

Data I'm working with:

sht\_07, sht\_10, sht\_15, sht\_20 ← add respective year columns

socio\_economic\_df, which is christophers df of all socioeconomic data

in creating dfs based off of features from socioeconomic df

Per 2007:

gdp\_df\_2007 will have all socioeconomic data for 2007. Add One Year GDP Growth Rate and Incumbent term GDP Growth rate using this function.

Implementation

```
def calculate_one_year_growth(df, column_name, election_year):
    df[column_name] = np.log(df[column_name] / df[column_name].shift(election_year))
    return df

[199] ✓ 0.0s
```

Generate + Code + Markdown

```
sht_07['Year'] = 2007
soc_gni = socioeconomic_df
soc_gni = calculate_one_term_growth(soc_gni, 'GNI (Per Capita US $)', 7)
soc_gni = calculate_one_year_growth(soc_gni, 'GNI (Per Capita US $)')

[211] ✓ 0.0s
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS JUPYTER

Repeat this for all features where incumbent term growth rate and one-year growth rate are necessary.

Repeat from "Per 2007" for all years.

Extract the information from the socioeconomic table where the Row = election year and merge it with the df for election information.

for example:

sht\_07 with socioeconomic [year] == 2007  
etc.

concatenate sht\_07, sht\_10, sht\_15, sht\_20

↑  
master table

$X_{train}$  = all socioeconomic info from master table

Predict voter share for PNM and UNC using columns from the master table

So

$Y_{pnm} = \text{master table}[\text{'PNM\_Votes'}]$

$Y_{unc} = \text{master table}[\text{'UNC\_Votes'}]$

End of Plan

Implementation might be different for models? But master table and training data remain the same