DS 2500 **Final Project**

B L A C K

LIVING IN AMERICA

Eva B, Ihunaya E, Amanda R, Racheal A

Overview

Introduction

- Our project aims to analyze where Black people in the U.S. will migrate to in the coming years in relation to where they were in the past.
- This provokes important questions such as: is this where Black people actually want to be living?
- Inspired by the "Negro Motorist Green Book"
- This project will attempt to take a peek into the future using polynomial regression.

Data Description

```python
# loading dataframe
df_state_demo = pd.read_csv('state_demographics.csv')
df_state_demo.head()
```

Python

	State	Population.Population Percent Change	Population.2014 Population	Population.2010 Population	Age.Percent Under 5 Years	Age.Percent Under 18 Years	Age.Percent 65 and Older	Miscellaneous.Percent Female	Et
0	Connecticut	-10.2	3605944	3574097	5.1	20.4	17.7	51.2	
1	Delaware	8.4	989948	897934	5.6	20.9	19.4	51.7	
2	District of Columbia	17.3	689545	601723	6.4	18.2	12.4	52.6	
3	Florida	14.2	21538187	18801310	5.3	19.7	20.9	51.1	
4	Georgia	9.6	10711908	9687653	6.2	23.6	14.3	51.4	

5 rows × 48 columns

```python
# loading dataframe
df_population = pd.read_csv('reformatted_percent_black_over_time2.csv.csv')

# replacing NaN values with -1
df_population = df_population.fillna("-1%")
df_population
```

	State/Territory	Year	Population Percentage
0	AL	1790	-1%
1	AL	1800	41.40%
2	AL	1810	29.00%
3	AL	1820	33.20%
4	AL	1830	38.50%
...	...	...	...
1351	WY	1980	0.70%
1352	WY	1990	0.80%
1353	WY	2000	0.80%
1354	WY	2010	0.80%
1355	WY	2018	1.30%

Black Living **in America**

Methodology

01 Polynomial Regression

Using Historical Population Data

- Used a polynomial regression to predict the black population of each state based on historical data
- For qualitative curiosity - visualize the movement trends in the black community

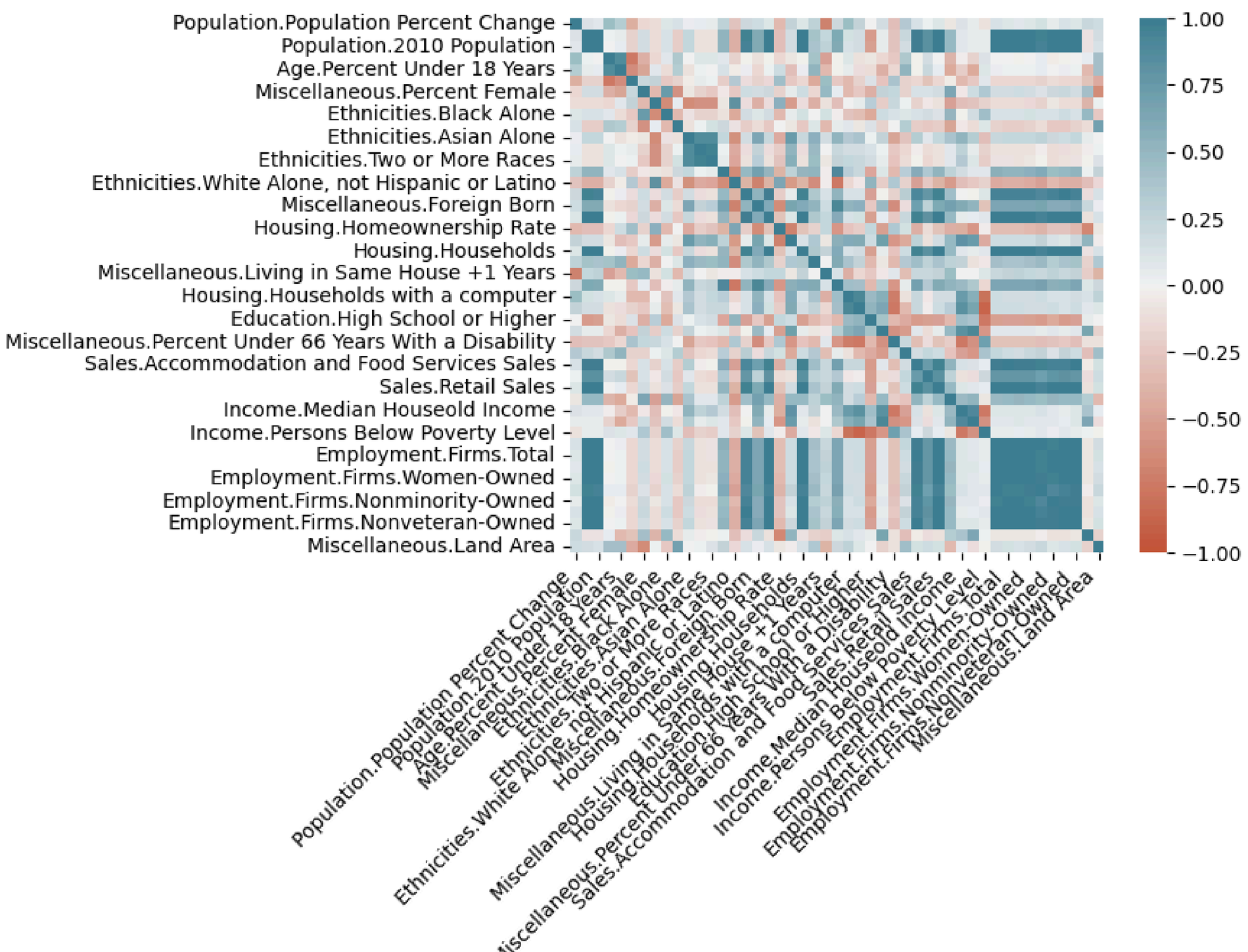
02 Linear, Lasso, Regression

Using State Demographics

- Feature engineering to prevent: overfitting, long computation time, etc
- The model assigned coefficients to each of the 46 initial features
- Removed all the ones with 0 as their coef
- r2 score of 91%

Black Living **in America**

Feat. Correlation *Before* Feat. Engineering



Black Living **in America**

Feat. Correlation *After* Feat. Engineering

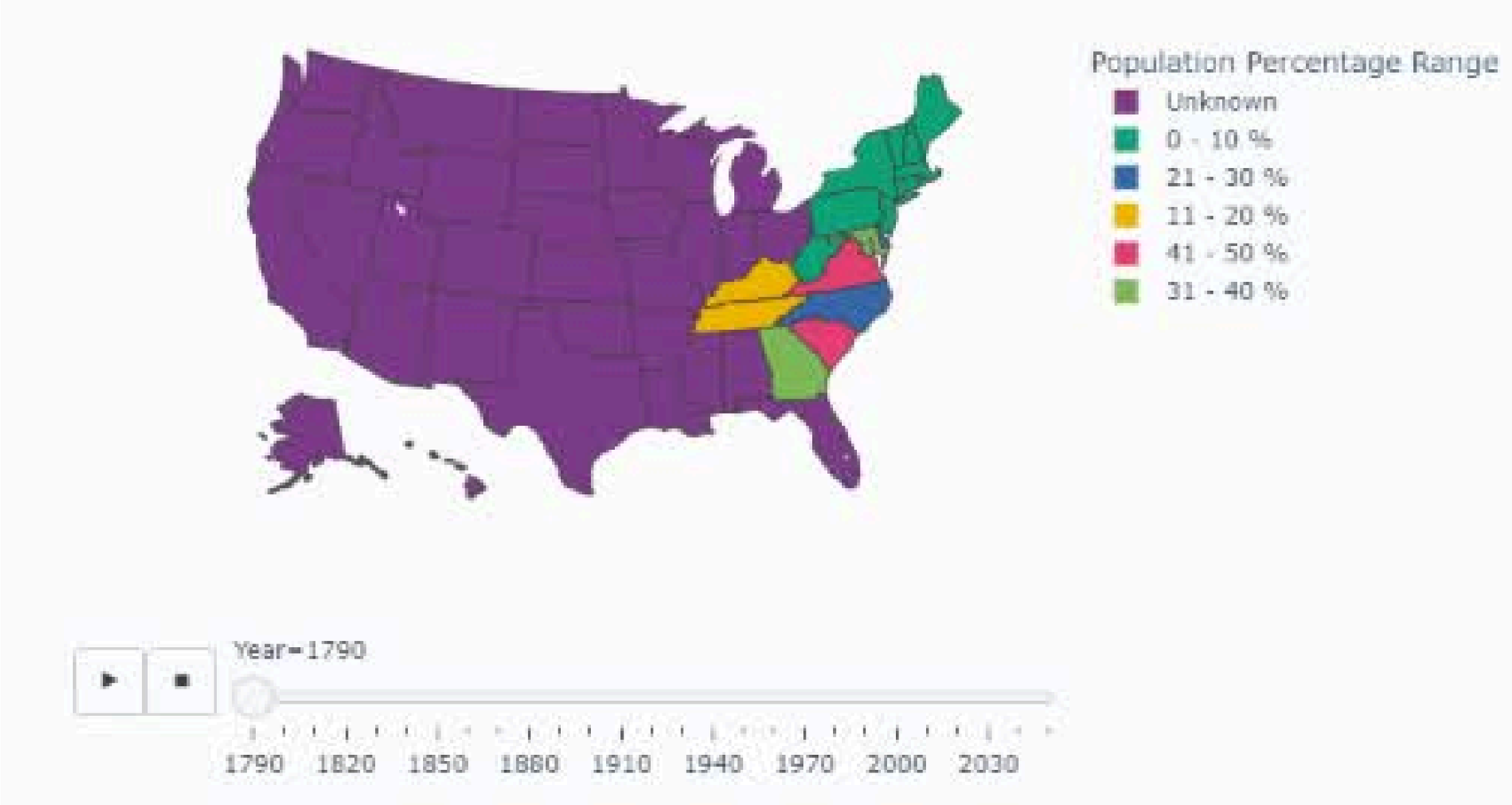
Black Living **in America**

Percent of Black Population by State (1790-2050)

Extrapolated States

Using our Model to Predict 2030-2050 Black population %s in:

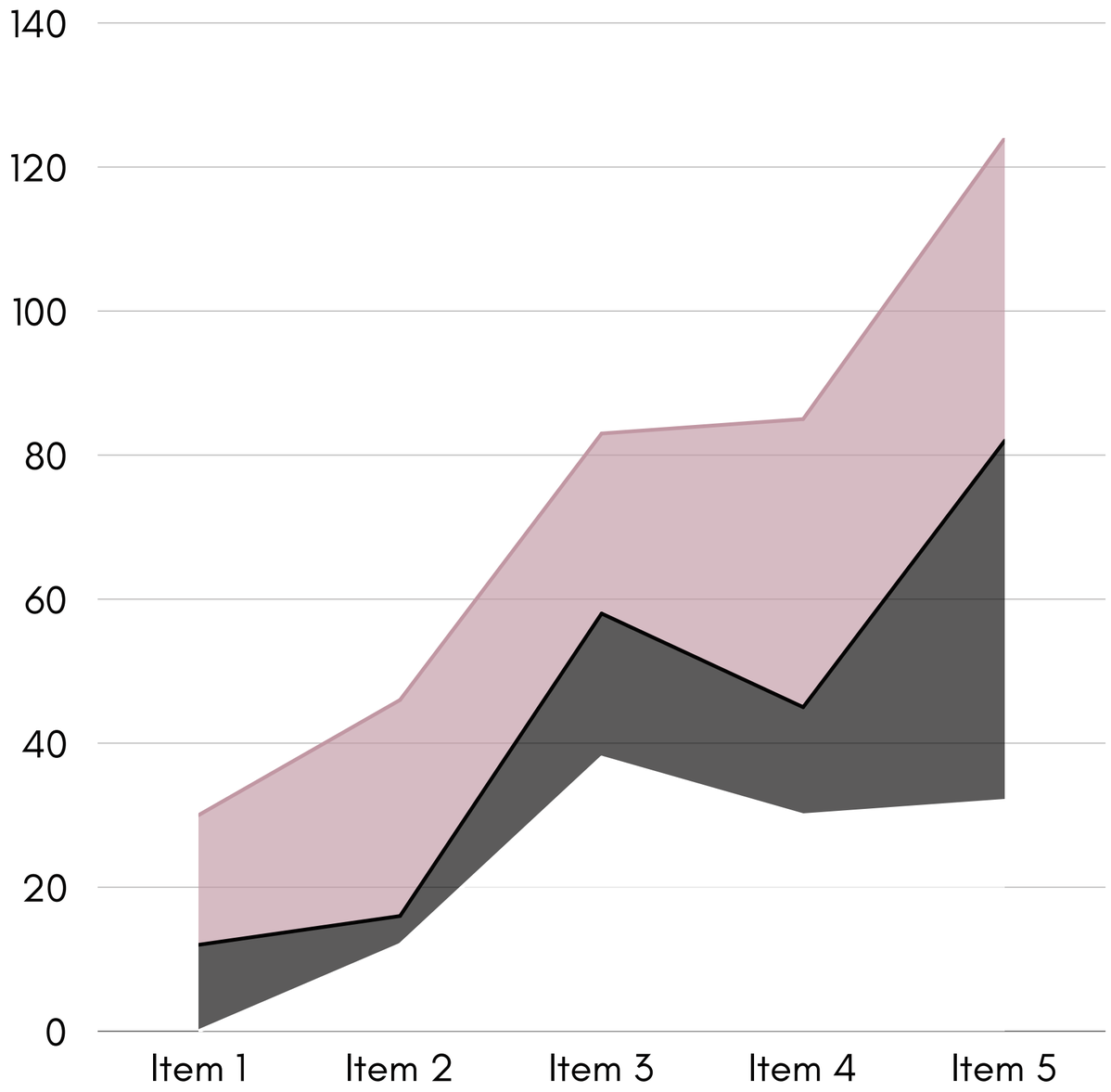
- California
- Texas
- Ohio
- Massachusetts



Black Living **in America**

Results

- We found the estimated Black population of each state for future years out to the year 2050
- Most important features:
  - number of minority-owned firms
  - median household income
  - population per square mile
- Wide variety of Black population projected percentages in America
  - Georgia is very high at 32%, and Montana at 0.6%
  - From our predicted polynomial regression model: These trends are likely to continue



just for aesthetic <3

Discussion

- The problem that we are investigating needs to be done in the context of the housing problem in America.
  - The most ethical implication is that where Black people **are** right now is not necessarily where they **want** to live.
- As a whole, data surrounding the Black experience in the world is hard to come by. More accurate and timely data needs to be crafted so that projects about the Black experience are easier to complete and more meaningful/validated.

Where Will *You* Live?