

Predicting E.U. GDP

Sarah Duncan and Hayley Chang



Introduction

We want to view **European data** and see how it **changes over time**. We looked at economical trends over time (GDP) and attempted to predict it using a variety of features. Originally we wanted global data but **narrowed our scope** to fit the deadline.

About our Data

Data Source

eurostat 

Data Collection

Collected by
European Union
statistical authorities.

Data Cleaning

- Merge Data
- Filter for E.U. countries
- Drop NaNs
- Limit year range
- Log transform

Observational Unit

Countries at
years over time
from 2011

Dependent Variable

Gross Domestic Product (GDP)

Input Features (9)

Consumer Price
Index

Housing Index

Immigrants

Population

Unemployment

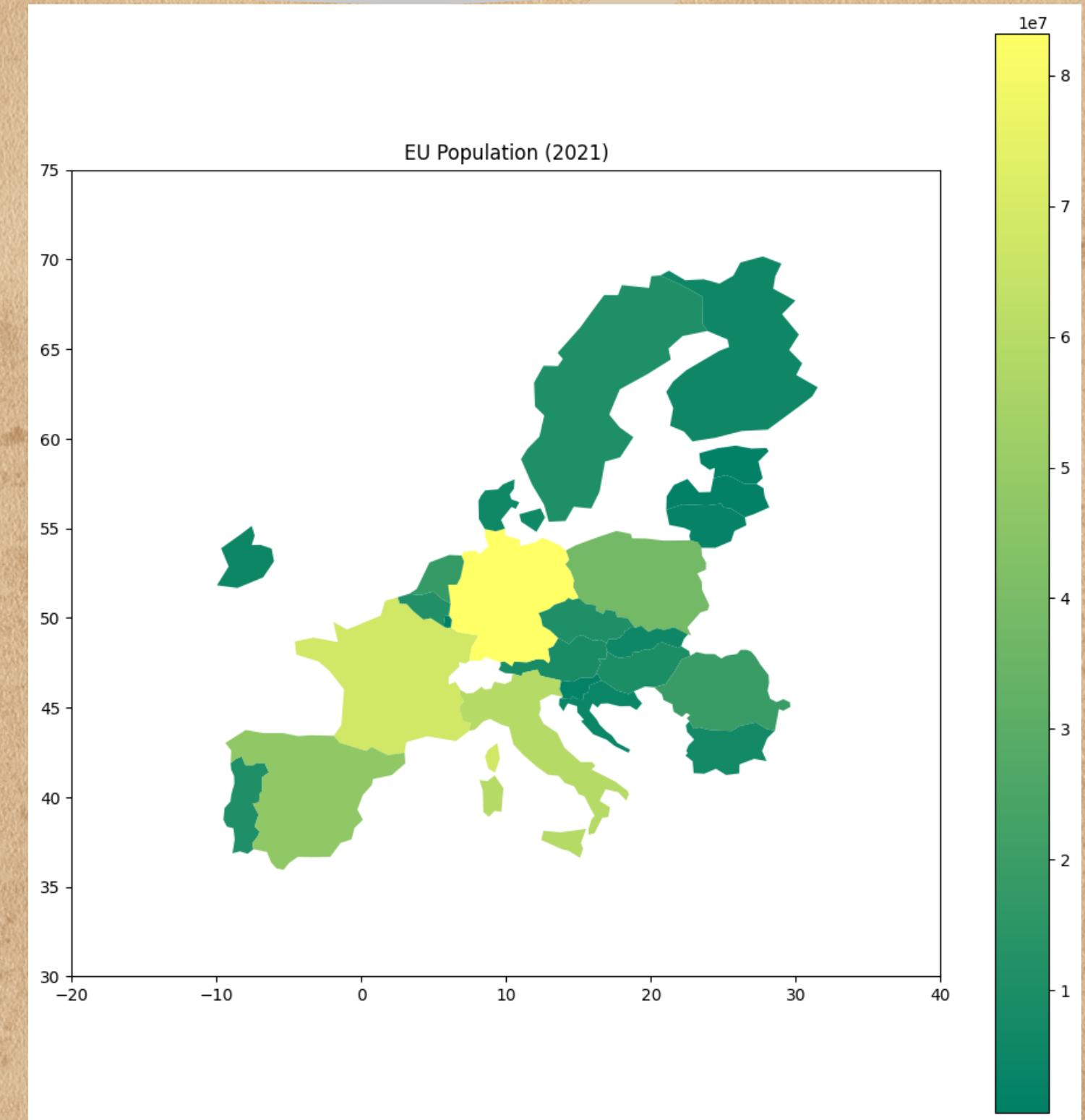
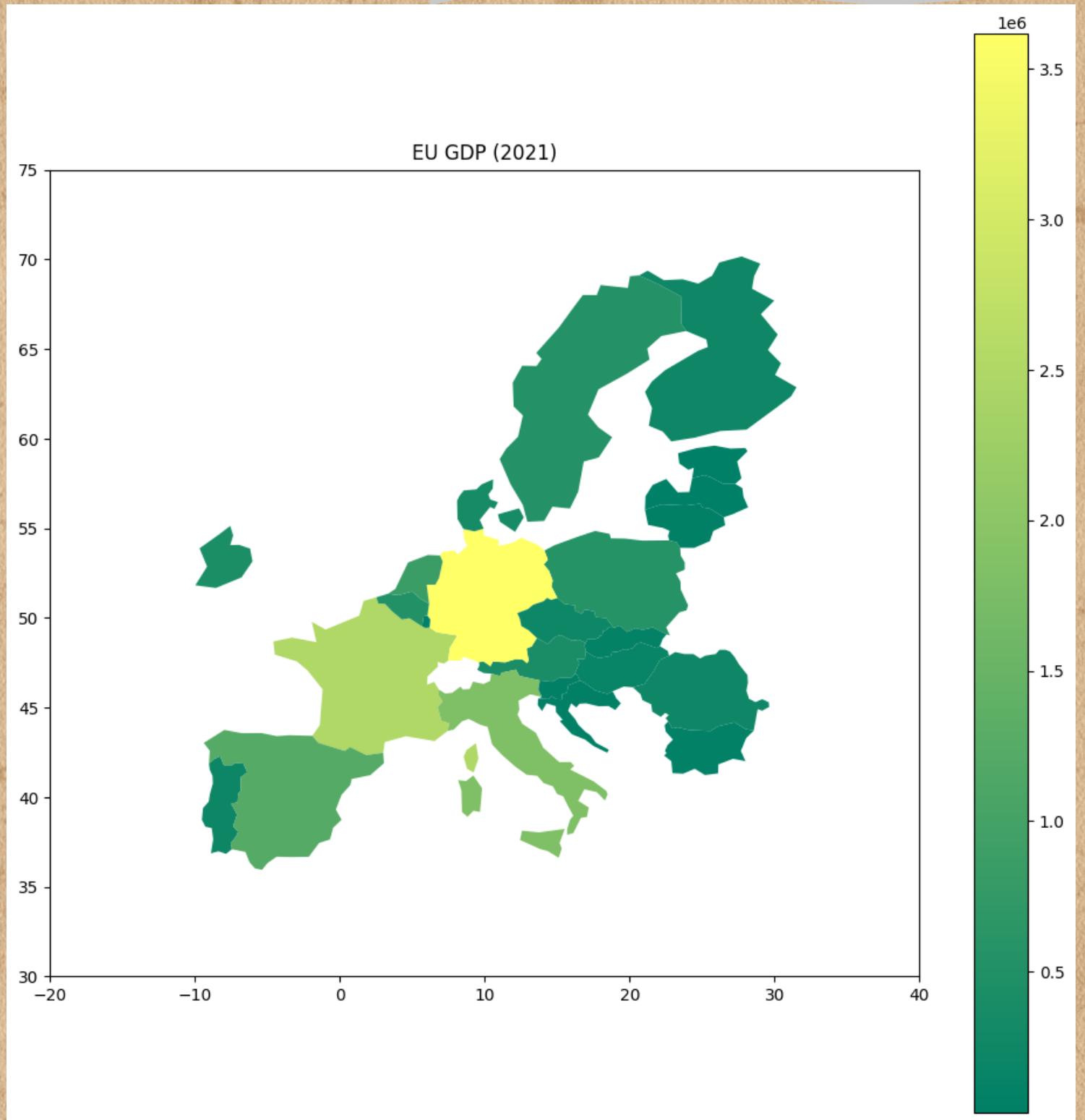
Emigrants

Imports

Total Deaths

Exports

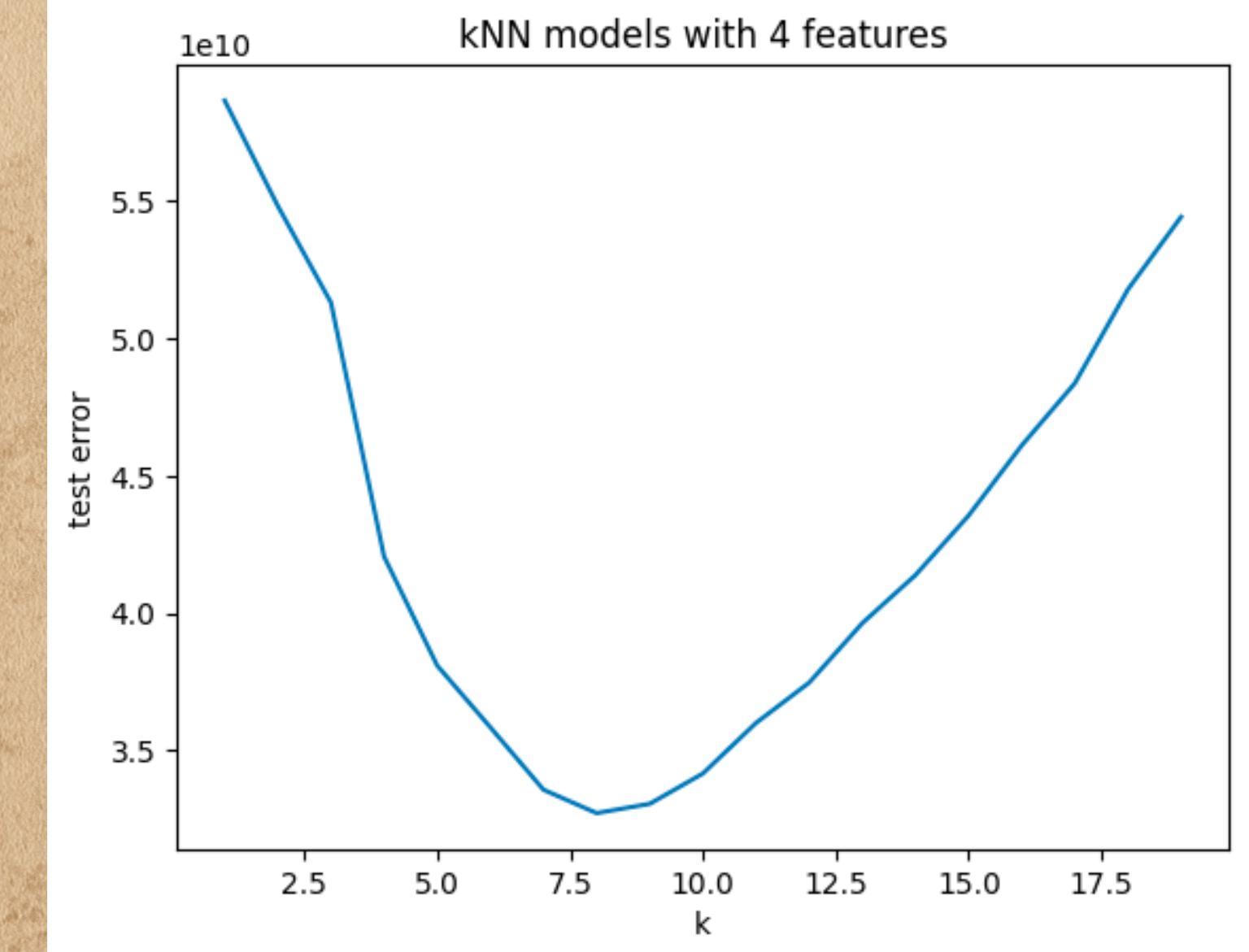
Target - GDP



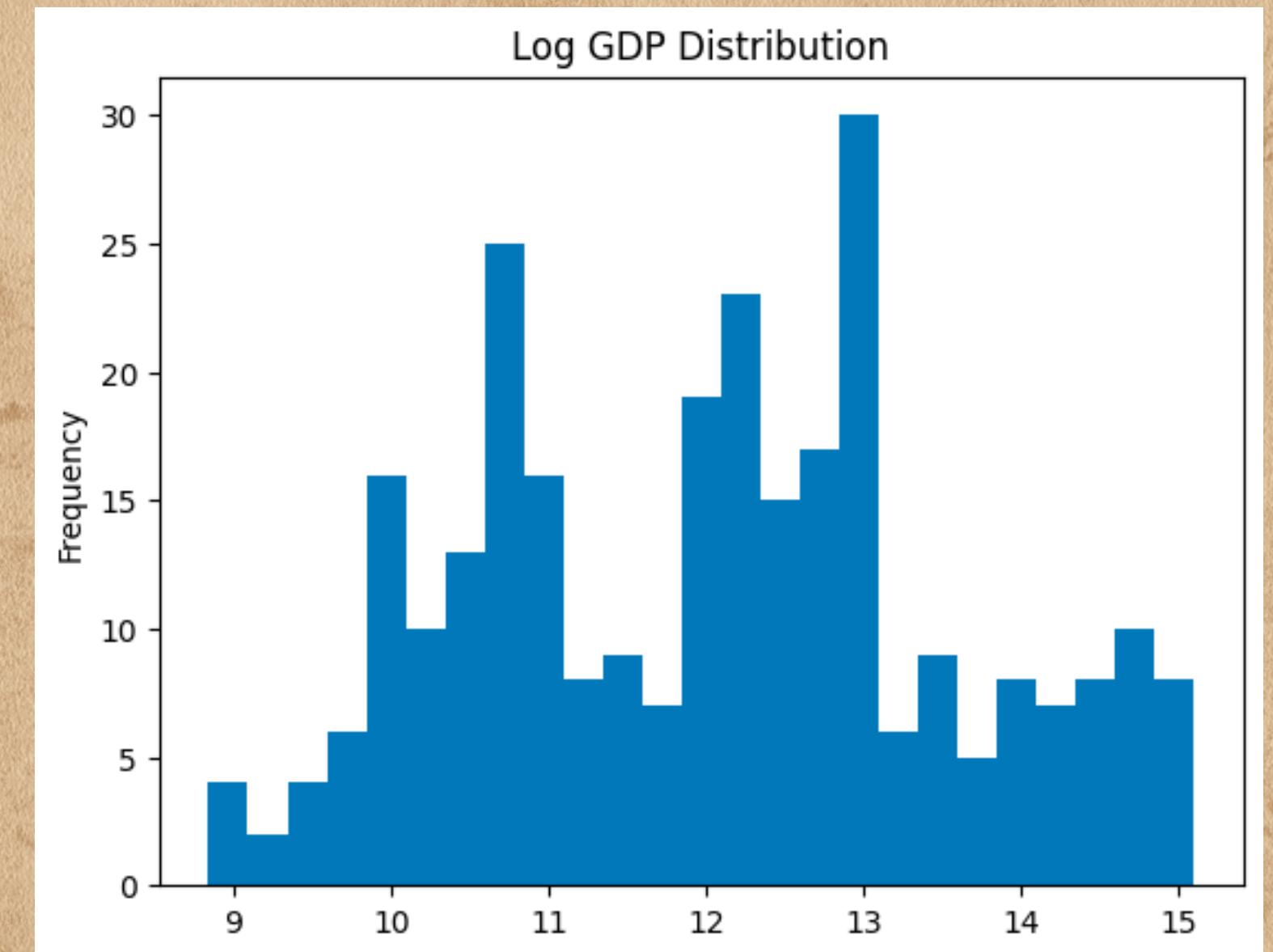
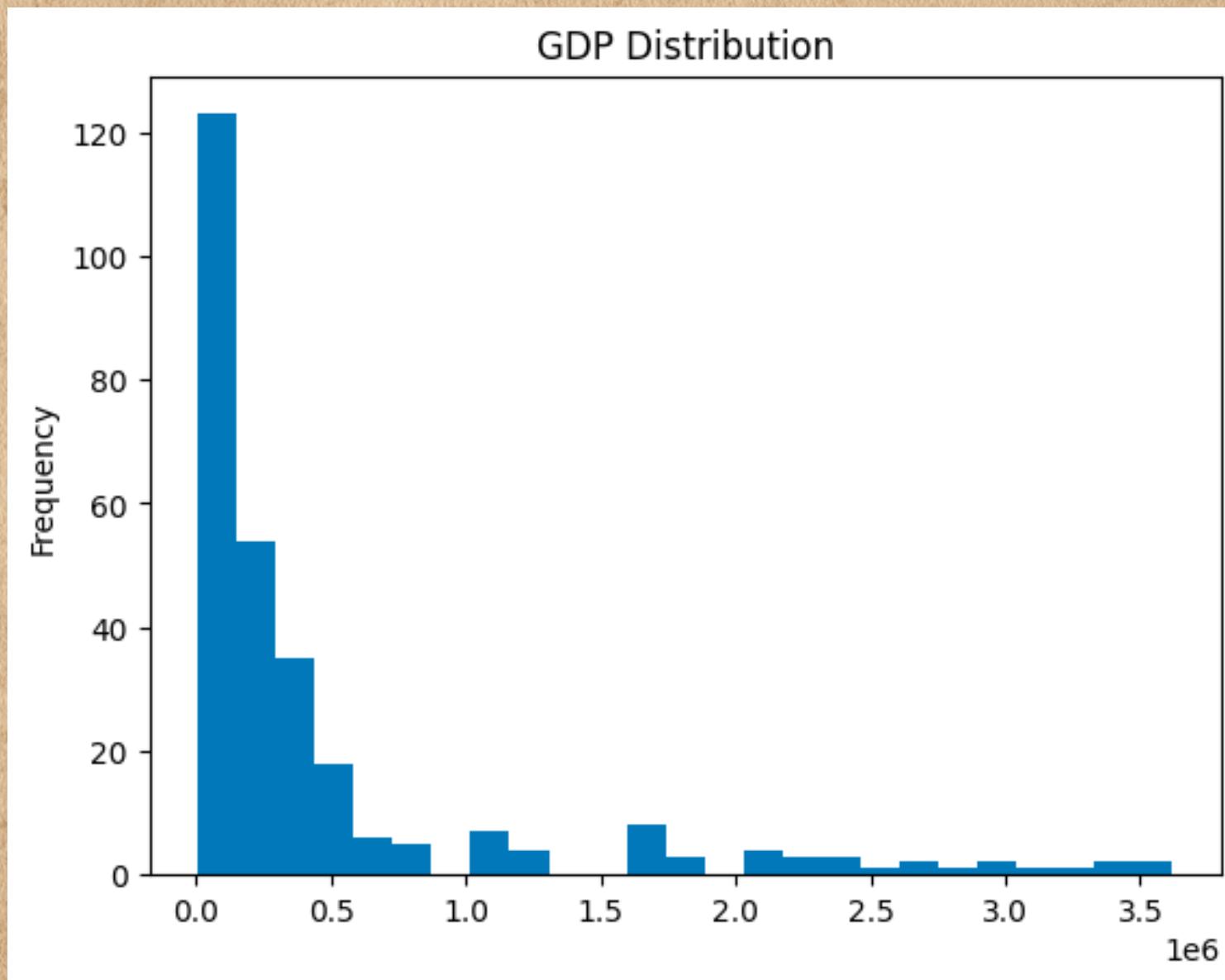
KNN Model

Best Features
['geo', 'Year', 'Population', 'Imports']

Best K = 8



Log Transforms



Linear Model

Best Features

- 'Year'
- 'CPI',
- 'log(Immigrants)'
- 'log(Population)'
- 'log(Unemployment)'
- 'log(Deaths)'
- 'log(Exports)'

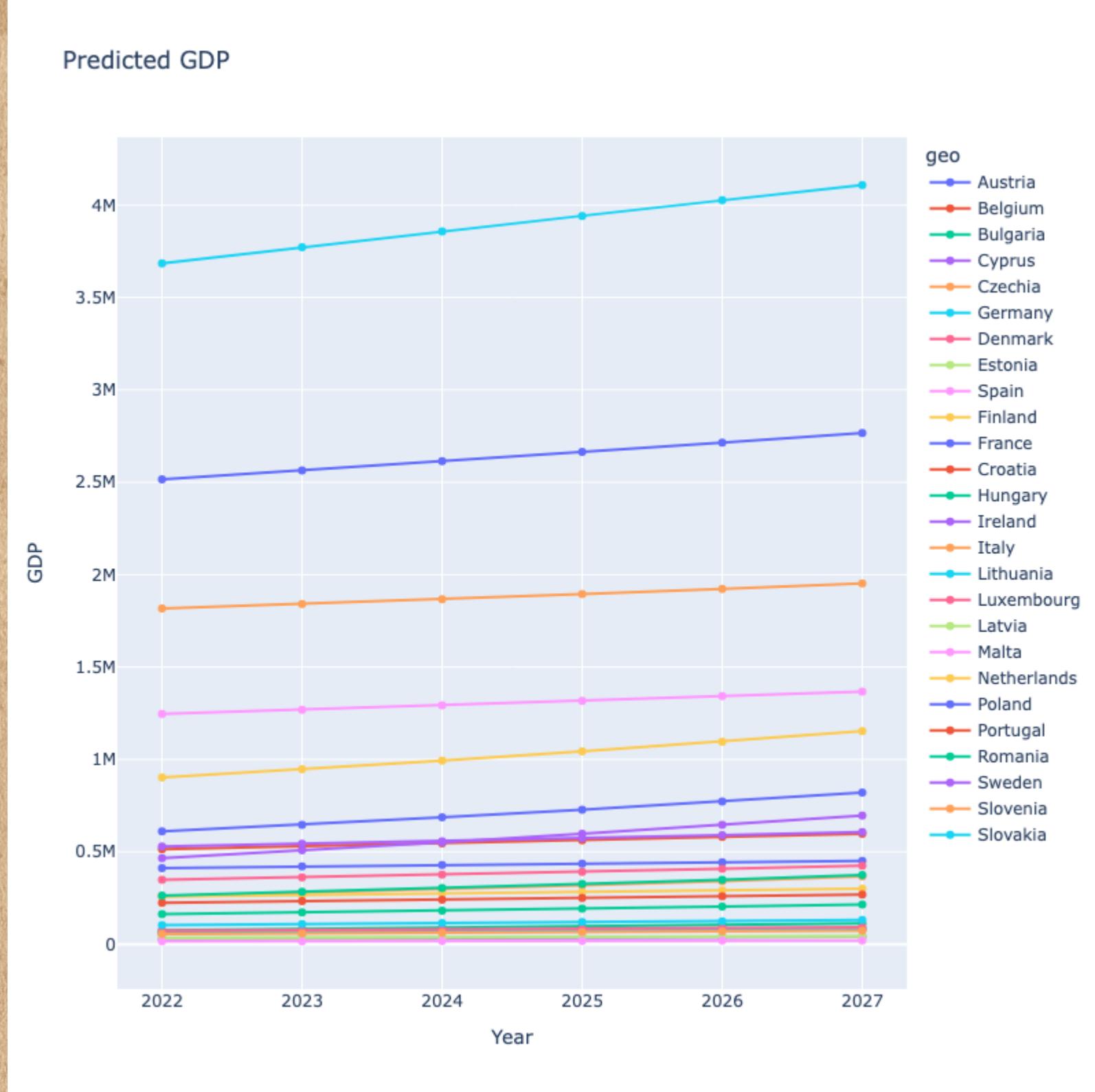
RMSE

Model Name	Base RMSE	Predicted RMSE - 2021
KNN Model	124842.114	145482.642
Linear Log Transform	0.274	0.264
Linear - No Transformation	42190.426	75970.98
Ensemble Model	42305.281	85751.335

Time Series

Fitting the Data

Polynomial Features
of Degree 1 and
Linear Regression



Linear Model Fit

The log transformed linear model predicted well!
RMSE = 0.251