Good morning everyone. My project is a data analysis type project that I titled 'The Proliferation of Dollar General in Distressed Communities, a Spotlight Analysis in Virginia.' I chose Virginia because I live nearby.

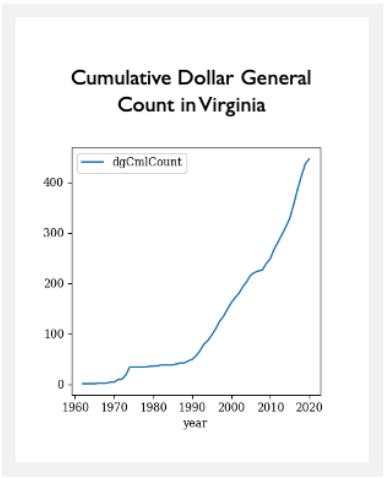
For those of you who don't know what Dollar General is, it is a small-box convenient store that sells various foods, snacks, cleaning supplies, housewares, et cetera for sometimes only $1, but usually less than $10.

They tend to be located in neighborhood and small community locations, which is different from other larger grocery retailers that tend to locate in urban shopping centers.

I was trying to find what the Chinese equivalent may be, and it turns out that many stores in China already offer cheaper items like Dollar General. This "budget store" concept may be unique to western culture, but if you think of any businesses that are similar, feel free to write it in the chat!

RESEARCH QUESTION & BACKGROUND

- 16,278 stores across America, almost 450 stores in Virginia
- Key operating strategy: capitalize on a "low-risk real estate growth model"
  - Trip-frequency
  - Alternative to large-box grocery stores
- **Do declining economic factors create attractive locations for new Dollar General stores?**
  - Hypothesis: lower incomes and declining work force leads to greater number of Dollar generals



Dollar General operates over sixteen thousand stores across the United States, and there are almost 450 stores in Virginia.I said earlier that they tend to be in neighborhood locations, and many times (not always), these communities can be struggling economically.

On the surface, it seems beneficial that Dollar General provides increased access to affordable food and other essential items to these areas. However, there is some research that suggests that Dollar General is actually contributing to such economic decline.

One of Dollar General's key operating strategies from their 2019 Annual Report, involves capitalizing on A low-risk-real-estate-growth-model. This involves obtaining a vast amount of small space stores in convenient locations. This strategy helps Dollar General drive trip-frequency by being closer to neighborhoods than large-box grocery stores.

I ask if declining economic factors in an area create attractive locations for new Dollar General store? And I hypothesize that lower incomes and lowering labor force leads to a greater number of Dollar Generals.

Table 3: Pairwise Correlation Matrix

	dgCount	povPcnt	realMedInc	lfPartRate	lf_size	logMedInc	logLf_size
dgCount	1.000000	0.098851	-0.119284	-0.031421	0.173329	-0.093381	0.423217
povPcnt	0.098851	1.000000	-0.809335	-0.739640	-0.283034	-0.857862	-0.360101
realMedInc	-0.119284	-0.809335	1.000000	0.656484	0.527488	0.985048	0.557318
lfPartRate	-0.031421	-0.739640	0.656484	1.000000	0.252061	0.704889	0.313625
lf_size	0.173329	-0.283034	0.527488	0.252061	1.000000	0.467236	0.720648
logMedInc	-0.093381	-0.857862	0.985048	0.704889	0.467236	1.000000	0.550274
logLf_size	0.423217	-0.360101	0.557318	0.313625	0.720648	0.550274	1.000000

RESULTS AND DESCRIPTIVE STATISTICS

- Small, but negative relationship with median income
- Relatively strong relationship with log(laborforce)
- Multicollinearity issue between poverty and median income

After creating the main data frame, I chose these variables to explain the number of Dollar Generals in the counties of Virginia: poverty rate, real median income, labor force participation rate, and size of the labor force. I also added log of income and log of labor force size, since it makes more sense to explain these as percentages.
Using a pairwise correlation matrix, I noticed right off the bat that Dollar General count had a very small, but [negative] relationship with median income. This makes sense intuitively. As incomes rise, you'd expect people to buy their groceries from higher quality stores.
There is also a relatively strong relationship between Dollar General count and log of labor force size. I wasn't expecting to see a relationship that strong the number of Dollar Generals.
Another observation is that poverty rate and median income are very strongly correlated, which could - will - cause a multicollinearity problem in my regression model that follows.

MODEL OUTPUT

- Surprising R-squared
- *log(median income):*
  - Dollar General goods are like inferior goods
  - Suggestive of target market?
- *log(laborforce):*
  - Contradicted my hypothesis
  - That's just business!

Table 6: Model 2 OLS Regression Results

Dep. Variable:	dgCount	**R-squared:**	0.332
Model:	OLS	**Adj. R-squared:**	0.319
Method:	Least Squares	**F-statistic:**	25.82
No. Observations:	133	**Log-Likelihood:**	-207.78
Df Residuals:	130	**Prob (F-statistic):**	7.86e-10
Df Model:	2		

	coef	std err	t	P> |t|	[0.025	0.975]
Intercept	25.6148	6.390	4.008	0.000	12.943	38.287
logMedInc	-3.2005	0.657	-4.875	0.000	-4.502	-1.899
logLF	1.2299	0.173	7.091	0.000	0.886	1.574

$$dgCount = 25.618 - 3.201log(realMedInc) + 1.230log(lf\_size)$$

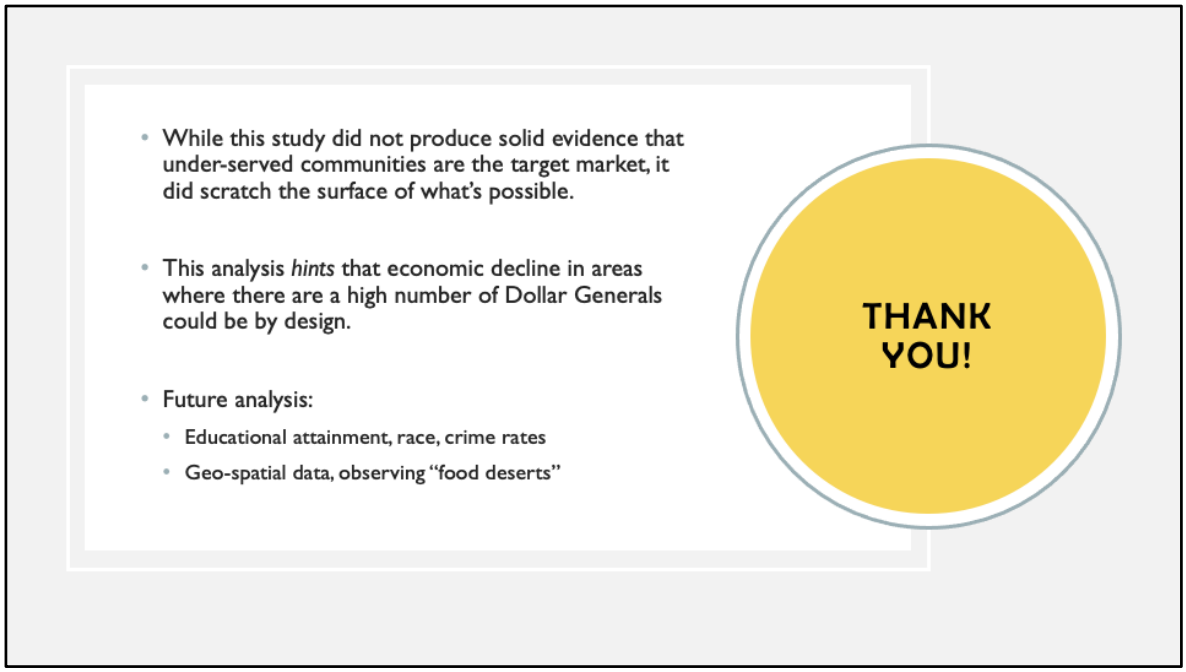
I ran a few models, and the best one involved only two variables that were statistically significant: log of median income and log of labor force size.

I had a realization pretty early on during data extraction that my variables were not going to capture the full explanation of the number of Dollar Generals, so when I saw n R-squared of .332, I was pleased. Unfortunately, the coefficients are so small, that they still don't paint a clear picture, but I was happy with this.

There is a negative coefficient with log of median income, which again makes sense intuitively, and strengthens the argument that Dollar General products are viewed as inferior goods.

While further analysis is definitely required, this relationship also suggests that Dollar General might be targeting areas with households on the lower end of the income scale.

There is a positive relationship with log of labor force, which contradicted my initial hypothesis. I assumed that areas with a declining labor force would have more Dollar Generals. This coefficient suggests otherwise. One potential explanation is that it's just business. As a county grows economically, it invites more business to the region.



To conclude… While this study did not produce solid evidence that Dollar General "preys on the poor," it did scratch the surface of what's possible to research with publicly available data.
This analysis hints that economic decline in areas where there are lots of Dollar Generals could be by design.
In the future, it would be helpful to add to this basic model other explanatory variables like educational attainment, race, and crime rates.
It would also be very interesting to look at this data geo-specially. The original data that I web scraped from Dollar General's location webpage gave the coordinates for each location. It would interesting to identify if Dollar General contributes to food deserts.
Thank you so much for this semester!