

Understanding the relationship between stocks of community wealth and food system

Elliot Hohn

Introduction

Community wealth quantification outline.

LAMP grant program description.

Hypothesis re: relationship between the two.

Methods

Plan right now is load county level community wealth data, and LAMP grant data, identify counties with low grant funding amounts (or maybe split into high, medium, and low?), then perform cluster analysis to determine common characteristics of the counties in each group, if any.

Environment setup

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.4.4      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr      1.0.2
```

```
-- Conflicts ----- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
```

```
x dplyr::lag()     masks stats::lag()
```

```
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(sf)
```

Linking to GEOS 3.11.0, GDAL 3.5.3, PROJ 9.1.0; sf_use_s2() is TRUE

```
library(openxlsx)
library(stringr)
library(corrplot)
```

corrplot 0.92 loaded

```
library(stringr)
```

Load community stocks data

Load data directly from CSU [github repo](#).

```
comm_stocks_url <- 'https://raw.githubusercontent.com/CSU-Local-and-Regional-Food-Systems/
comm_stocks <- read.csv(comm_stocks_url)

# Function to add leading 0 to 4-digit fips values in comm_stocks
add_leading_zero <- function(x) {
  # Check if the number has 4 digits
  ifelse(nchar(x) == 4, paste0("0", x), x)
}

comm_stocks$fips <- sapply(comm_stocks$fips, add_leading_zero)
print(head(comm_stocks))
```

	fips	county_name	state_name	category	topic_area	year
1	01001	Autauga County	Alabama	Community Characteristics	Infrastructure	2007
2	01003	Baldwin County	Alabama	Community Characteristics	Infrastructure	2007
3	01005	Barbour County	Alabama	Community Characteristics	Infrastructure	2007
4	01007	Bibb County	Alabama	Community Characteristics	Infrastructure	2007
5	01009	Blount County	Alabama	Community Characteristics	Infrastructure	2007
6	01011	Bullock County	Alabama	Community Characteristics	Infrastructure	2007
	variable_name		value			
1	highway_km		0.12519374			

```

2 highway_km 0.06419832
3 highway_km 0.01535104
4 highway_km 0.04835747
5 highway_km 0.05661337
6 highway_km 0.02661536

```

```

# ggplot(comm_stocks, aes(x = value)) +
#   geom_histogram(fill = 'red', colour = 'black') +
#   facet_wrap(variable_name ~ .)

print(unique(comm_stocks$variable_name))

```

```

[1] "highway_km" "foodbev_est_CBP"
[3] "est_CBP" "broad_11"
[5] "broad_16" "pc1b_manufacturing"
[7] "pc2b_infrastructure" "create_indus"
[9] "racial_div" "pub_lib"
[11] "create_jobs" "museums"
[13] "pc1c_artsdiversity" "pc2c_creativeindustries"
[15] "localgovfin" "owner_occupied"
[17] "deposits" "pc1f"
[19] "ed_attain" "food_secure"
[21] "insured" "primary_care"
[23] "health_factors" "health_outcomes"
[25] "pc1h_healtheducation" "pc2h_medicalfoodsecurity"
[27] "natamen_scale" "prime_farmland"
[29] "conserve_acre" "acre_FSA"
[31] "acre_NFS" "pc1n_naturalamenitiesconservation"
[33] "pc2n_farmland" "pvote"
[35] "nccs" "assn"
[37] "respn" "pc1s_nonprofitsocialindustries"
[39] "pc2s_publicvoiceparticipation"

```

Check correlations between variables

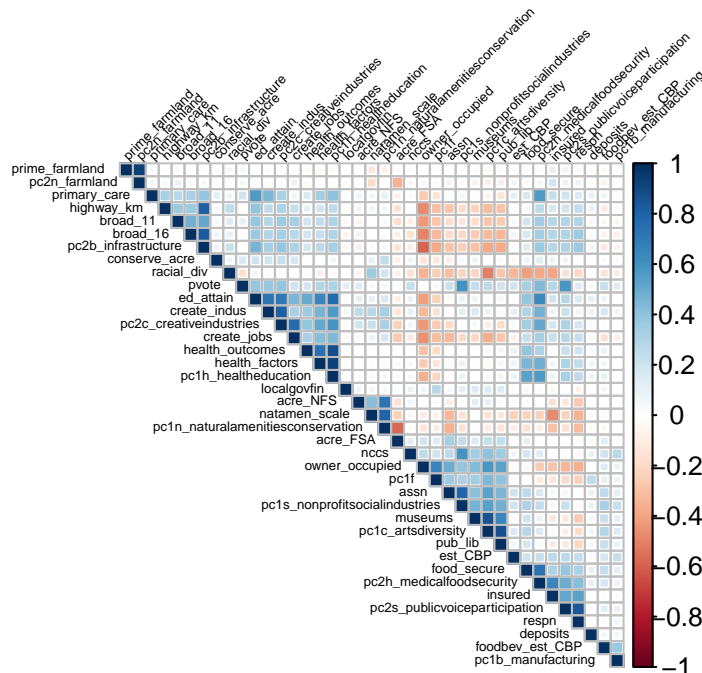
There are different years associated with the rows of data...how should I deal with this? Filter to only the newest data? Use the average? The grants data will also have a date associate with it, but I was thinking those grants would all be summed to get a “total amount invested” value.

I think for now I'm going to use the mean across all years for any given variable. A quick look at the data indicates that the values don't shift drastically from year to year, unsurprisingly.

```
# reshape data from tall to wide to give each variable its own column
comm_stocks_wide <- comm_stocks %>%
  select(fips, year, variable_name, value) %>%
  group_by(fips, variable_name) %>%
  summarise(value = mean(value)) %>%
  ungroup() %>%
  pivot_wider(names_from = variable_name, values_from = value)
```

`summarise()` has grouped output by 'fips'. You can override using the
`.groups` argument.

```
# correlation matrix
correlation <- cor(comm_stocks_wide[, 2:ncol(comm_stocks_wide)], use = "complete.obs")
corrplot(correlation, method = 'square', tl.cex = 0.4, tl.col = 'black', tl.srt = 45,
  type = 'upper', order = "hclust")
```



Load counties shapefile and join to community stocks data

```
counties <- st_read('data/cb_2021_us_county_20m/cb_2021_us_county_20m.shp') %>%  
  filter(!STATE_NAME %in% c('Alaska', 'Hawaii', 'Puerto Rico')) %>%  
  select(GEOID, geometry) %>%  
  st_transform(5070)
```

Reading layer `cb_2021_us_county_20m' from data source

`/Users/elliott/Documents/repos/fas-capstone/data/cb_2021_us_county_20m/cb_2021_us_county_20m.shp'
using driver `ESRI Shapefile'

Simple feature collection with 3221 features and 12 fields

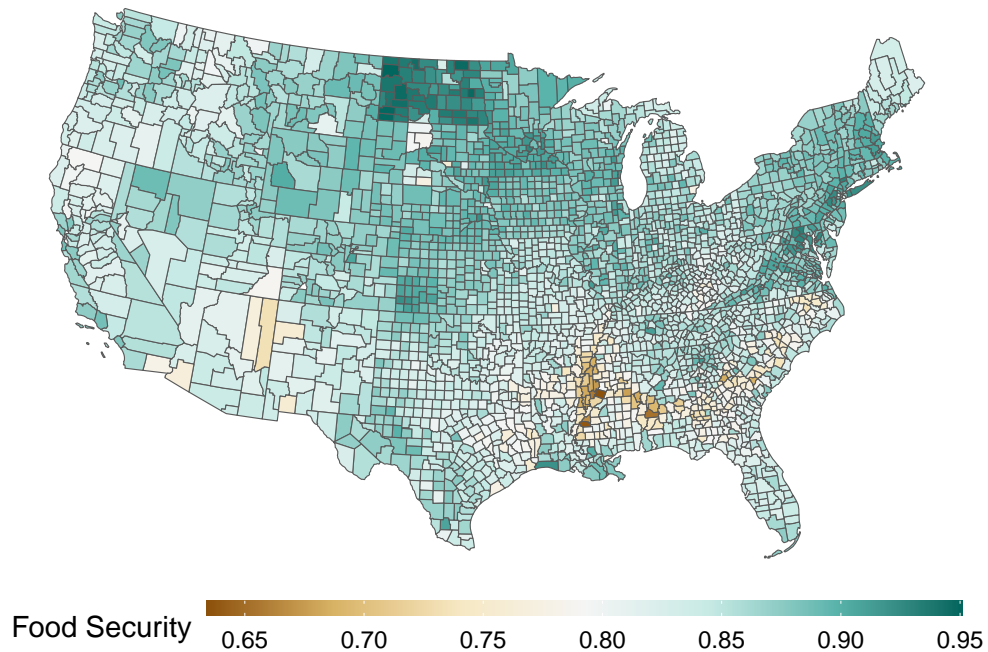
Geometry type: MULTIPOLYGON

Dimension: XY

Bounding box: xmin: -179.1743 ymin: 17.91377 xmax: 179.7739 ymax: 71.35256

Geodetic CRS: NAD83

```
comm_stocks_geo <- merge(counties, comm_stocks_wide, by.x = 'GEOID', by.y = 'fips')  
  
ggplot(comm_stocks_geo) +  
  geom_sf(aes(fill = food_secure)) +  
  scale_fill_distiller(  
    name="Food Security",  
    palette = "BrBG",  
    direction = 1  
  ) +  
  theme_void() +  
  theme(legend.position = "bottom",  
        legend.key.width = unit(2,"cm"),  
        legend.key.height = unit(0.2,"cm")  
  )
```



```
summary_stats <- comm_stocks %>%
  group_by(variable_name) %>%
  summarize(
    mean = mean(value),
    median = median(value),
    sd = sd(value),
    max = max(value),
    min = min(value)
  )

print(summary_stats)
```

A tibble: 39 x 6

	variable_name	mean	median	sd	max	min
	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	acre_FSA	0.0148	0.00330	0.0262	0.254	0
2	acre_NFS	0.0476	0	0.124	0.936	0
3	assn	1.40	1.27	0.717	7.08	0
4	broad_11	0.267	0.00871	0.353	1	0
5	broad_16	0.635	0.73	0.323	1	0
6	conserve_acre	0.0143	0.00392	0.0278	0.284	0
7	create_indus	161.	139.	111.	1779.	0

```

8 create_jobs      0.160      0.150      0.0584      0.502 0.00540
9 deposits         22.5       16.9       51.1      2363.      0
10 ed_attain       0.262      0.235      0.119      1.22 0.0244
# i 29 more rows

```

Load LAMP data

Data downloaded directly from URL on the LAMP Navigator [webpage](#).

```

lamp_url <- 'https://www.ams.usda.gov/sites/default/files/media/LAMPDatasetandDataDictiona

lamp <- read.xlsx(lamp_url, sheet = 'LAMP Dataset 2006 to 2023')
print(head(lamp))

```

		Organization	City.2	
1		FRESHFARM Markets, Inc.	Washington	
2		Jesus Provides Our Daily Bread	Chico	
3		County of Riverside Office of Economic Development	Riverside	
4	San Luis Obispo County Community College District dba Cuesta	San Luis Obispo		
5		Sierra Kings Health Care District	Reedley	
6		Historic St. Andrews Waterfront Partnership, Inc.	Panama City	
	State	Zip	Congressional.District	Project.Congressional.District
1	DC	20010	1	1
2	CA	95928	1	1
3	CA	92501	25	25
4	CA	93405	24	24
5	CA	93654	21	21
6	FL	32401	2	2
		Entity.Type	Year	
1		Nonprofit Corporation	2023	
2		Nonprofit Corporation	2023	
3		Local Government	2023	
4	Public/State Controlled Institution of Higher Education	2023		
5		Local Government	2023	
6		Nonprofit Corporation	2023	
	Unique.Project.Number	Grant.Type	Project.Type	Project.Count
1	2023FMPPDC1	FMPP Capacity Building	1	
2	2023FMPPCA2	FMPP Capacity Building	1	
3	2023FMPPCA3	FMPP Capacity Building	1	
4	2023FMPPCA4	FMPP Capacity Building	1	
5	2023FMPPCA5	FMPP Capacity Building	1	

6	2023FMPPFL6	FMPP Capacity Building	1
	Match.Amount.(Final)	Award.Amount	Total.Project.Cost
1	125089.4	500000.0	625089.4
2	65150.0	250000.0	315150.0
3	45245.0	180980.0	226225.0
4	52000.0	207046.5	259046.5
5	69512.0	114900.0	184412.0
6	55132.0	220000.0	275132.0

1
2 Harvesting Hope: Addressing L
3 Enriching and nourishing the agricultural communities in the
4
5 Reedley Wellness Farmer's Market; creating an attractive mar
6 Execute the "Farm to Town" initiative to include, in part, a new farmers market in Panama C

1
2 Butte County is burdened by some of the highest food insecurity rates in the state and nat
3
4
5 benefit to local producers (with additional support for underrepresented producers), local co

6 The Historic St. Andrews Waterfront Partnership, in collaboration with
The Market at St. Andrews- and the proposed new Glenwood Farmers Market as reliable and flour

	Project.Start.Date	Project.End.Date	Aggregation	Agritourism
1	9/30/2023	9/29/2026	0	0
2	9/30/2023	46294	0	0
3	9/30/2023	9/29/2026	0	1
4	9/30/2023	9/29/2026	0	0
5	9/30/2023	46294	0	0
6	9/30/2023	9/29/2026	1	0

	Farm.to.Institution/Wholesale	Food.Safety	Food.Waste	Grantwriting
1	0	0	0	0
2	0	0	0	0
3	0	0	1	0
4	0	0	0	0
5	0	0	0	0
6	1	0	0	0

	Infrastructure	Marketing.and.Promotion	Organic	Other	Processing	Production
1	0	1	0	0	0	0
2	1	1	1	0	0	0
3	0	1	1	0	0	0
4	0	1	0	0	0	0

5	0	1	0	0	0	0
6	1	1	0	0	0	0

	Recruitment	Season.Extension	Training	Transportation	Value.Added
1	1	1	1	0	1
2	1	1	1	0	1
3	1	0	1	0	1
4	1	0	0	0	0
5	1	0	1	0	0
6	0	0	1	1	0

	Cooperative.Development	Diversification	Education.Outreach	Food.Insecurity
1	0	1	0	0
2	0	1	0	0
3	0	1	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0

	Food.Hub	Resource.Development	Strategic.Planning	Value.Chain.Coordination
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0

	Not.Specified
1	0
2	0
3	0
4	0
5	0
6	0

1	Marketing and Promotion, Recruitment, Season Extension, Training,
2	Infrastructure, Marketing and Promotion, Organic, Recruitment, Season Extension, Training,
3	Agritourism, Food Waste, Marketing and Promotion, Organic, Recruitment, Training,
4	Marketing and Promotion, Organic, Recruitment, Training,
5	Marketing and Promotion, Organic, Recruitment, Training,
6	Aggregation, Farm to Institution/Wholesale, Infrastructure, Marketing and Promotion,

	StoryMap.Link	Seed.Nof.Success.Highlights
1	No StoryMap	No
2	No StoryMap	No
3	No StoryMap	No
4	No StoryMap	No
5	No StoryMap	No

Load city lat/lon dataset

Loaded directly from [GitHub repo](#).

Going to try to join to cities from LAMP dataset.

```
url <- 'https://raw.githubusercontent.com/kelvins/US-Cities-Database/main/csv/us_cities.csv'

city_loc <- read.csv(url)
```

Prep data for merge by renaming some columns, converting city names to all lower case in both data sets to ensure clean join. Note: there are some cities in the city_loc dataset that have more than one row, I think because they are bisected by county lines, so I dealt with that short list of issues by grouping them by city and state and selecting the first row in the group.

Also, there are three remaining rows in the LAMP data that don't have cities listed, so there are no coords for them, but I could track those down and fix them manually.

```
lamp <- lamp %>%
  rename(
    state = State,
    city = City.2
  ) %>%
  mutate(city = str_squish(str_trim(tolower(city)))) %>%
  filter(!state %in% c('HI', 'AK', 'PR'))

city_loc <- city_loc %>%
  select(-c(ID, COUNTY)) %>%
  rename(
    city = CITY,
    state = STATE_CODE
  ) %>%
  mutate(city = str_squish(str_trim(tolower(city)))) %>%
  group_by(city, state) %>%
  filter(row_number() == 1) %>%
  ungroup()

# check how many cities in LAMP data are in city_loc data
print(length(unique(lamp$city)))
```

[1] 960

```
print(length(unique(lamp$city) %in% city_loc$city))
```

[1] 960

```
# do the join
lamp <- merge(lamp, city_loc, by = c('city', 'state'), all.x = TRUE)

# convert to sf object
lamp_sf <- lamp %>%
  drop_na(LATITUDE, LONGITUDE) %>%
  st_as_sf(coords = c("LONGITUDE", "LATITUDE")) %>%
  st_set_crs(4326) %>%
  st_transform(st_crs(comm_stocks_geo))
```

Add LAMP grant points to the map

```
ggplot(comm_stocks_geo) +
  geom_sf(aes(fill = food_secure)) +
  geom_sf(
    data = lamp_sf,
    size = 1.5,
    alpha = .5,
    color = 'purple',
    fill = 'purple'
  ) +
  scale_fill_distiller(
    name="Food Security",
    palette = "BrBG",
    direction = 1
  ) +
  theme_void() +
  theme(legend.position = "bottom",
        legend.key.width = unit(2,"cm"),
        legend.key.height = unit(0.2,"cm")
  )
```

