# Diverse thinking in biological anthropology and data science: methods and mentoring

Amy Shapiro, Nana Technologies

## Introduction

Diversity in background and experience is often a core goal when making admissions and hiring decisions, since it is well documented that diversity promotes innovation and creativity. However, it can be hard to switch fields or know that experience in one field may carry over into supporting experience in another field.

In this poster, I summarize what I learned being mentored by Kaye that allowed me to switch fields and what I use myself when mentoring. I also present research I conducted as a data scientist at Yelp that relies on my training in biological methods to demonstrate one such example.

## How to be a mentor

### Be visible

- Watching how Kaye navigated teaching, leading, and interacting with colleagues taught me so much. Simply modeling behavior we want to promote demonstrates inclusivity, particularly for women, minorities, and those on non-traditional tracks.

### Be available: offer your time early and often

- While many academics focus solely on their own work, Kaye always made time to meet with students and was always responsive. As someone who was not taught how to ask for help, having that support was crucial.

### Focus on experiential learning

- Kaye championed including students in field and museum research. “On the job” learning through experience, even through failure, was how I learned best and something I encourage in those I mentor.

### Encourage differences, even in the face of rejection

- I faced a lot of funding challenges in part due to methods that were not standard. Kaye supported me despite this and helped me balance new methods with validation.

### Ensure access for all

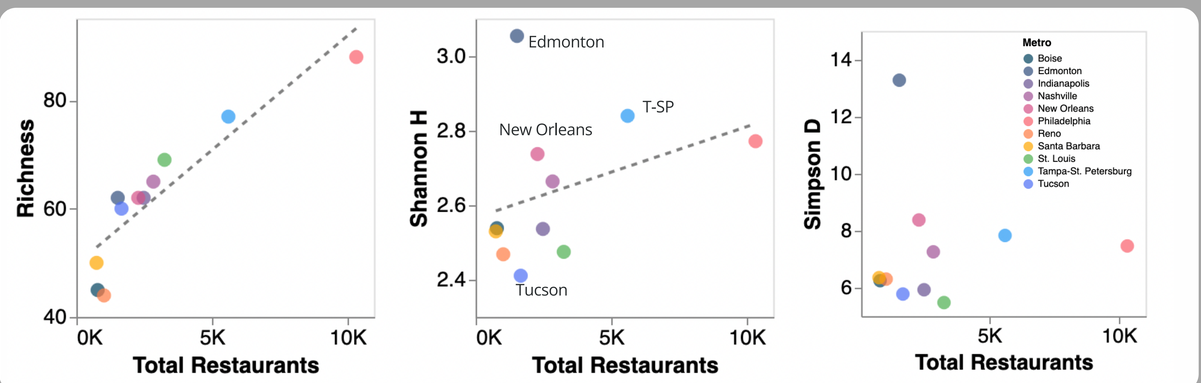
- Part of inclusion is ensuring everyone can access research and tools needed. I try to use open source tools for research and distribution of results.

## Research Example: Metro Cuisine Diversity

I explored the cuisine diversity of 11 North American metros using the publicly accessible Yelp dataset. The analysis resembles an ecological diversity analysis, where I treat the metros as sites, the cuisines as species, and the restaurants as individuals sampled.

### Methods

- Sampled over 18k restaurants from 11 North American metros
- Classified each restaurant as a single cuisine
- Calculated richness (the number of unique cuisines) and diversity indices (which incorporate richness and evenness, how evenly the restaurants are distributed across the cuisine types) for each metro
- Compared the number of restaurants in each metro to the indices




## Results

- Richness correlates strongly with the number of restaurants in each metro.
- Most metros are dominated by the most common cuisine types. The top 15 cuisines in a metro account for 85-90% of all restaurants.
- Metros that are dominated by a few cuisine types, such as Tucson (American and Mexican make up 55%), have the lowest diversity, while metros with lower dominance, such as Edmonton (Canadian and Italian make up 27%) or regional cuisines, such as New Orleans (Cajun/Creole) and Tampa-St. Petersburg (Cuban and Greek), have the highest diversity.

See the full analysis + code on my github page

Interested in Data Science or have a data or analytical problem? Reach out to me at **amesshaps@gmail.com**

Thanks to Amy Rector and Irene Smail for organizing this symposium, and most of all thanks to Kaye for all she has taiught me