“Segmenting Major Donors”

Charlie Gardner

Regis University, MSDS 696

YouTube Presentation: <https://youtu.be/sVtfwk_qzVs>

**Abstract:**

My final project utilizes internal data from the Winston-Salem Foundation’s Development and Donor Services on gifts made to the Foundation over the past 30 years. This analysis identifies several potential barriers, areas for further exploration, and opportunities for intervention.

**Data background, sources, and transformation:**

Each year the Winston-Salem Foundation awards scholarships to students enrolling in post-secondary education. In 2021, 865 students applied for scholarships and the foundation awarded $1.49 million to 398 applicants. The staff that manages this student aid program expressed a desire to evaluate its reach and impact with a racial equity lens.

Previously in the class, I explored different racial equity implications of between different types of scholarship funds and about which students receive awards based upon race, GPA, and financial need. Overall, I concluded that the internal process of awarding scholarships does align with the foundation’s equity values, but there are still disparities in that arise from the racial demographics of applicants. For this final project, I will look specifically at potential barriers that depress the number of BIPOC students applying and will limit my analysis to the 629 applicants from Forsyth County. The dataset includes applicants’ home address, race, and high school, as well as how they heard about the scholarships and if they are a first-generation college student.

To supplement this WSF internal data, I used a variety of resources from the US Census Bureau, including Tiger/Line shapefiles, geocoding services, and data tables. I identified the GEOIDs for the 93 census tracts in Forsyth County and used GeoPandas to join the shapefiles with demographic data from the 2019 American Community Survey 5-Year estimates (the most recent that had census tract data for Forsyth County). I also used the Census’s geocoding services to batch process applicant’s home addresses and high school addresses to coordinate points. I also used Excel’s tables and pivot tables to tabulate, sort, and filter various rows to simplify generating visualizations both in Excel and Matplotlib.

**Research question:**

When focusing on WSF applicant the data from Forsyth County, Latinx students are significantly underrepresented in terms of applying for scholarships. This is illustrated by using the local school district’s racial demographics as a benchmark. Most demographics apply for scholarships at a rate similar to their proportion in the school district, however Latinx students apply for scholarships at half the rate. Latinx students comprise 28.4% of the school district, but only represent 14.1% of the WSF scholarship applicants.

Chart, bar chart

Description automatically generated

This significant gap directly translates to a disparity in the number of Latinx students receiving scholarships. If WSF wants to improve its racial equity impact, it needs to address the question: What barriers do Latinx students face when apply for WSF scholarships?

**Identifying patterns in Forsyth County Applicants:**

To begin identifying what impedes Latinx students in applying for scholarships, we can look at strong patterns among the students that do apply, especially in where they live and how they heard about the scholarship program.

The following visualization maps out Forsyth County and overlays it with a heat map of where applicants live and a bubble plot of where they attend high school. To create this plot, I used the US Census to pull shapefiles of the tracts in Forsyth County and their geocoding services to find the coordinates of the applicants’ home and school addresses. I also used Excel pivot tables to tabulate the number of applicants per each high school.

Map

Description automatically generated

As evident from the plot, most local applicants live clustered in the Western side of the county, and the high schools that have the most applicants are also in the Western side. Six out of the top 7 high schools with the most applicants are on the West side. Inversely all high schools with single digit applicants (8 in total) are in the Eastern side. The exception to this is Atkins High school which is tied for the second greatest number of applicants 74.

The following bar chart visualizing where applicants heard about the scholarship also point to a strong clustering of applicants. The two primary ways that applicants hear about WSF scholarships are from high school guidance counselors and from friends and family. These two far surpass all other means, yet in both categories, Latinx applicants are underrepresented. They comprise only 16% of those who indicated they learned about the scholarship program through a guidance counselor, and only 10% from friends and family. On the other hand, White students are most likely to hear about the scholarships through guidance counselors or from a friend/family member.

Chart

Description automatically generated

**Mapping applicants by referring source:**

In search for additional patters and clusters, I created a pair of maps that sorted the applicants by how they heard about the scholarship opportunities and recreated the weighted bubble plot per high school. To add another dimension to the map, I overlayed this scatterplot over a choropleth chart that illustrates which neighborhoods have the largest percent Latinx residents. In Forsyth County, the majority of the Latinx community lives clustered in the Southeast side, with multiple neighborhoods being having Spanish speakers make up between 40% to 50% of their residents.

Diagram, schematic

Description automatically generated

By plotting the high schools over the choropleth map, it is evident that the guidance counselors in schools that are in/near neighborhoods with high percentage of Latinx students refer the fewest number of applicants to WSF scholarships. Again, most of the schools with the highest referring guidance counselors are on the West side, where there are fewer Spanish speakers. At this point I could have recreated this choropleth map highlighting other Census data on race, income, education level, home-ownership, poverty, and there would have been a similar pattern because due to historic and discriminatory policies, the West side is wealthier and Whiter than the East side. So this map highlights disparities in the community at large, and that often wealthier schools are better resourced. However, the anomaly is that the high school counselor with the most referrals is at Atkins High, which is on the East side and whose students are predominantly BIPOC.

Diagram

Description automatically generated

This next choropleth similarly plots applicants per high school, but this one is parsed out as those who heard about the scholarships from friends and family. Additionally, I reinserted the heat map with home addresses. This indicates that most of the family that shared about the scholarship with the students live in neighborhoods with few Spanish speaking residents. In fact, several high schools in/near neighborhoods with large Latinx populations did have any applicants that heard about WSF scholarships through friends and family. Even Atkins high school had relatively few students who heard about the program through friends and family. On the other hand, Reagan High school had 42 students apply who heard about the program through friends/family, which is almost double the next. Raegan is also the wealthiest high school in the community and not near a neighborhood with a large Latinx population.

**Conclusion:**

These maps illustrate strong correlations between where you live/what high school you attend and applying for scholarships, especially if you further parse out how applicants heard about the program. While high school guidance counselors are important for encouraging students to apply, it is also clear that the importance of friends and family should not be overlooked.

This insight is reinforced by analyzing the applicants who will be first-generation college students. Of the 629 applicants from Forsyth County, 159 are first-generation (total of 25%). However, Latinx are overrepresented in this classification, with 61% of applicants being first generation college students. This is likely a significant reason that fewer Latinx students apply for scholarships. In contrast, White applicants are much more likely to have family members who have been to college before (86%).

Chart, bar chart

Description automatically generated

Overall White students are the least likely to be first generation, most likely to hear about the opportunity through guidance counselors or through a friend/family member, and most likely to apply for scholarships. The inverse is true for Latinx students. These patterns indicate that social networks are very important for encouraging students to apply for scholarships. Barriers to application include where one lives, which school someone attends, and which language is spoken at home.

While correlation doesn’t indicate causation, these patterns do point towards some potential hypothesis that could be tested in order to increase Latinx scholarship applications. One intervention would be to target outreach to guidance counselors in at high schools in/near Lantix neighborhoods. Another intervention would be to create marketing material in Spanish and engage Latinx cultural organizations to engage the families of the students and Latinx community overall.

**References:**

US Census Bureau. “Demographic Data Table” 2019 *ACS American Community Survey 5-YR Estimates*. American Community Survey Office. <https://www.census.gov/data/developers/data-sets/acs-5year.html>.

US Census Bureau, Geography Division. “Geocoding Services”. <https://geocoding.geo.census.gov/>

US Census Bureau, Geography Division, TIGER/Line Shapefiles, “ACS\_2019\_5YR\_TRACT\_37.gdb” <https://www.census.gov/cgi-bin/geo/shapefiles/index.php>

Winston-Salem Foundation. “2021 Student Aid Applications Data Set.”