**Goal 1**

Identify clusters of regions based on a set of factors related to demographics and the anchor institutions for regional benchmarking.

* Characterize regions:
  + by their anchor institutions
  + the size and dynamics of the anchors relative to the community
  + the “quality” of anchors
    - Variables for all are provided but I reduced significantly in my model because of issues in analysis – but maybe you can do better.
      * Converting variables (percentages/per capita etc.) is OK.
      * If all factors can’t be included, we may want to focus on one of the dimensions above

Notes on cluster methods

* No optimal number of clusters but I think less is better for explanation
* No preferred method. I used this as a guide - http://www.philender.com/courses/multivariate/notes2/cluster.htmlbut but use whatever is helpful.

**Demographic factors**

|  |  |
| --- | --- |
| **Variable** |  |
| totpop\_19 | Total population in 2019 |
| popchange | 2010-2019 change, in percentage |
| medage | Median age, in years |
| poc | Person of color, in percentage |
| pov | People living in poverty, in percentage |
| youngpov | Population under 18 years living in poverty, in percentage |
| oldpov | Population over 65 years living in poverty, in percentage |
| labfor | Population in the labor force, in percentage |
| emp | People employed, in percentage |
| mbsaocc | People employed in management business, science and arts, in percentage |
| serocc | People employed in service occupations, in percentage |
| salocc | People employed in sales and office occupations, in percentage |
| ncmocc | People employed in natural resources, construction, and maintenance, in percentage |
| ptmocc | People employed in production, transportation, and material moving occupations |
| houincome | Median household income, in dollars |
| wfh | People working from home, in percentage |
| pcinc | Percapita income, in dollars |
| medearn | Median earnings, in dollars |
| nohealthin | Population with no health insurance coverage -civilian non institutionalised |
| highed | Population with at least bachelor's degree, in percentage |
| forborn | Population foreign-born, in percentage |
| noenglish | Population that speak English less than very well, in percentage |
| ownerocc | Population that is owner, in percentage |
| renterocc | Population that is renter, in percentage |

net\_mig Net domestic migration

**Anchor Institution Factors**

highered\_emp\_qcew Higher education employment QCEW

highered\_inc\_qcew Higher Education income QCEW

highered\_estab\_qcew Higher Education establishments QCEW

hospital\_emp\_qcew Hospital employment QCEW

hospital\_inc\_qcew Hospital income QCEW

hospital\_estab\_qcew Hospital establishments QCEW

inst\_highered\_endowment Higher education institution endowments for the region National Assoc of College and University Business Officers

inst\_highered\_nihresearchfunding Higher education NIH research funding for the region NIH

nst\_highered\_nsfresearchfunding Higher education NSF research funding for the region NSF

inst\_ipeds\_enrollment\_all Higher education enrollment for the region IPEDS

inst\_ipeds\_noncarnegie Higher education institutions without Carnegie Classification IPEDS

inst\_ipeds\_associatescolleges Higher education institutions-associates degree granting IPEDS

inst\_ipeds\_baccassoccollege Higher education institutions-bachelors degree granting IPEDS

inst\_ipeds\_baccalaureatecolleges Higher education institutions-bachelors degree granting IPEDS

inst\_ipeds\_masterscollegeuni Higer education institutions-masters degree granting IPEDS

inst\_ipeds\_doctuniallother Higher education institutions-doctoral degree granting IPEDS

inst\_ipeds\_doctoralunihighrese Higher education institutions-doctoral degree high research IPEDS

inst\_ipeds\_specialfocus Higher education institutions-special focus IPEDS

inst\_ipeds\_tribalcolleges Higher education institituions-tribal colleges IPEDS

inst\_ipeds\_undergradgrad Higher education institutions-undergraduate enrollment IPEDS

inst\_ipeds\_totalgradenrollment Higher education institutions-graduate enrollment IPEDS

inst\_ipeds\_completions Higher education institutions-completions IPEDS

inst\_ipeds\_pellawards Higher education institutions-total Pell grant amount awarded IPEDS

inst\_ipeds\_pellnumberofstudents Higher education institutions-total number of Pell grant recipients IPEDS

inst\_hosp\_ahacommunityhospitals Hospitals-Community Hospitals AHA

inst\_hosp\_ahabeds Hospitals-Number of beds in community hospitals AHA

inst\_hosp\_ahaadmissions Hospitals-Number of admissions in community hospitals AHA

inst\_hosp\_ahasurgicaloperations Hospitals-number of surgical procedures in community hospitals AHA

inst\_hosp\_nihresearchfunding Hospitals-NIH research funding for the region AHA

inst\_hosp\_medpayments Hospitals-Medicare payments for the region AHA

inst\_hosp\_medbeneficiaries Hospitals-Medicare beneficiaries for the region AHA

**I identified a 30 cluster solution it is not bad but I don’t have a lot of confidence in it**

[cluster 30 analysis | Tableau Public](https://public.tableau.com/app/profile/unocpar/viz/cluster30analysis/Dashboard1?publish=yes)

**Goal 2**

Conduct analysis (regression) to determine if and how the clusters explain economic variation in the region.

* Regional GDP should be fine
* Interested in the impact of clusters on all of them but simple and straightforward is fine

**Economic Outcomes**

impact\_gva\_direct Direct gross value added for hospitals and higher ed institutions Implan

impact\_share\_allind\_gva Share of gross value added from hospitals and higher ed institutions for the region Calculation

impact\_share\_allind\_emp Share of employment from hospitals and higher ed institutions for the region Calculation

impact\_share\_allind\_inc Share of income from hospitals and higher ed institutions for the region Calculation

regional\_gva\_allind Regional gross value added for all industries Bureau of Economic Analysis

regional\_emp\_allind Regional employment for all industries Bureau of Economic Analysis

regional\_inc\_allind Regional income for all industries Bureau of Economic Analysis

index\_real\_gdp\_19 2019: Chain-type indexes for real GDP

current\_dollar\_gdp\_19: 2019 Current Dollar GDP

real\_gdp\_19: 2019 Real GDP

index\_real\_gdp\_21: 2021 Chain-type indexes for real GDP

current\_dollar\_gdp\_21: 2021 Current Dollar GDP

real\_gdp\_21: 2021 Real GDP

percapita\_personal\_income\_19: 2019 Per capita personal income

total\_income\_pc\_19: 2019 total per capita personal income

percapita\_personal\_income\_21: 2021 Per capita personal income

total\_income\_pc\_21: 2021 Total per capita personal income

**Deliverables**

* As clsx. all data points as used in the cluster analysis and the cluster solution (see dashboard map)
* Descriptive statistics for all factors used by cluster (see dashboard summary)
* Formatted table output of goal 2