Summer 8-18-2021

Client Access Feature Engineering for the Homeless Community of the City of Portland

Oswaldo Ceballos Jr
University of Texas at Austin

Client Access Feature Engineering for the Homeless Community: Portland

By: Oswaldo Ceballos

CENTRAL CITY CONCERN

HOMES HEALTH JOBS

About/Background

Organization: Central City Concern (CCC)

Services include:

- Housing
- Recovery
- Health Care
- Jobs

Rundown

Goal: Feature Engineering and Exploratory Analysis on existing CCC client data

Existing Methods of Approach: Methodology is not uncommon; outcomes vary

Motivation: Serve the homeless community in Portland

Methodology (Data Science approach)

1. Become Familiar with CCC client data
2. Merge CSV’s effectively
3. Feature Extraction/Encoding: Line of Business, Program Name
4. Explore Client Characteristics and Engagement Patterns

(If time permitted): Unique client table & Decision Tree Analysis

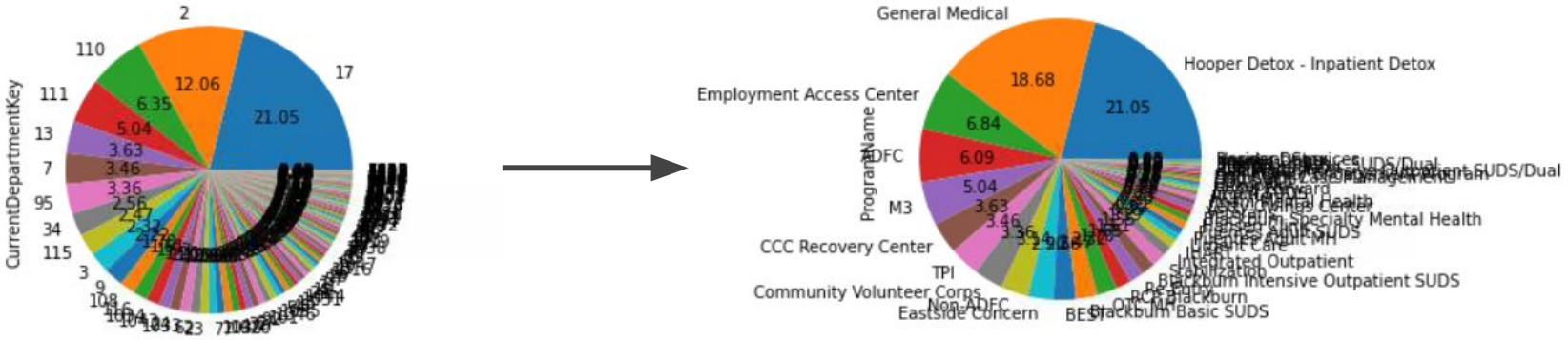
Becoming Familiar with CCC Client Data

- Daily Meetings
- Entry and feature recognition

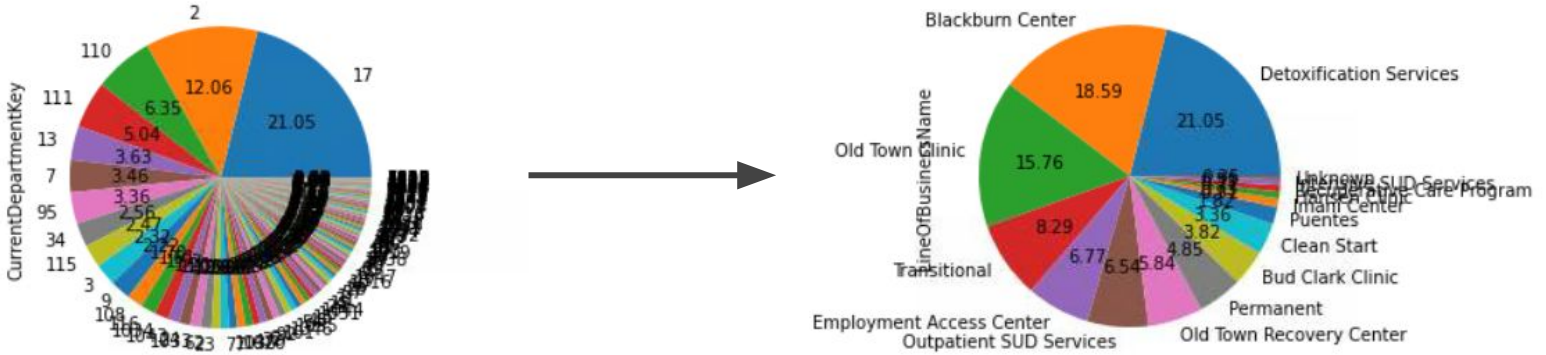
Merging Datasets/ Feature Engineering

- **Task**: Effectively merge datasets via ‘client keys’
- Dealing with a large dataset

Example: Active Client Program Enrollments



Merging Datasets/ Feature Engineering



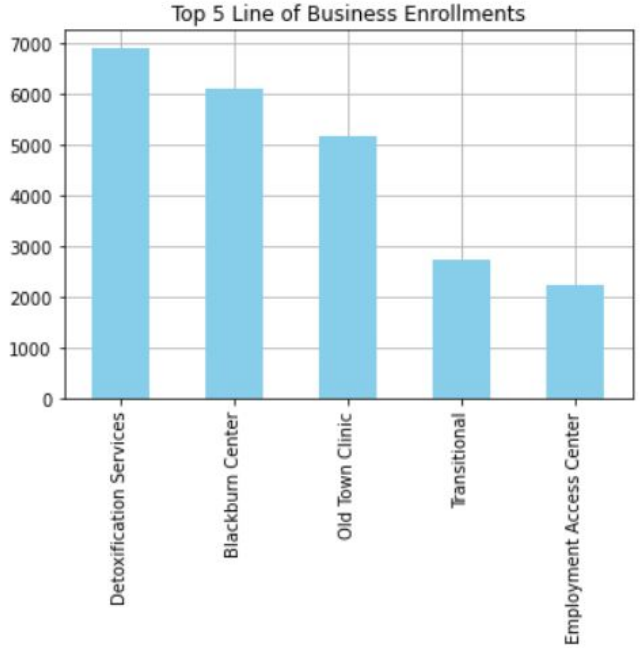
Exploring Client Engagement Patterns

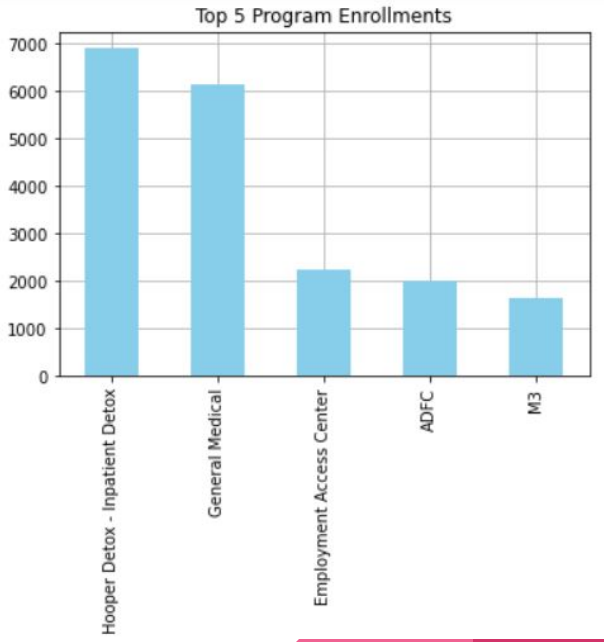
- Focus on top client enrollments

Questions of Interest:

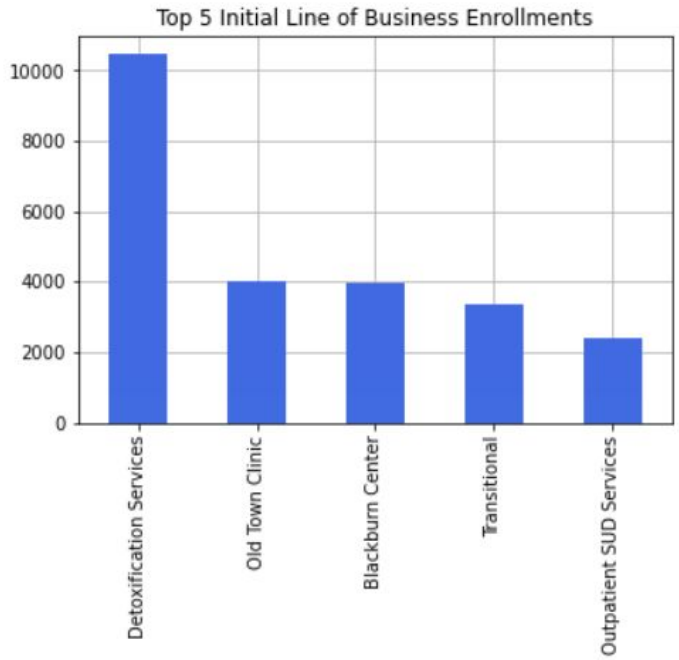
- Which programs have the most enrollments?
- Which lines of business have the most enrollments?
- Which T0 programs results in the most enrollments?
- What happens when you slice any of the questions above by demographics?

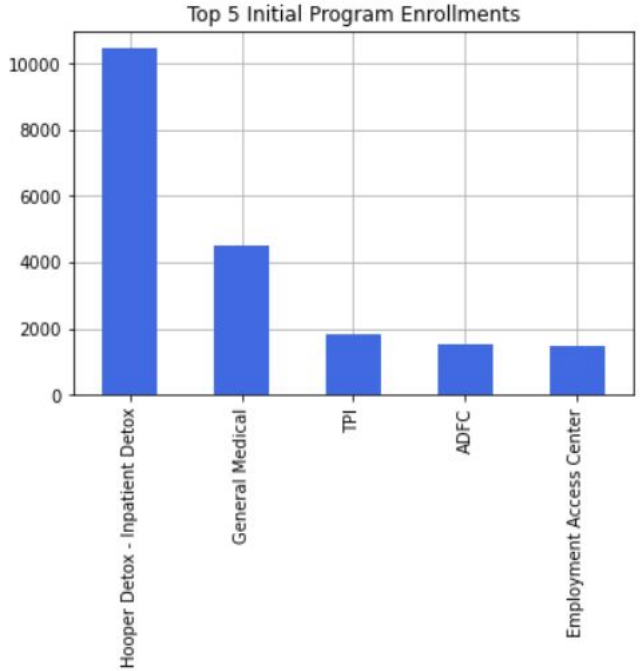
Exploring Client Engagement Patterns: Top Enrollments





Exploring Client Engagement Patterns: Top Initial Enrollments



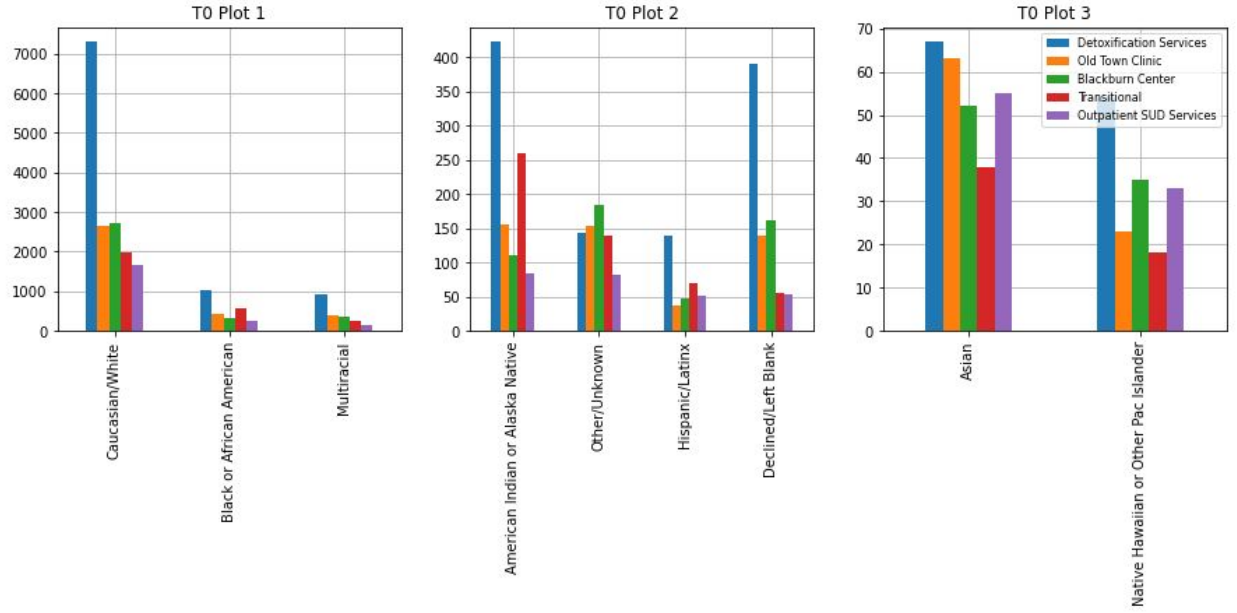


Exploring Client Engagement Patterns: Top Initial Enrollments by Race

- **T0 Line of Business**

```
Plots of Top 5 t0_LineOfBusiness Names by: Primary Race
NOTE: Scaling for each plot is different to compensate for relative comparison

<matplotlib.legend.Legend at 0x1e00cc56580>
```

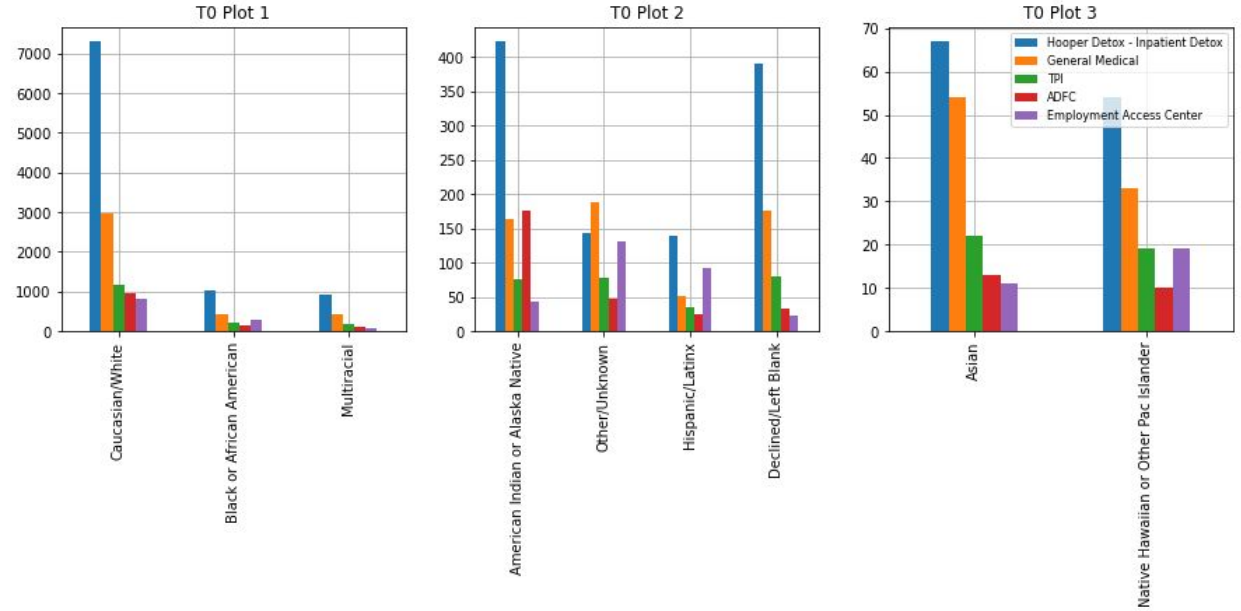


Exploring Client Engagement Patterns: Top Initial Enrollments by Race

- **T0 Programs**

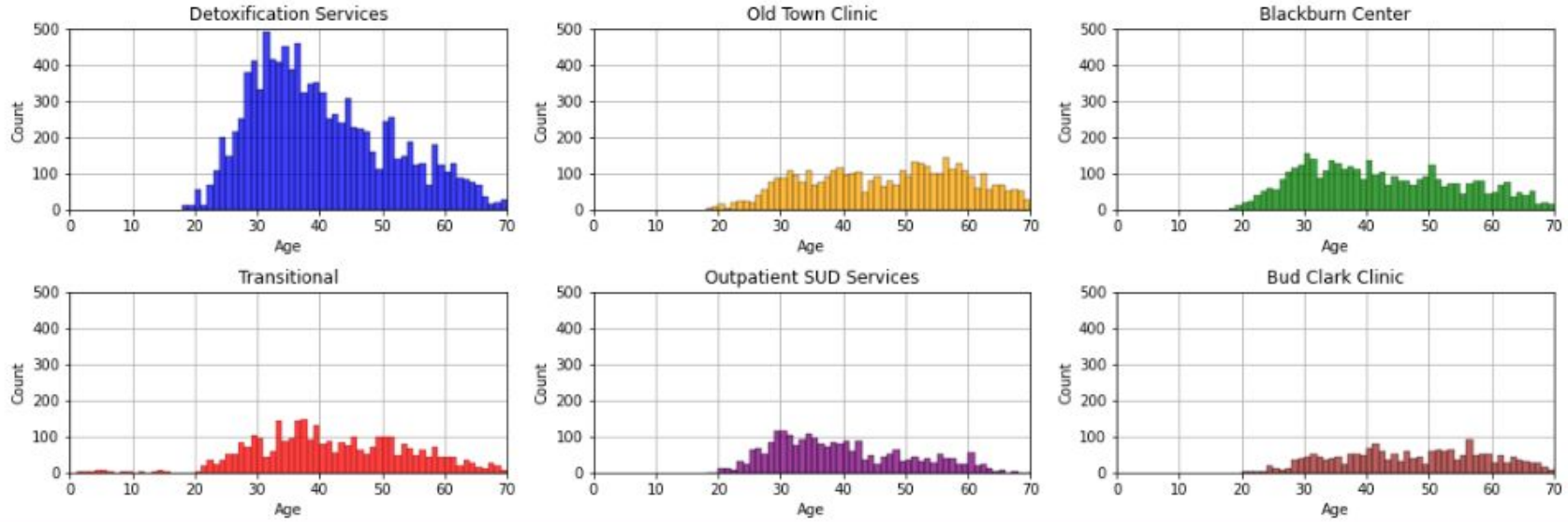
```
Plots of Top 5 t0_Program Names by: Primary Race
NOTE: Scaling for each plot is different to compensate for relative comparison

<matplotlib.legend.Legend at 0x1e00d0c5400>
```



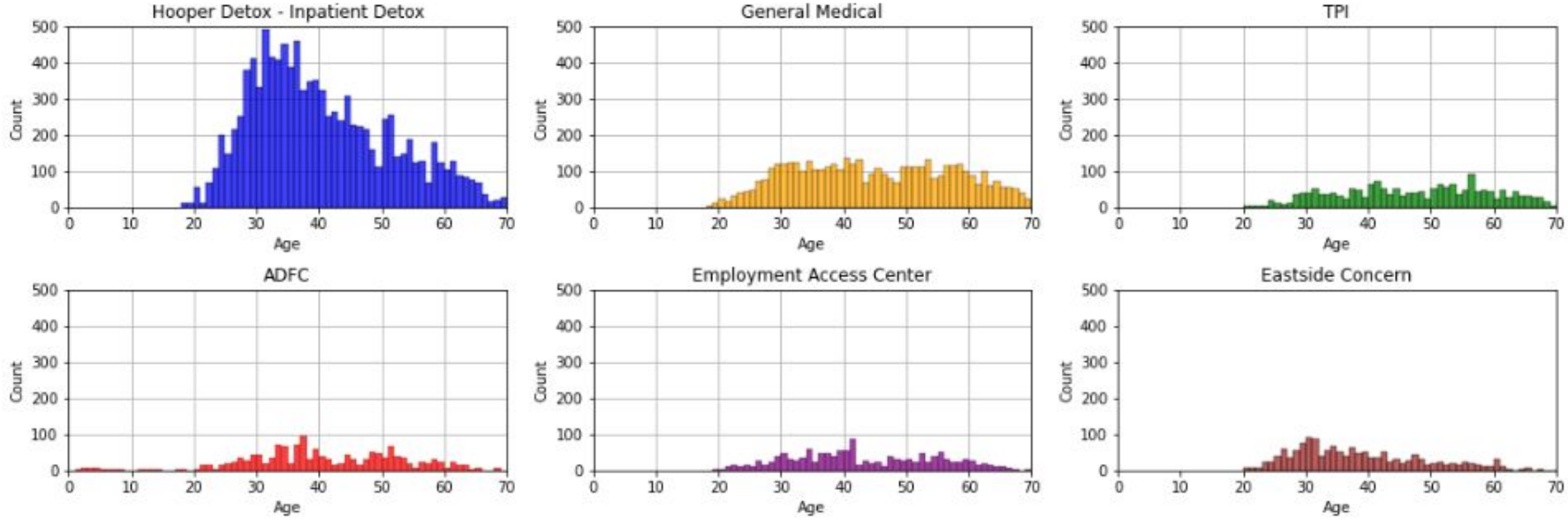
Exploring Client Engagement Patterns: Top Initial Enrollments by Age

Histogram plots for AGE by: Top 6 first index (t0) lines of businesses
NOTE: Scaling differs by subplot depending on relative counts of subplots

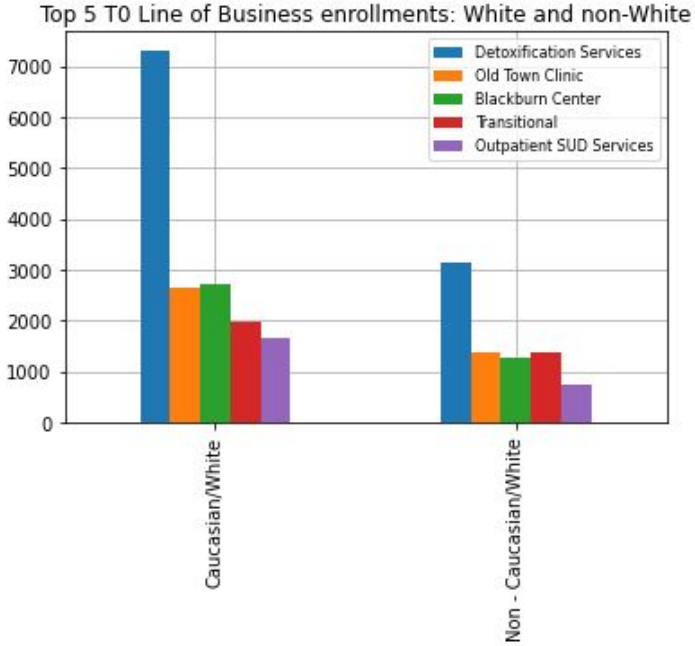


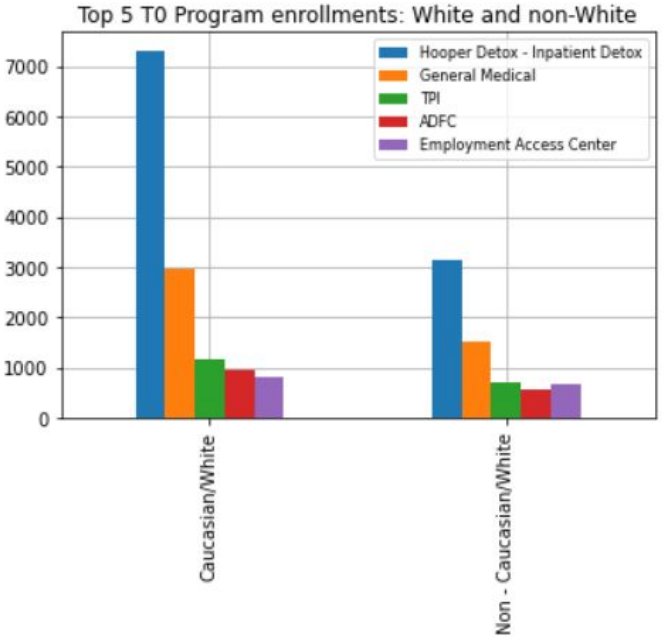
Exploring Client Engagement Patterns: Top Initial Enrollments by Age

Histogram plots for AGE by: Top 6 first index (t0) program names enrollments
NOTE: Scaling differs by subplot depending on relative counts of subplots



Noticeable Takeaways





Conclusion/Takeaways

- Newly Merged datasets
- Top enrollment patterns
- Distribution recognition

Future Next Steps (if time permitted)?

Acknowledgements

- AltREU PSU
- Central City Concern (CCC)