SAMSUNG

Together for Tomorrow! Enabling People

Education for Future Generations

Forecasting The Future: Climate Change & Its impacts

Facilitator

Mohamed Elshaghnoby

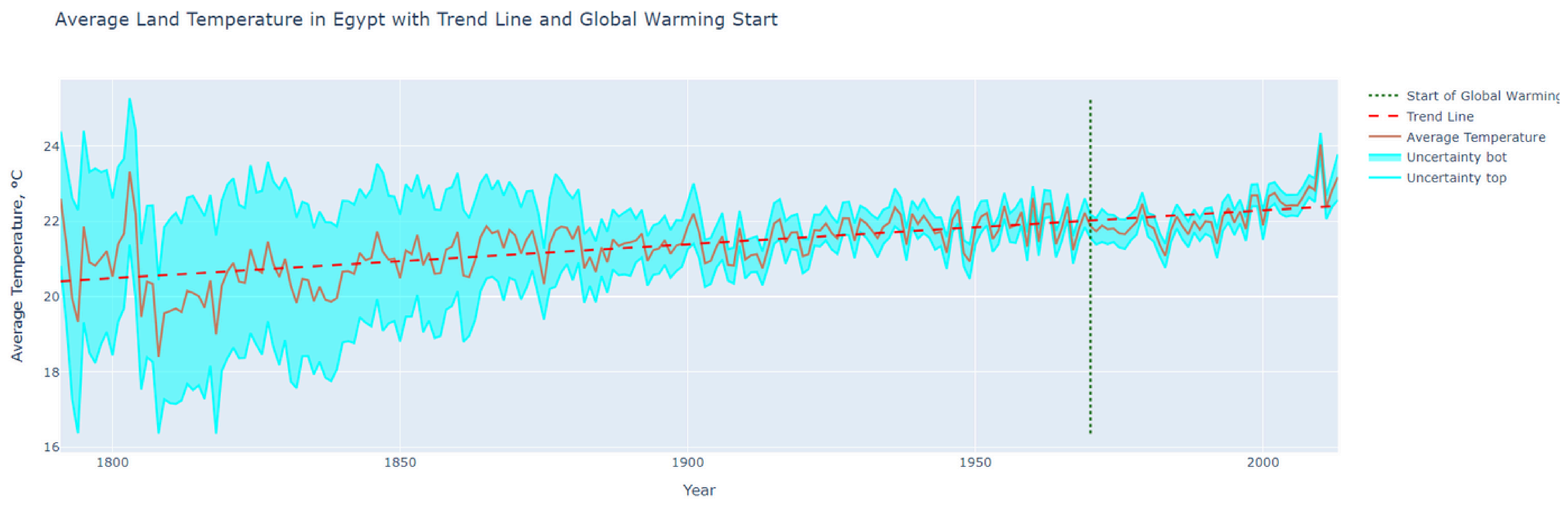
Team 8

- Eyad Mohamed
- Hossam Osama
- Shahd Saad

Agenda

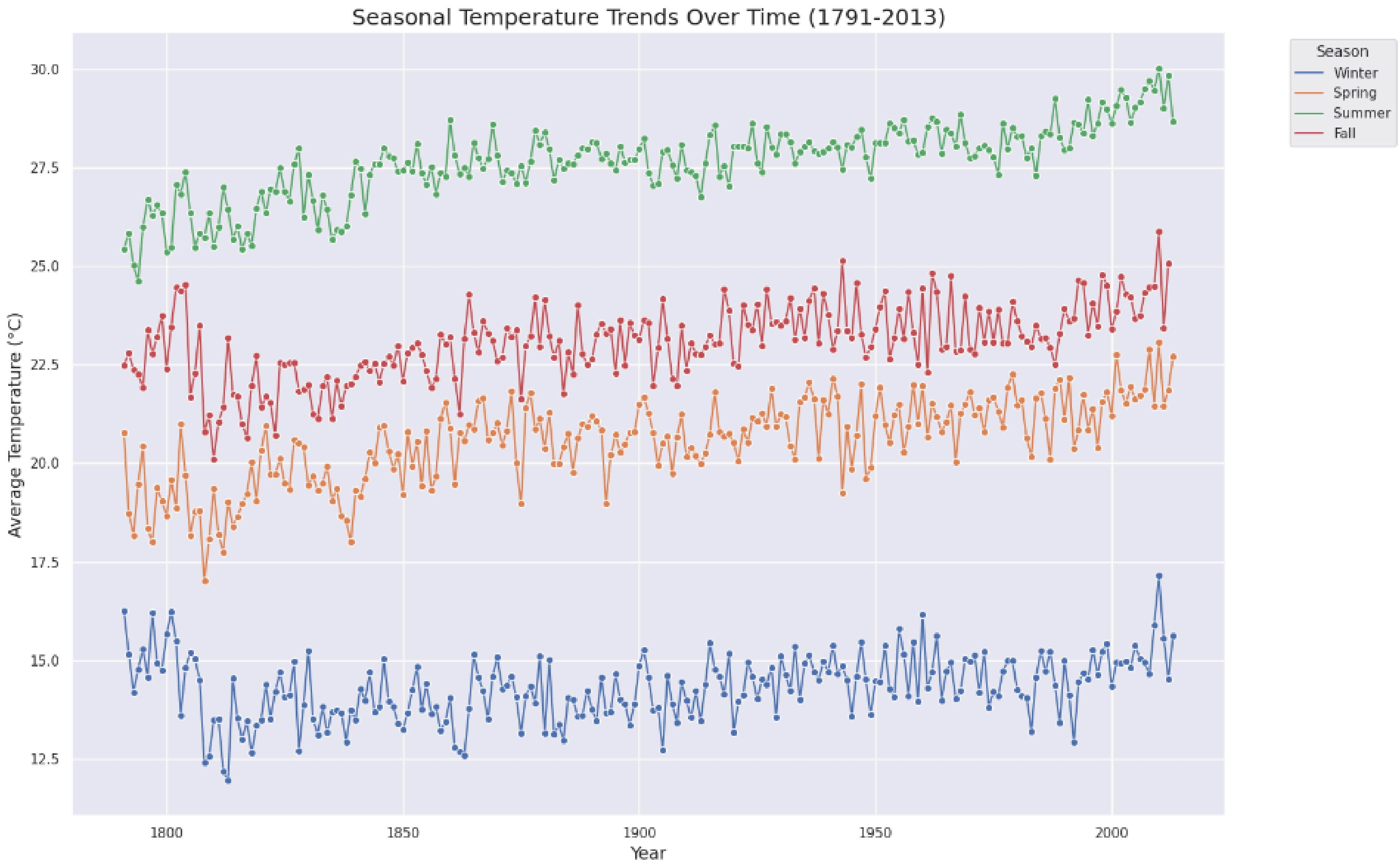
- Climate Change Impacts
  - Economics & Commerce
  - Agriculture
  - Health
- Data Source & Preprocessing
  - Handling Missing Values (NULLS)
  - Feature Engineering
  - Categorical Encoding
  - Handling Outliers
  - Data Transformation
- Predictive Modeling
  - Upsampling
  - Clustering
  - Regression model
- Q&A
- References

Tempreature effects:



The average temperature has been rising, with a notable increase around 1970, coinciding with the peak of the industrial revolution and the expansion of factories, which contributed significantly to global warming

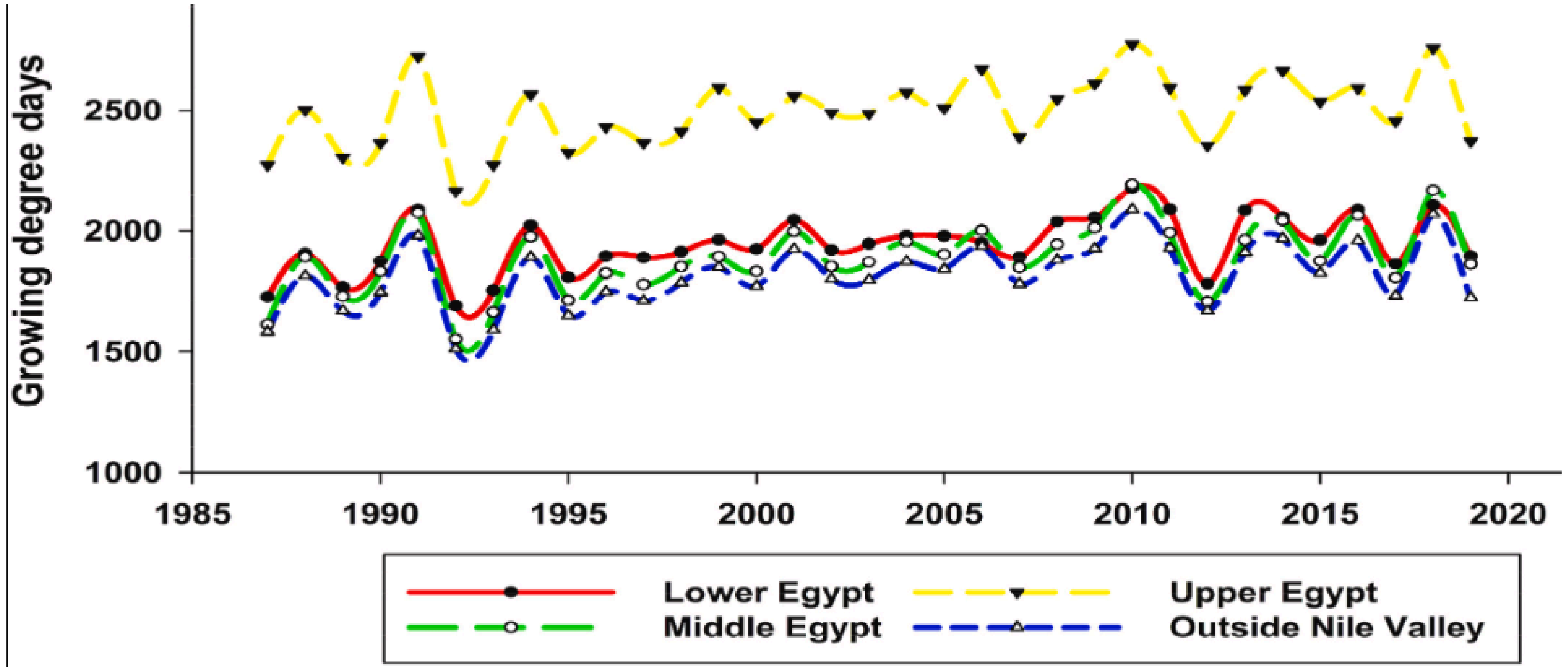
Same can be noticed with the seasonal analysis of the temperature increase

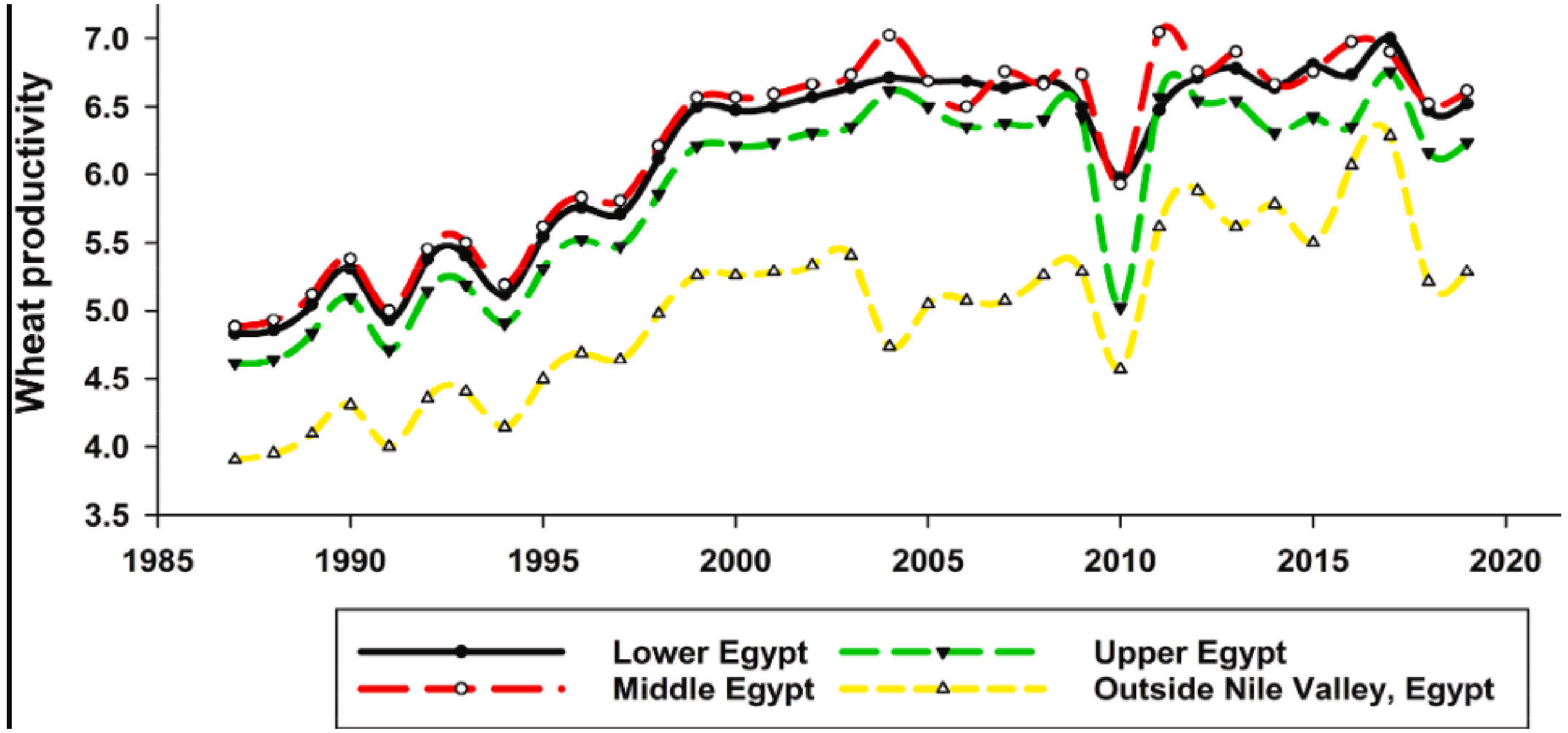


Climate Change effects on Economy

- Extreme heat events, which are projected to double by 2050 and increase seven-fold by 2100, will negatively impact soil moisture, leading to more sandstorms and dust storms, damaging agricultural productivity.
- Sea level rise and seawater intrusion into the Nile are already depleting the river's freshwater supply, while extreme temperatures and irregular precipitation are likely to increase surface water evaporation, worsening droughts and crop failure.
- Welfare losses in agriculture by 2060 are estimated between 40 to 234 billion EGP.
- Food prices could rise by 16–68%, impacting affordability and access to basic needs.

A decrease by 8-47% by 2060 is projected, and that also effect the employment potential by 39%





Health effects:

- Climate change causes 37% of global heat-related deaths, with a 70% increase in deaths among those over 65 in the last two decades.
- Rising particulate matter and heat stress in Egypt could lead to 2,000–5,000 additional deaths and 20–48 billion EGP in economic losses annually.
- WHO reports 2 billion people lack safe drinking water, and 600 million suffer from foodborne illnesses, with climate change worsening risks for children under 5.
- In 2020, 98 million more people faced food insecurity than the 1981–2010 average, worsening hunger in Africa and Asia.

Commerce effects:

- The Suez Canal handles 10–12% of global trade, contributing 2–3% to Egypt's GDP and providing vital foreign currency.
- Coastal inundation risks, especially in Port Said and the Suez Canal Container Port, could disrupt cargo operations due to rising sea levels.
- Extreme heat and wind events are projected to increase, affecting infrastructure, navigation, and port operations in the Suez Canal.
- Property at risk from sea-level rise in the Nile Delta is valued at 7–16 billion EGP, threatening Egypt's real estate and commercial sectors.

Water scarcity:

- With the Nile as Egypt's main water source, overreliance on it and upstream dam projects create vulnerability, necessitating more efficient water use and international cooperation.

Funding for adaptation:

- The high cost of implementing climate adaptation measures requires significant investment, and Egypt needs more international support to fund these initiatives.

Public awareness and capacity building:

- There is a need for greater awareness among farmers and citizens about climate risks, as well as training programs to implement climate-smart practices.

Technological advancements:

- Egypt needs access to modern technology and innovation to improve agricultural resilience and enhance early warning systems for extreme weather events.

Climate-resilient agriculture:

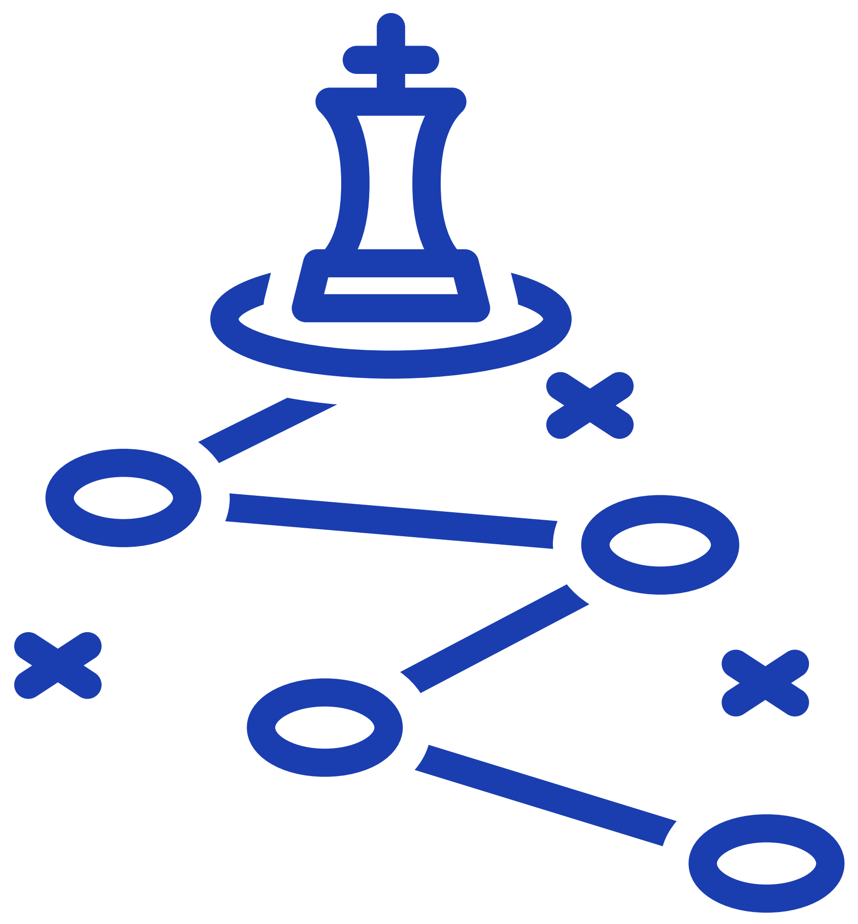
- The government promotes water-efficient irrigation techniques and drought-resistant crops to adapt to changing climate conditions in agriculture.

Coastal protection initiatives:

- Egypt is building sea walls and implementing coastal zone management strategies to protect vulnerable areas, such as the Nile Delta, from rising sea levels.

Sustainable transportation:

- Egypt is investing in public transportation infrastructure, like electric buses and expanding metro systems, to reduce greenhouse gas emissions from the transport sector.



Data Preprocessing

Feature Engineering

Date Time Features:

- Extracted time-based information (e.g., year, month) to capture seasonal and long-term climate trends.

Climate Regions Feature:

- Grouped cities into specific climate zones to analyze regional variations in temperature and other weather patterns.

Lag Features:

- Introduced previous time period data to detect trends and improve model predictions by incorporating historical climate information.

Handling Missing Values (NULLS)

Removing Missing Records:

- Eliminated rows with missing values to ensure data consistency, though it may reduce dataset size.

Statistical Imputation:

- Filled missing values using simple statistics like the mean or median to maintain data completeness.

KNN Imputation:

- Used the nearest neighbors' values to replace missing data based on similarity to other observations.

Iterative Imputation:

- Repeatedly predicted missing values using a model, updating estimates iteratively for more accurate imputation.

Categorical Encoding

- Season
- City
- Climate Region

Handling Outliers

- Removing outliers data points
- Quantile-based Flooring and Capping
- Mean/Median Imputation
- Not Handling Outliers

Dropping Columns

Country:

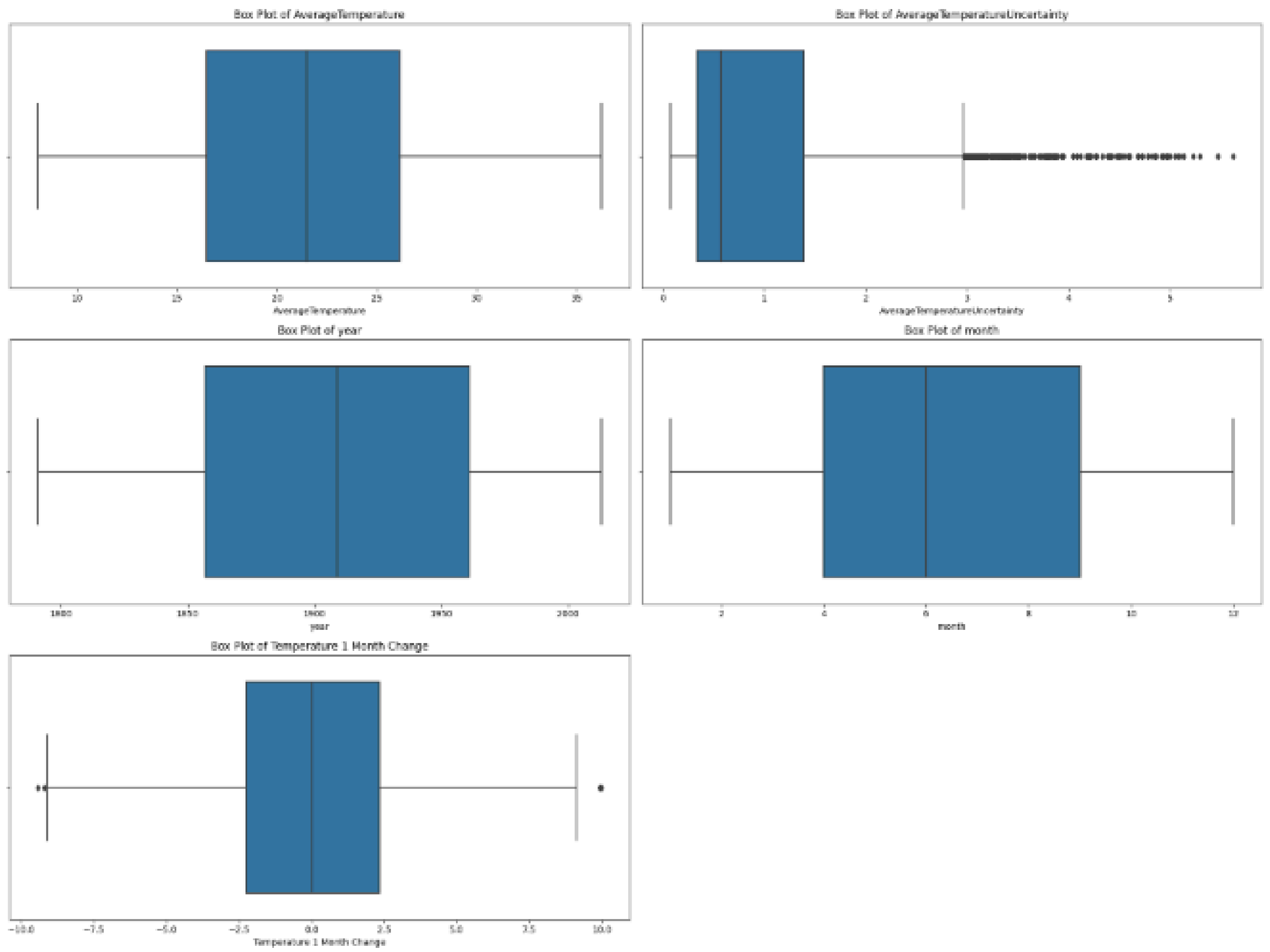
- since we are focusing on Egypt

Latitude & Longitude

- since we have already extracted a new feature (Climate Region)

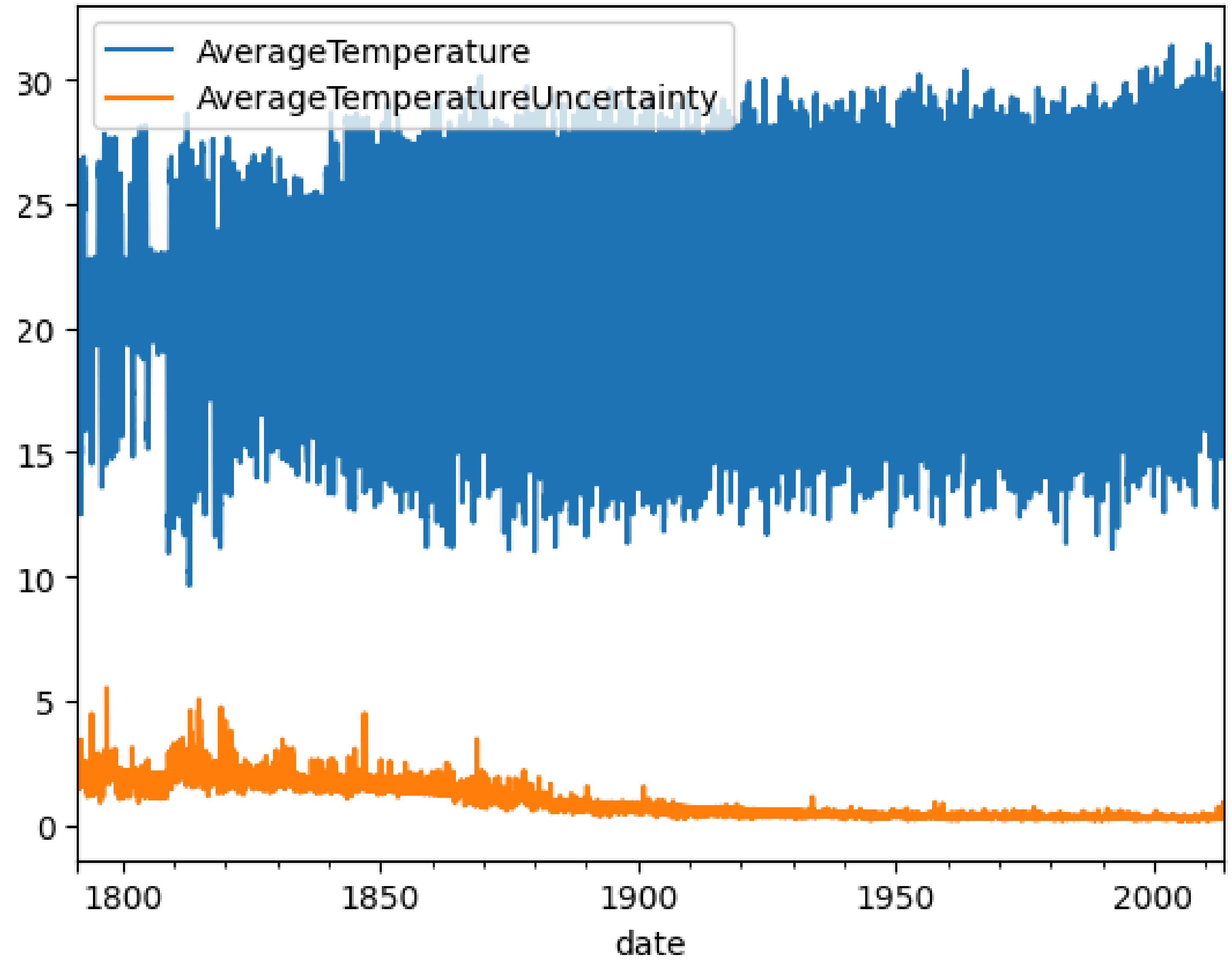
Outliers Detection

Handled outliers using box plots to identify and remove extreme values, ensuring data consistency and reducing skewness.

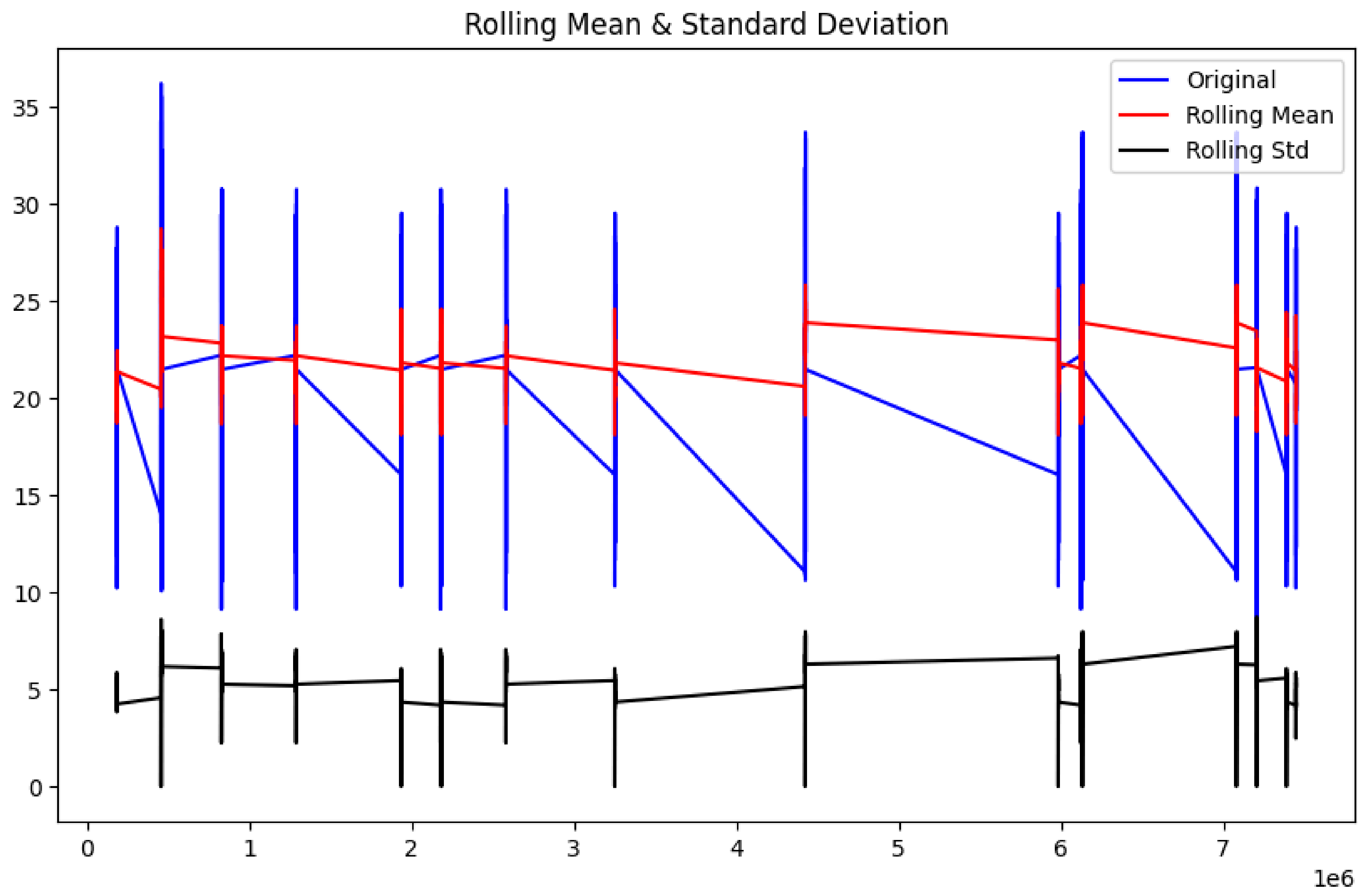


Upsampling:

- Increasing the data frequency so we can forecast on a daily basis instead of monthly.



Check Stationarity

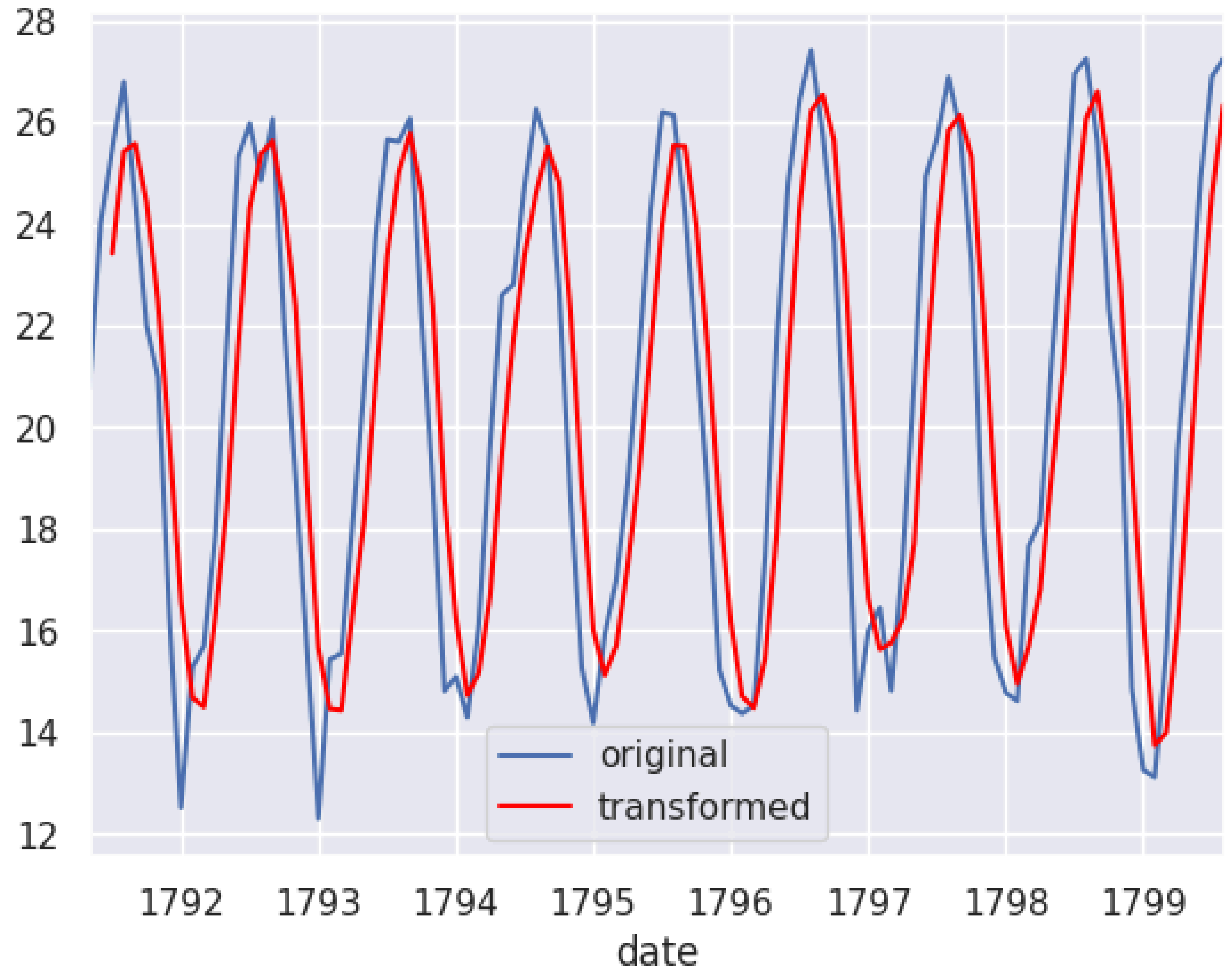



```
Results of Dickey-Fuller Test:
Test Statistic                  -6.451219e+00
p-value                          1.521811e-08
#Lags Used                       5.700000e+01
Number of Observations Used      4.748000e+04
Critical Value (1%)             -3.430488e+00
Critical Value (5%)             -2.861601e+00
Critical Value (10%)            -2.566802e+00
dtype: float64
```

The results of the Dickey-Fuller test indicate that the `AverageTemperature` time series is stationary.

Moving Average Smoothing : tail-rolling average transform as data Preparation

- Reduced noise to highlight the underlying signal of causal processes.
- Removed both trend and seasonal components, making the time series stationary.
- Checked for stationarity to ensure stability in the time series.



Modeling

Clustering:

Split the dataset by grouping cities based on their average temperature

After trying different models with tuning, KMeans proved to be the most effective for:

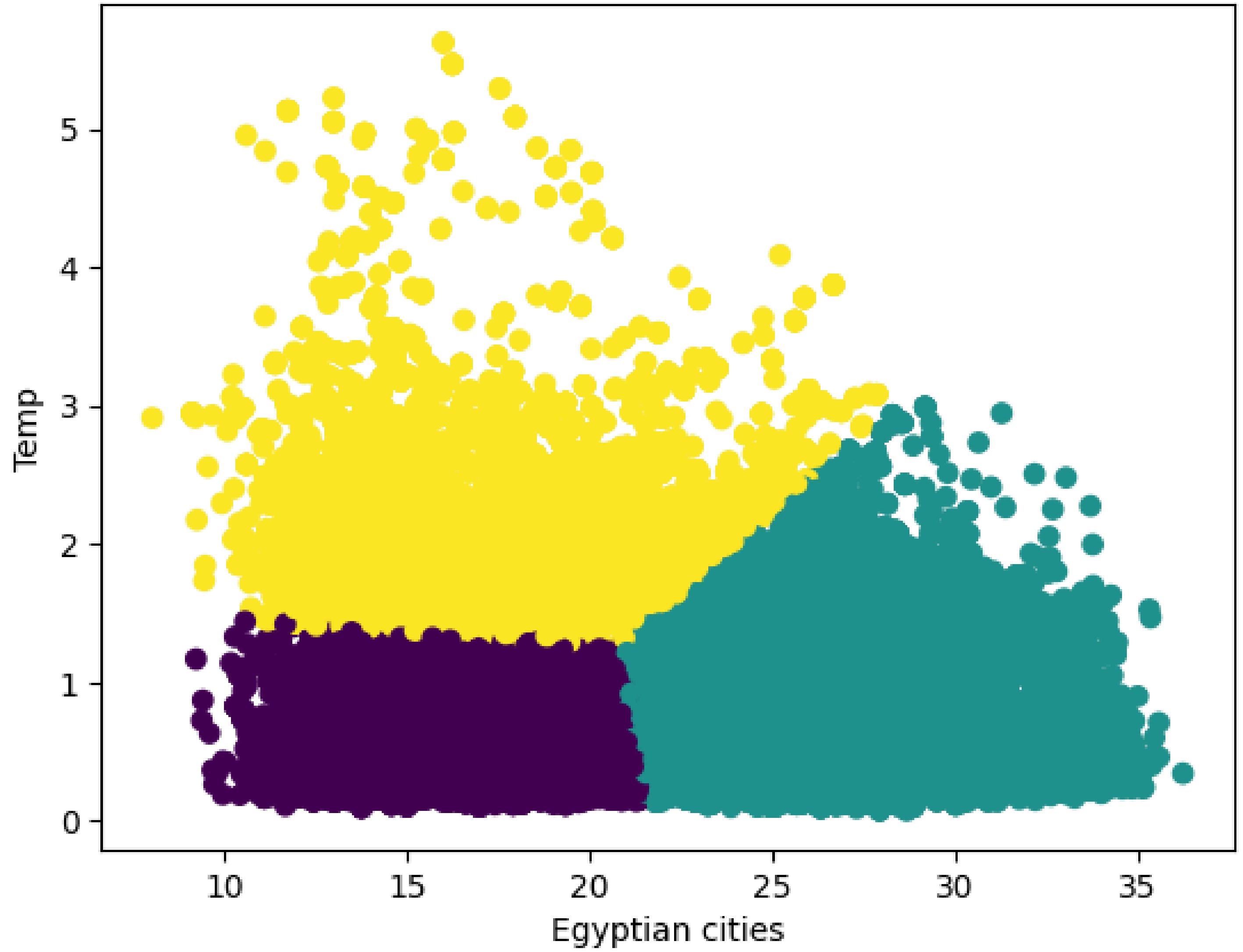
- 1- Faster than other models.
- 2- Good visualization.
- 3- High silhouette score.

Modeling

Silhouette Score: 0.7443

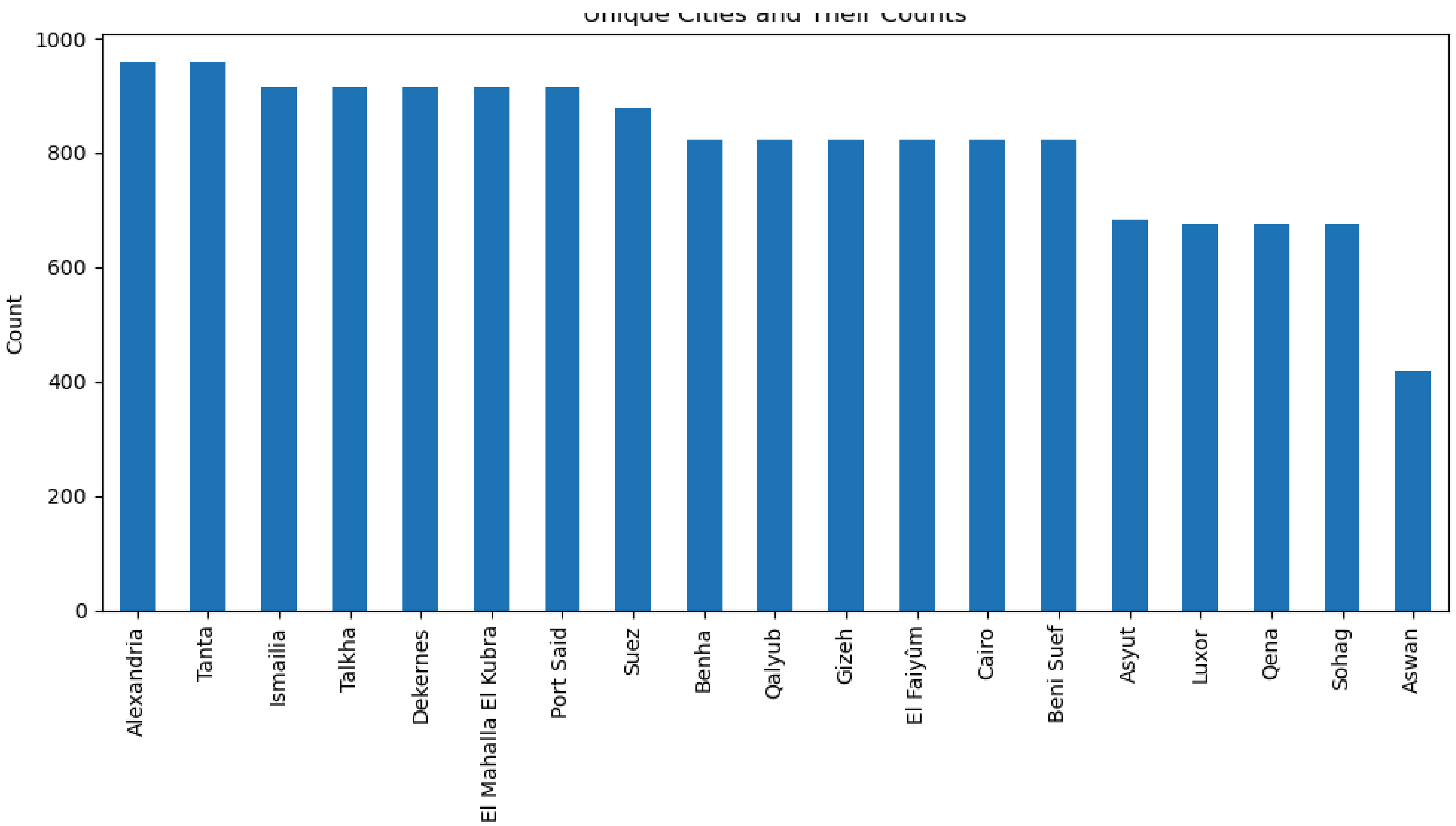
Cities clustered into 3 regions

a. South of Egypt: High average temperature.
b. Middle of Egypt: Moderate average temperature.
c. North of Egypt: Low average temperature.

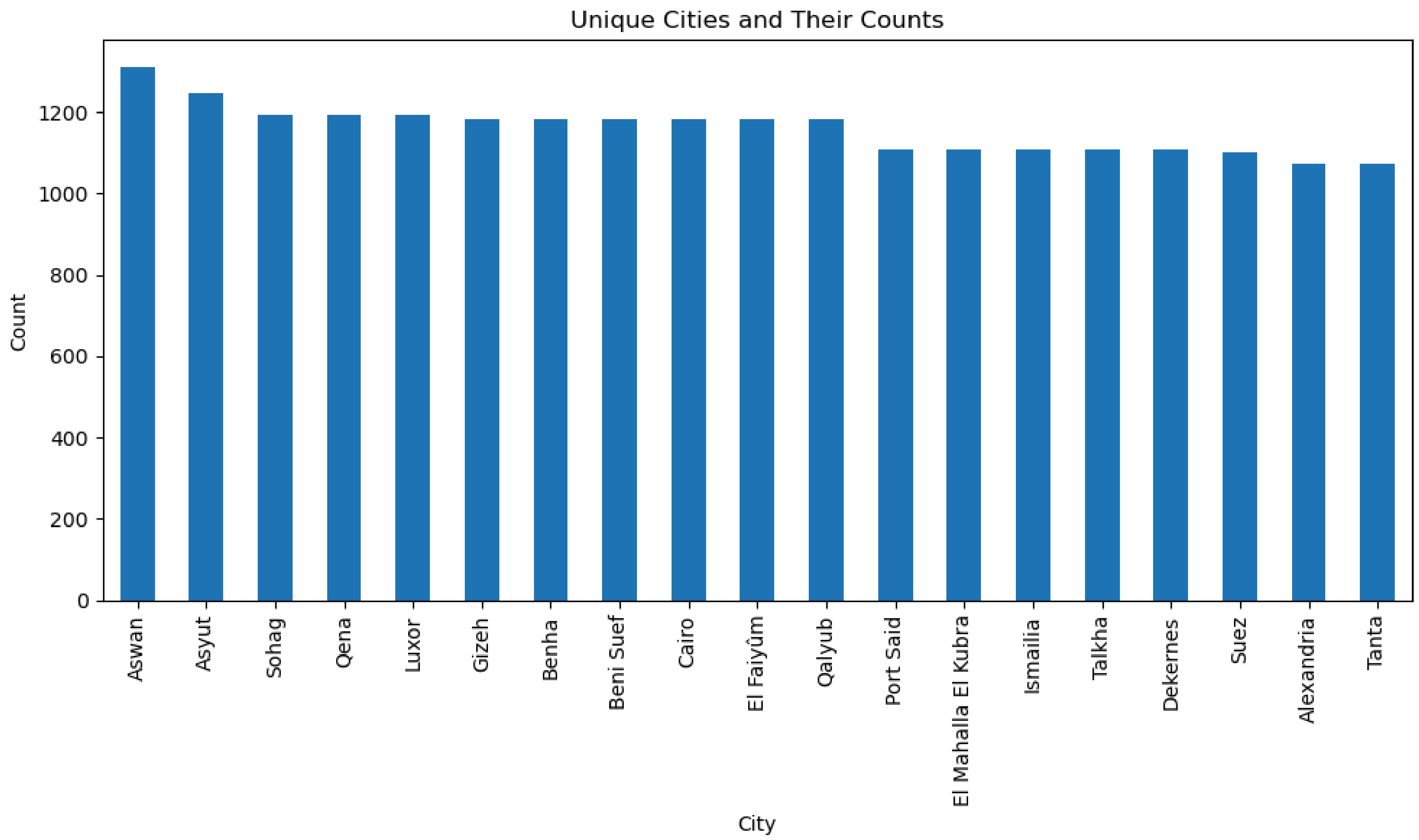


Modeling

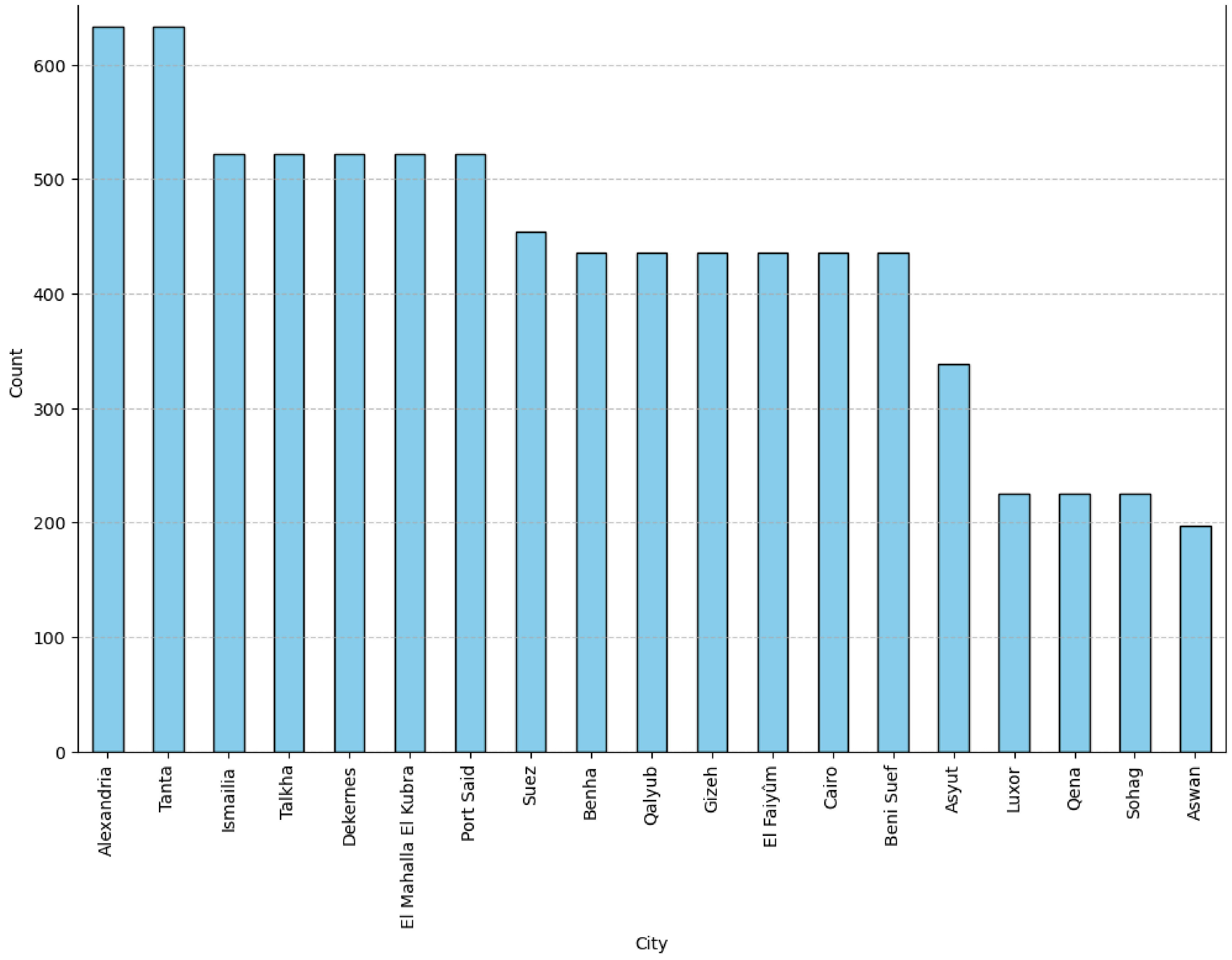
North of Egypt



South of Egypt



Middle of Egypt



By selecting a representative city from each cluster, we can forecast for each region:

1. North: Alexandria
2. Middle: Cairo
3. South: Aswan

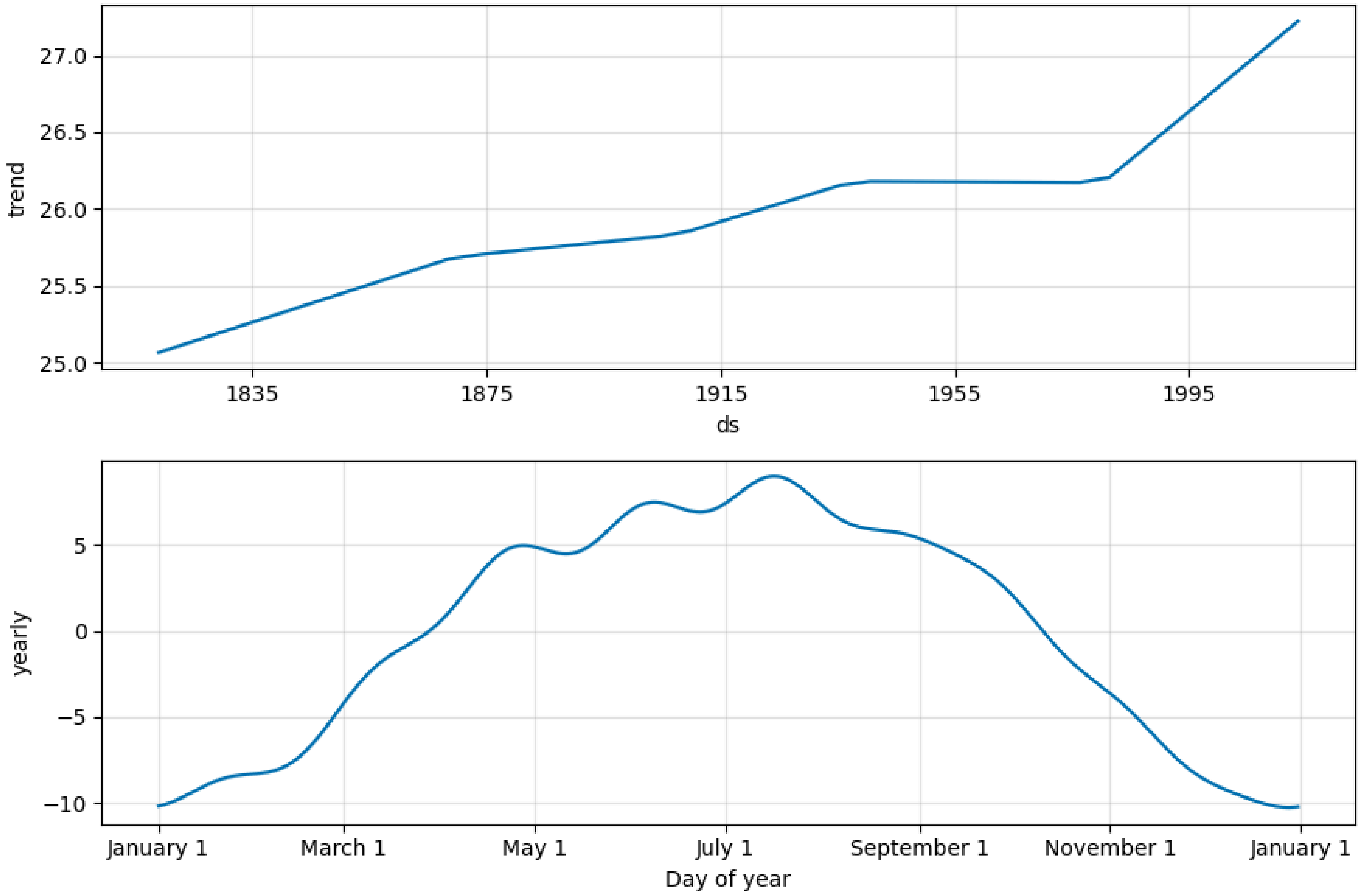
Modeling

For all the cities, we observed that the data exhibits certain features:

- Trend: Indicates the presence of global warming.

- Seasonality: Represents normal cyclical patterns.

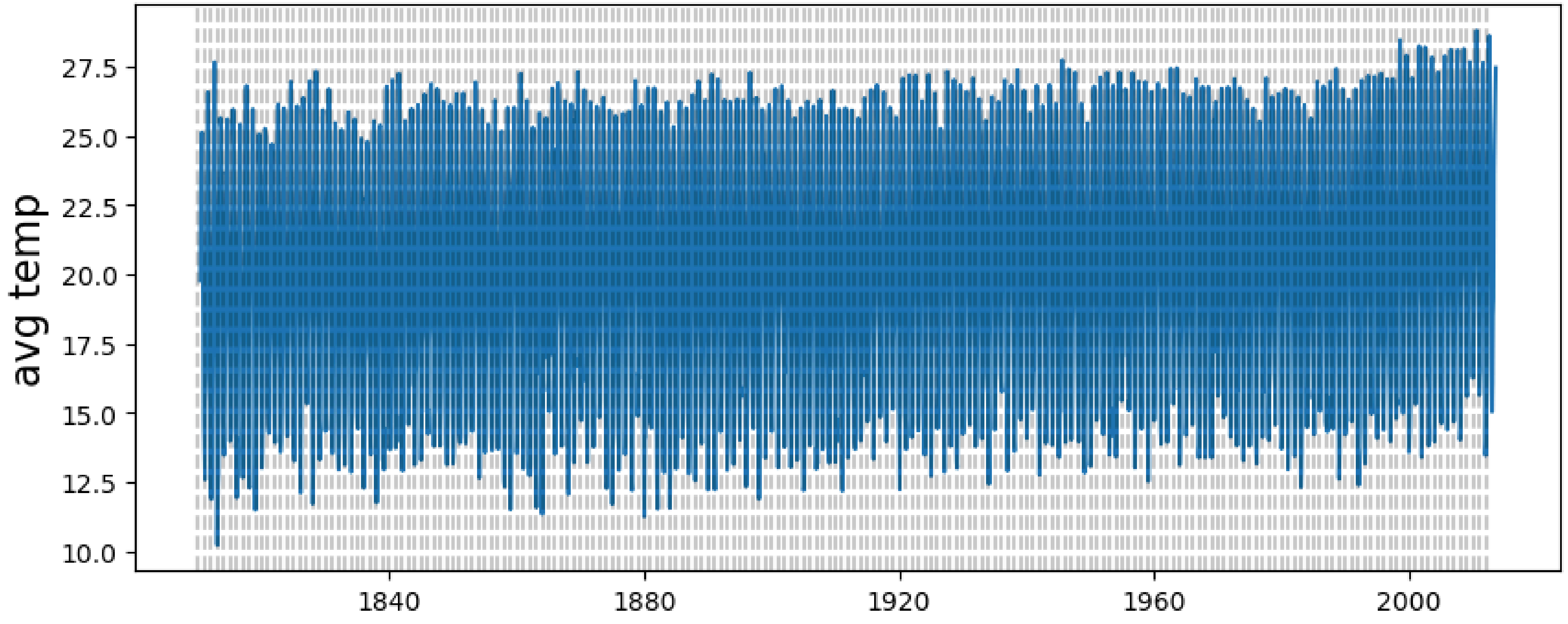
These components were analyzed using the Prophet model.

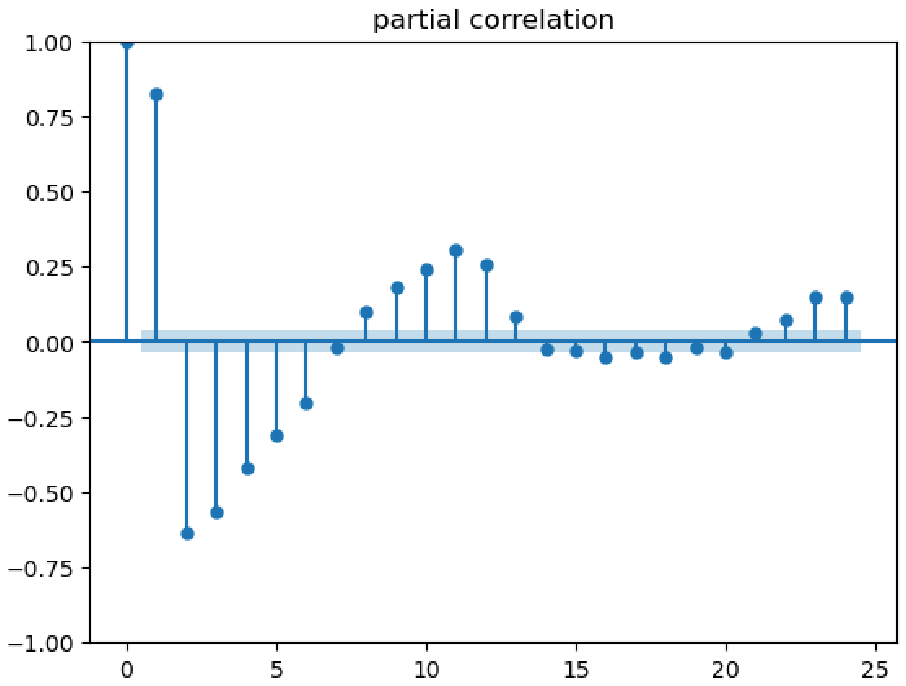


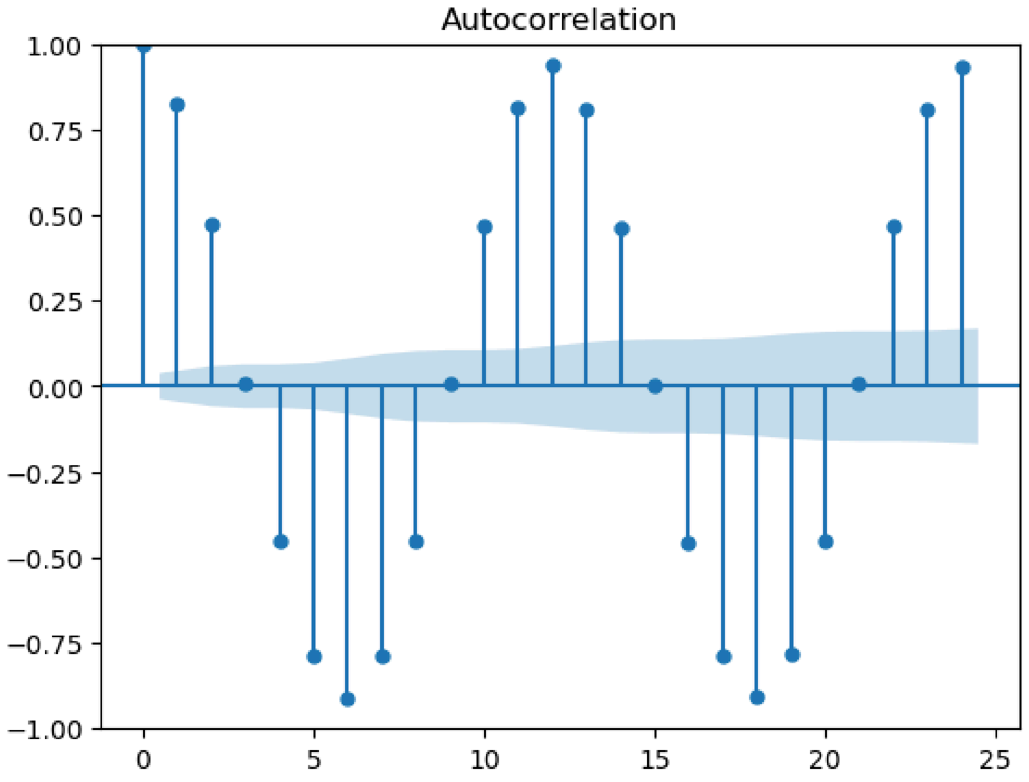
Modeling

The data was modeled on the selected cities by:

- Addressed the trend by applying a 3-period difference
- Used the SARIMAX model to handle seasonality.

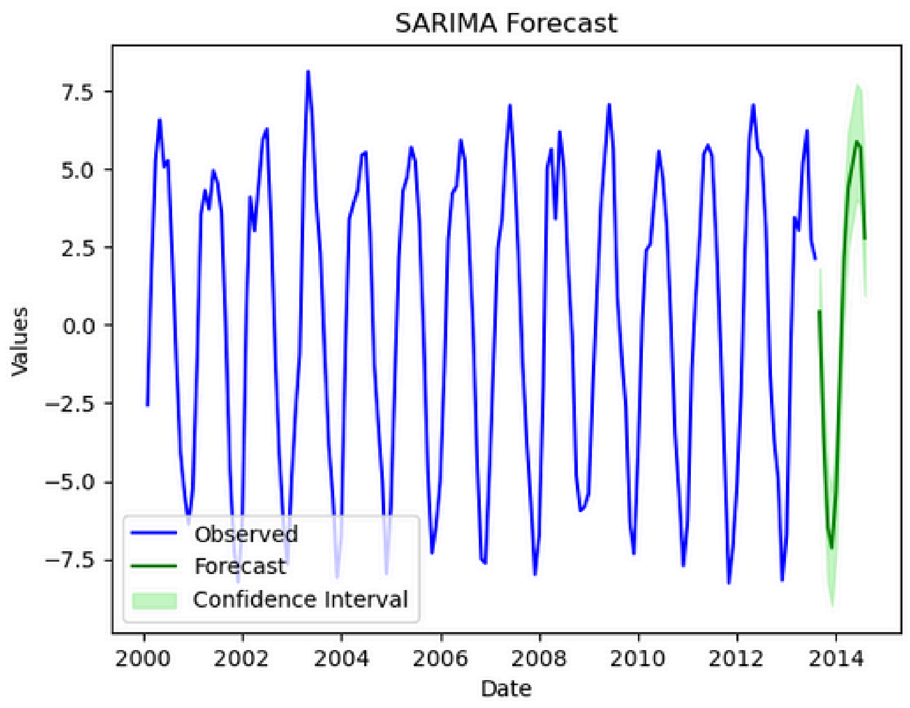




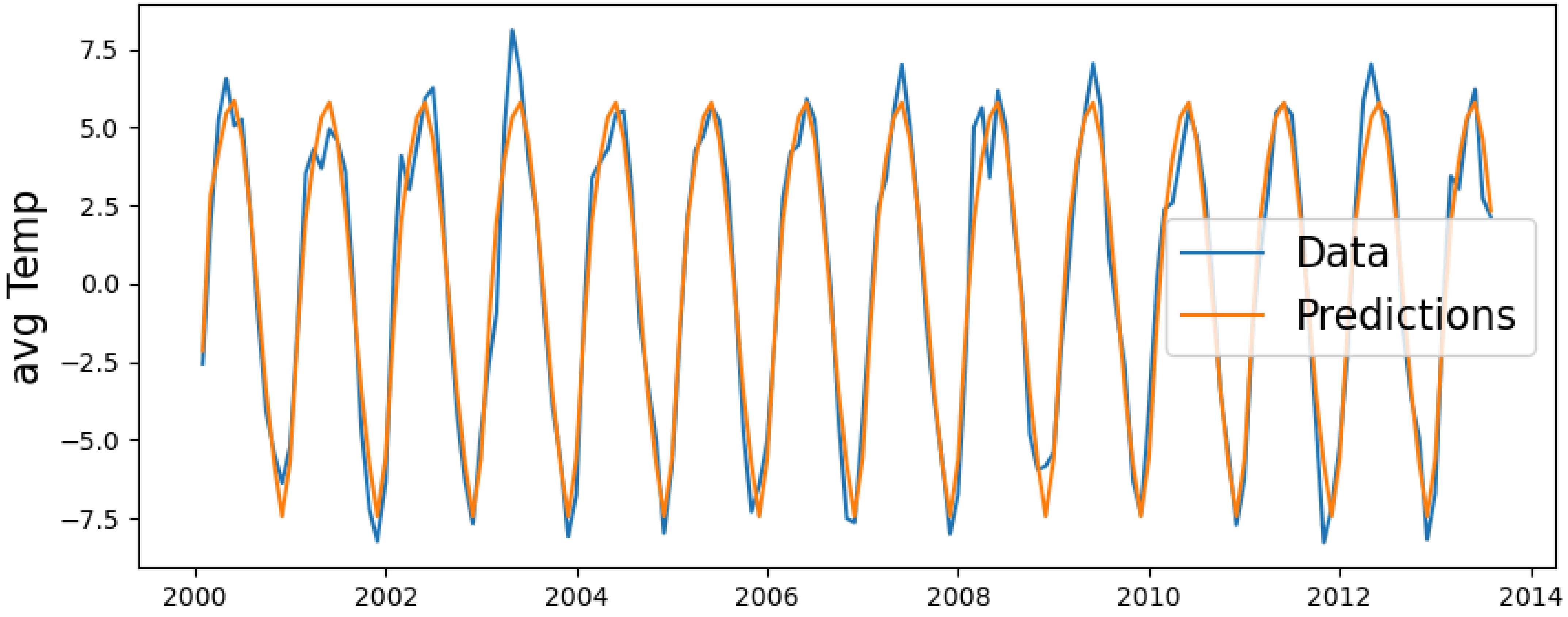


Alexandria

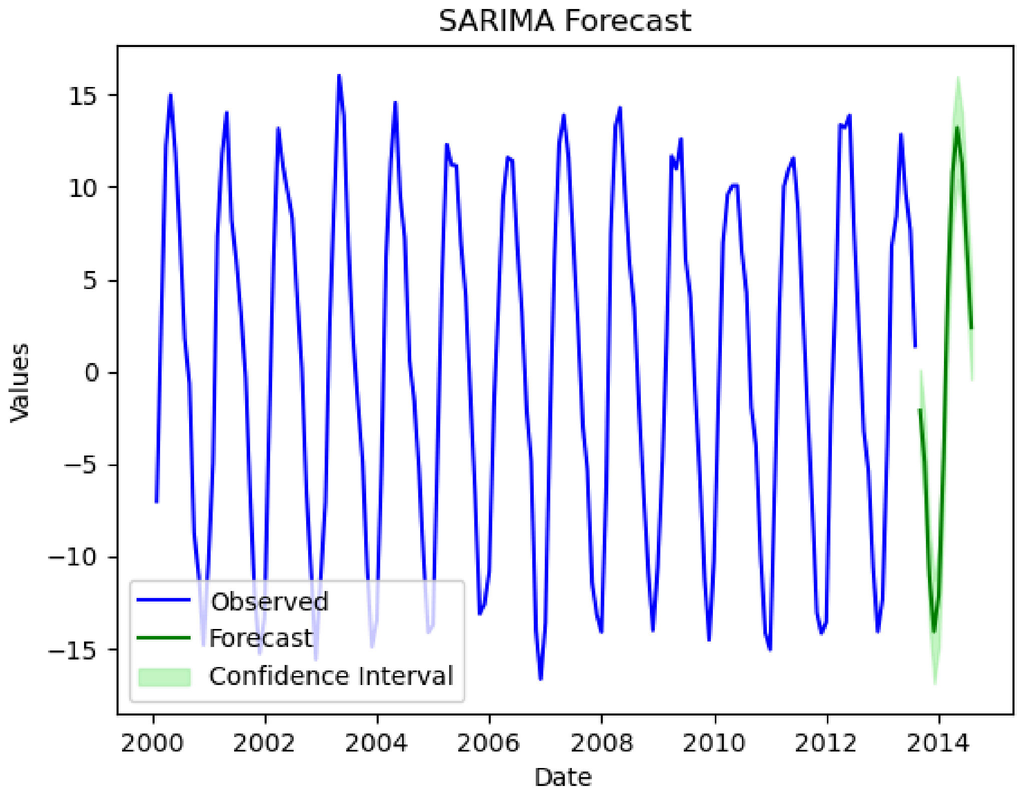
Forecast of the next year

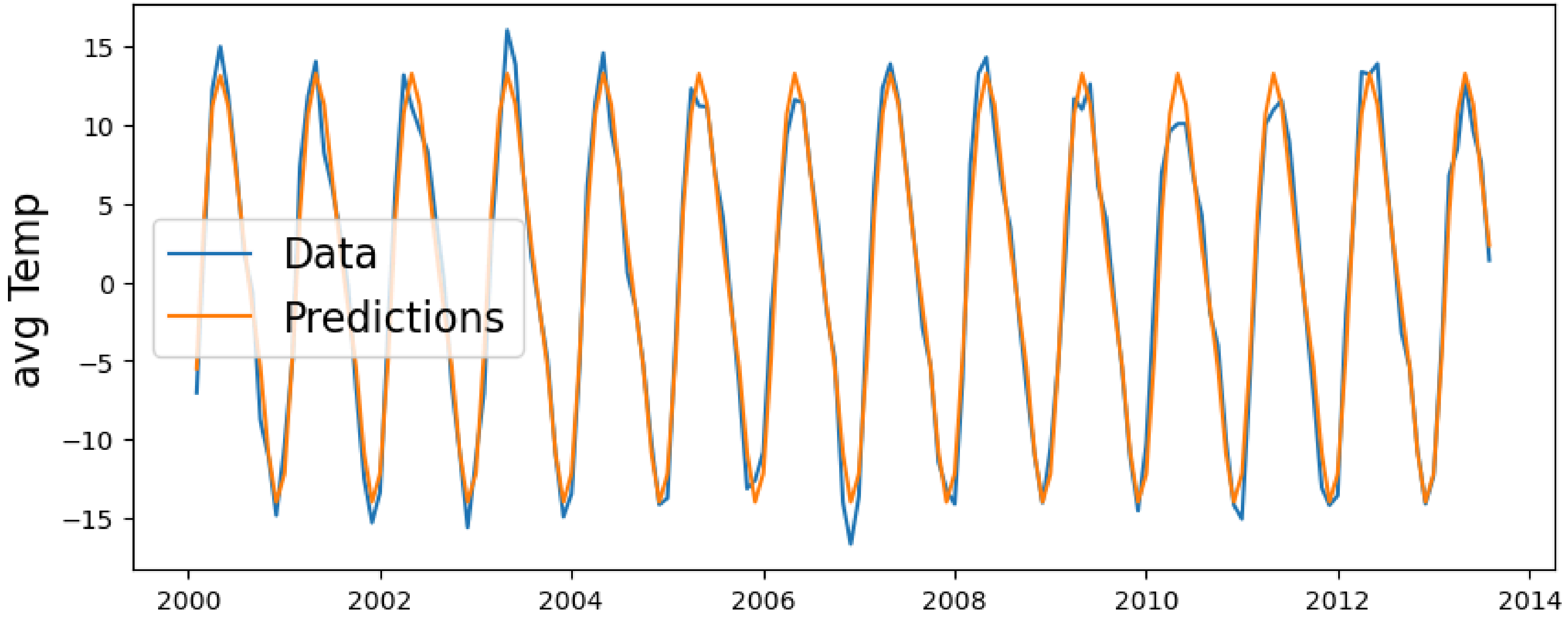


prediction with our test data

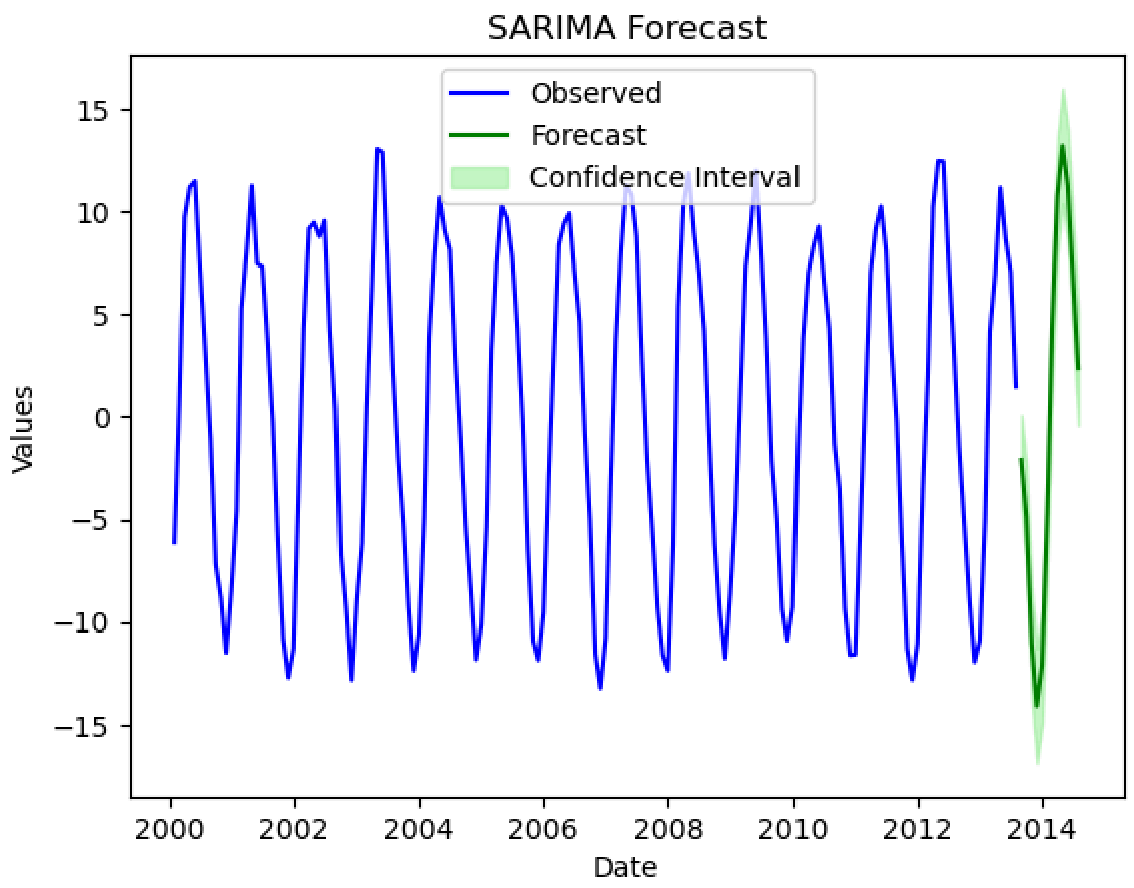


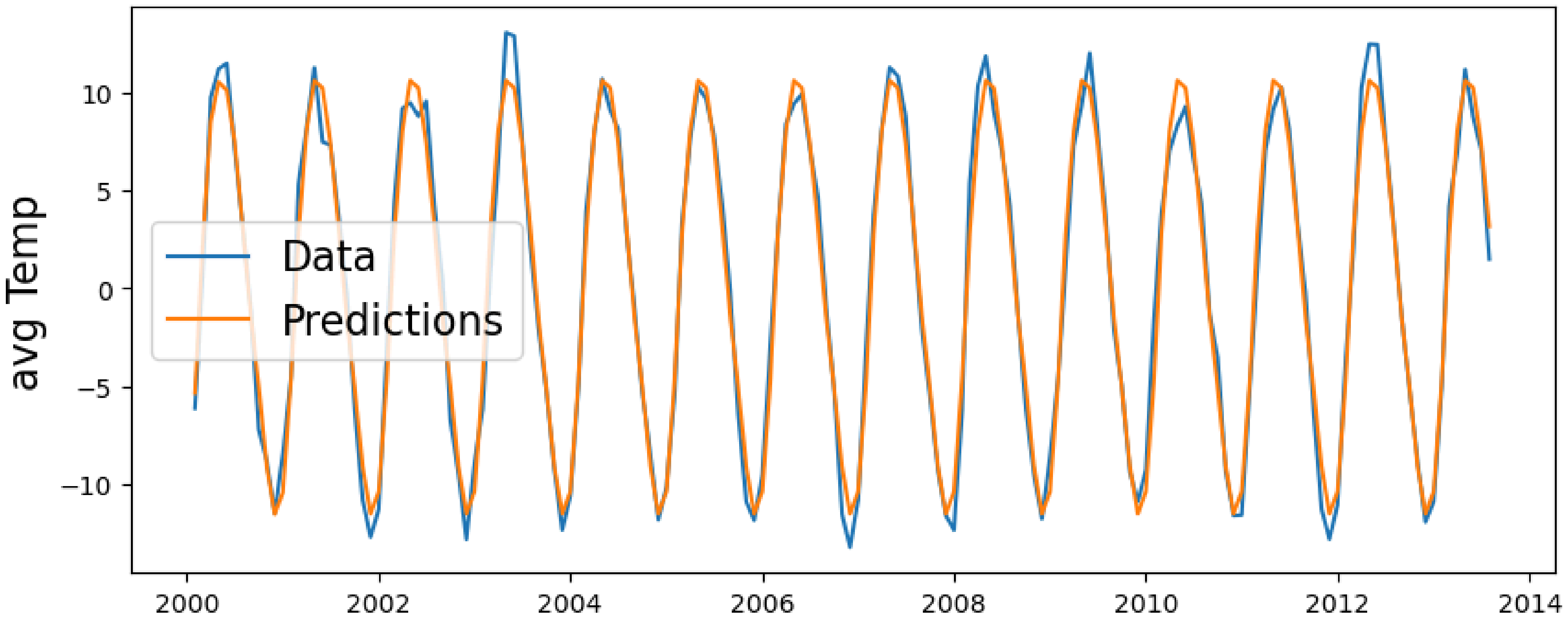
Cairo


SARIMA Forecast
Values
Date
Observed
Forecast
Confidence Interval
15
10
5
0
−5
−10
−15
2000
2002
2004
2006
2008
2010
2012
2014


avg Temp
Data
Predictions
15
10
5
0
−5
−10
−15
2000
2002
2004
2006
2008
2010
2012
2014

Aswan


SARIMA Forecast
Observed
Forecast
Confidence Interval
15
10
5
0
−5
−10
−15
Values
2000
2002
2004
2006
2008
2010
2012
2014
Date

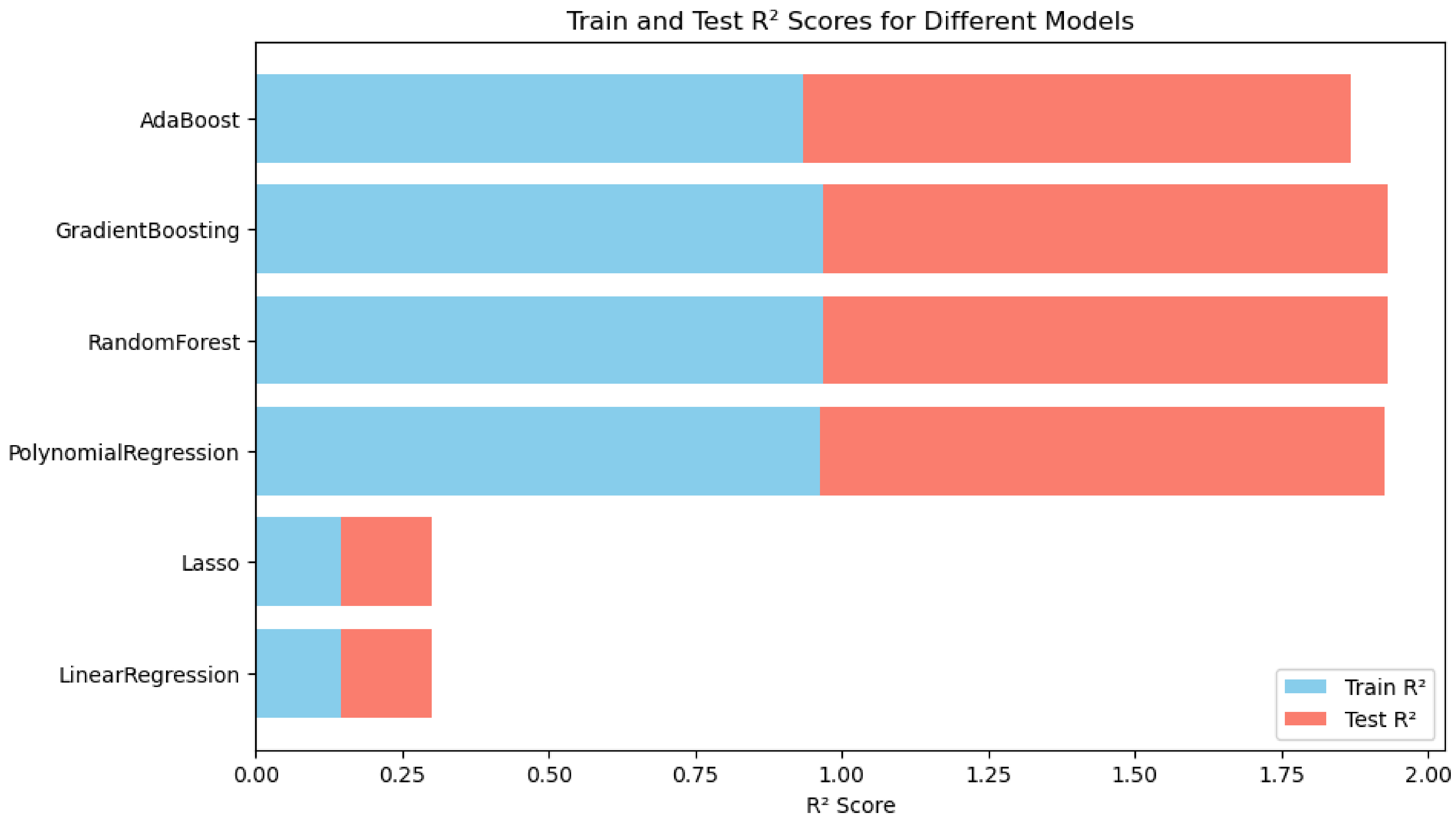

Data
Predictions
avg Temp
10
5
0
−5
−10
2000
2002
2004
2006
2008
2010
2012
2014

Prediction of Average Temperature Based on Month, Longitude, and Latitude:

- Linear Regression:
  - Best Params: {}
  - Train R^2: 0.1468
  - Test R^2: 0.1533
- Lasso Regression:
  - Best Params: {'alpha': 0.01}
  - Train R^2: 0.1468
  - Test R^2: 0.1533
- Polynomial Regression:
  - Best Params: {'poly__degree': 4}
  - Train R^2: 0.9635
  - Test R^2: 0.9623
- Random Forest:
  - Best Params: {'max_depth': 20, 'n_estimators': 50}
  - Train R^2: 0.9667
  - Test R^2: 0.9659
- Gradient Boosting:
  - Best Params: {'learning_rate': 0.1, 'max_depth': 5, 'n_estimators': 150}
  - Train R^2: 0.9667
  - Test R^2: 0.9659
- AdaBoost:
  - Best Params: {'learning_rate': 1.0, 'n_estimators': 150}
  - Train R^2: 0.9333
  - Test R^2: 0.9346

Modeling

Best Model: RandomForest with Test R^2 Score: 0.9659



SAMSUNG

Together for Tomorrow!

Enabling People

Education for Future Generations