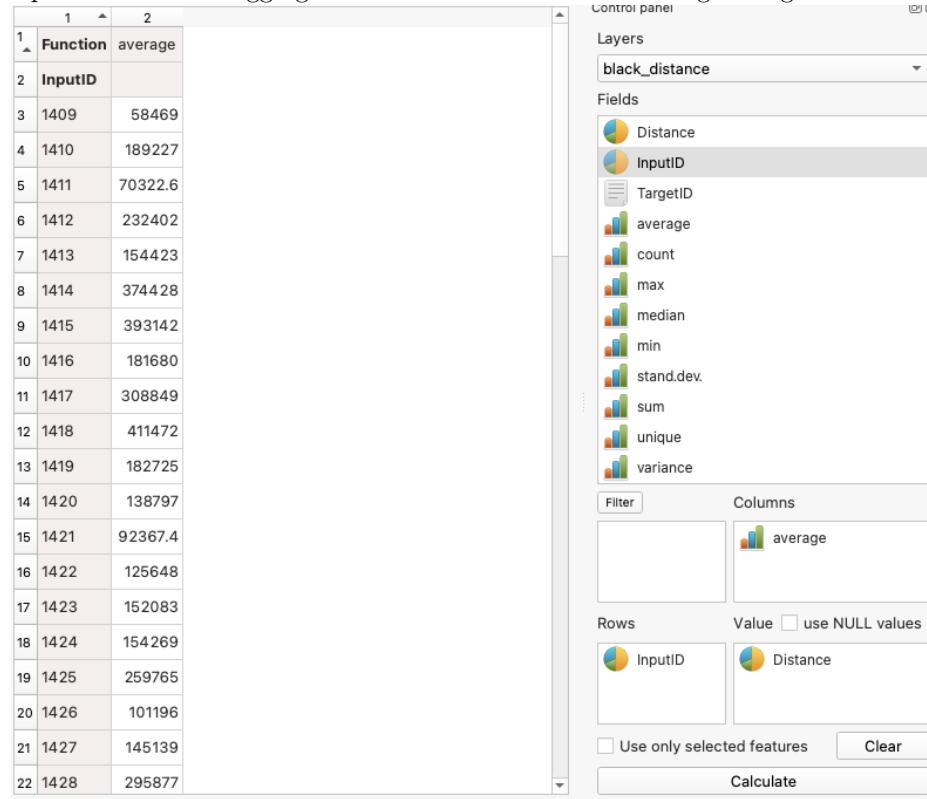
InputID	TargetID	Distance
1409	Children's Ho...	51861.52120...
1409	Einstein Medi...	46296.4492...
1409	Geisinger Me...	116161.3640...
1409	Hospital of th...	51725.92837...
1409	Lankenau Ho...	45259.4458...
1409	Lehigh Valley...	23524.53916...
1409	Penn State M...	101209.2390...
1409	St.Christoph...	49803.2485...
1409	Temple Univ...	48602.0702...
1409	Thomas Jeff...	53402.10974...
1409	UPMC Harris...	118497.8631...
1409	Children's Ho...	31160.71575...
1409	Einstein Medi...	26514.36759...
1409	Geisinger Me...	135000.8511...
1409	Hospital of th...	31027.24972...
1409	Lankenau Ho...	24542.3347...

Input ID is each dot density in a county polygon, the target ID is the transplant center point value. The distance is the length between each point in meters.

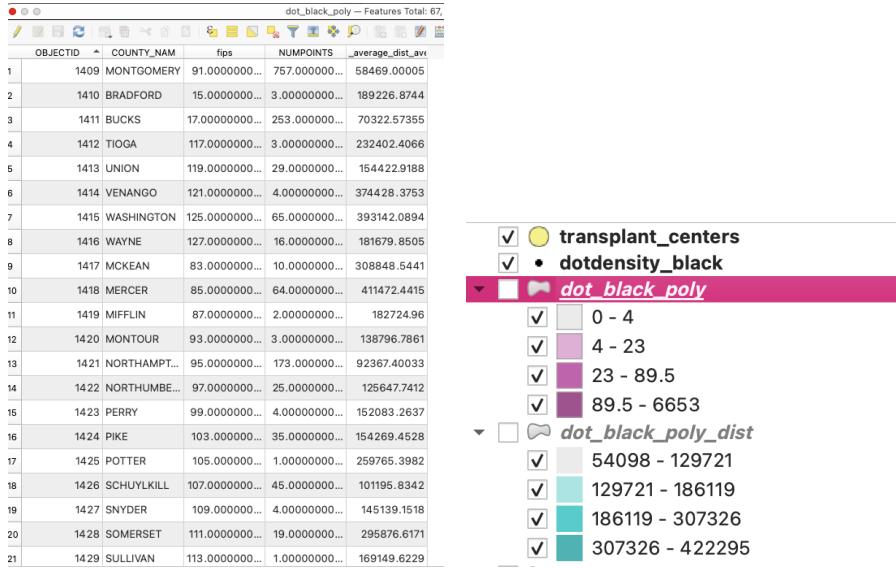
4.1.5 Summary Statistics

Average distance was produced using GroupStats plugin. Distance between all the dot density points and transplant centers were added and divided by total

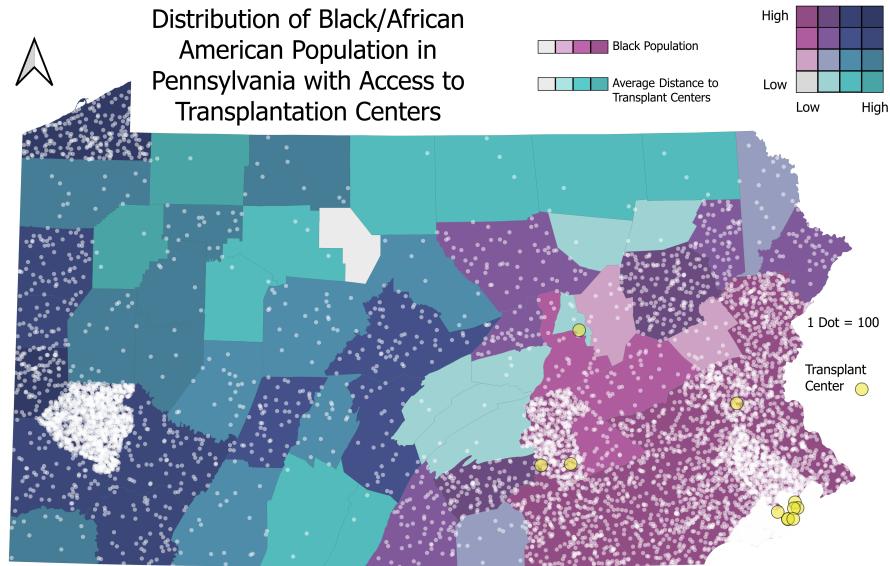
dot density points in that county. ObjectID is the unique ID for each county. GroupStats allows for aggregation of common ID for measuring averages.

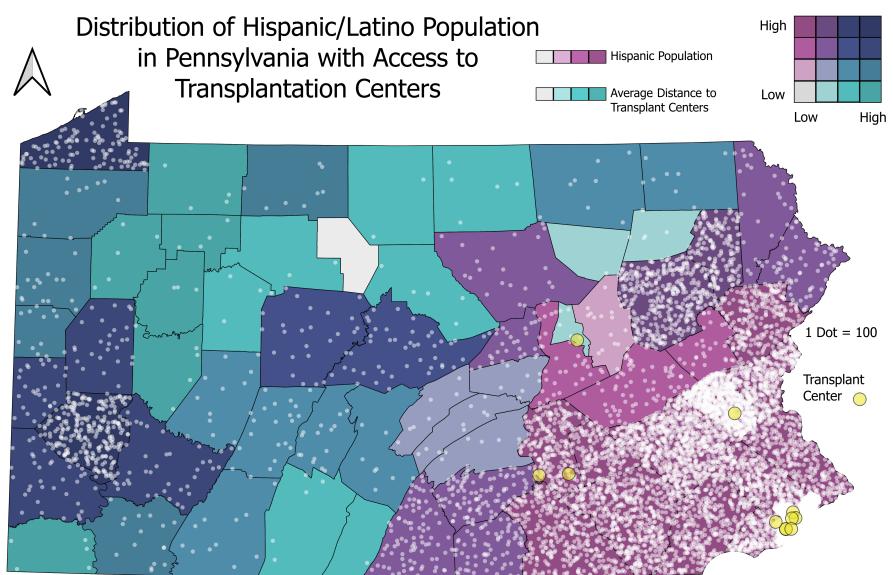
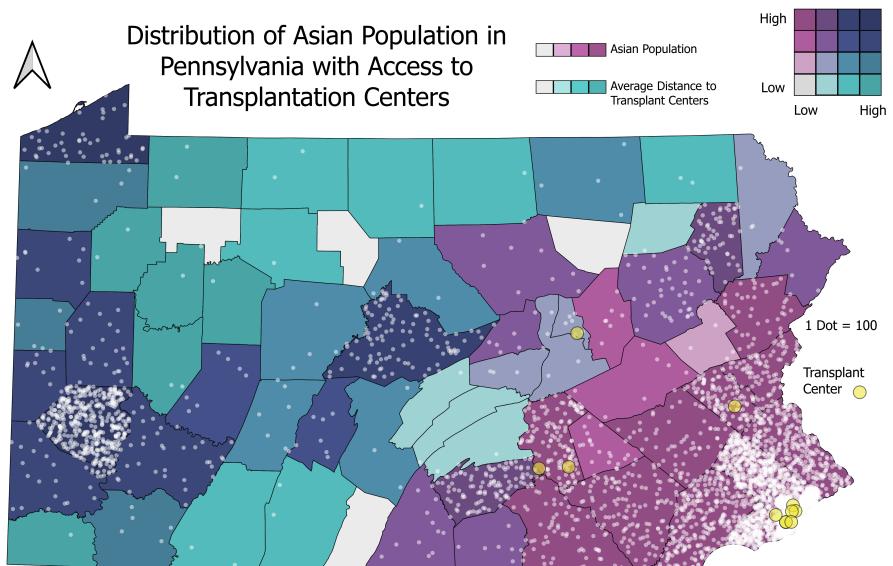


The average values are exported as CSV and re-imported to be joined with the dot density layer.



4.1.6 Bivariate Analysis





QGIS does not have a symbology setting for bi-variables. Layer had to be duplicated to make two kinds of symbology and blend graduate colors together for a bivariate effect.

Methodologies were repeated for Asian and Hispanic/Latino populations.

5 Conclusion

The layouts for Pro and QGIS were similar, QGIS has null values (white polygon) because of the distance matrix. Some populations were not accounted for since it was insignificant. There was about 37 percent difference in distance measured, QGIS had longer distance most likely due to all the dot density points calculated against transplant center dot values.

Although the transplant centers were within clusters of minority populations, it fails to cover two third's of Pennsylvania's total surface area. There are some densely populated areas with minorities that are on the west side that has very limited access to these centers. For organ donating criteria, some of those areas can not be reached if hospitals prioritize location proximity rather than measurement of sickness.

6 Reference

<https://www.donors1.org>

Williams DR, Priest N, Anderson NB. Understanding associations among race, socioeconomic status, and health: Patterns and prospects. *Health Psychol.* 2016 Apr;35(4):407-11. doi: 10.1037/he0000242. PMID: 27018733; PMCID: PMC4817358.