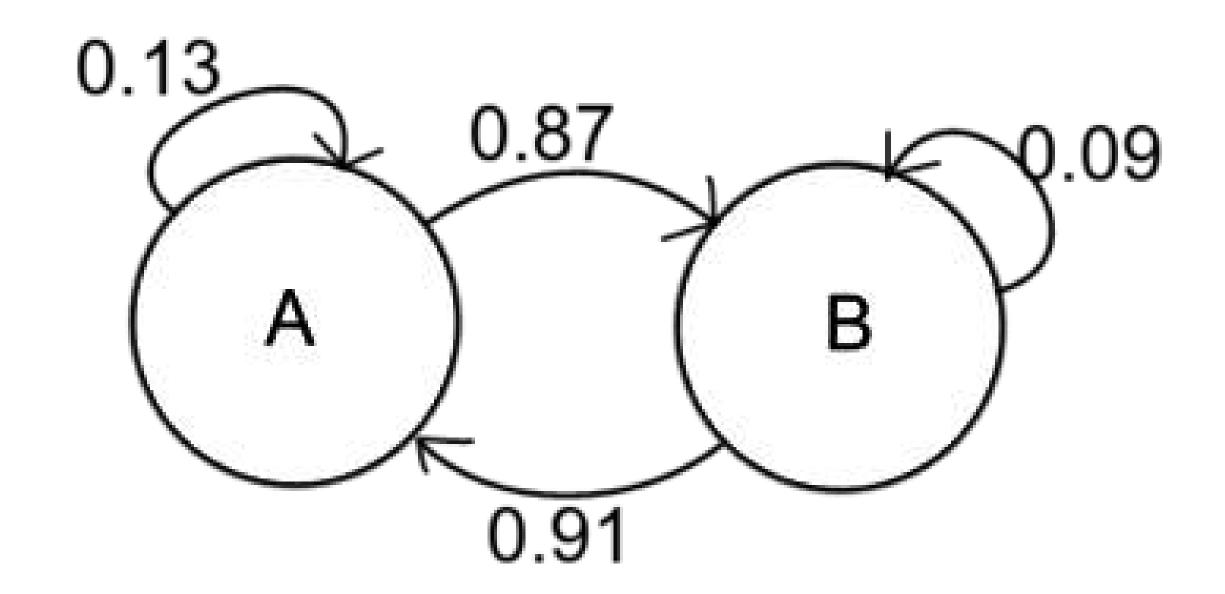
MScA 32010 Linear Algebra - Course Project

## Migration Flows in Cook County, IL from 2012 to 2016

\_\_ Markov Chains Analysis



Group Members:

Daniela Matinho, Tamer Abousoud Yuling Gu, Jenny Zhihan Wang

### AGENDA

O1 Executive Summary
Overview
Problem Statement

Assumptions

- O2 Markov Chains Analysis

  Approach and Proposed Solution

  Migration Evolution Result
- O3 Summary and Future of Work

PART 1

## **Executive Summary**

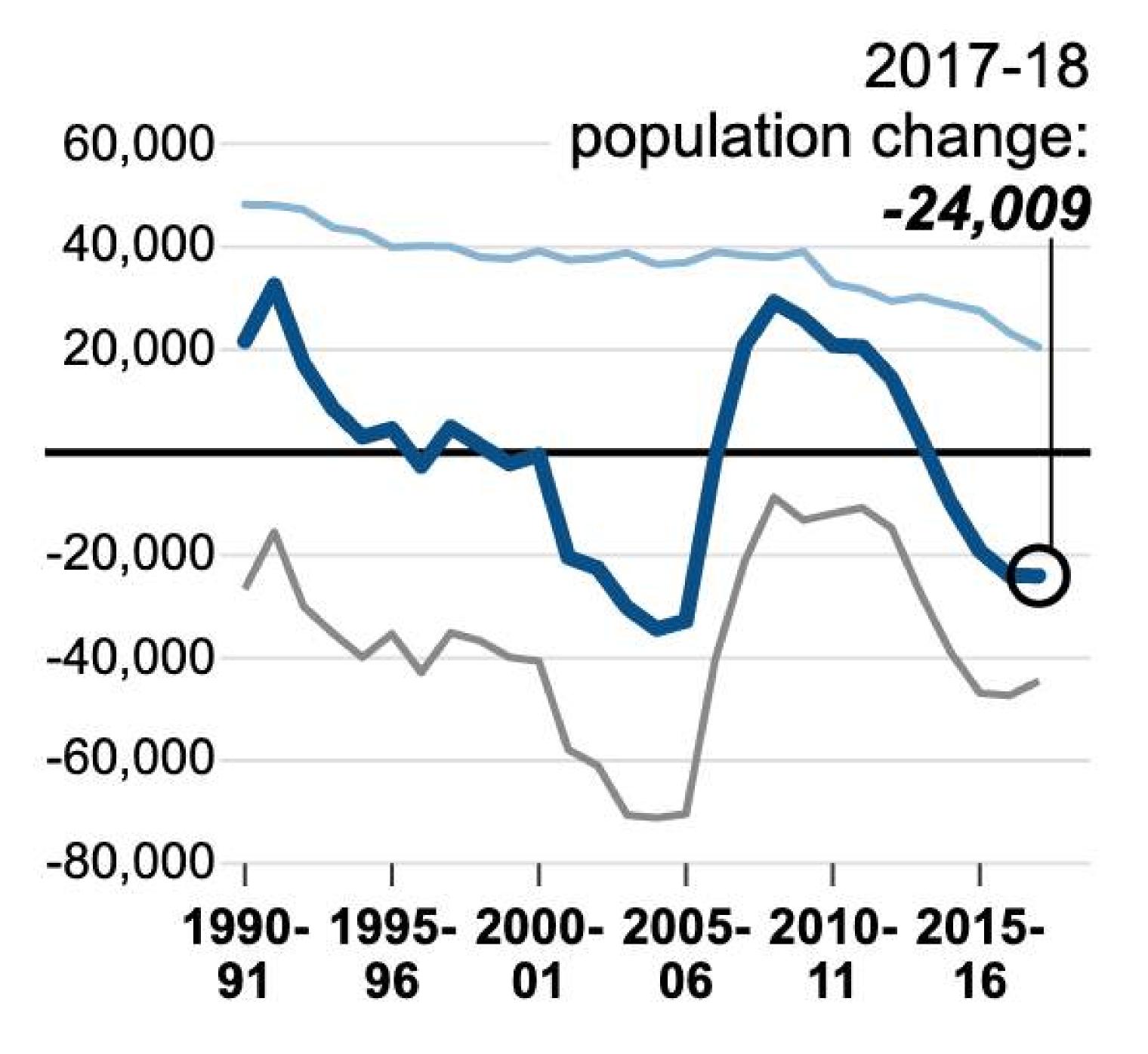
## Overview

Cook County is the 2<sup>nd</sup> most populous county in the U.S. after Los Angeles County. However, Cook County's net migration has been **negative for at least 27 years**, meaning more people moved away than moved to the area. The newest data put the current rate of net migration loss at 8.6 per 1,000 people.

#### Potential Reasons

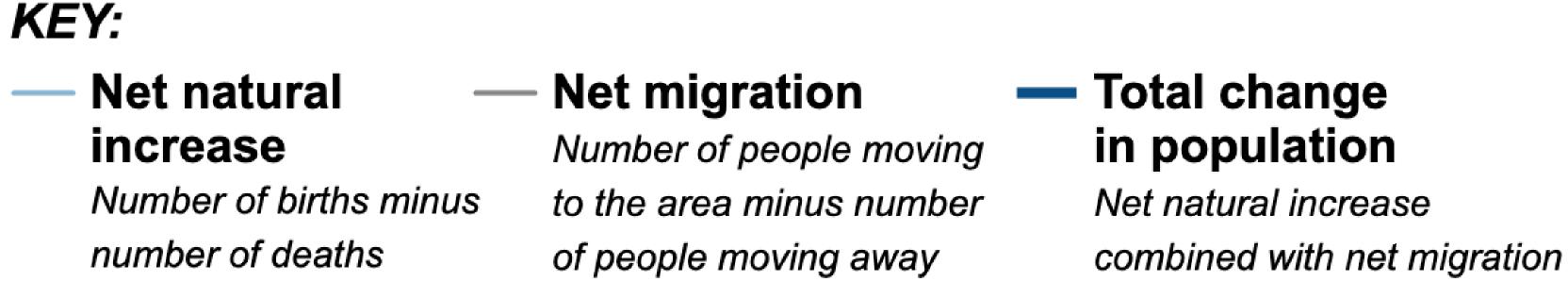
- Real Estate Increase
- Falling Birth Rate
- High Crime Rate
- Government Dysfunctional
- Taxes increase
- Severe Weather in Winter

#### Cook County



SOURCE: U.S. Census Bureau, Population Division

#### COMPONENTS OF POPULATION CHANGE From 1990 to 2018

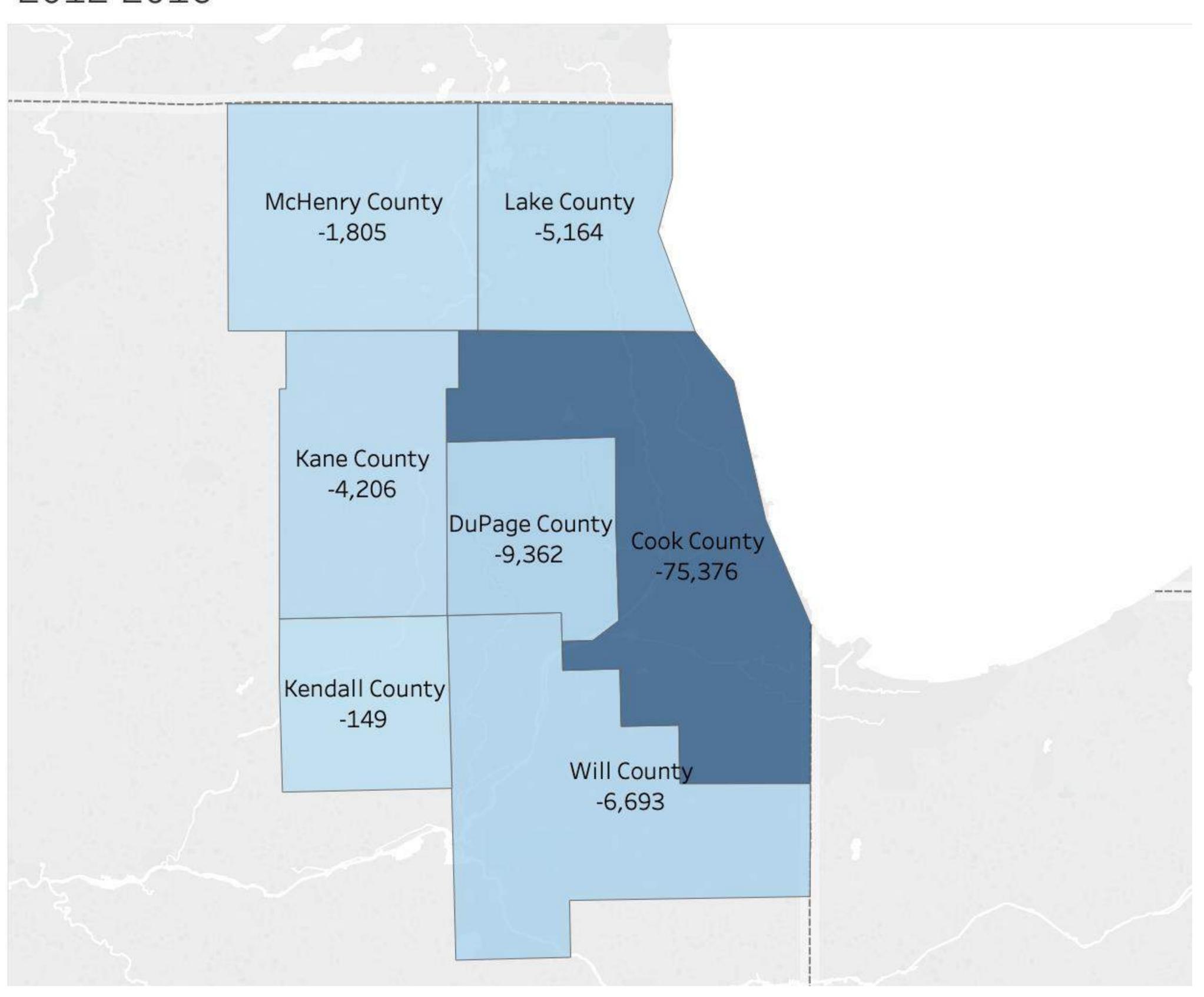


## Problem Statement

Can we visualize the migration pattern for the Cook County?

- Number of inflows and outflows per year
- Top origins and destinations

Cook County and Collar Counties Experienced Negative Net Out-Migration 2012-2016





## Project Overview

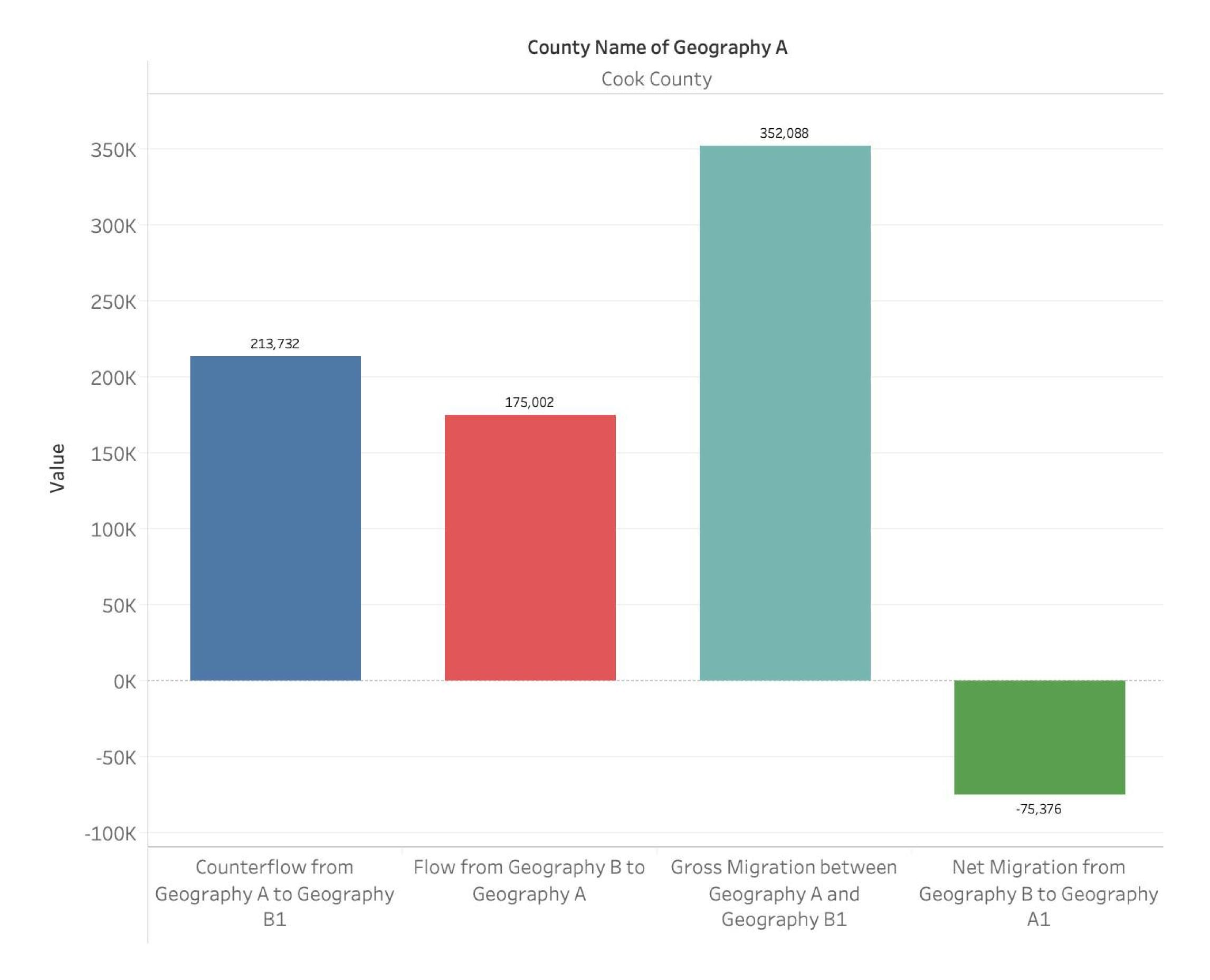
## Data Source

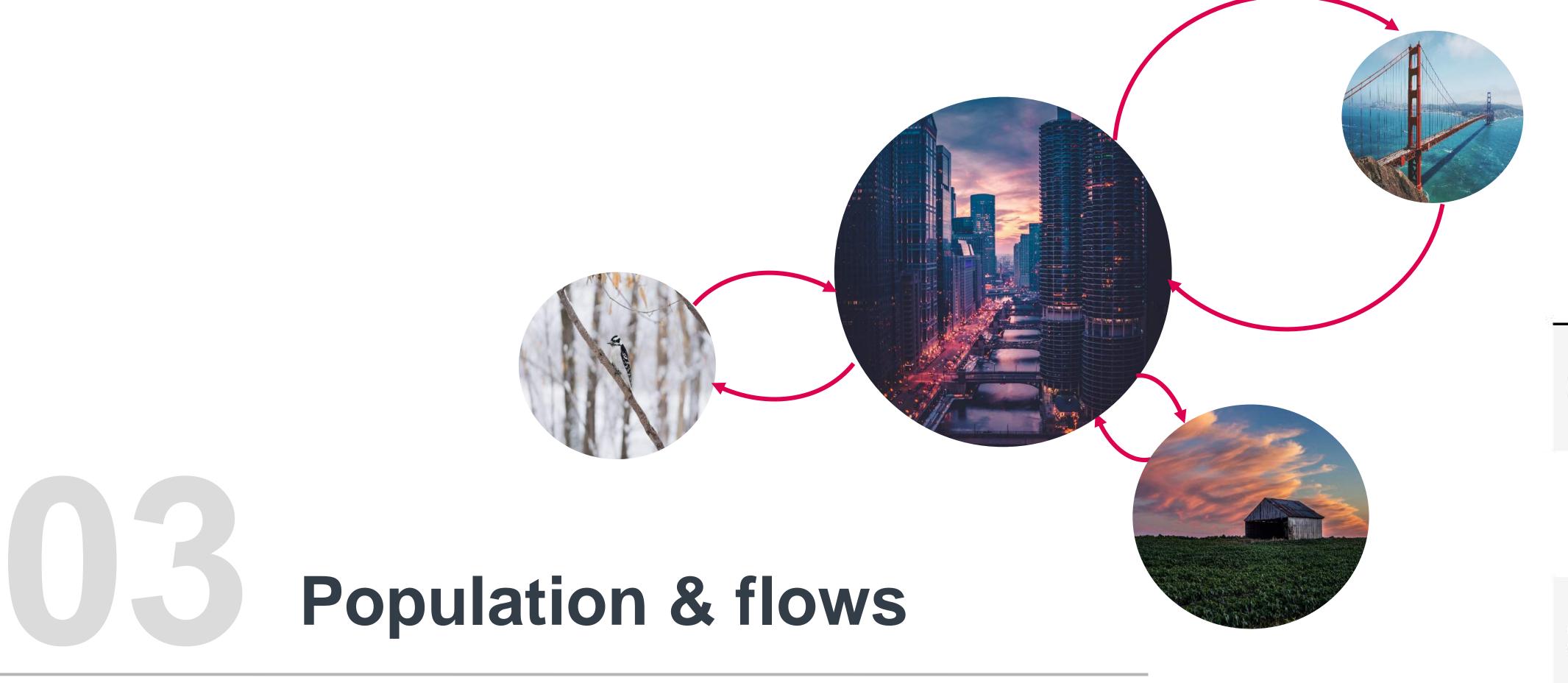
The project is based on migration flows in Cook County - Estimation from 2012 to 2016 – Data Source US Census Bureau.

The dataset only includes basic counts and no characteristics.

# Domestic County-to-County Migration Flows

The focus is on domestic migration because data from international migration flows is unavailable.





Merging the inflows and outflows of Cook County with the population data of each county. Top 80% of migration regions are listed, along with the remaining 20% summed up as All Other States.

550		County	Source or Destination	Population of Current County	Inbound Migration	Outbound Migration	Total Migration	Population of Dest County
	0	Cook County	Alabama	5162969	269	522	791	2441727
	1	Cook County	Alaska	5162969	531	227	758	431201
	2	Cook County	Arizona	5162969	1920	4052	5972	6144828
	3	Cook County	Arkansas	5162969	327	655	982	1148269
	4	Cook County	California	5162969	8040	12587	20627	35899876

	County	Source or Destination	Population of Current County	Inbound Migration	Outbound Migration	Total Migration	Population of Dest County	CumPercentage
0	Cook County	Illinois	5162969.0	51742	79186	130928	7095539	0.386884
1	Cook County	Indiana	5162969.0	7491	17688	25179	4956197	0.461286
2	Cook County	California	5162969.0	8040	12587	20627	35899876	0.522237
3	Cook County	Wisconsin	5162969.0	5450	9547	14997	4781577	0.566553
4	Cook County	Texas	5162969.0	4428	8901	13329	19658267	0.605939
5	Cook County	Florida	5162969.0	5363	7944 Cook Al	13307 ahama	17780368	0.645260

80%-20%- PART 1

PART 2

PART 3

**Executive Summary** 

**Markov Chain Analysis** 

Migration Evolution

# Markov Chain Analysis

# Assumptions

## Stability

Assume the population in Cook County is stable

## Migration flows

There are two possible migration stages taken into consideration: Inflow or Outflow from Cook County

County Migration

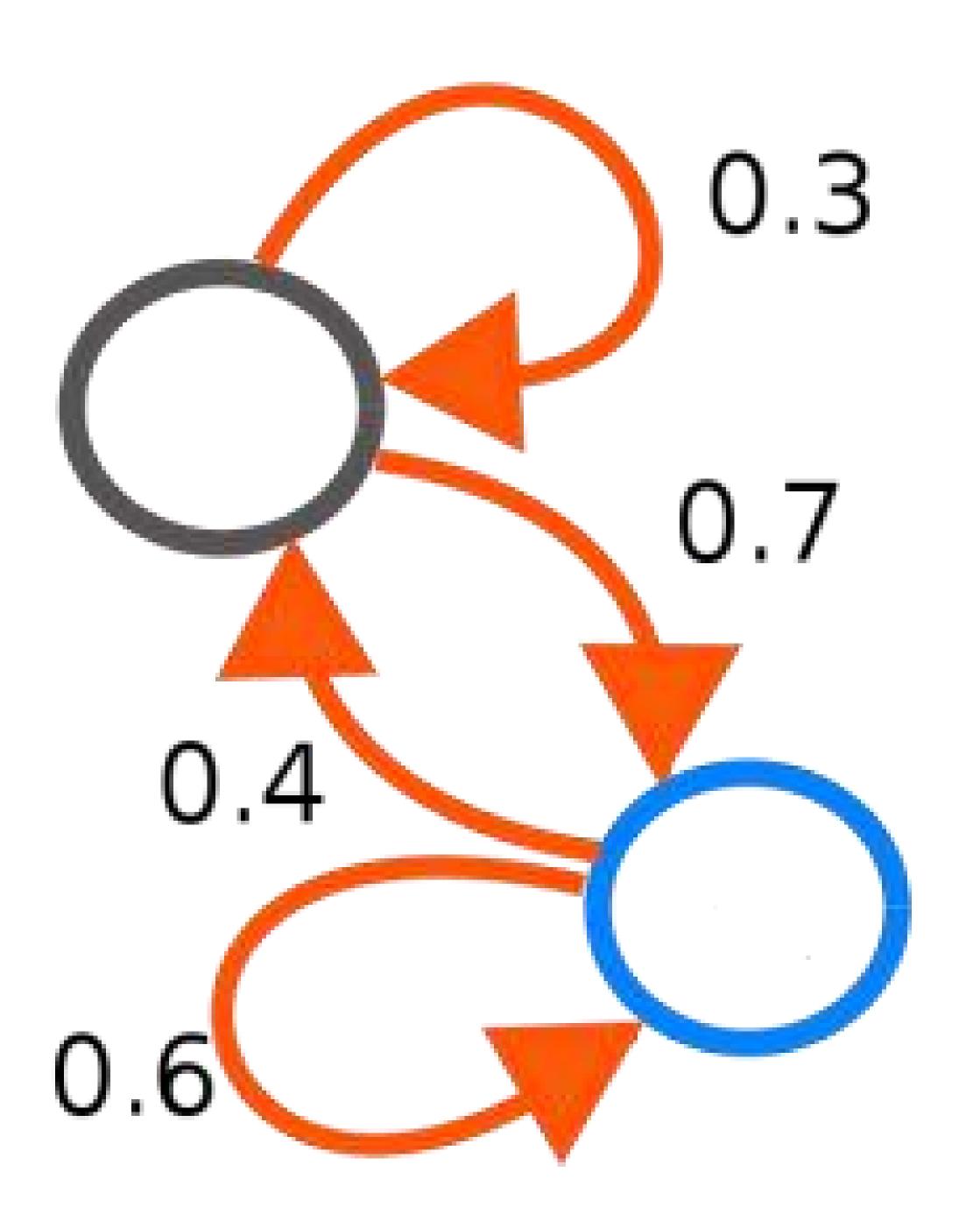
Analyze the migration flows (in and out) from US states to Cook County

## Introduction

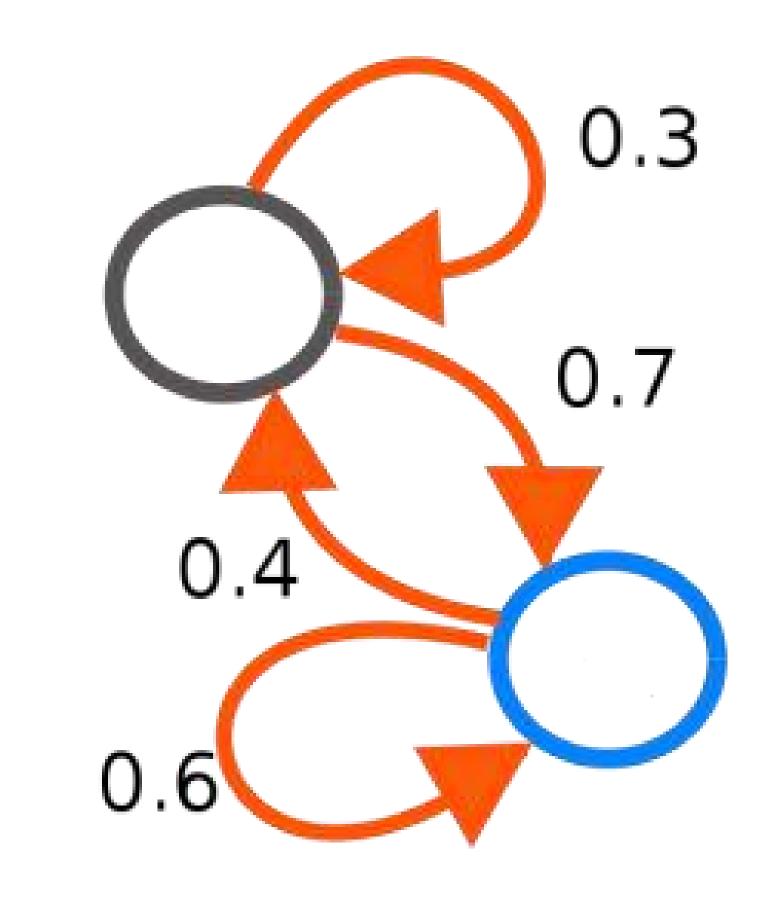
In this model, we will use a Markov chain to simulate the migration flows between Cook County and other locations. A simple chain would have the form on the right:

#### Recall: Markov Chain

Sequence of possible events in which the probability of each event is independent from the previous events – no memory from the past



Using Cook County as point "A", and any other location as "B", the result is a 2×2 Markov transition matrix for Cook County and *each* of the other locations in the data-frame with each column representing the probability of moving to the other location vs not moving to the other location:

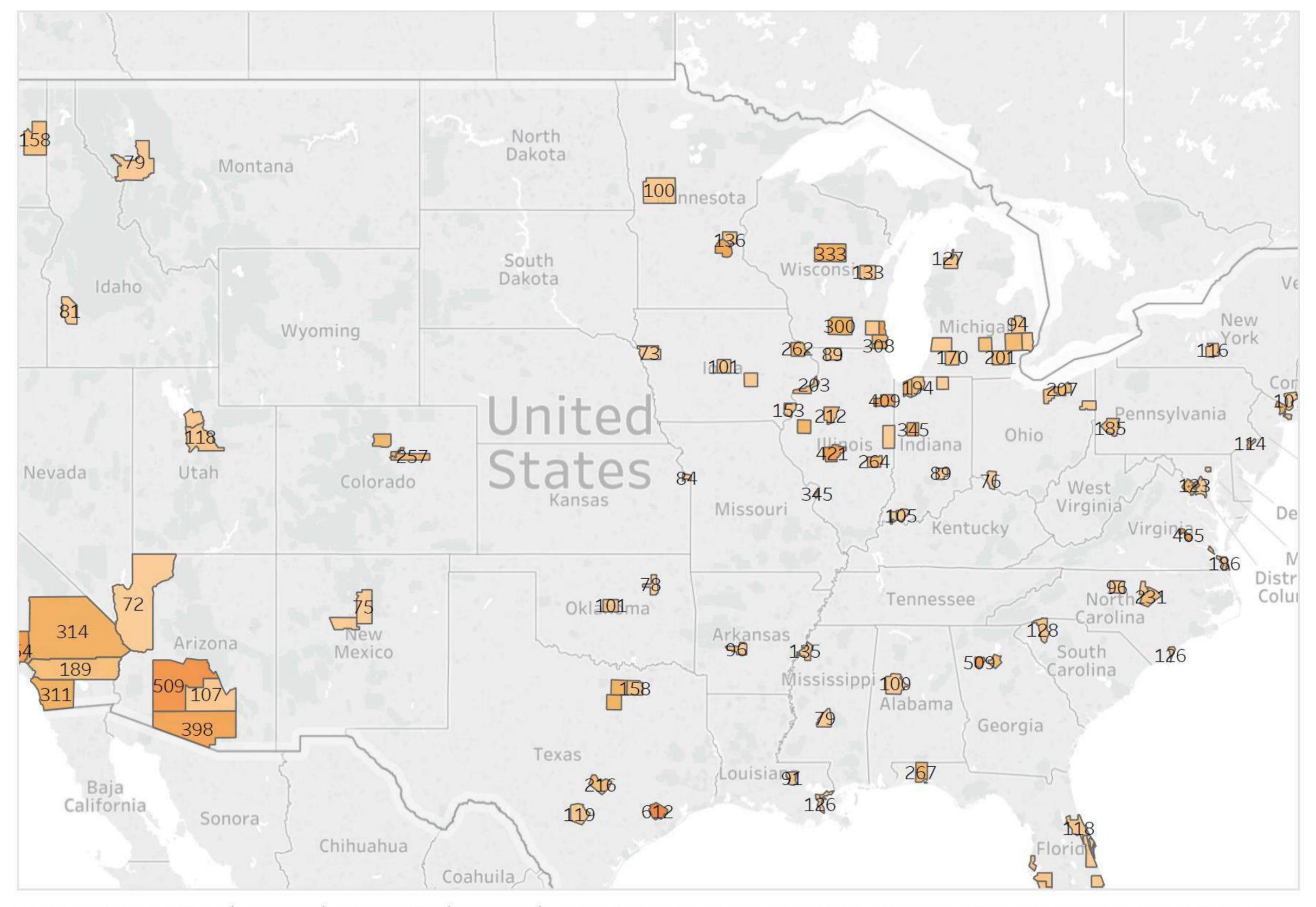


	Cook County	Other Location
Cook County	1 - P(Other Cook Co)	P(Cook Co Other)
Other Location	P(Other Cook Co)	$1 - P(Cook\ Co Other)$

Therefore: Population in n year = A<sup>n</sup> • P<sub>0</sub>

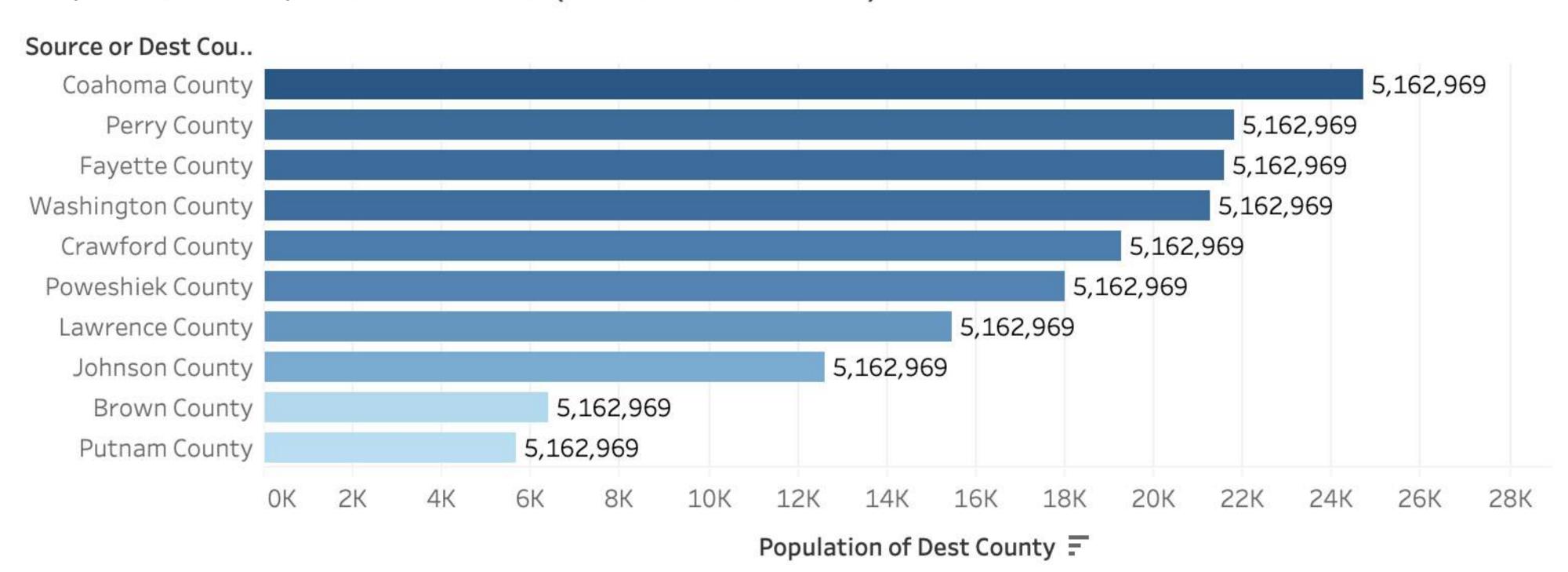
Pois the initial population

# Outbound Migration Geographic Allocation Snapshot

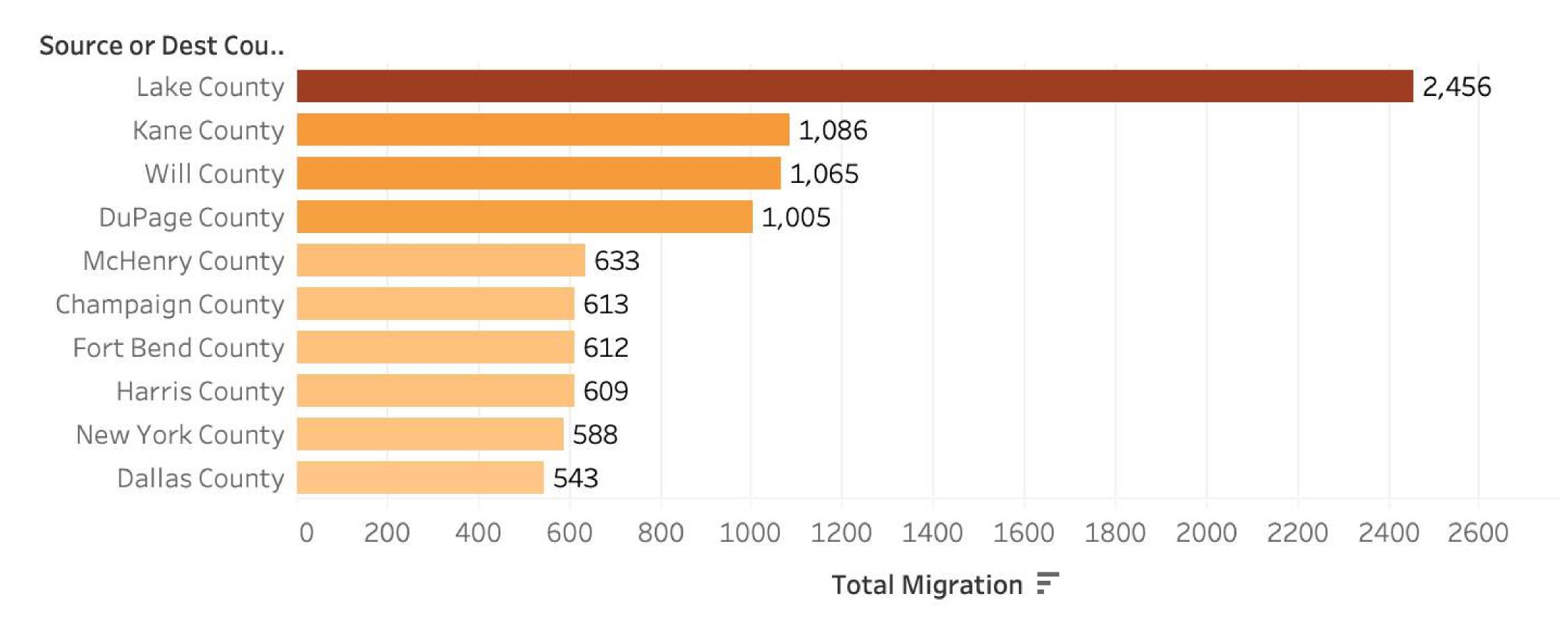


Map based on Longitude (generated) and Latitude (generated). Color shows sum of Total Migration. The marks are labeled by sum of Total Migration. Details are shown for Source or Dest County. The data is filtered on Total Migration, which ranges from 71 to 700 and keeps Null values.

#### Population Top 10 Counties (other than Cook)



#### Top 10 Outbound Migration Counties



$$A = SAS^{-1}$$

After setting up the Markov matrix, we can solve for the eigenvalues and eigenvectors and diagonalize our matrix:

$$\mathbf{A} = S \Lambda S^{-1}$$

It is now easy to calculate multiple years because  $\Lambda^n$  is simply the eigenvalues raised to the power of n.

The Python program for the above analysis is as show on the right:

1)

Define a function to build the transition matrix:

2)

Define a function to decompose the Markov matrix and apply the transformation:

```
def Transformation(A, n):
    eigVal_A, eigVec_A = LA.eig(A)
    LambdaMatrix = np.diag(np.real(eigVal_A))

SMatrix = eigVec_A
    SMatrixInv = LA.inv(SMatrix)
    A_transformed = SMatrix @ np.power(LambdaMatrix, n) @ SMatrixInv #1-year power 1, 5-ye

return(A_transformed)
```

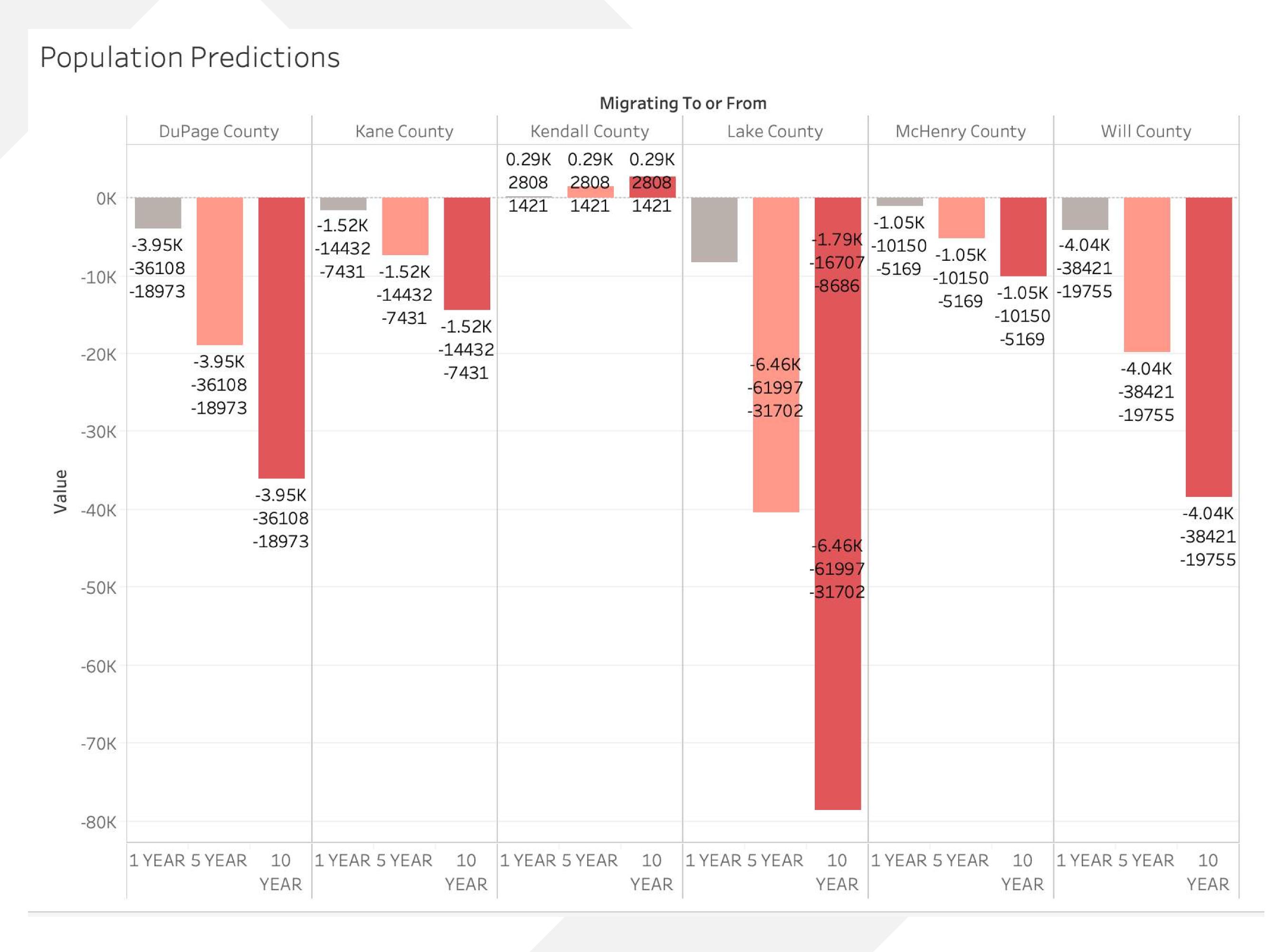
Run the predictions and add them to our new data frame PopulationPredictions:

skip code...

#### PARI 3

## Migration Evolution

#### Collar Counties Migration Predictions (in 1-5-10 years)



This table summarizes the net change from migrations to and from that could happen over **one**, **five** and **ten** years assuming the rate of inflows/outflows remains constant over these periods. Overall change in Cook County Population:

	Changes	New Population
1 Year	-63,479	5,099,490
5 Year	-313,657	4,849,312
10 Year	-618,307	4,544,662

# Migration Predictions on State Level (in 1-5-10 years)

# The top 5 states that have the biggest effect on the population of Cook County (in 10 years)

Net	Population	Change

1 Year 10 Years 5 Years

#### (Migrating To or From, State)

Alaska	261.0	2592.0	1301.0
All Other States	-4181.0	-40470.0	-20604.0
Arizona	-1919.0	-19115.0	-9579.0
California	-4420.0	-44066.0	-22069.0
Colorado	-1987.0	-19827.0	-9926.0

**Net Population Change** 

10 Years

#### (Migrating To or From, State)

Illinois	-234822.0
Indiana	-93191.0
California	-44066.0
Texas	-43062.0
All Other States	-40470.0

# Population distribution for Cook County

	Cook County	Illinois other than Cook County	Indiana	California	•••••	Other States
Cook County	1-P(Other Cook Co)	Inflow Probability	Inflow Probability	Inflow Probability		Inflow Probability
Illinois other than Cook County	Inflow Probability	1-Outflow Probability	0	0		0
Indiana	Inflow Probability	0	1-Outflow Probability	0		0
California	Inflow Probability	0	0	1-Outflow Probability		0
	*****					
Other States	Inflow Probability	0	0	0		1-Outflow Probability

Assume peoope from other state in cook county all has same distribution at the initial state

Initial Stage of the population distribution

# Distribution of the population in % (1-5-10-Infinite)

	Population Distribution	1 Year	5 Years	10 Years	Infinite
0	Original Cook County	92.371707	78.936705	65.089109	1.554856
1	Illinois other than Cook County	1.771245	7.006800	12.281032	3.270253
2	Indiana	0.628199	1.815938	3.051530	3.524344
3	California	0.533631	1.382752	2.272584	16.925826
4	Wisconsin	0.476767	1.118088	1.786540	2.522507
5	Texas	0.465016	1.065397	1.694542	11.900553
6	Florida	0.447179	0.982780	1.543782	7.931639
7	Michigan	0.418496	0.849524	1.299648	2.441819
8	New York	0.392935	0.731936	1.086989	5.026690
9	Ohio	0.383435	0.687814	1.006182	2.675072
10	Arizona	0.374726	0.647723	0.933612	3.905428
11	Iowa	0.362619	0.591045	0.828754	0.809298
12	Georgia	0.354487	0.553958	0.762459	1.544683
13	All Other States	1.019560	3.629540	6.363237	35.967031



## Conclusions

How Linear Algebra Drive the Project and Research



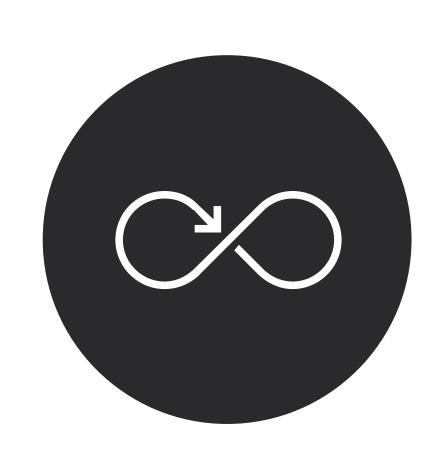
#### Markov Chain Analysis

Envisions the population evolution on the county-to-county level.



#### Simplify the Calculation Process

By implementing linear algebra approaches, the process of understanding the dataset and solve our problem statements becomes more efficient through Python.



#### Other States Growth

By implementing the same methodology used in this project, we can migrate our research to the State-to-state level.

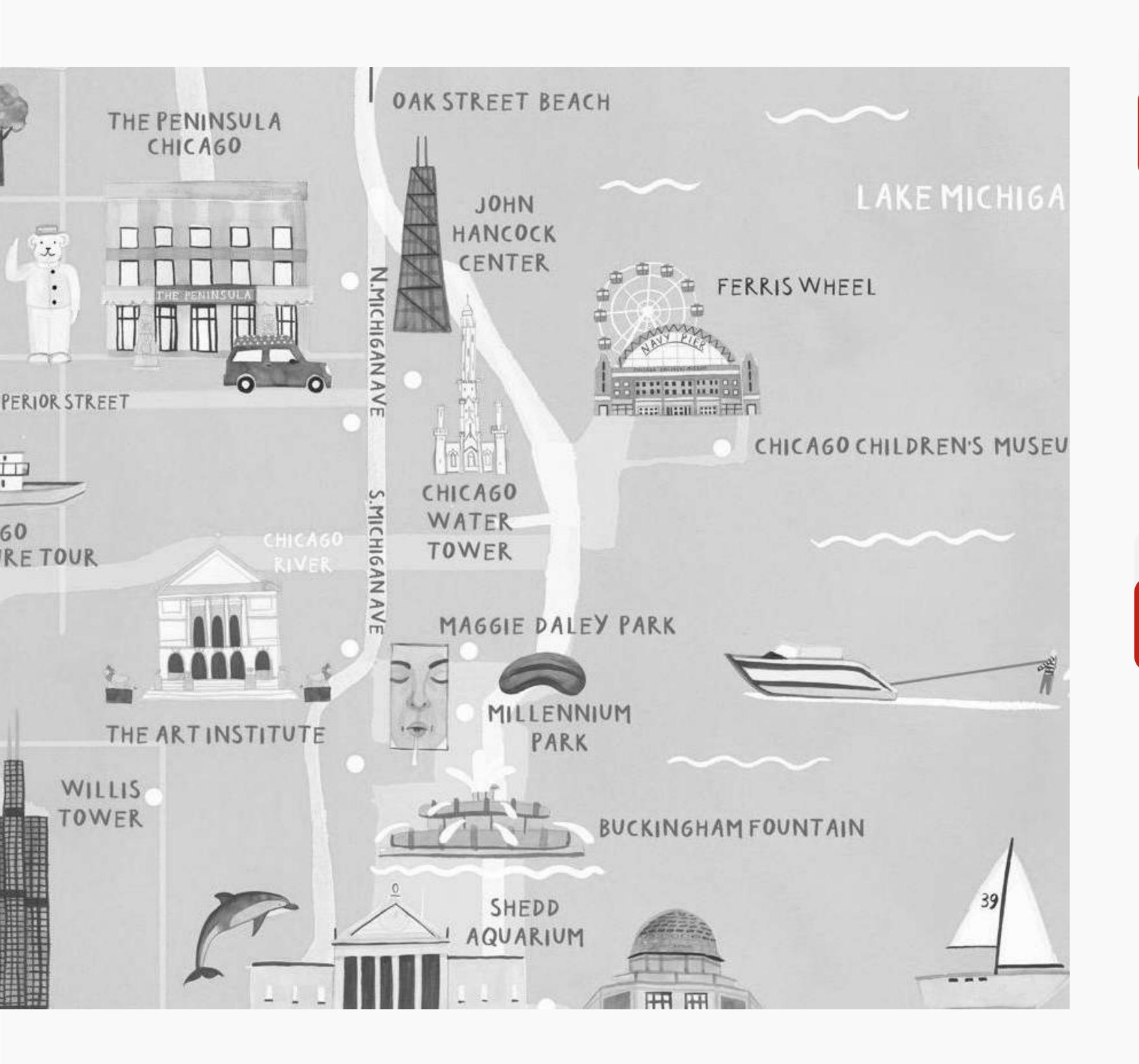


## Deepen the Understanding of 32010 Course Material

Strength the linear algebra knowledge by implementing the methodology and leveraging the power of Python in the group project.

## Future Works

#### — proposals based on the assumptions and business needs



#### **Further Back Test**

Use older datasets (e.g 2000-2011) to back test the project outcome and explain the differences.

## Using the Result from the Utility Function to drive insights

E.g. Based on the migration evolution trend, Chicago can propose new policy to encourage local business hiring out of state talents to relocate.

#### **Utility Function**

Use PCA approach to research and create dataset to find the key components that impact the population migration.

## Measure Population Inflation Impacts to Retail Business

E.g. Forecast the production and needs to optimize inventory based on the dynamic population distribution.

## Project Contribution



#### Daniela Matinho

- Research and collect the project background information
- Presentation design
- Project data analysis and solution proposal
- Calculation validation





#### Tamer Abousoud

- Metadata cleaning and ingestion
- Research design with Python
- Markov chain Migration Evolution Analysis
- Python code enhancement





#### Yuling Gu

- Project research and data acquisition
- Python programming for project analytics
- Population evolution data analysis
- Linear Algebra calculation support





#### Jenny Zhihan Wang

- Python data visualization
- Tableau reports and graphic design
- Presentation design
- Future of work with Linear
   Algebra knowledge proposal



## Reference

Zamudio, M. I. (n.d.). Migration To Other States The Biggest Factor In Illinois' Population Decline. Retrieved May 29, 2019, from (link)

02 Data Source (<u>link</u>)

Cook count population (link)

Attached Jupyter Notebook with Python code for project data analysis

Cook\_County\_Immigration\_Flow\_with\_Markov\_Chains.html



