Hierarchical Clustering Summary Report- Event Attendance

Project Overview

Objective:Hierarchical clustering is a method used to group event attendees based on important traits. This helps us find common characteristics among participants, allowing us to customize our messages and outreach efforts to boost attendance.

Data Description

Dataset: The dataset includes anyone who registered and attended an event.. The events included campus tours, accepted student days and open houses from 2021-2024.

- **Source**
- **Variables Used**: 7-
    1. Applicant
    2. Inquiry
    3. DO to Biomed Target Marketing
    4. First Generation
    5. Income Economically Disadvantaged
    6. Health Professional Shortage Area
    7. URIM
- **Number of Observations(event attendees):** 235
- **Data Preprocessing**: All data was transformed into binary data(yes, no). Some attendees were filtered out that did not have certain data points. This was caused by these fields not being required on event registrations and/or inquiry.

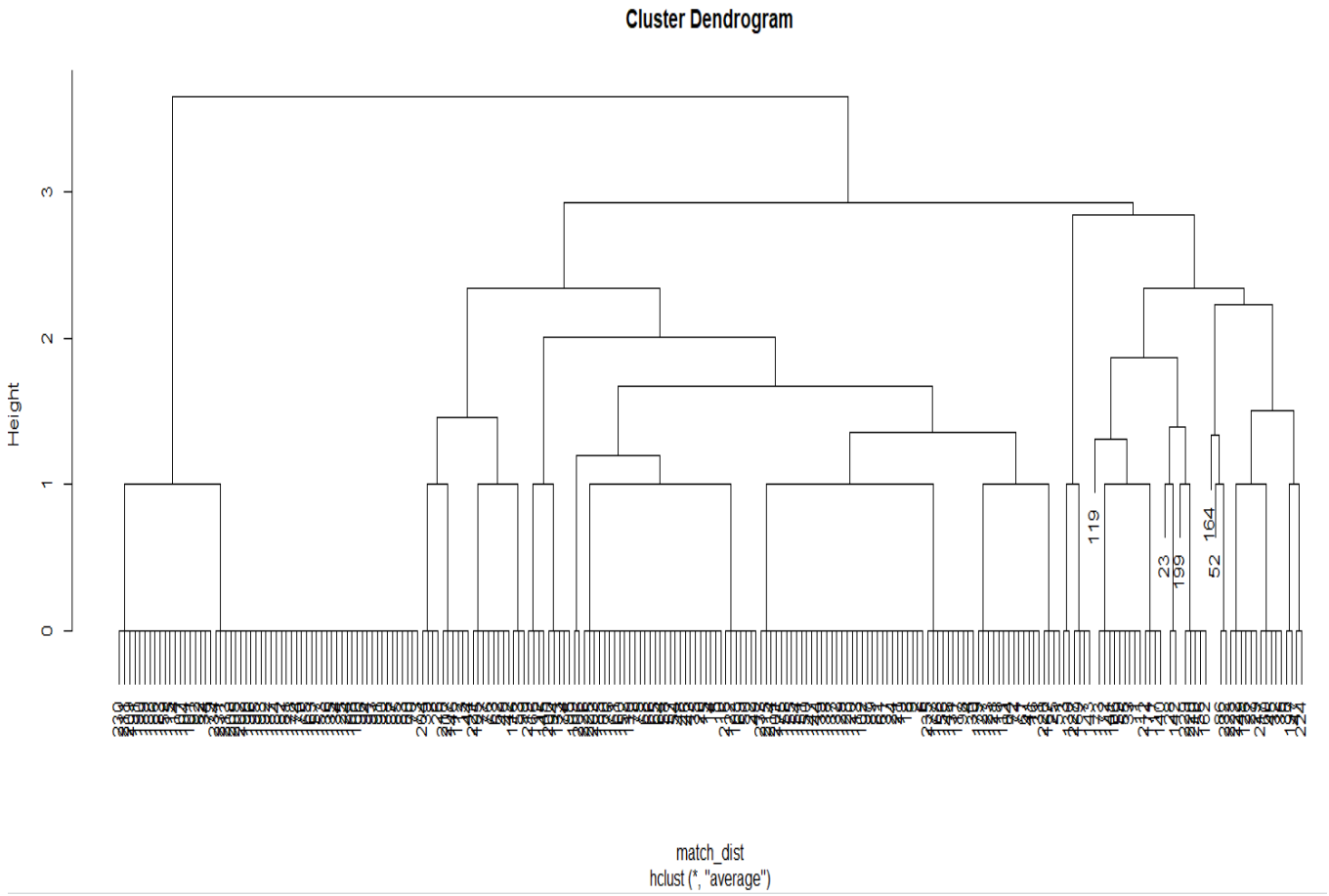
Methodology

Clustering Approach:
I performed agglomerative hierarchical clustering in Rstudio (This is a bottom-up approach where each data point starts as its own cluster, and pairs of clusters are merged as you move up the hierarchy.) using Manhattan distance(absolute differences between data points) and group average linkage(average similarity among all of the possible pairs of observation between 2 clusters and dividing by all of the possible pairs.). The dendrogram was generated to visualize the clustering structure.

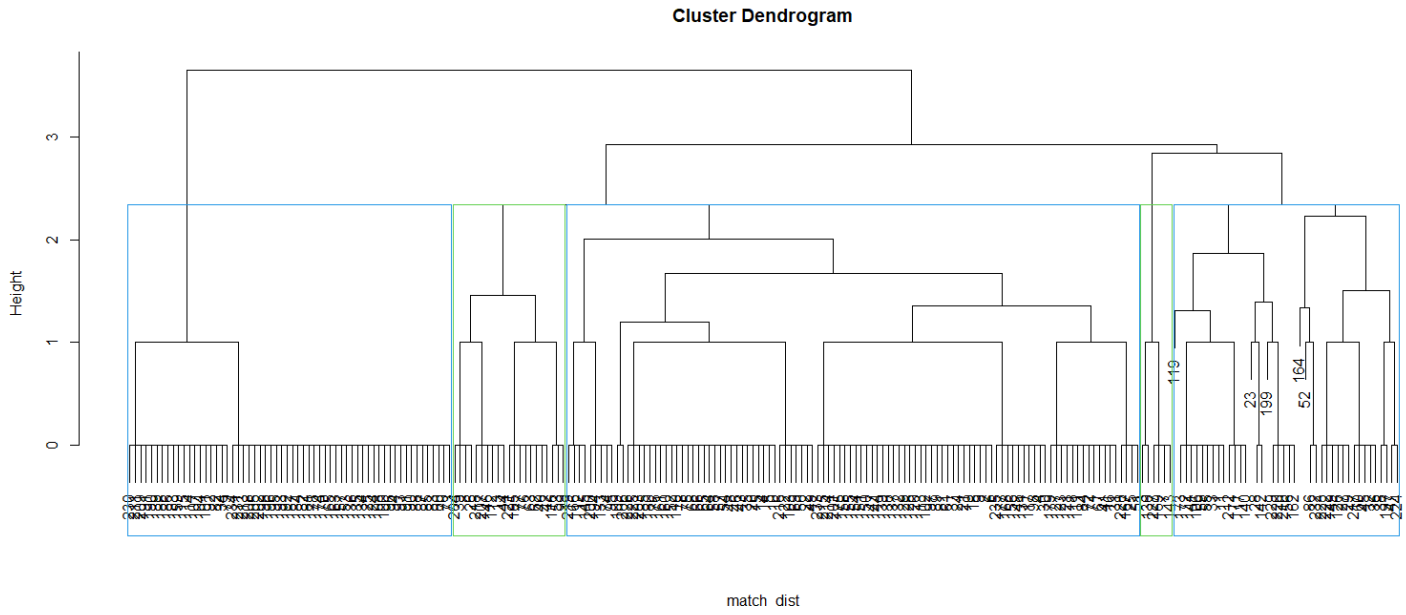
Results

Dendrogram

The dendrogram illustrates the hierarchical relationships between data points collected on application and/or inquiry form, event registration. Clusters are formed at different height levels, indicating the similarity between responses.



Cluster Analysis- Out of all observations of event attendees 175 were applicants, 60 were inquiries, 48 were part of the DO To Biomed targeted marketing, 52 come from a disadvantaged income, 64 live in a medically underserved area, 96 are URIMs, 42 are first generation students.



Summary of Clusters for Event Attendees(5 clusters are outlined above)

Cluster 1: Economically Disadvantaged First-Generation Applicants

- **Observations**: 42
- **Targeted Marketing**: 7 participants (17% of the cluster) were part of the DO to Biomed targeted marketing initiative.
- **Economic Background**: 28% come from economically disadvantaged households.
- **First-Generation Status**: 38% are first-generation college students.
- **Underrepresented Minorities**: 13% identify as Underrepresented in Medicine (URIM).
- **Geographic Disadvantage**: 61% reside in medically underserved areas.

Cluster 2: Diverse Applicants with Moderate Economic Status

- **Observations**: 106
- **Targeted Marketing**: 23 participants (22% of the cluster) were part of the DO to Biomed targeted marketing initiative.
- **Economic Background**: 10% come from economically disadvantaged households.
- **First-Generation Status**: 0% are first-generation college students.
- **Underrepresented Minorities**: 34% identify as URIM.
- **Geographic Disadvantage**: 16% reside in medically underserved areas.

Cluster 3: High Engagement in Targeted Marketing

- **Observations**: 6
- **Targeted Marketing**: 100% of participants were part of the DO to Biomed targeted marketing initiative.
- **Economic Background**: 100% come from economically disadvantaged households.
- **First-Generation Status**: 66% are first-generation college students.
- **Underrepresented Minorities**: 100% identify as URIM.
- **Geographic Disadvantage**: 0% reside in medically underserved areas.

Cluster 4: Underserved Applicants

- **Observations**: 21
- **Targeted Marketing**: 10 participants (47% of the cluster) were part of the DO to Biomed targeted marketing initiative.
- **Economic Background**: 33% come from economically disadvantaged households.
- **First-Generation Status**: 0% are first-generation college students.
- **Underrepresented Minorities**: 0% identify as URIM.
- **Geographic Disadvantage**: 100% reside in medically underserved areas.

Cluster 5: General Inquiries with URIM Representation

- **Observations:** 60
- **Targeted Marketing:** NA
- **Economic Background**: NA
- **First-Generation Status:** NA
- **Underrepresented Minorities**: 31% identify as URIM.
- **Geographic Disadvantage**: NA

Comparative Insights

- **Targeted Marketing Participation:** Cluster 3 shows the highest engagement with targeted marketing initiatives, with 100% participation, while Cluster 5 lacks specified information on participation.
- **Economic Disadvantage:** Cluster 3 is fully composed of economically disadvantaged individuals, while Cluster 1 also has a significant proportion (28%). Cluster 2 has the lowest representation of economically disadvantaged individuals (10%).
- **First-Generation Status:** Clusters 1 and 3 include significant portions of first-generation students (38% and 66%, respectively), while Clusters 2, 4, and 5 do not provide information on first-generation students.
- **URIM Representation:** Cluster 3 stands out with 100% URIM representation. Cluster 5 also has notable URIM representation at 31%, while Cluster 4 shows no representation in this category.
- **Medically Underserved Areas:** Cluster 4 is unique in that all its members live in medically underserved areas, while Cluster 3 has no members from such areas.

To get more people to attend our events, we can focus on the right candidates by looking at the traits of different groups we've identified. Here are some important insights from the analysis that can help us plan our outreach strategy:

Key Insights for Targeting Candidates

1. Leverage Targeted Marketing Programs

- **Focus on Engagement**: Clusters with higher participation in the DO to Biomed targeted marketing initiative (like Clusters 1, 3, and 4) indicate that this targeted marketing strategy was effective. Increase outreach efforts in these clusters, emphasizing the benefits of participation and engagement opportunities available through the events. At information sessions (in person or virtual) we could mention the retargeting and reiterate to be on the lookout for more information or send out a survey earlier if they wish to participate. If possible, we would need an earlier outreach due to the academic calendar change.

2. Address Economic Disadvantages

- **Financial Incentives**: Given that a significant portion of candidates in Clusters 1 and 3 come from economically disadvantaged backgrounds, we may consider providing financial incentives . Maybe we can highlight entering to win a scholarship(donor funds), application waivers(Couple with Fin Aid Presentation), coupon codes, drawings for donated gas cards from local vendors or gift certificates from restaurants. Maybe use

donor funds to cover travel expenses to visit campus. Highlight these opportunities in marketing materials.
- **Supportive Resources**: Offer resources that help economically disadvantaged students navigate event participation, such as information on travel arrangements and accommodation options. Showcase resources on our campus, resource workshop for applicants and/or accepts. Show the available resources and how our students use them. Interview tips , use pre med brochure to tailor the event(create a slide deck)

3. Focus on First-Generation College Students

- **Tailored Messaging**: Clusters with first-generation college students (especially Clusters 1 and 3) may require tailored messaging that addresses their unique challenges and needs. Develop marketing materials that highlight success stories of past first-generation ambassadors that attended an event and how the event can help them achieve their academic and professional goals.
- **Mentorship Opportunities**: Implement mentorship programs during events that pair first-generation students with other current students with similar backgrounds. This may be a stand alone event or part of an open house/campus tour/accepted student day. Promote these opportunities as part of the event agenda to encourage attendance.

4. Engage Underrepresented Minorities (URIMs)

- **Diversity-Focused Outreach**: Clusters 3 and 5 show a significant presence of URIMs. Tailor outreach efforts to emphasize diversity and inclusivity in the event, showcasing diverse speakers, panels, and networking opportunities that resonate with URIM attendees. We may think of getting a local doctor to do a virtual session with us talking about their journey to medicine.
- **Community Partnerships**: Partner with and strengthen our relationships with organizations that support URIMs to facilitate communication and increase event visibility. Tailored messaging to org contacts in Slate that is ongoing.

5. Utilize Geographic Targeting

- **Local Outreach**: For clusters residing in medically underserved areas (notably Cluster 4), focus on local outreach efforts. Consider hosting events in locations that are easily accessible to these candidates(host events on college campuses, like ABAC, VSU and ASU), or offer more virtual events(info sessions, meet the students etc) options to broaden participation without geographical constraints.
- **Highlight Community Impact**: Emphasize how the event will address issues pertinent to those living in medically underserved areas, such as healthcare accessibility, career opportunities in these fields, and resources available to them.

Conclusion

By leveraging these insights, our outreach efforts can be tailored to effectively engage specific segments of the applicant/prospect/inquiry pool, ultimately increasing event attendance. A complex approach that considers economic background, first-generation status, URIM representation, and geographic location will help ensure that our marketing and recruitment strategies resonate with the diverse characteristics of the candidates in our pools. Implementing these strategies may enhance attendance but also help foster a more inclusive and supportive environment at our events.