From Runways to Revenue:

A Data-Driven Exploration of Logan Airport

The issues:

- How have the numbers of passengers traveling through Logan Airport changed each year? Are there specific trends or patterns?
- In which month and year did Logan Airport experience the most traffic? What might be causing these changes? What patterns have been seen over time in the quantity of foreign flights arriving at Logan Airport?
- How do hotel occupancy rates and average daily rates vary year by year? During which months are hotels most and least occupied, and when are the rates highest and lowest?
- How has the total number of jobs in the region changed annually? What can these changes tell us about the area's economic health?
- How has the unemployment rate evolved each year? What might be influencing these shifts?
- Are there any noticeable changes in the labor force participation rate over time? What could be driving these changes?
- How has the development of new construction projects (measured in square feet) varied each year? What does this say about economic development?
- How does number of projects correlate with the overall job market health?

- How has the total development cost of pipeline projects changed over the years? What might this indicate about the region's investment in infrastructure?

- Which factors are the strongest predictors of employment growth in the region?

- Can we forecast the number of passengers for the upcoming years?

Findings:

- Logan Airport's passenger traffic has steadily increased year over year, suggesting a strong increase in the region's desire for air travel.
- The airport experienced its highest traffic during specific months and years, likely influenced by seasonal tourism patterns, business travel, and an increase in flight routes, particularly international ones.
- The region's hotel business has demonstrated year-over-year variations in average daily rates and occupancy rates, demonstrating the industry's adaptability to a range of variables such local events, economic situations, and seasonal demand.
- The overall number of jobs in the region has increased annually, indicating a strong and expanding economy. This indicates a solid trend in job growth.
- The unemployment rate has been on a general decline, suggesting improvements in the job market and possibly reflecting broader economic growth.
- The labor force participation rate fluctuates over time, suggesting changes in the makeup of the workforce that may be brought about by shifting economic conditions or demographic transitions.

- The development of new construction projects, measured in square feet, has varied annually, providing insight into the economic development and confidence in the region.
- The number of jobs created by construction projects seems to align with the overall job market health, indicating the construction sector's significant impact on the regional economy.
- The overall development cost of pipeline projects has fluctuated throughout time, indicating shifts in regional infrastructure investment and economic priorities.
- Certain economic indicators have emerged as strong predictors of employment growth, offering insights into factors that could drive regional economic development.
- It is anticipated that Logan Airport's passenger volume will continue to rise in the future, requiring strategic planning for resource management and infrastructure to handle this increase.

Discussions:

The sustained increase in passenger traffic at Logan Airport indicates a robust and growing demand for air travel, which has significant implications for airport infrastructure and resource management. This trend suggests the need for careful planning and development to accommodate future growth, potentially requiring expansion of facilities, enhancement of operational efficiency, and reinforcement of transportation networks connecting to the airport.

The identified peak traffic periods, influenced by seasonal tourism and business travel, highlight the importance of strategic resource allocation during these times. This could involve adjusting flight schedules, staffing, and airport services to efficiently manage the influx of passengers during peak months. Additionally, the growth in international flights suggests an expanding global role for the airport, potentially increasing its economic impact through enhanced tourism and business travel.

The hotel industry is very susceptible to external forces, as seen by the variations in daily rates and occupancy rates. This means that hoteliers must enhance their service offerings and use flexible pricing methods in order to take advantage of periods of high demand and prepare for slower ones. In a market that is constantly changing, this flexibility is essential to preserving profitability and competitiveness.

Regarding employment, the consistent increase in jobs within the region indicates a thriving economic environment. This may be partly attributed to the airport's influence, which acts as a catalyst for job creation in sectors such as retail, hospitality, and transportation. The link between airport activity and employment suggests the importance of targeted educational programs to equip the local workforce with the skills needed to fill new roles, particularly those associated with the airport's expansion and technological advancements.

The decline in the unemployment rate, coupled with fluctuations in labor force participation, highlights the importance of understanding the underlying factors driving these changes. It is crucial to consider the role of automation, changing industry demands, and the gig economy in shaping the employment landscape.

The fluctuations in the number of jobs generated and the square footage of newly developed projects show growth and confidence in the economy. It also calls into question the long-term viability and planning of these kinds of developments. The state of the construction industry is a good indicator of the state of the economy, and its success ripples out to other industries.

Furthermore, the regression analysis performed as part of this study sheds light on the relationship between various economic indicators. The results underscore the interdependencies within the region's economy, where factors such as hotel occupancy rates, average daily rates, and passenger numbers at Logan Airport interact to influence job creation and economic growth.

The forecasting models, such as the SARIMA model used, provide valuable projections for passenger traffic, offering a basis for proactive planning. Accurate forecasting is vital for managing the expected growth, ensuring that the airport and associated businesses are well-prepared to meet future demands.

Appendix A -Methodology:

To address the various issues related to Logan Airport and its economic impact, a comprehensive methodology was employed, encompassing data collection, statistical analysis, and forecasting techniques. Here's an overview of the methodology used:

1. **Data Collection:** The first step involved is gathering relevant data. We have collected the data from the link https://data.boston.gov/dataset/economic-indicators-legacy-portal (Analyze Boston web portal).

2. **Data Description:** The dataset has 84 entries with 19 columns ranging from year 2013 to 2019. This included passenger traffic numbers at Logan Airport, international flight statistics, hotel occupancy and rate data, regional employment figures, construction project details, and pipeline investment costs.

3. **Variable Creation:**

logan_passengers	Number of domestic and international passengers at Logan Airport
logan_intl_flights	Total international flights at Logan Airport

hotel_occup_rate	Hotel occupancy for Boston
hotel_avg_daily_rate	Hotel average daily rate for Boston
total_jobs	Total number of jobs
unemp_rate	Unemployment rate for Boston
labor_force_part_rate	Labor rate for Boston
pipeline_unit	Number of units approved
pipeline_total_dev_cost	Total development cost of approved projects
pipeline_sqft	Square feet of approved projects
pipeline_const_jobs	Construction jobs
Year	Year
Month	Month of the year
foreclosure_pet	Foreclosure house petitions
foreclosure_deeds	Foreclosure house deeds
med_housing_price	Median housing sales price
housing_sales_vol	Number of houses sold
new_housing_const_permits	New housing construction permits
new-affordable_housing_permits	New affordable construction permits

4. **Descriptive Analysis**: Initial explorations of the data involved descriptive statistics to understand basic trends and patterns. This included year-over-year comparisons and month-to-month analyses to identify fluctuations and trends in passenger traffic, hotel industry performance, and job market changes.

5. **Time Series Analysis**: For forecasting and analyzing time-dependent trends, a Seasonal Auto Regressive Integrated Moving Average (SARIMA) model was employed, especially suited for time series data exhibiting seasonal patterns. The SARIMA model was chosen for its ability to handle both non-stationarity and seasonality in the data, making it ideal for forecasting future passenger traffic at Logan Airport. The model parameters were determined based on the Autocorrelation Function (ACF) and Partial Autocorrelation Function

(PACF) plots, alongside stationarity tests like the Augmented Dickey-Fuller (ADF) test.

I. **Autocorrelation Function (ACF)**:
   - **Description**: ACF measures the correlation between a time series and its lagged values. It helps to identify the extent of the correlation between observations at different time lags.
   - **Use in Project**: We used ACF to determine the moving average (MA) component of the SARIMA model. Significant spikes in the ACF plot at specific lags indicated how many past values in the series influence the current value, guiding the choice of the 'q' parameter in the model.

II. **Partial Autocorrelation Function (PACF)**:
   - **Description**: PACF measures the correlation between a time series and its lagged values, but after eliminating the influence of earlier lags. It helps to identify the direct relationship between an observation and its lag.
   - **Use in Project**: PACF was utilized to identify the autoregressive (AR) component of the SARIMA model. Significant spikes at certain lags in the PACF plot suggested the number of past values ('p' parameter) that should be used in the model.

III. **Augmented Dickey-Fuller (ADF) Test**:
   - **Description**: The ADF test is a statistical test used to check for stationarity in a time series. It tests the null hypothesis that a unit root is present in a time series sample, which would indicate non-stationarity.
   - **Use in Project**: We applied the ADF test to determine whether the passenger traffic data (and other time series data) were stationary. Since SARIMA models require stationarity, any indication of a unit root (non-stationarity) from the ADF test would mean we need to difference the data to make it stationary.

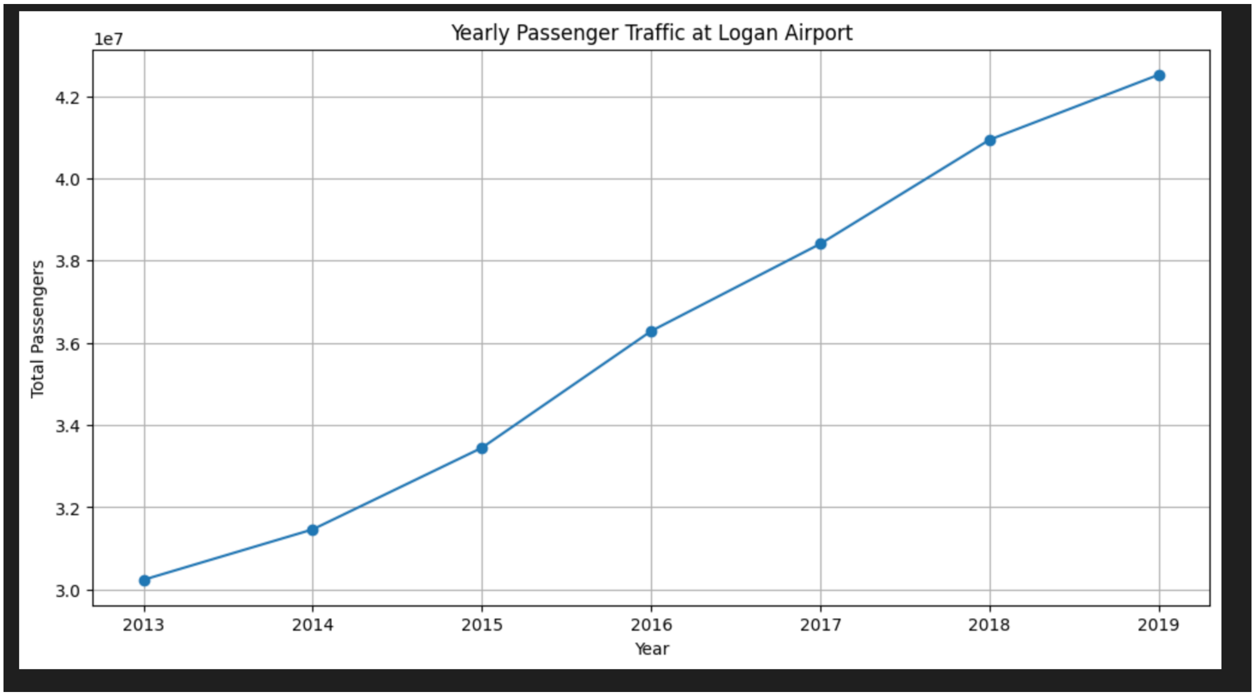
6. **Seasonal Analysis**: To understand the seasonal trends in hotel occupancy and rates, a seasonal analysis was conducted. This involved identifying specific months with peak and off-peak demands.

7. **Correlation and Regression Analysis**: To explore relationships between different economic indicators, such as the correlation between passenger traffic and hotel industry performance or the factors predicting employment growth, correlation analysis and regression modeling were used. These statistical methods helped in understanding the strength and nature of the relationships between different variables.

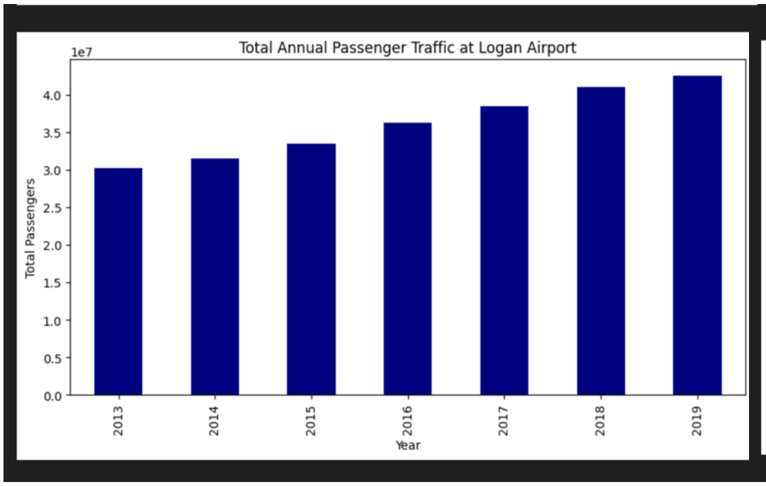
8. **Predictive Modeling**: Predictive models were developed to forecast future trends, especially in passenger traffic at Logan Airport. These models were based on historical data and aimed to provide insights into future patterns.

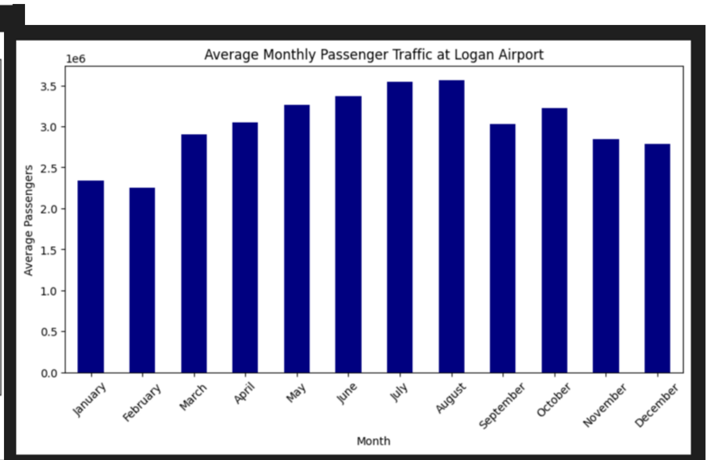
9. **Data Visualization**: Throughout the analysis, data visualization techniques were used to present findings in an understandable and accessible manner. This included charts and graphs to illustrate trends, patterns, and relationships in the data.

Appendix B - Results:

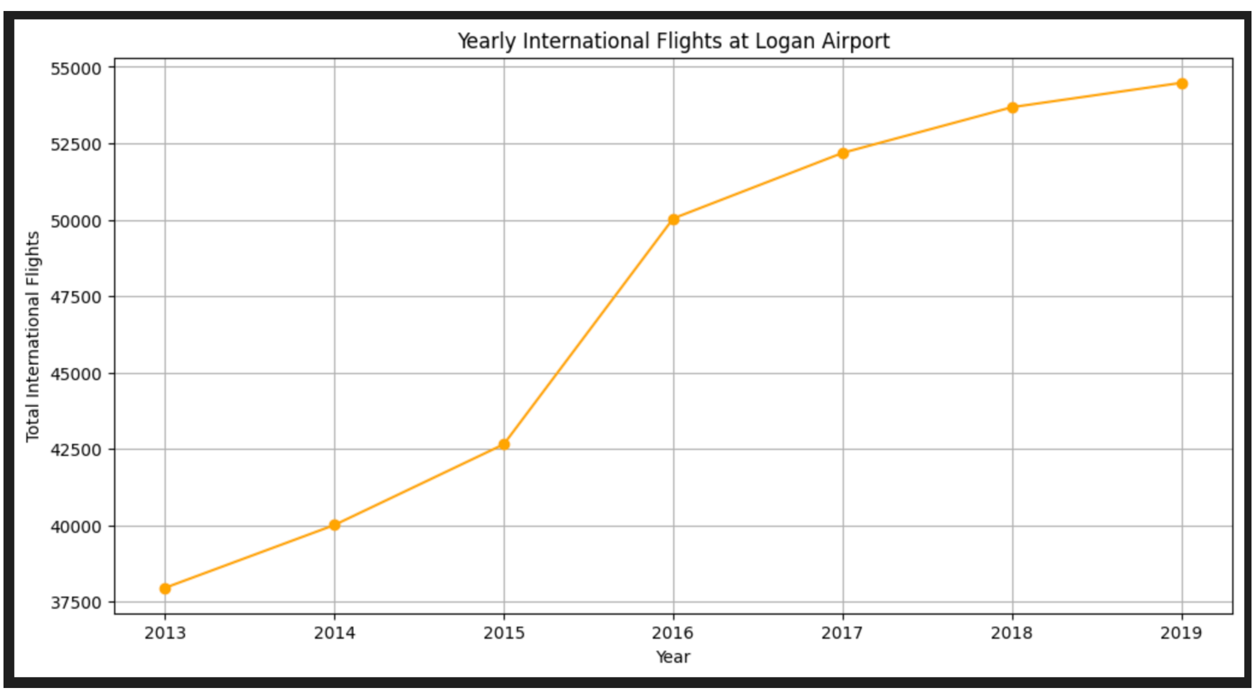


The analysis of passenger traffic trends at Logan Airport revealed a consistent increase over the years. The data showed a clear upward trajectory, indicating a growing demand for air travel in the region. This sustained growth trend suggests that Logan Airport is becoming increasingly important as a travel hub, reflecting the region's expanding economic and cultural connectivity.

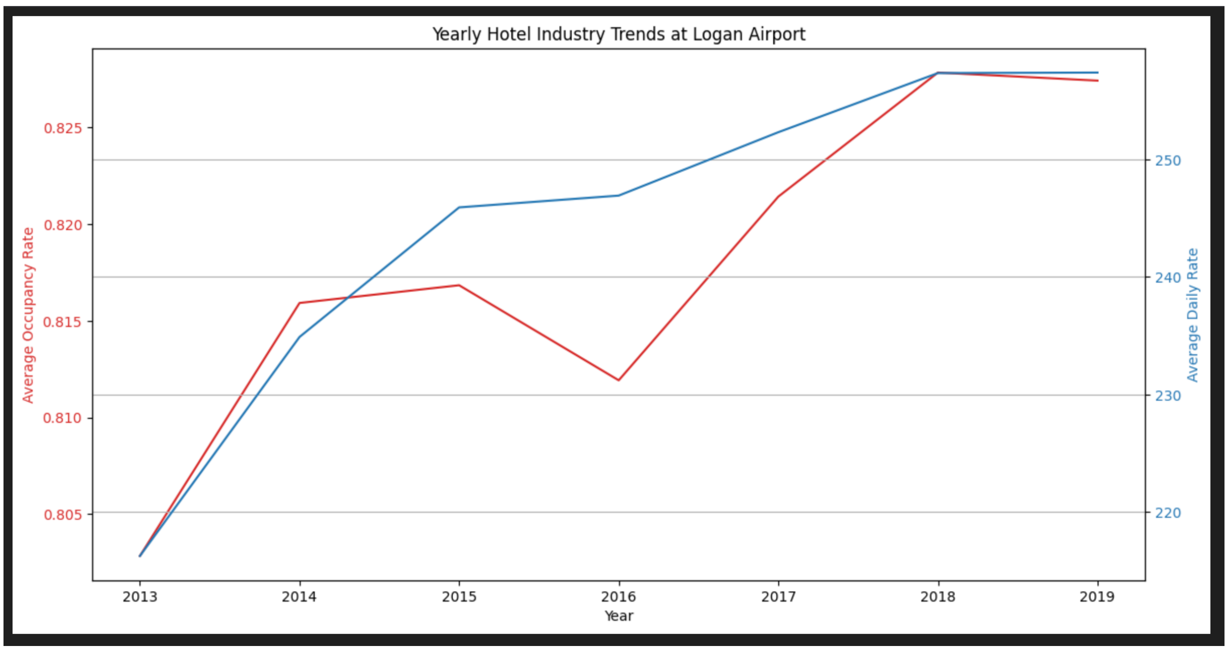


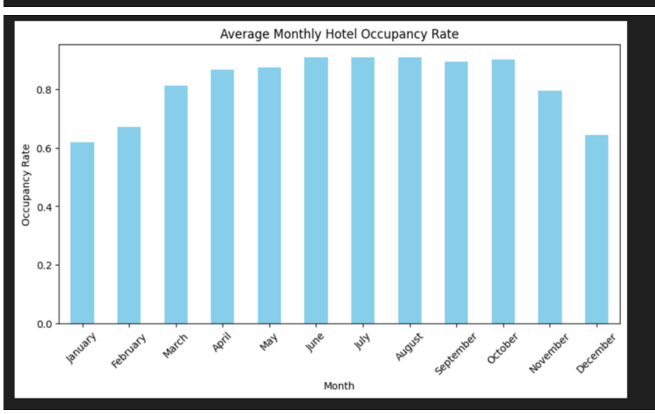


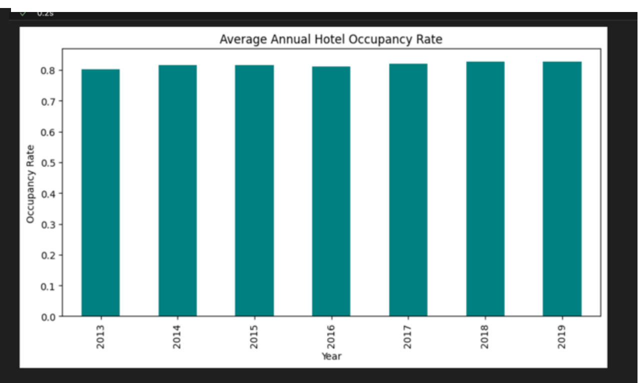
The plots indicated specific months and years when Logan Airport experienced its highest traffic. These peak periods were likely influenced by seasonal factors such as holidays and summer tourism, as well as business travel patterns. The data also suggested a possible impact of new international flight routes contributing to these peak traffic times. The highest month in which the average monthly passenger traffic at Logan Airport is July and least month is February. Similarly, the highest year in which the total annual passenger traffic at Logan Airport is 2019 and least is 2013.



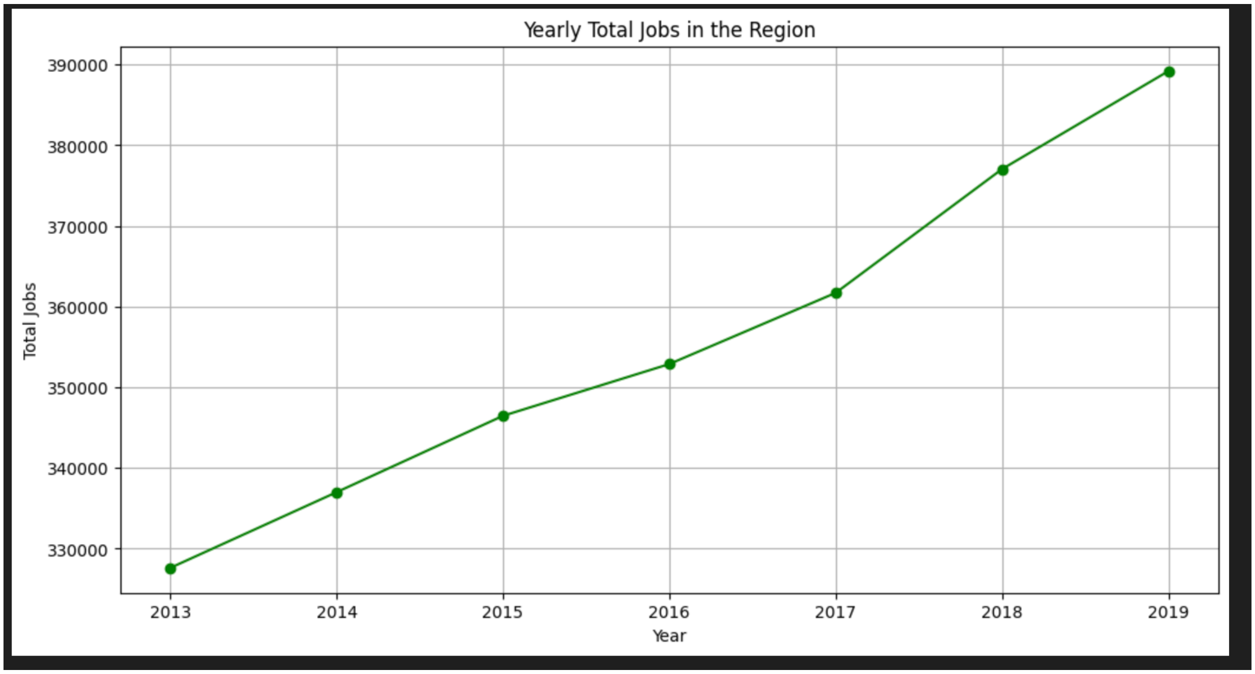
The trends in international flights showed an upward movement, indicating an increase in the number of international flights arriving at Logan Airport. This increase aligns with the airport's growing role in global air travel and suggests a broader economic and cultural engagement with international destinations.



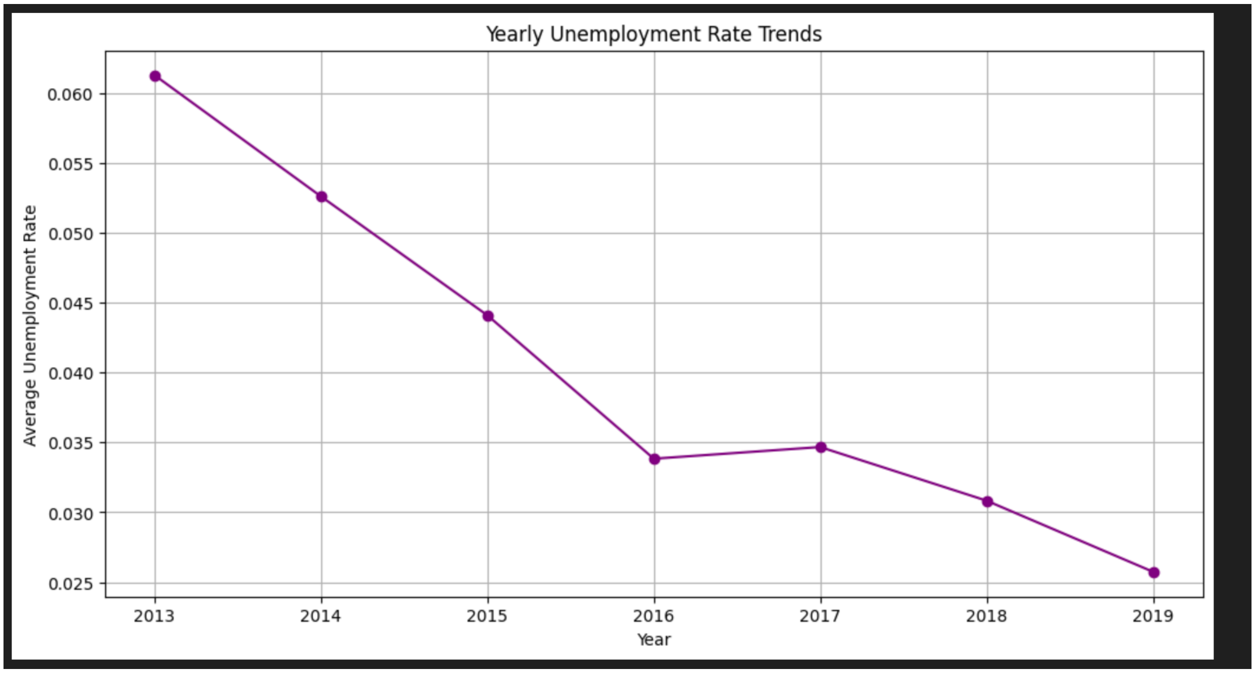




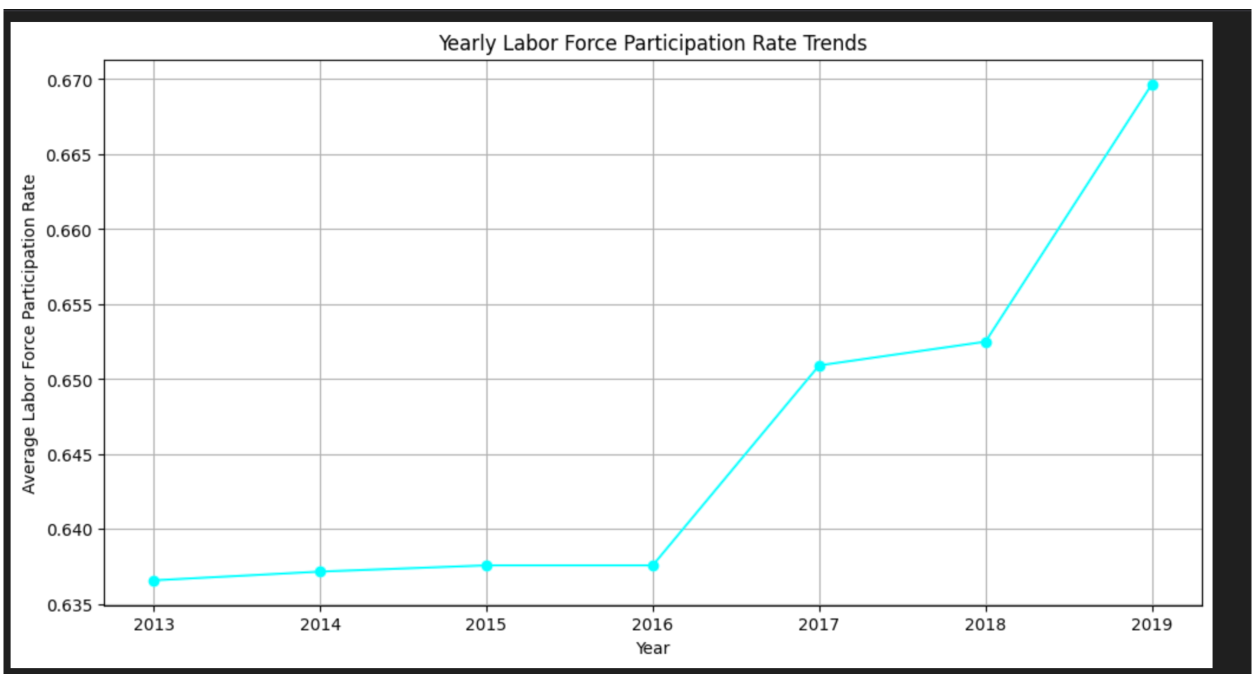
The trends in international flights showed an upward movement, indicating an increase in the number of international flights arriving at Logan Airport. This increase aligns with the airport's growing role in global air travel and suggests a broader economic and cultural engagement with international destinations. We can see that the average monthly hotel occupancy rate is highest in June and least in January.



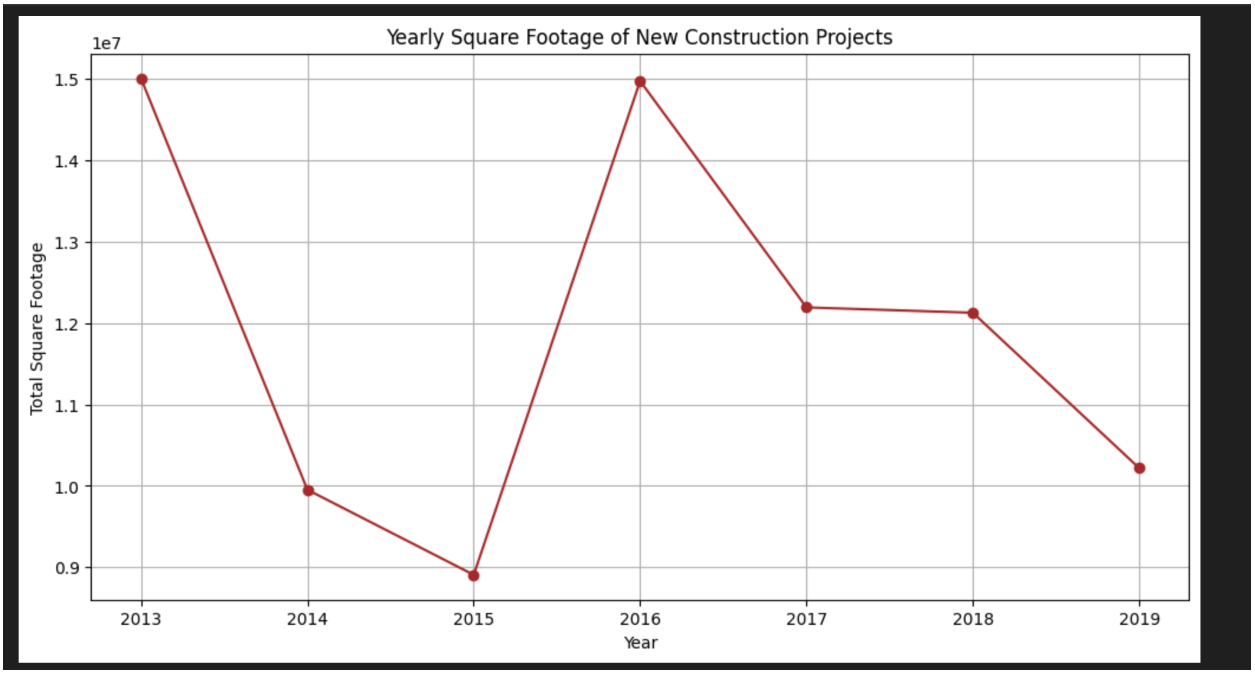
The job market analysis showed a positive trend, with an annual increase in the total number of jobs in the region. This consistent growth indicated a strong and expanding local economy, with the potential for continued job creation and economic development.



The unemployment rate data exhibited a general downward trend over the years. This decline in unemployment is indicative of an improving job market and could reflect broader economic growth, increased job opportunities, and possibly successful employment policies.



Over time, there was fluctuation in the labor force participation rate. Variations in this rate may be a sign of changing industry demands, shifting demographics, or shifting economic conditions that impact the labor market participation of the workforce.



The analysis of new construction projects, measured in square feet, showed annual variations. These fluctuations provided insights into the region's economic development, with certain years exhibiting higher construction activity, possibly reflecting periods of economic confidence and growth.

Correlation Matrix:

	Year	logan_passengers	hotel_occup_rate	hotel_avg_daily_rate	total_jobs
Year	1.000000	0.996234	0.891263	0.919236	0.992367
logan_passengers	0.996234	1.000000	0.867567	0.893352	0.987963
hotel_occup_rate	0.891263	0.867567	1.000000	0.907686	0.899390
hotel_avg_daily_rate	0.919236	0.893352	0.907686	1.000000	0.890556
total_jobs	0.992367	0.987963	0.899390	0.890556	1.000000

The coefficients for **logan_passengers**, **hotel_occup_rate**, and **hotel_avg_daily_rate** are particularly noteworthy.

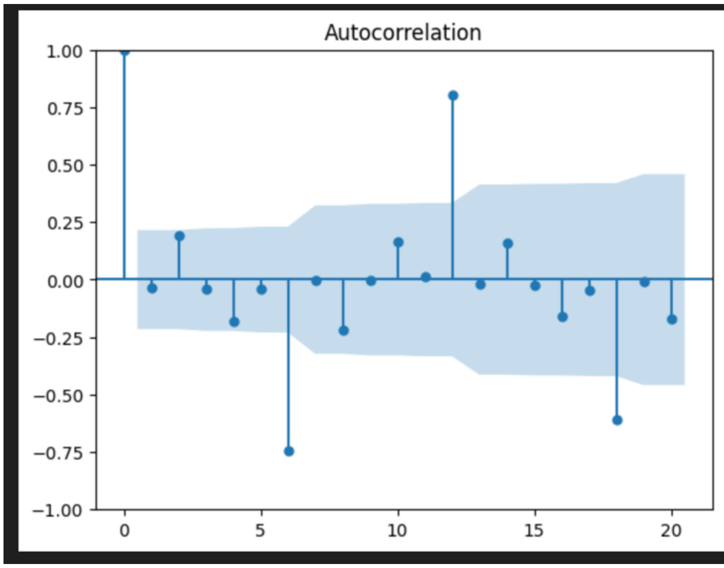
- The positive coefficient for **logan_passengers** (0.0430) suggests that an increase in the number of passengers is associated with an increase in total jobs. This is intuitive as more passengers can lead to more jobs in the airport and related industries such as tourism and services.

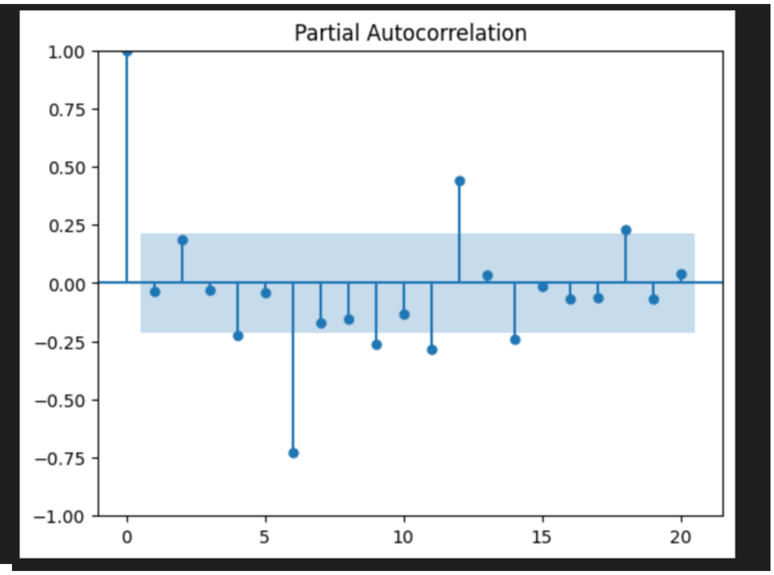
- Conversely, the coefficient for **hotel_occup_rate** is negative (-1.581e+05), indicating an inverse relationship between hotel occupancy rates and job numbers. This could suggest that as hotels are more occupied, perhaps because of higher visitor numbers, there could be a downturn in other job sectors, or it could be a statistical anomaly due to multicollinearity or other data issues.
- The **hotel_avg_daily_rate** has a small positive coefficient (89.1268), which was not statistically significant at the 0.05 level (p-value of 0.076), suggesting a weak relationship between average daily rates of hotels and job numbers. This implies that the average rate of hotel stays has a negligible direct effect on job creation.

```
...                            OLS Regression Results
==============================================================================
Dep. Variable:             total_jobs   R-squared:                       0.782
Model:                            OLS   Adj. R-squared:                  0.774
Method:                 Least Squares   F-statistic:                     95.94
Date:                Sat, 09 Dec 2023   Prob (F-statistic):           2.02e-26
Time:                        21:57:51   Log-Likelihood:                -889.33
No. Observations:                  84   AIC:                             1787.
Df Residuals:                      80   BIC:                             1796.
Df Model:                           3
Covariance Type:            nonrobust
========================================================================================
                           coef    std err          t      P>|t|      [0.025      0.975]
----------------------------------------------------------------------------------------
const                 3.338e+05   9350.265     35.697      0.000    3.15e+05    3.52e+05
logan_passengers         0.0430      0.003     14.881      0.000       0.037       0.049
hotel_occup_rate     -1.581e+05   2.14e+04     -7.393      0.000   -2.01e+05   -1.16e+05
hotel_avg_daily_rate    89.1268     49.630      1.796      0.076      -9.640     187.893
==============================================================================
Omnibus:                       11.448   Durbin-Watson:                   1.150
Prob(Omnibus):                  0.003   Jarque-Bera (JB):               15.173
Skew:                           0.615   Prob(JB):                     0.000507
Kurtosis:                       4.680   Cond. No.                     6.46e+07
==============================================================================
```

The regression model, which predicts the total number of jobs (**total_jobs**), shows a strong R-squared value of 0.782. This indicates that approximately 78.2%

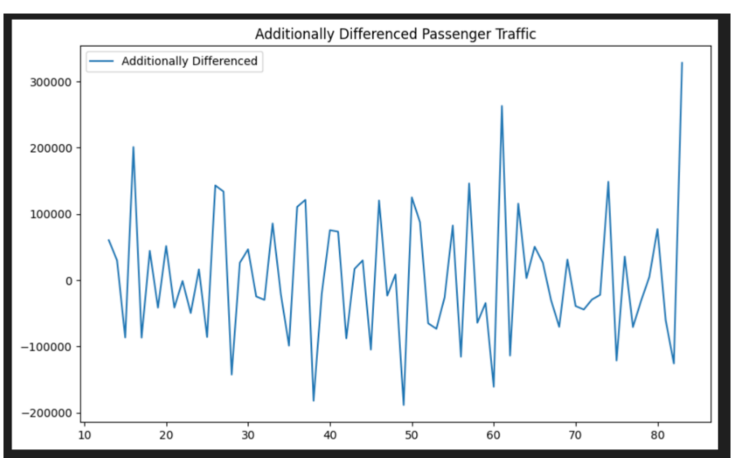
of the variability in the regional job numbers can be explained by the model's inputs, which is quite high in social science research, implying a strong model fit.





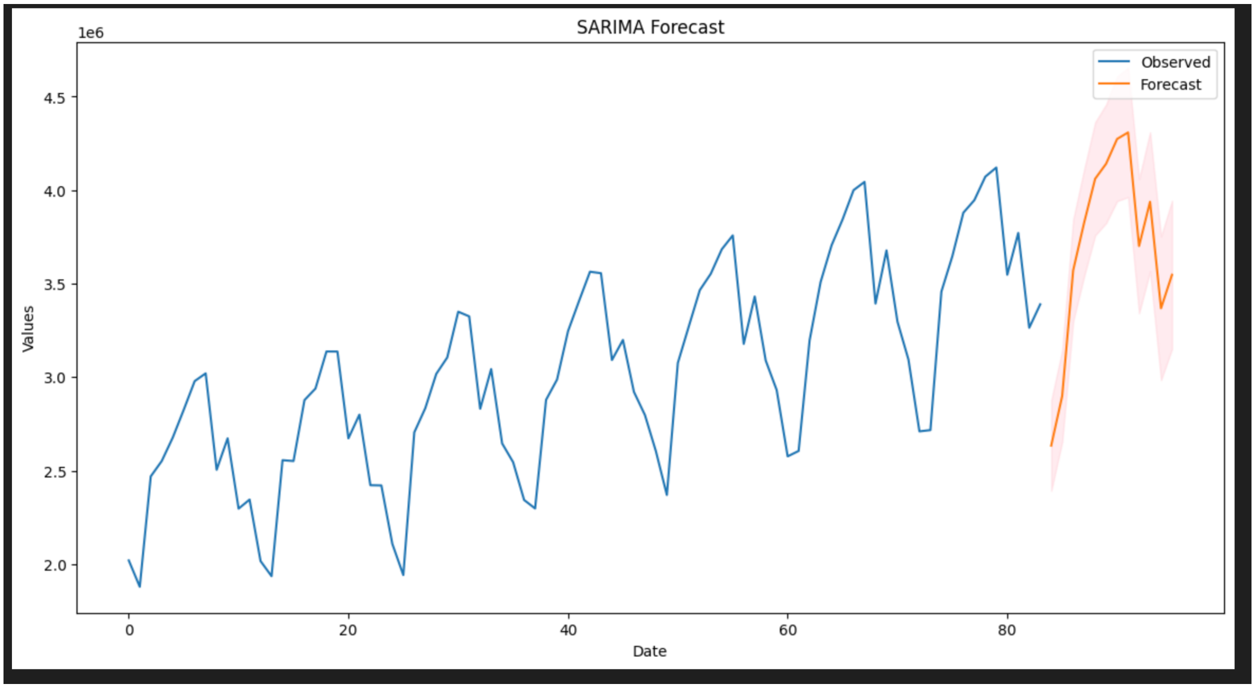
In ACF plot, if the autocorrelation values are significant (crossing the blue confidence interval bands) at the first few lags and then cut off, this would indicate an AR(p) component, where 'p' is the last significant lag. If the plot shows a slow decay, this would imply an MA(q) component in the data. Here q is 1.

Similarly, in PACF plot, a significant spike at the first few lags followed by a cutoff would suggest an AR(p) component, where 'p' is the last significant lag. If the PACF plot tails off, this would imply an MA(q) component. So p is 1.




```
ADF Statistic: -3.734964
p-value: 0.003647
Critical Values:
        1%: -3.527
        5%: -2.904
        10%: -2.589
The differenced series is stationary.
```

We can see that p-value is less than 0.05. So the differenced series is stationary.



The forecast we got from our SARIMA model tells us a lot about how busy Logan Airport will probably be in the future. It seems like the number of people coming and going from the airport will keep following the same ups and downs we've seen before. We have used the values p=q=P=Q=1 by ACF and PACF plots. And d=1 & D=2 as we differenced the seasonal twice and non-seasonal once.

Appendix C- Coding:

```
import pandas as pd
import matplotlib.pyplot as plt

# Load your data
data = pd.read_csv("/Users/tysonmukesh/Desktop/MTH-522/Project-3/economic-indicators.csv")

# Ensure 'Year' and 'logan_passengers' columns are correctly formatted
data['Year'] = pd.to_datetime(data['Year'], format='%Y').dt.year

# Aggregate passenger data by year
annual_passengers = data.groupby('Year')['logan_passengers'].sum()

# Plotting the trend
plt.figure(figsize=(12, 6))
plt.plot(annual_passengers, marker='o', linestyle='-')
plt.title('Yearly Passenger Traffic at Logan Airport')
plt.xlabel('Year')
plt.ylabel('Total Passengers')
plt.grid(True)
plt.show()
```

```python
import pandas as pd
import matplotlib.pyplot as plt

# Load your data
data = pd.read_csv("/Users/tysonmukesh/Desktop/MTH-522/Project-3/economic-indicators.csv")

# Ensure 'Year' and 'logan_intl_flights' columns are correctly formatted
data['Year'] = pd.to_datetime(data['Year'], format='%Y').dt.year

# Aggregate international flight data by year
annual_intl_flights = data.groupby('Year')['logan_intl_flights'].sum()

# Plotting the trend
plt.figure(figsize=(12, 6))
plt.plot(annual_intl_flights, marker='o', linestyle='-', color='orange')
plt.title('Yearly International Flights at Logan Airport')
plt.xlabel('Year')
plt.ylabel('Total International Flights')
plt.grid(True)
plt.show()
```

```python
# Aggregate hotel data by year
annual_hotel_data = data.groupby('Year').agg({
    'hotel_occup_rate': 'mean',
    'hotel_avg_daily_rate': 'mean'
})

# Plotting the trends
fig, ax1 = plt.subplots(figsize=(12, 6))

color = 'tab:red'
ax1.set_xlabel('Year')
ax1.set_ylabel('Average Occupancy Rate', color=color)
ax1.plot(annual_hotel_data.index, annual_hotel_data['hotel_occup_rate'], color=color)
ax1.tick_params(axis='y', labelcolor=color)

ax2 = ax1.twinx()  # instantiate a second axes that shares the same x-axis
color = 'tab:blue'
ax2.set_ylabel('Average Daily Rate', color=color)
ax2.plot(annual_hotel_data.index, annual_hotel_data['hotel_avg_daily_rate'], color=color)
ax2.tick_params(axis='y', labelcolor=color)

fig.tight_layout()  # otherwise the right y-label is slightly clipped
plt.title('Yearly Hotel Industry Trends at Logan Airport')
plt.grid(True)
plt.show()
```

```python
# Aggregate job data by year
annual_jobs = data.groupby('Year')['total_jobs'].mean()

# Plotting the trend
plt.figure(figsize=(12, 6))
plt.plot(annual_jobs.index, annual_jobs, marker='o', linestyle='-', color='green')
plt.title('Yearly Total Jobs in the Region')
plt.xlabel('Year')
plt.ylabel('Total Jobs')
plt.grid(True)
plt.show()
```

```python
# Aggregate unemployment rate data by year
annual_unemp_rate = data.groupby('Year')['unemp_rate'].mean()

# Plotting the trend
plt.figure(figsize=(12, 6))
plt.plot(annual_unemp_rate.index, annual_unemp_rate, marker='o', linestyle='-', color='purple')
plt.title('Yearly Unemployment Rate Trends')
plt.xlabel('Year')
plt.ylabel('Average Unemployment Rate')
plt.grid(True)
plt.show()
```

```python
# Aggregate labor force participation rate data by year
annual_labor_force_part = data.groupby('Year')['labor_force_part_rate'].mean()

# Plotting the trend
plt.figure(figsize=(12, 6))
plt.plot(annual_labor_force_part.index, annual_labor_force_part, marker='o', linestyle='-', color='cyan')
plt.title('Yearly Labor Force Participation Rate Trends')
plt.xlabel('Year')
plt.ylabel('Average Labor Force Participation Rate')
plt.grid(True)
plt.show()
```

```python
# Aggregate square footage data by year
annual_sqft = data.groupby('Year')['pipeline_sqft'].sum()

# Plotting the trend
plt.figure(figsize=(12, 6))
plt.plot(annual_sqft.index, annual_sqft, marker='o', linestyle='-', color='brown')
plt.title('Yearly Square Footage of New Construction Projects')
plt.xlabel('Year')
plt.ylabel('Total Square Footage')
plt.grid(True)
plt.show()
```

```python
# Aggregate construction job data by year
annual_const_jobs = data.groupby('Year')['pipeline_const_jobs'].sum()

# Plotting the trend
plt.figure(figsize=(12, 6))
plt.plot(annual_const_jobs.index, annual_const_jobs, marker='o', linestyle='-', color='blue')
plt.title('Yearly Construction Job Creation Trends')
plt.xlabel('Year')
plt.ylabel('Total Construction Jobs Created')
plt.grid(True)
plt.show()
```

```python
# Prepare the data for analysis
data['Year'] = pd.to_datetime(data['Year'], format='%Y').dt.year
annual_data = data.groupby('Year').agg({
    'logan_passengers': 'sum',
    'hotel_occup_rate': 'mean',
    'hotel_avg_daily_rate': 'mean',
    'total_jobs': 'mean'
}).reset_index()

# Correlation matrix
correlation_matrix = annual_data.corr()
print('Correlation Matrix:\n', correlation_matrix)
```

```python
# Prepare the independent variables (X) and the dependent variable (y)
X = data[['logan_passengers', 'hotel_occup_rate', 'hotel_avg_daily_rate']]
y = data['total_jobs']

# Add a constant to the model (the intercept)
X = sm.add_constant(X)

# Fit the regression model
model = sm.OLS(y, X).fit()

# Get the summary of the regression
summary = model.summary()
print(summary)
```

```python
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
data = pd.read_csv("/Users/tysonmukesh/Desktop/MTH-522/Project-3/economic-indicators.csv")

from statsmodels.tsa.stattools import adfuller

# Perform the Augmented Dickey-Fuller test to test for stationarity
# First difference of the non-seasonal data
data['logan_passengers_diff'] = data['logan_passengers'].diff()

# Drop the NA values introduced by differencing
data_diff_dropped = data.dropna()

# Perform the ADF test again on the differenced data
adf_test_diff = adfuller(data_diff_dropped['logan_passengers_diff'])

print('ADF Statistic: %f' % adf_test_diff[0])
print('p-value: %f' % adf_test_diff[1])
print('Critical Values:')
for key, value in adf_test_diff[4].items():
    print('\t%s: %.3f' % (key, value))

# If the p-value is less than 0.05, we reject the null hypothesis (the series is stationary)
if adf_test_diff[1] < 0.05:
    print("The differenced series is stationary.")
else:
    print("The differenced series is not stationary.")

# Plot the ACF and PACF on the first differenced data
plot_acf(data_diff_dropped['logan_passengers_diff'].dropna())
plt.show()

plot_pacf(data_diff_dropped['logan_passengers_diff'].dropna())
plt.show()
```

```python
import pandas as pd
from statsmodels.tsa.statespace.sarimax import SARIMAX
import matplotlib.pyplot as plt

# Load your data
data = pd.read_csv("/Users/tysonmukesh/Desktop/MTH-522/Project-3/economic-indicators.csv")

# Define the order and seasonal_order for your SARIMA model (use your chosen parameters here)
order = (1, 1, 1)  # Example: (p, d, q)
seasonal_order = (1, 2, 1, 12)  # Example: (P, D, Q, s)

# Fit the SARIMA model
model = SARIMAX(data['logan_passengers'],
                order=order,
                seasonal_order=seasonal_order,
                enforce_stationarity=False,
                enforce_invertibility=False)
results = model.fit()

# Plot the diagnostics
results.plot_diagnostics(figsize=(15, 12))
plt.show()

# Forecast the next 12 steps ahead in future
forecast_steps = 12
forecast = results.get_forecast(steps=forecast_steps)

# Get confidence intervals of forecasts
conf_int = forecast.conf_int()

# Plot the time series and forecasts of its future values
plt.figure(figsize=(14, 7))
plt.plot(data.index, data['logan_passengers'], label='Observed')
plt.plot(forecast.predicted_mean.index, forecast.predicted_mean, label='Forecast')
plt.fill_between(forecast.predicted_mean.index,
                 conf_int.iloc[:, 0],
                 conf_int.iloc[:, 1], color='pink', alpha=0.3)
plt.title('SARIMA Forecast')
plt.xlabel('Date')
plt.ylabel('Values')
plt.legend()
plt.show()
```

```python
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
data = pd.read_csv("/Users/tysonmukesh/Desktop/MTH-522/Project-3/economic-indicators.csv")

from statsmodels.tsa.stattools import adfuller

# Perform the Augmented Dickey-Fuller test to test for stationarity
# First difference of the non-seasonal data
data['logan_passengers_diff'] = data['logan_passengers'].diff(12).diff(12)

# Drop the NA values introduced by differencing
data_diff_dropped = data.dropna()

# Perform the ADF test again on the differenced data
adf_test_diff = adfuller(data_diff_dropped['logan_passengers_diff'])

print('ADF Statistic: %f' % adf_test_diff[0])
print('p-value: %f' % adf_test_diff[1])
print('Critical Values:')
for key, value in adf_test_diff[4].items():
    print('\t%s: %.3f' % (key, value))

# If the p-value is less than 0.05, we reject the null hypothesis (the series is stationary)
if adf_test_diff[1] < 0.05:
    print("The differenced series is stationary.")
else:
    print("The differenced series is not stationary.")

# Plot the ACF and PACF on the first differenced data
plot_acf(data_diff_dropped['logan_passengers_diff'].dropna())
plt.show()

plot_pacf(data_diff_dropped['logan_passengers_diff'].dropna())
plt.show()
```

Contributions:

- **Mukesh Kumar Karanam Rameshbabu –** Worked on Findings, coding, Discussions, Methods and Results.
- **Sai Sudhamsh Kamisetty –** Worked on issues, coding, graphs and results.
- **Rohith Rasi Reddy –** Worked on initial cleaning of data using excel and coding.
- **Anish Krishna Kalisetti –** Worked on results and report preparation.