

Evolving Employment Landscape: An Analysis of New York State's Job Market (1990 - Present)



IE6600 – Computation & Visualization for Analytics

Presented By: Half Coders

Navisha Shetty

Prateek Shetty

Rishab Radesh

Shalini Dutta

Table of Contents

Sl no	Description	Page No.
1.	Introduction	3
2.	Raw Data	3
3.	Data Processing	4
4.	Data Visualizations – Line Graphs	5
5.	Data Visualizations – Line Graphs: Employment Trend over the years in top 5 Sectors	6
6.	Data Visualizations – Line Graphs: Employment Trend over the years in bottom 5 Sectors	8
7.	Data Visualizations – Bar Graphs: Employment Trend across top 5 & bottom 5 sectors in the year 2022	10
8.	Data Visualizations – Line Graphs: Employment Trend across all years for top 5 areas	12
9.	Data Visualizations – Line Graphs: Employment Trend across all years for bottom 5 areas	14
10.	Data Visualizations – Bar Graphs: Employment Trend across top 5 & bottom 5 areas in the year 2022	16
11.	Data Visualizations – Pie Chart: Employment Trend of Top 10 sectors in New York State	18
12.	Data Visualizations – Pie Chart: Employment Trend of Top 10 sectors in Hamilton County	20
13.	Data Visualizations – Heat Map	22
14.	Data Visualizations – Heat Map: Employment Variations Over Time in Service-Providing Sector	22
15.	Data Visualizations – Single Line Graph: Employment Variations Over Time in Service-Providing Sector Seasonal Employment Trends in Manufacturing in 2022	24
16.	Unique Findings	25
17.	Conclusion	26
18.	References	26

Introduction:

The significance of understanding employment trends in a bustling metropolis like New York City is paramount, especially as these trends offer a window into the city's economic health and societal shifts. Leveraging the power of data analysis and visualization, this presentation aims to dissect the multifaceted employment landscape of New York City from 1990 to the present. Utilizing Python for data processing and analysis, coupled with effective visualization tools, we delve into the Current Employment Statistics dataset obtained from Data.gov. Our exploration will highlight the evolving job market, identify the top sectors driving employment, and reveal the impact of economic and policy changes on the workforce. By navigating through these employment patterns, we aim to provide a comprehensive understanding of the job dynamics in New York City, aiding stakeholders in making informed decisions for future growth and resilience.

RAW Data:

The raw data consists of 482584 rows × 7 columns.

Area	Area Name	Series	Title	Year	Month	Current Employment
93561	New York City	65621400	Outpatient Care Centers	1990	4	9300
93561	New York City	65621400	Outpatient Care Centers	1990	3	8600
93561	New York City	65621400	Outpatient Care Centers	1990	2	8500
93561	New York City	65621400	Outpatient Care Centers	1990	1	8400
93561	New York City	65621400	Outpatient Care Centers	1990	0	9100
93561	New York City	65621600	Home Health Care Services	1990	12	16400
93561	New York City	65621600	Home Health Care Services	1990	11	16200
93561	New York City	65621600	Home Health Care Services	1990	10	16200
93561	New York City	65621600	Home Health Care Services	1990	9	15300
93561	New York City	65621600	Home Health Care Services	1990	8	15200
93561	New York City	65621600	Home Health Care Services	1990	7	15200
93561	New York City	65621600	Home Health Care Services	1990	6	15500
93561	New York City	65621600	Home Health Care Services	1990	5	15400
93561	New York City	65621600	Home Health Care Services	1990	4	15100
93561	New York City	65621600	Home Health Care Services	1990	3	15500
93561	New York City	65621600	Home Health Care Services	1990	2	15600
93561	New York City	65621600	Home Health Care Services	1990	1	15600
93561	New York City	65621600	Home Health Care Services	1990	0	15600
93561	New York City	65622000	Hospitals	1990	12	145600
93561	New York City	65622000	Hospitals	1990	11	145400
93561	New York City	65622000	Hospitals	1990	10	144200
93561	New York City	65622000	Hospitals	1990	9	144900
93561	New York City	65622000	Hospitals	1990	8	144200
93561	New York City	65622000	Hospitals	1990	7	143500
93561	New York City	65622000	Hospitals	1990	6	143400
93561	New York City	65622000	Hospitals	1990	5	142700
93561	New York City	65622000	Hospitals	1990	4	142300
93561	New York City	65622000	Hospitals	1990	3	141100
93561	New York City	65622000	Hospitals	1990	2	139900
93561	New York City	65622000	Hospitals	1990	1	140200
93561	New York City	65622000	Hospitals	1990	0	143100
93561	New York City	65622100	General Medical and Surgical Hospitals	1990	12	133200
93561	New York City	65622100	General Medical and Surgical Hospitals	1990	11	133000
93561	New York City	65622100	General Medical and Surgical Hospitals	1990	10	132000
93562	Orange-Rockland-Westchester Metro Area	00000000	Total Nonfarm	1998	0	620400
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	12	528000
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	11	523500
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	10	520600
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	9	518000
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	8	519200
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	7	520000
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	6	521400
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	5	514000
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	4	507300
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	3	503700
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	2	498800
93562	Orange-Rockland-Westchester Metro Area	05000000	Total Private	1998	1	497900
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	0	514400
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	12	75900
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	11	76200
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	10	76300
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	9	76400
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	8	76100
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	7	75500
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	6	76100
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	5	75200
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	4	74100
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	3	72700
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	2	71800
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	1	71500
93562	Orange-Rockland-Westchester Metro Area	06000000	Goods Producing	1998	0	74800
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	12	559800
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	11	555200
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	10	550300
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	9	544400
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	8	545800
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	7	549200
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	6	553400
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	5	546400
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	4	539900
93562	Orange-Rockland-Westchester Metro Area	07000000	Service-Providing	1998	3	538100

Data Processing:

```
In [27]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
data = pd.read_csv("/Users/shalinidutta/Desktop/Current_Employment_Statistics__Beginning_1990.csv")
```

```
In [28]: data
```

```
Out[28]:
```

	Area	Area Name	Series	Title	Year	Month	Current Employment
0	93561	New York City	65621400	Outpatient Care Centers	1990	4	9300
1	93561	New York City	65621400	Outpatient Care Centers	1990	3	8600
2	93561	New York City	65621400	Outpatient Care Centers	1990	2	8500
3	93561	New York City	65621400	Outpatient Care Centers	1990	1	8400
4	93561	New York City	65621400	Outpatient Care Centers	1990	0	9100
...
482579	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	4	41000
482580	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	3	40600
482581	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	2	40400
482582	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	1	39600
482583	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	0	39200

482584 rows × 7 columns

- Top 5 Sectors:

Based on the data set, below are the top 5 sectors across New York City from 1990 – 2022

- Information
- Service-Providing
- Financial Activities
- Trade, Transportation and Utilities
- Professional and Business Services

```
In [32]: def top_values(df, column_name, top_n=5):
    # Use value_counts to count occurrences and sort the values
    occurrences_count = df[column_name].value_counts().sort_values(ascending=False)

    # Select the top n occurrences
    top_occurrences = occurrences_count.head(top_n)

    return top_occurrences

# Specify the column for which you want to find the top occurrences
target_column = 'Title'

# Get the top 5 values with the maximum occurrences in the specified column
top_5_values = top_values(df, target_column, top_n=5)

# Print or use the top values as needed
print(f"Top 5 values with the maximum occurrences in column 'Title':")
print(top_5_values)
```

Top 5 values with the maximum occurrences in column 'Title':

Information	17640
Service-Providing	17640
Financial Activities	17640
Trade, Transportation, and Utilities	17640
Professional and Business Services	17640

Name: count, dtype: int64

- Bottom 5 Sectors:

Based on the data set, below are the bottom 5 sectors across New York City from 1990 – 2022

- Information
- Service-Providing
- Financial Activities
- Trade, Transportation and Utilities

V. Professional and Business Services

```
In [33]: def least_values(df, column_name, least_n=5):
    # Use value_counts to count occurrences and sort the values
    occurrences_count = df[column_name].value_counts().sort_values(ascending=True)

    # Select the top n occurrences
    least_occurrences = occurrences_count.head(least_n)

    return least_occurrences

# Specify the column for which you want to find the top occurrences
target_column = 'Title'

# Get the top 5 values with the maximum occurrences in the specified column
least_5_values = least_values(df, target_column, least_n=5)

# Print or use the top values as needed
print(f"Top 5 values with the maximum occurrences in column 'Title':")
print(least_5_values)
```

Top 5 values with the maximum occurrences in column 'Title':

Title	count
Local Government excluding Educational Services	168
Specialty Food Retailers	246
Office Administrative Services	246
Business, Professional, Labor, Political, and Similar Organizations	246
Motor Vehicle Parts Manufacturing	272

Name: count, dtype: int64

Data Cleaning:

The dataset remains unchanged, ensuring reliability for analysis and insights.

```
In [29]: def clean_dataset(data):
    # Handling missing values
    data = data.dropna() # Remove rows with any missing values
    # Removing duplicates
    data = data.drop_duplicates()
    return data

df = clean_dataset(data)
```

```
In [30]: df
```

```
Out[30]:
```

	Area	Area Name	Series	Title	Year	Month	Current Employment
0	93561	New York City	65621400	Outpatient Care Centers	1990	4	9300
1	93561	New York City	65621400	Outpatient Care Centers	1990	3	8600
2	93561	New York City	65621400	Outpatient Care Centers	1990	2	8500
3	93561	New York City	65621400	Outpatient Care Centers	1990	1	8400
4	93561	New York City	65621400	Outpatient Care Centers	1990	0	9100
...
482579	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	4	41000
482580	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	3	40600
482581	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	2	40400
482582	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	1	39600
482583	93562	Orange-Rockland-Westchester Metro Area	90931611	Local Government Educational Services	1990	0	39200

482584 rows x 7 columns

Data Visualizations:

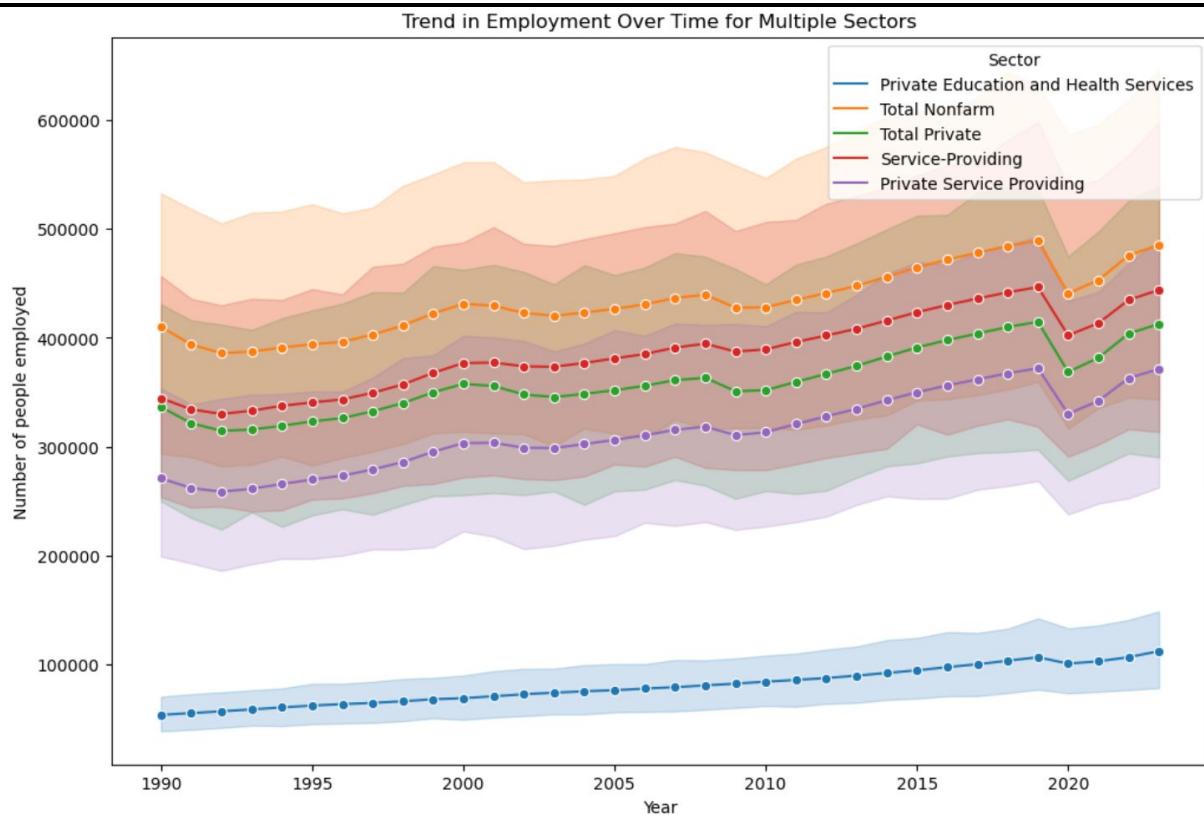
1. Line & Bar Graphs:

- Employment Trend over the years in top 5 Sectors:

```
In [37]: sectors_of_interest = ['Total Nonfarm', 'Service-Providing', 'Total Private', 'Private Service Providing', 'Private Education and Health Services']
# Filter the DataFrame for the specified sectors
sector_data = df[df['Title'].isin(sectors_of_interest)]
# Create a line plot to visualize trends over time for multiple sectors
plt.figure(figsize=(12, 8))
sns.lineplot(x='Year', y='Current Employment', hue='Title', data=sector_data, marker='o')

# Set plot labels and title
plt.title('Trend in Employment Over Time for Multiple Sectors')
plt.xlabel('Year')
plt.ylabel('Number of people employed')
plt.legend(title='Sector')

# Show the plot
plt.show()
```



The chart illustrates the employment trends over time for top 5 performing sectors within New York State. The data spans from 1990 to 2022, providing a long-term view of employment changes. The below are the interpretations from the chart.

i. Private Education and Health Services:

This sector shows a robust and steady upward trend, indicating significant growth in employment over the years. The broad shaded area around the trend line suggests there's a considerable degree of variability or confidence interval in the data.

ii. Total Nonfarm:

This line aggregates the total employment across all nonfarm sectors. It shows an upward trend, which indicates overall growth in employment, though with some periods of stagnation or minor decline.

iii. Total Private:

Reflecting the total private sector employment, this line follows a similar upward trend as the Total Nonfarm, showing growth in private sector jobs.

iv. Service-Providing:

The upward trend in this sector is indicative of the growing importance of service-oriented jobs within the state's economy.

v. Private Service Providing:

This sector, which represents a subset of the Service-Providing category that excludes government services, also demonstrates a clear upward trend, underscoring the growth in private service jobs.

Insights and Results:

- i. The substantial growth in Private Education and Health Services suggests an increasing demand for these services, potentially driven by demographic changes such as an aging population and a focus on higher education.
- ii. The general upward trend in Total Nonfarm and Total Private indicates economic growth and diversification in the state's job market.
- iii. The consistent increase in both Service-Providing and Private Service Providing sectors points to a shift toward a service-based economy, aligning with national and global economic trends.
- iv. The variability shown by the shaded areas indicates that while the overall trend is positive, there are fluctuations that may correspond to economic cycles, policy changes, or other external factors.

The graph illustrates the long-term growth and structural shifts in the New York State job market, with a particular emphasis on the expansion of service-oriented and private sector employment.

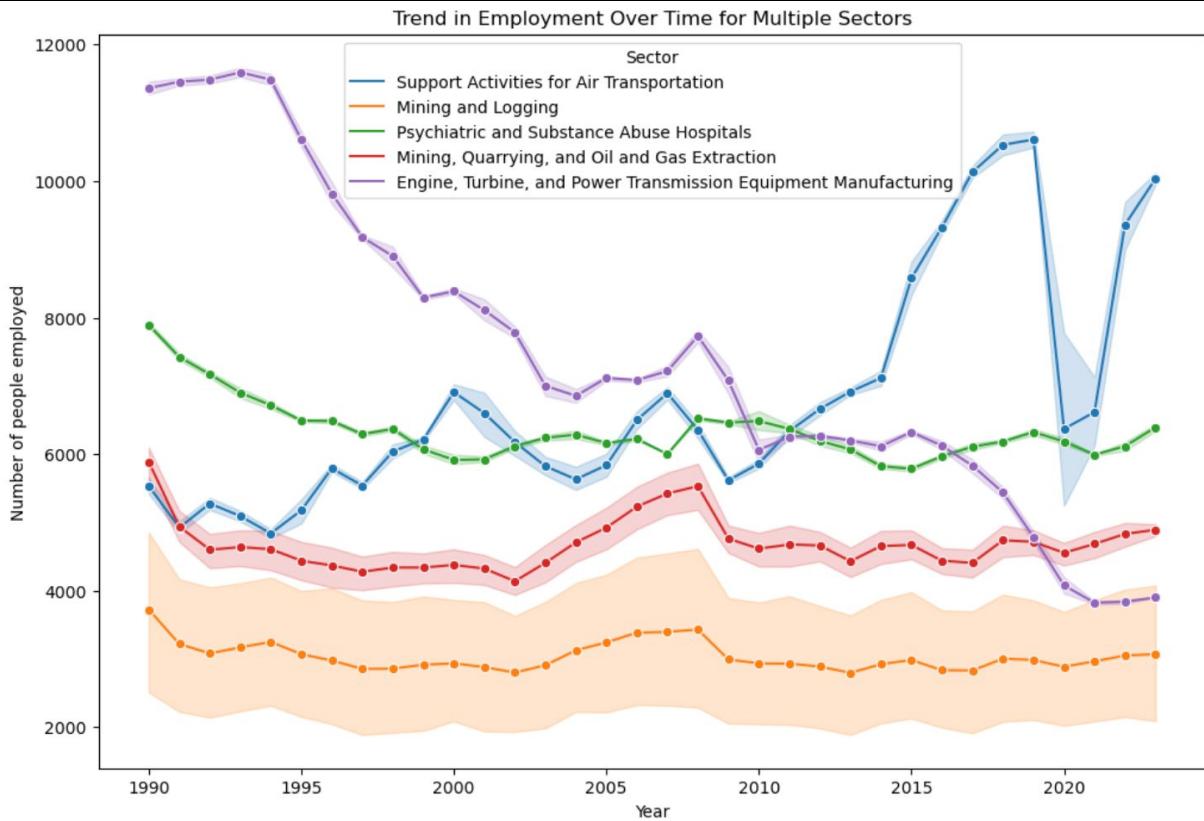
- Employment Trend over the years in bottom 5 Sectors:

```
In [38]: sectors_of_interest = ['Engine, Turbine, and Power Transmission Equipment Manufacturing', 'Support Activities for Air Transportation', 'Mining and Logging', 'Psychiatric and Substance Abuse Hospitals', 'Mining, Quarrying, and Oil and Gas Extraction']
sector_data = df[df['Title'].isin(sectors_of_interest)]

# Create a line plot to visualize trends over time for multiple sectors
plt.figure(figsize=(12, 8))
sns.lineplot(x='Year', y='Current Employment', hue='Title', data=sector_data, marker='o')

# Set plot labels and title
plt.title('Trend in Employment Over Time for Multiple Sectors')
plt.xlabel('Year')
plt.ylabel('Number of people employed')
plt.legend(title='Sector')

# Show the plot
plt.show()
```



The provided chart illustrates the employment trends over time for several sectors that are among the lower performing in terms of the number of people employed in New York State. The chart covers data from approximately 1990 to 2022. Below are the interpretations from the chart.

- i. Support Activities for Air Transportation:

This sector shows considerable volatility over time, with significant peaks and troughs. Notably, there's a sharp decline and subsequent recovery around the 2020 mark, which may be indicative of the impact of the COVID-19 pandemic on the air transportation industry.

- ii. Mining and Logging:

The employment in this sector has generally been declining over the years, which might reflect the decreasing importance or depletion of natural resources, or a shift towards more sustainable practices.

- iii. Psychiatric and Substance Abuse Hospitals:

The trend in this sector is relatively stable with a gentle upward trend, suggesting a consistent demand for mental health and substance abuse services.

iv. Mining, Quarrying, and Oil and Gas Extraction:

This sector's employment levels show a gradual decline with some variability, indicating possible impacts from economic factors, policy changes, or shifts in energy production methods.

v. Engine, Turbine, and Power Transmission Equipment Manufacturing:

This sector exhibits the most dramatic changes, with sharp increases and decreases in employment. The steep decline around 2020, followed by a recovery, again might be associated with the impact of the COVID-19 pandemic and subsequent economic recovery efforts.

Insights and Results:

- i. The significant fluctuations in the Support Activities for Air Transportation sector highlight its sensitivity to external shocks, such as global pandemics or economic downturns.
- ii. The decline in Mining and Logging, and Mining, Quarrying, and Oil and Gas Extraction, may signal a shift away from these industries within the state's economy, possibly due to environmental concerns or reduced demand.
- iii. The stability of employment in Psychiatric and Substance Abuse Hospitals indicates ongoing needs in these areas of healthcare.
- iv. The volatile nature of employment in Engine, Turbine, and Power Transmission Equipment Manufacturing suggests that this sector may be highly responsive to technological changes and global market conditions.

The graph highlights the challenges and opportunities within these less robust sectors in New York State, offering a detailed understanding of the factors influencing employment trends and the potential need for targeted economic support or workforce development.

- Employment Trend across top 5 & bottom 5 sectors in the year 2022:

```
In [39]: sectors_of_interest = ['Total Nonfarm', 'Service-Providing', 'Total Private', 'Private Service Providing', 'Private
# Filter the DataFrame for the specified sectors
sector_data = df[df['Title'].isin(sectors_of_interest)]
target_year = 2022

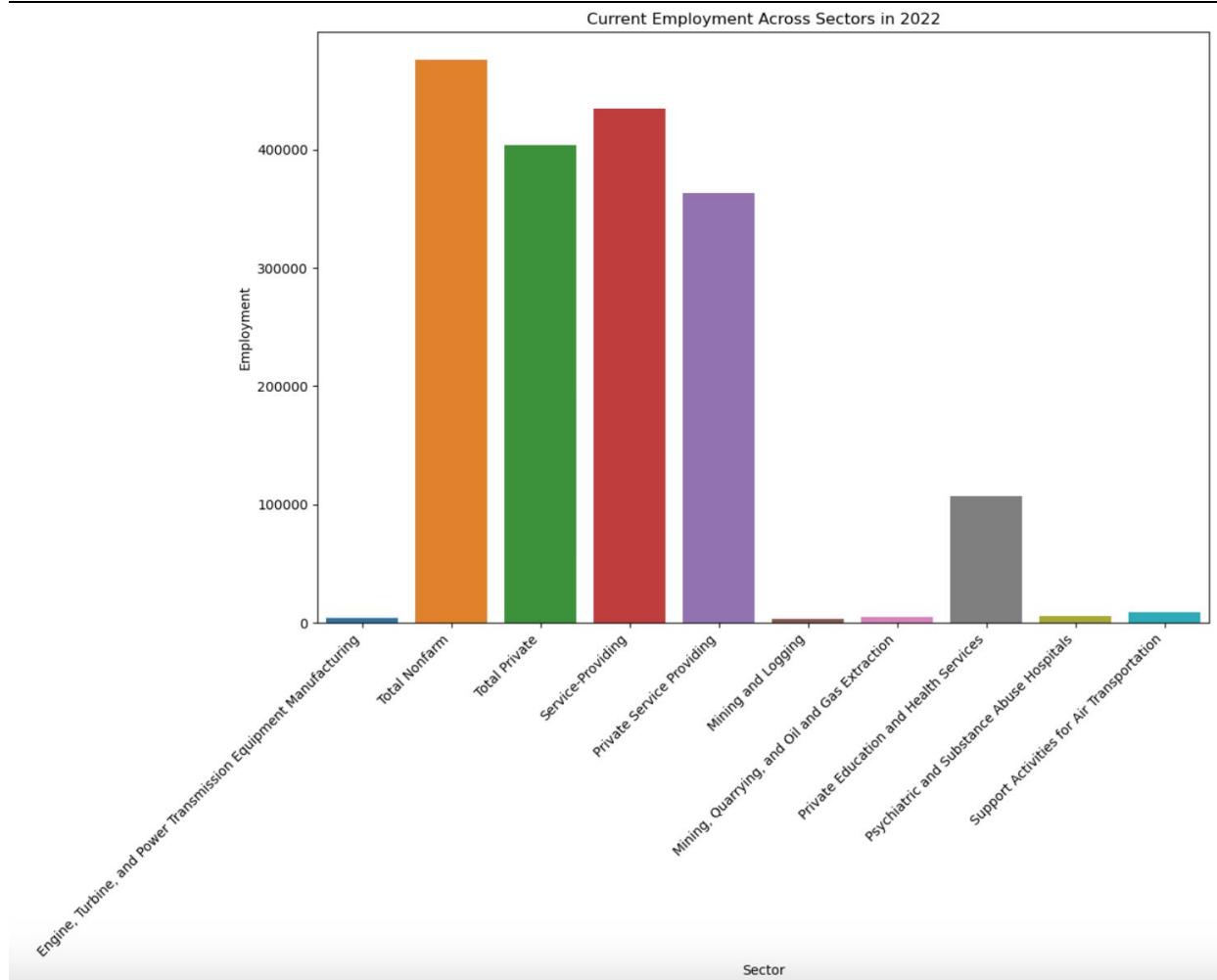
# Filter the DataFrame for the specified year or month
filtered_data = df[(df['Year'] == target_year) & df['Title'].isin(sectors_of_interest)]
# If you want to filter by month as well, uncomment the following line:
# filtered_data = filtered_data[filtered_data['Month'] == target_month]

# Create a bar chart to visualize current employment in different sectors
plt.figure(figsize=(12, 8))
sns.barplot(x='Title', y='Current Employment', data=filtered_data, ci=None)

# Set plot labels and title
plt.title('Current Employment Across Sectors in {target_year}')
plt.xlabel('Sector')
plt.ylabel('Employment')

# Rotate x-axis labels for better readability (optional)
plt.xticks(rotation=45, ha='right')

# Show the plot
plt.show()
```



This chart shows the current employment levels across top & bottom 5 sectors in New York State for the year 2022, represented as a bar chart. The chart enables us to compare employment figures across different sectors. Below are the interpretations from the chart:

- i. Total Nonfarm:
This sector has the highest employment, signifying that the nonfarm sector is the largest employer in the state. The nonfarm category typically includes a wide range of industries, excluding agricultural jobs.
- ii. Total Private:
Following closely, the Total Private sector shows a slightly lower level of employment than Total Nonfarm, which suggests that the private sector is a major driver of employment in the state.
- iii. Service-Providing:
The Service-Providing sector is also one of the leading sources of employment, which is consistent with trends toward a service-based economy.
- iv. Private Service Providing:
As a subset of the broader Service-Providing category, this sector also represents a significant portion of employment, indicating the strength of the private service industry.
- v. Mining and Logging:
This sector has a lower level of employment compared to the leading sectors, which could be due to the limited presence of these industries in the state's economy.
- vi. Mining, Quarrying, and Oil and Gas Extraction:
Similarly, this sector shows lower employment levels, further indicating that resource extraction is not a primary employment source within New York State.
- vii. Private Education and Health Services:
This sector has a modest level of employment, suggesting these services play a role in the state's job market.
- viii. Psychiatric and Substance Abuse Hospitals:
The employment in this sector is relatively low, which may reflect the specialized nature of these jobs.
- ix. Support Activities for Air Transportation:
This sector has the least employment among those listed, which could be due to the niche requirements of the industry.

Insights and Results:

- i. The dominance of the Total Nonfarm, Total Private, and Service-Providing sectors points to a diverse and robust job market, with a clear orientation toward services rather than manufacturing or resource extraction.
- ii. The substantial difference in employment numbers between these top sectors and those related to natural resources or specific health services highlights the economic structure and job distribution in the state.
- iii. Understanding the sector-specific employment levels can aid policymakers and business leaders in assessing the health of various industries and in making informed decisions about where to focus economic development efforts.

The graph provides stakeholders with a snapshot of the employment distribution across sectors in New York State for the year 2022, offering insights into the current state of the job market.

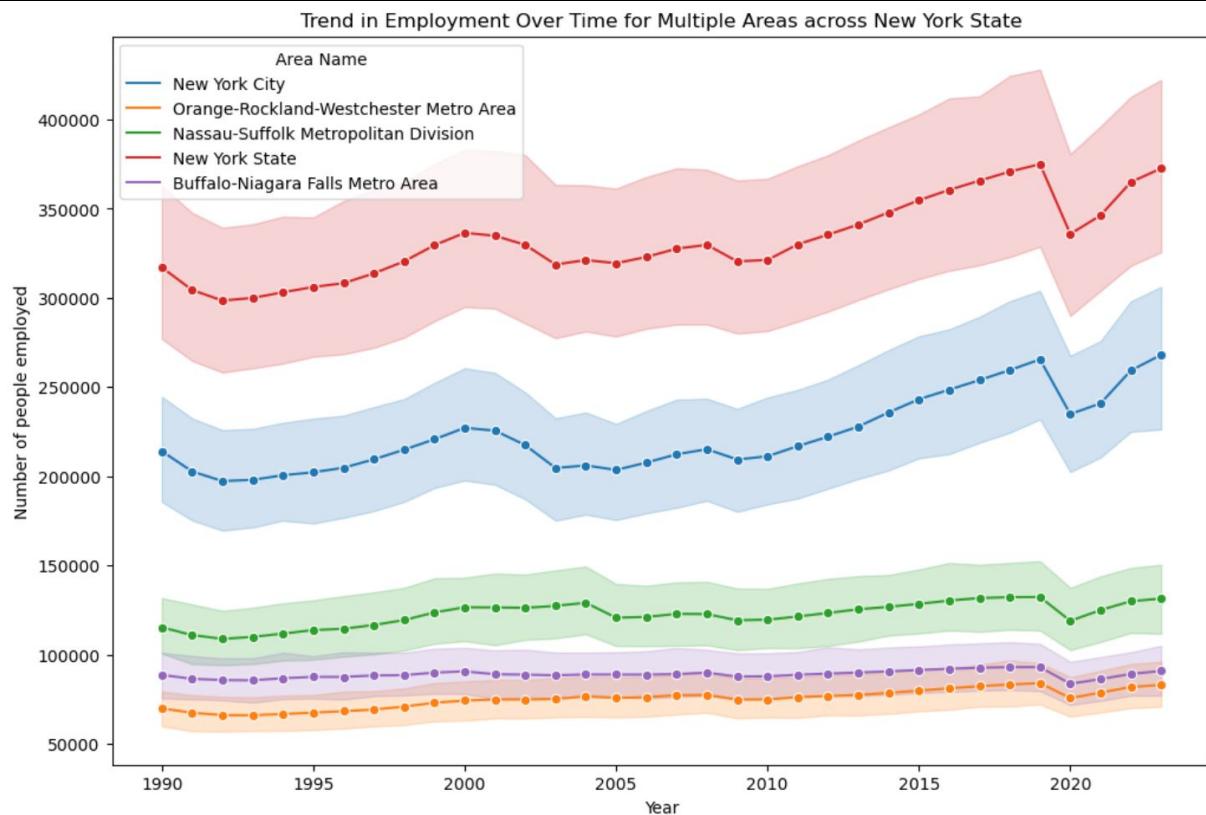
- Employment Trend across all years for top 5 areas:

```
In [42]: areas_of_interest = ['New York State', 'New York City', 'Nassau-Suffolk Metropolitan Division', 'Orange-Rockland-Westchester Metro Area', 'Buffalo-Niagara Falls Metro Area']
# Filter the DataFrame for the specified sectors
area_data = df[df['Area Name'].isin(areas_of_interest)]

# Create a line plot to visualize trends over time for multiple sectors
plt.figure(figsize=(12, 8))
sns.lineplot(x='Year', y='Current Employment', hue='Area Name', data=area_data, marker='o')

# Set plot labels and title
plt.title('New York Expanding Job Hubs: Employment Trend in Top Career Centers, 1990-2020')
plt.xlabel('Year')
plt.ylabel('Number of people employed')
plt.legend(title='Area Name')

# Show the plot
plt.show()
```



The chart depicts the trend in employment over time for multiple areas across New York State, focusing on the top-performing regions in terms of job numbers. Below are the interpretations from the chart:

- i. New York City:
As the line with the highest level of employment, New York City shows an overall increasing trend, with some variability. The shaded area indicates the level of uncertainty or variability in the data, but the general upward trend is clear, emphasizing New York City's role as the state's economic powerhouse.
- ii. Orange-Rockland-Westchester Metro Area:
This region also displays an upward trend in employment over time, although with a smaller workforce compared to New York City. The trend suggests growth in job opportunities in this metro area.

iii. Nassau-Suffolk Metropolitan Division:

The employment trend for this area shows stability with a slight upward trajectory, indicating consistent job growth over the years.

iv. Buffalo-Niagara Falls Metro Area:

Representing the smallest workforce among the regions shown, the Buffalo-Niagara Falls area has a relatively stable employment level with minor fluctuations over time.

v. New York State:

The trend line for the entire state shows an aggregate view, combining all areas. It indicates a steady increase in employment, with the shaded area suggesting variability across different regions within the state.

Insights and Results:

- i. New York City's significant and growing employment levels underscore its status as an economic hub, attracting diverse industries and offering a wide range of job opportunities.
- ii. The Orange-Rockland-Westchester Metro Area and Nassau-Suffolk Metropolitan Division both exhibit growth, although at different scales, reflecting regional economic development and possibly the expansion of suburban job markets.
- iii. The stability in the Buffalo-Niagara Falls Metro Area may suggest a mature economy with established industries, although the lack of substantial growth could point to potential areas for economic stimulus.
- iv. The overall increase in state employment demonstrates economic expansion across New York State, but the variability suggests that growth is unevenly distributed among regions.

The graph provides a comprehensive view of the employment trends in New York State's most populous areas, highlighting regional economic strengths and potential areas for development.

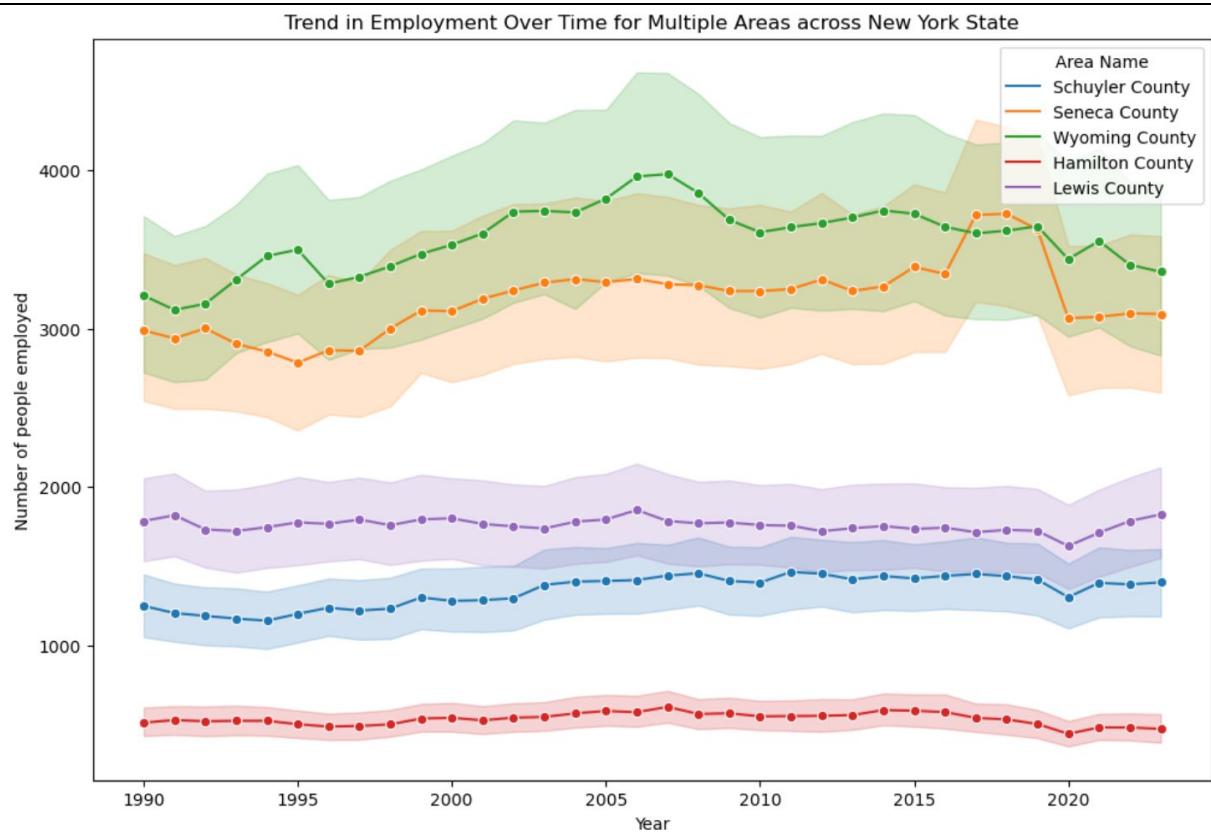
- Employment Trend across all years for bottom 5 areas:

```
In [43]: areas_of_interest = ['Wyoming County', 'Seneca County', 'Lewis County', 'Schuyler County', 'Hamilton County']
# Filter the DataFrame for the specified sectors
area_data = df[df['Area Name'].isin(areas_of_interest)]

# Create a line plot to visualize trends over time for multiple sectors
plt.figure(figsize=(12, 8))
sns.lineplot(x='Year', y='Current Employment', hue='Area Name', data=area_data, marker='o')

# Set plot labels and title
plt.title('New York Declining Job Hubs: Employment Trend in Bottom Career Centers, 1990–2020')
plt.xlabel('Year')
plt.ylabel('Number of people employed')
plt.legend(title='Area Name')

# Show the plot
plt.show()
```



The chart shows the trend in employment over time for several areas across New York State that are among the bottom performers in terms of the number of people employed. Below are the interpretations from the chart:

- i. Schuyler County:
This area shows a relatively stable employment level with minor fluctuations over the years. The shaded area indicates a degree of uncertainty or variability in the data.
- ii. Seneca County:
This county also exhibits stability in employment levels with slight increases and decreases over the period, but no significant growth or decline is apparent.
- iii. Wyoming County:
The trend for Wyoming County shows a gradual but inconsistent decrease in employment, suggesting potential challenges in job retention or creation.

iv. Hamilton County:

Employment levels in Hamilton County are relatively low, with some variability over time but no clear long-term trend.

v. Lewis County:

This county has the lowest employment numbers among the areas depicted, with a trend that shows minor fluctuations but overall low levels of employment.

Insights and Results:

- i. The stability observed in Schuyler and Seneca Counties suggests a consistent but limited job market, potentially reflecting the economic activities and opportunities available in these regions.
- ii. The gradual decrease in employment in Wyoming County could be a sign of economic challenges or a declining industry that may require attention from policymakers and economic development efforts.
- iii. The low and variable employment in Hamilton and Lewis Counties highlights the limited job opportunities in these areas, which might be due to their rural nature, smaller populations, or economic structures that do not support significant job growth.
- iv. Across these bottom-performing areas, the employment levels are considerably lower than in the more populous regions of the state, underscoring the economic disparities within New York State.

The graph provides insights into the employment trends in less economically vibrant areas of New York State, highlighting the need for targeted interventions to foster job growth and economic development.

- Employment Trend across top 5 & bottom 5 areas in the year 2022:

```
In [44]: areas_of_interest = ['New York State', 'New York City', 'Nassau-Suffolk Metropolitan Division', 'Orange-Rockland-Westchester Metro Area', 'Wyoming County', 'Hamilton County', 'Lewis County', 'Schuyler County', 'Seneca County']
# Filter the DataFrame for the specified sectors
area_data = df[df['Area Name'].isin(areas_of_interest)]
target_year = 2022

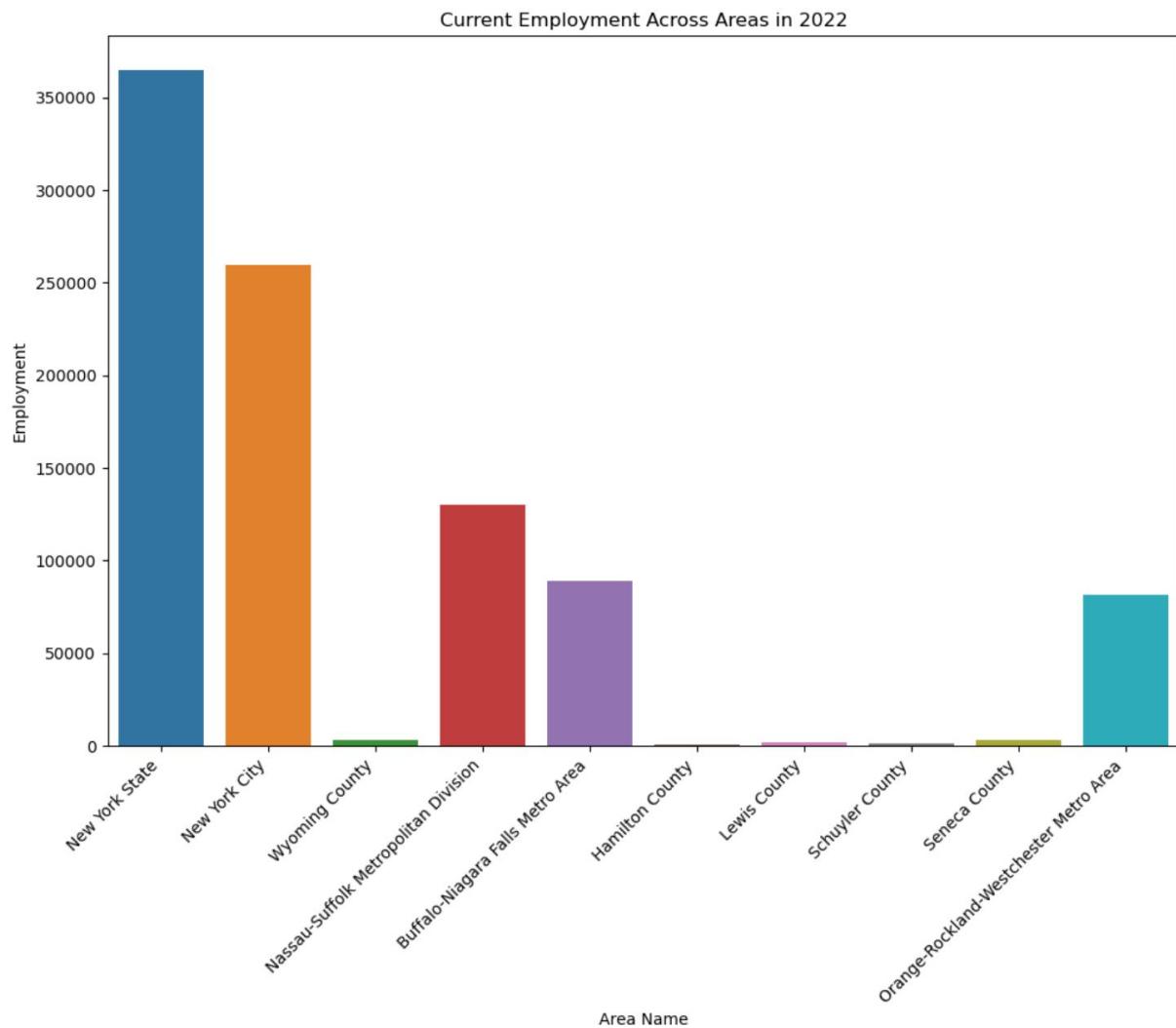
# Filter the DataFrame for the specified year or month
filtered_data = df[(df['Year'] == target_year) & df['Area Name'].isin(areas_of_interest)]
# If you want to filter by month as well, uncomment the following line:
# filtered_data = filtered_data[filtered_data['Month'] == target_month]

# Create a bar chart to visualize current employment in different sectors
plt.figure(figsize=(12, 8))
sns.barplot(x='Area Name', y='Current Employment', data=filtered_data, ci=None)

# Set plot labels and title
plt.title(f'Current Employment Across Areas in {target_year}')
plt.xlabel('Area Name')
plt.ylabel('Employment')

# Rotate x-axis labels for better readability (optional)
plt.xticks(rotation=45, ha='right')

# Show the plot
plt.show()
```



The chart presents current employment figures for the top and bottom five performing areas in New York State for the year 2022. Below are the interpretations from the chart:

- i. New York City:

Clearly the highest employer with a significant margin, indicating its role as the economic center of the state.

ii. New York State:

This bar represents the total employment across the entire state, providing context for the scale of employment in New York City relative to the entire state.

iii. Nassau-Suffolk Metropolitan Division:

Shows substantial employment, although less than half that of New York City, indicating it as a significant economic area within the state.

iv. Buffalo-Niagara Falls Metro Area:

Demonstrates a moderate level of employment, which is less than the Nassau-Suffolk Metropolitan Division but still significant.

v. Orange-Rockland-Westchester Metro Area:

Shows a similar level of employment to the Buffalo-Niagara Falls Metro Area, marking it as another important economic region in the state.

Moving to the bottom performing areas:

i. Wyoming County:

The employment level is significantly lower compared to the top areas, reflecting the economic scale of this region.

ii. Hamilton County:

Shows even lower employment, which could be due to various factors like population size, industrial presence, or economic opportunities.

iii. Lewis County:

Has a comparable employment level to Hamilton County, indicating similar economic conditions or challenges.

iv. Schuyler County:

The employment level is slightly higher than Hamilton and Lewis Counties but still among the lowest in the state.

v. Seneca County:

Has the highest employment among the bottom five areas but is significantly lower than the top performing areas.

Insights and Results:

- i. New York City's employment figures far outstrip other regions, reinforcing its status as an economic powerhouse with a diverse job market.
- ii. The Nassau-Suffolk and Buffalo-Niagara Falls areas also play substantial roles in the state's economy but have less than half the employment figures of New York City.
- iii. The bottom five areas, predominantly rural with smaller populations, have much lower levels of employment, highlighting the economic disparities within the state.
- iv. These disparities underline the potential need for targeted economic development and job creation strategies in less populated or rural counties.

The graph provides a stark comparison between the most and least economically active areas in terms of employment within New York State for 2022. This data can inform discussions on economic policy, workforce development, and resource allocation.

- Employment Trend of Top 10 sectors in New York State:

```
In [45]: import pandas as pd
import matplotlib.pyplot as plt

# Load the dataset
file_path = '/Users/shalinidutta/Desktop/Current_Employment_Statistics__Beginning_1990.csv'
employment_data = pd.read_csv(file_path)

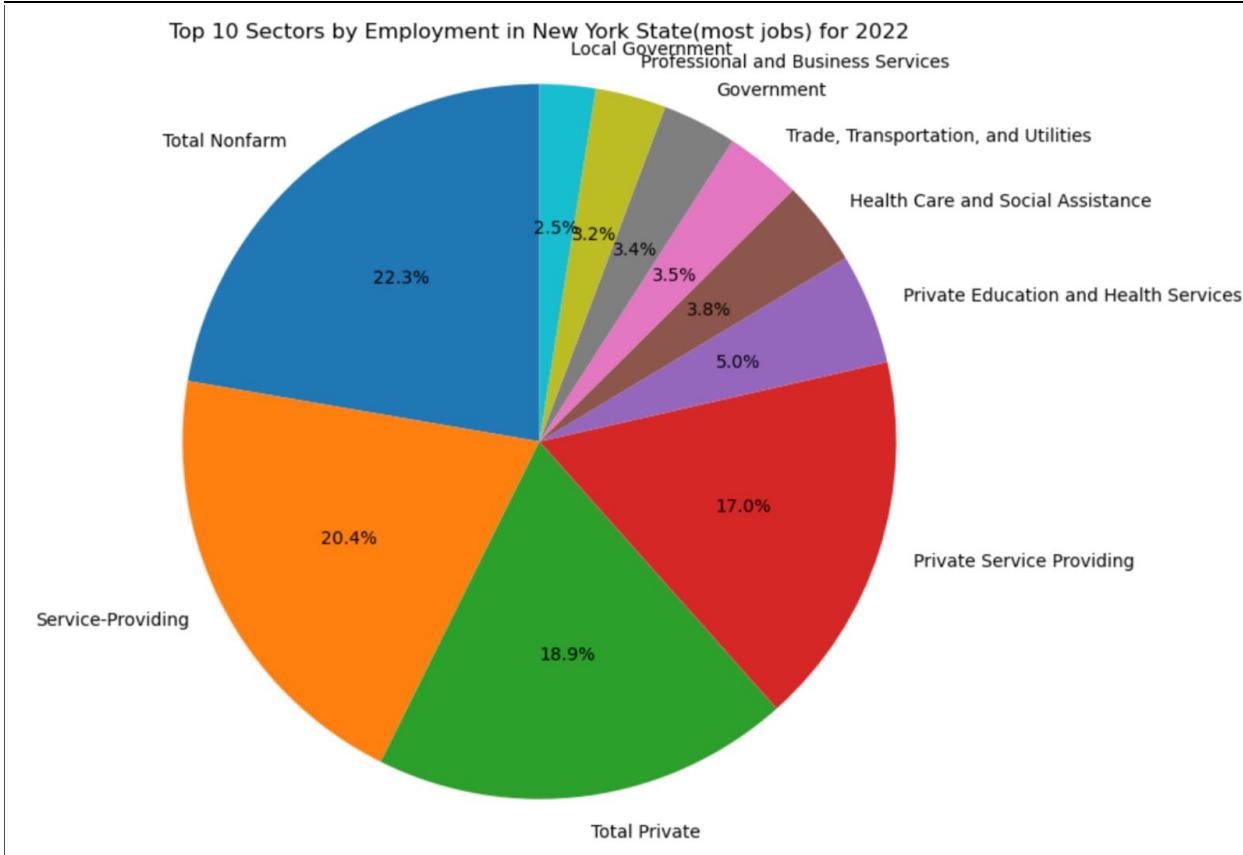
# Assuming the dataset contains 'Year', 'Area', 'Sector' (here assumed to be 'Title'), and 'Employment' (assumed to
# Adjust the year and area of interest as needed
year_of_interest = 2022
area_of_interest = 'New York State'

# Correcting column names based on assumptions
filtered_data = employment_data[(employment_data['Year'] == year_of_interest) &
                                  (employment_data['Area Name'] == area_of_interest)]

# Group by 'Sector' (assumed to be 'Title') to get the total employment per sector and sort to get the top 10 sector
sector_employment = filtered_data.groupby('Title')['Current Employment'].sum().sort_values(ascending=False).head(10)

# Plotting the pie chart for the top 10 sectors
plt.figure(figsize=(10, 8))
plt.pie(sector_employment, labels=sector_employment.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Top 10 Sectors by Employment in {area_of_interest} for {year_of_interest}')
plt.axis('equal') # Equal aspect ratio ensures that the pie chart is circular.

# Show the pie chart
plt.show()
```



The pie chart illustrates the proportion of employment by sector in New York State for the year 2022. Each sector's share of total employment is represented by a slice of the pie. Below are the interpretations from the chart:

- i. Service-Providing:

The largest sector, comprising 22.3% of employment, highlights the significance of service-oriented jobs in the state's economy.

- ii. Total Private:
Making up 20.4% of employment, this suggests a strong presence of private enterprises and businesses in the job market.
- iii. Government:
At 18.9%, the government sector is a major employer, reflecting jobs in public services, administration, and state-run institutions.
- iv. Health Care and Social Assistance:
This sector accounts for 17.0% of employment, emphasizing the importance of health and social services in the state's job landscape.
- v. Private Service Providing:
With 5.0%, this indicates that a significant portion of service jobs are within the private sector.
- vi. Private Education and Health Services:
At 3.8%, this slice shows the role of private educational and healthcare institutions as employers.
- vii. Trade, Transportation, and Utilities:
Accounting for 3.5% of employment, this sector is key to the state's infrastructure and commerce.
- viii. Professional and Business Services:
Comprising 3.4%, this reflects the jobs in professional, scientific, and technical services, as well as business management.
- ix. Local Government:
This makes up 2.5% of employment, which includes employment by local government bodies outside of state-level government jobs.
- x. Total Nonfarm:
This sector is listed at 2.2%, but this might be an error in the labeling of the chart, as "Total Nonfarm" typically refers to all employment excluding farm work, and it should be much larger. It is possible that this label is meant to represent a different, specific nonfarm sector.

Insights and Results:

- i. The service sector, both private and public, dominates the employment landscape in New York State, which aligns with broader trends towards a service-based economy.
- ii. Health care and social assistance, as well as education, are significant sources of employment, likely due to the state's large population and the increasing demand for health services.
- iii. The substantial government employment reflects New York's role as both a state capital and home to many government agencies and public institutions.
- iv. Sectors like trade, transportation, and utilities, as well as professional and business services, though smaller, are vital components of the state's economy, facilitating other sectors' operations and contributing to overall economic health.

This distribution of employment by sector provides valuable insight into the diverse and multifaceted job market of New York State for 2022, which shed light on the economic structure and labor trends.

- Employment Trend of Top 10 sectors in Hamilton County:

```
 In [46]: import pandas as pd
import matplotlib.pyplot as plt

# Load the dataset
file_path = '/Users/shalinidutta/Desktop/Current_Employment_Statistics__Beginning_1990.csv'
employment_data = pd.read_csv(file_path)

# Create filters for New York State and Hamilton County
ny_filter = (employment_data['Year'] == 2022) & (employment_data['Area Name'] == 'New York State')
hamilton_filter = (employment_data['Year'] == 2022) & (employment_data['Area Name'] == 'Hamilton County')

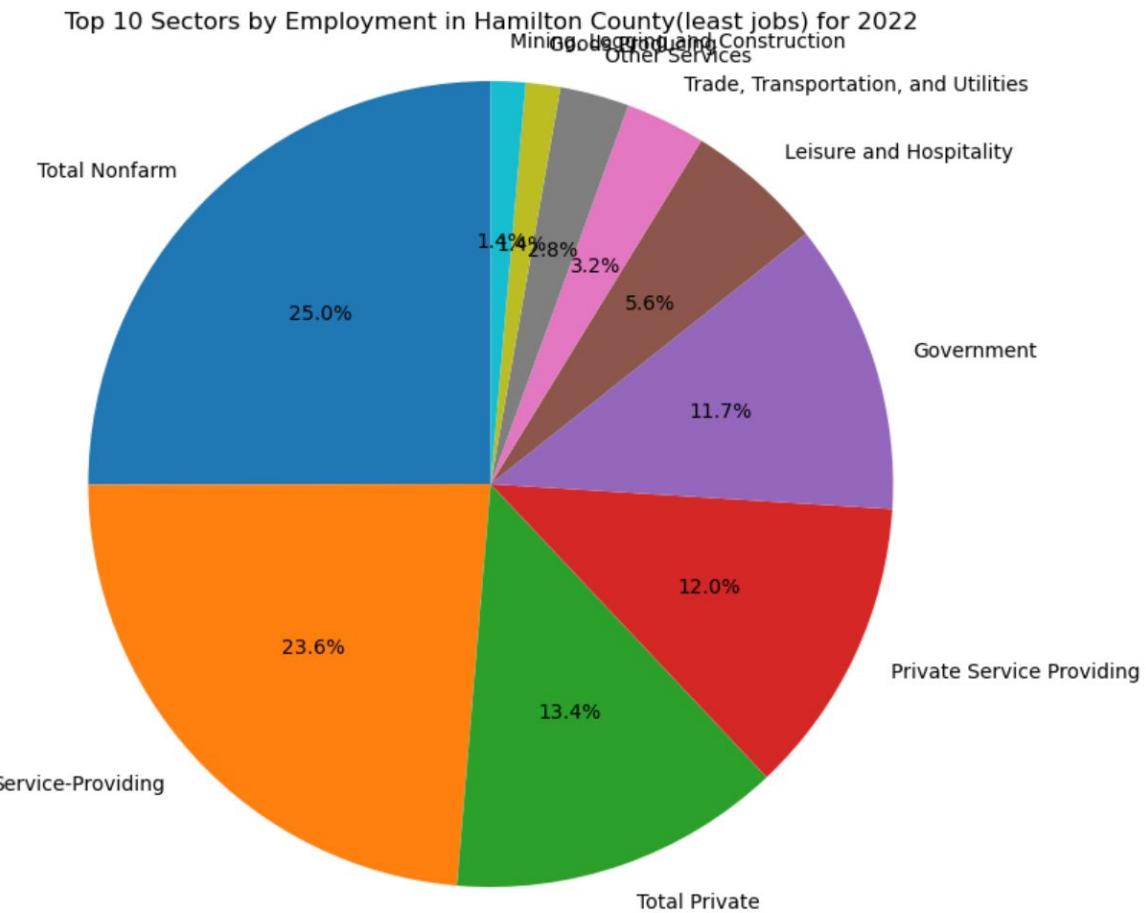
# Filter the data
ny_data = employment_data[ny_filter]
hamilton_data = employment_data[hamilton_filter]

# Group by sector and sort to get top 10 sectors
ny_sectors = ny_data.groupby('Title')['Current Employment'].sum().sort_values(ascending=False).head(10)
hamilton_sectors = hamilton_data.groupby('Title')['Current Employment'].sum().sort_values(ascending=False).head(10)

# Plot NY pie chart
plt.figure(figsize=(10,8))
plt.pie(ny_sectors, labels=ny_sectors.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Top 10 Sectors by Employment in New York State(most jobs) for 2022')
plt.axis('equal')

# Plot Hamilton pie chart
plt.figure(figsize=(10,8))
plt.pie(hamilton_sectors, labels=hamilton_sectors.index, autopct='%1.1f%%', startangle=90)
plt.title(f'Top 10 Sectors by Employment in Hamilton County(least jobs) for 2022')
plt.axis('equal')

# Show the plots
plt.show()
```



The pie chart shows the distribution of employment across the top 10 sectors in Hamilton County, New York State, for the year 2022. This county is noted for having one of the lowest

employment figures in the state. The chart details the percentage of total employment that each sector contributes. Below are the interpretations from the chart:

i. Service-Providing:

This sector comprises the largest share of employment at 25.0%, highlighting the importance of service jobs in Hamilton County.

ii. Total Private:

The second-largest sector at 23.6% reflects the combined employment across all private industries, indicating a robust private sector.

iii. Government:

With 13.4%, government jobs account for a significant portion of employment, likely including local, state, and possibly federal positions.

iv. Private Service Providing:

This sector, which includes private companies providing services, makes up 12.0% of employment, showing the importance of private services in the county's job market.

v. Leisure and Hospitality:

With 11.7%, this sector's share suggests that tourism and hospitality are important for the local economy, which is characteristic of areas with natural attractions or recreational offerings.

vi. Trade, Transportation, and Utilities:

Accounting for 5.6% of employment, this sector is essential for the functioning of the county's infrastructure and commerce.

vii. Mining, Quarrying, and Oil and Gas Extraction:

Although a smaller share at 3.2%, the presence of this sector indicates some level of natural resource extraction activities in the county.

viii. Construction:

Making up 2.9%, construction jobs contribute to the development and maintenance of infrastructure.

ix. Other Services:

At 2.8%, this category likely includes various sub-sectors providing miscellaneous services not captured in other categories.

x. Total Nonfarm:

This is marked at 1.4% and is likely mislabeled, as the term "Total Nonfarm" typically encompasses a much broader range of employment. This label might refer to a specific nonfarm sector that is not part of the other listed categories.

Insights and Results:

- i. The significant share of Service-Providing and Private Service Providing sectors indicates that service-oriented jobs are vital to Hamilton County's economy.
- ii. The prominence of the Leisure and Hospitality sector is indicative of the county's reliance on tourism and related industries for employment opportunities.
- iii. Government jobs form a crucial part of Hamilton County's employment landscape, which may be related to public services and administration being significant employers in less populated areas.
- iv. The smaller shares of sectors such as Mining and Construction reflect the presence of these industries but on a smaller scale compared to service-related sectors.

The pie chart depicts employment structure of Hamilton County, which has one of the lowest job counts in New York State, and to highlight the economic sectors that are key to its job market.

2. Heat Map:

- Employment Variations Over Time in Service-Providing Sector:

```
In [47]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Load the dataset
file_path = '/Users/shalinidutta/Desktop/Current_Employment_Statistics__Beginning_1990.csv'
df = pd.read_csv(file_path)

# Specify the sector of interest
sector_of_interest = 'Service-Providing'

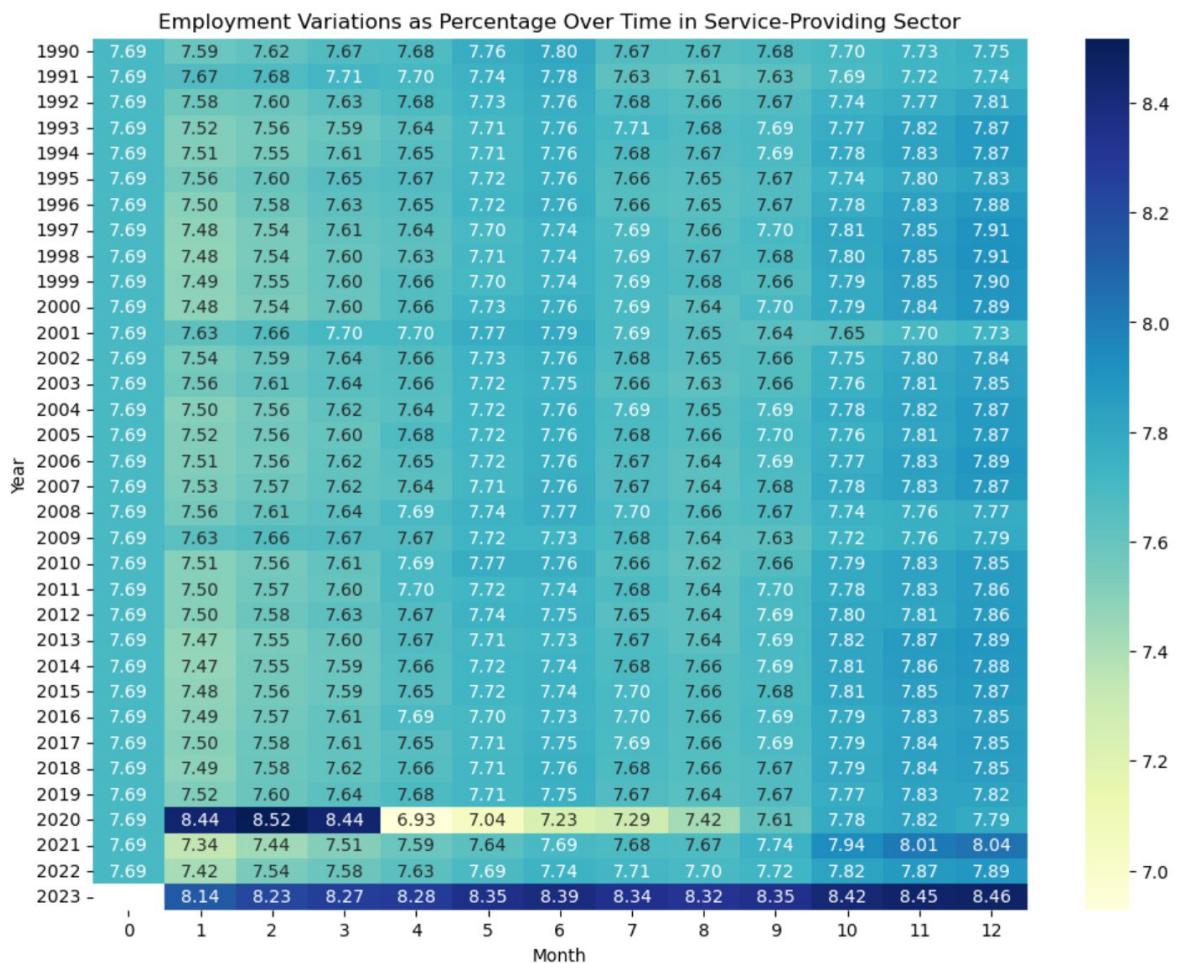
# Filter data for the sector of interest
filtered_data = df[df['Title'] == sector_of_interest]

# Create a pivot table with 'Year' and 'Month' as indices and 'Current Employment' as values, summing up for each year
heatmap_data = filtered_data.pivot_table(values='Current Employment', index='Year', columns='Month', aggfunc='sum')

# Convert the numbers to percentages of their row total (i.e., total annual employment for each year)
heatmap_data_percentage = heatmap_data.apply(lambda x: x/x.sum()*100, axis=1)

# Plotting the heatmap with percentages
plt.figure(figsize=(12, 9))
sns.heatmap(heatmap_data_percentage, cmap='YlGnBu', annot=True, fmt=".2f")
plt.title(f'Employment Variations as Percentage Over Time in {sector_of_interest} Sector')
plt.ylabel('Year')
plt.xlabel('Month')

# Show the heatmap
plt.show()
```



The heatmap displays employment variations as a percentage over time in the service-providing sector of New York State from 1990 to 2023. Each cell represents the percentage of employment for a particular month and year, with the color intensity reflecting the magnitude of the employment percentage. Below are the interpretations from the chart:

i. Color Gradient:

The color gradient ranges from light blue to dark blue, with lighter colors indicating lower employment percentages and darker colors indicating higher percentages.

ii. Annual Trends:

Each row represents a year, and the gradual change in color from left to right within a row indicates the seasonal variations in employment percentages throughout that year.

iii. Seasonal Patterns:

There is a visible pattern of seasonal fluctuations within each year, with certain months consistently showing higher employment percentages, likely reflecting seasonal hiring trends in the service sector.

iv. Long-Term Trends:

Across the years, there are shifts in the overall color intensity, indicating changes in the employment levels over time. For example, there are periods with generally lighter or darker shades, suggesting fluctuations in employment over the decades.

v. Notable Changes:

In 2020, there is a noticeable increase in employment percentages, indicated by a darker shade of blue, which might reflect the impact of the COVID-19 pandemic and subsequent economic changes. Similarly, in 2023, the employment percentage is markedly higher, suggesting a significant change in employment levels.

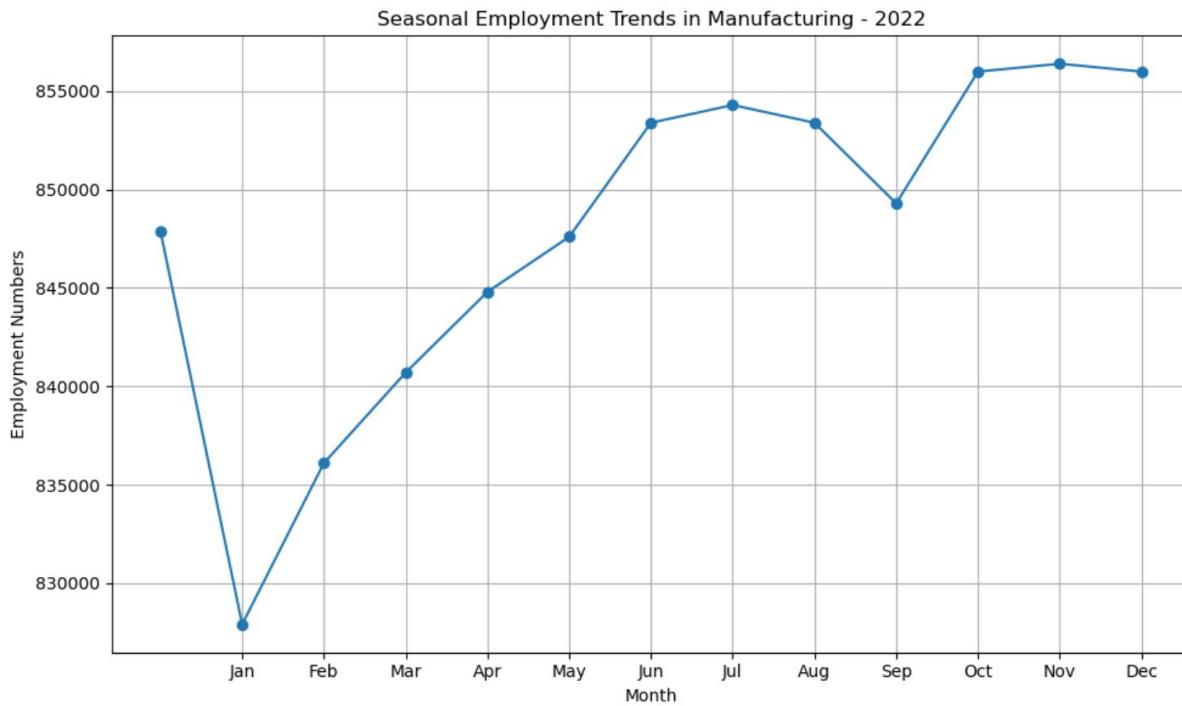
Insights and Results:

- i. The service-providing sector in New York State experiences clear seasonal variations in employment, which may be associated with industries like tourism, retail, and recreation that are influenced by seasonal demand.
- ii. Over the three decades, there are periods of increase and decrease in employment levels, which could correlate with economic cycles, policy changes, or industry-specific developments.
- iii. The significant increases in employment percentages in 2020 and 2023 warrant further investigation to understand the underlying causes, such as economic recovery efforts, shifts in consumer behavior, or changes in service industry dynamics.

The heat map provides a visual and analytical perspective on the employment variations in New York State's service-providing sector and how they have evolved over the past three decades.

3. Single Line Graph:

- Seasonal Employment Trends in Manufacturing in 2022:



The line graph depicts the monthly employment trends in the manufacturing sector of New York State for the year 2022. Below are the interpretations from the chart:

i. Overall Trend:

The graph shows fluctuations in manufacturing employment throughout the year, with a general increase from the beginning of the year leading into the middle months, followed by a dip and then another rise towards the end of the year.

ii. Seasonal Fluctuations:

There is a sharp drop in employment from January to February, followed by a steady increase through May. Employment peaks in June, dips slightly in July, and reaches the highest point in August. This could reflect seasonal production cycles or demand for manufactured goods.

iii. Fall and Winter Trends:

After August, there's a sharp decline in September, which could be due to a variety of factors such as end of summer production schedules, followed by a recovery in October and stability through November and December.

Insights and Results:

- The manufacturing sector in New York State displays a seasonal pattern, with employment typically peaking in the summer months and having a notable decrease in early fall.
- The sharp decrease in September might warrant further investigation to understand specific industry dynamics or external factors contributing to this decline.
- The relative stability in employment numbers towards the end of the year suggests that the sector may be influenced by consistent holiday season demand or year-end production quotas.

- iv. These insights can be included in the project report and presentation to highlight the seasonality and potential volatility within the manufacturing sector's employment trends, and to inform decision-makers about the times of year that might require more workforce planning or resource allocation.

Unique Findings:

```
import sweetviz as sv

# Analyze a single dataset
report = sv.analyze(df)
report.show_html() # Generates and opens a HTML report
```

Sweet viz is an open-source Python library that is used to create visually appealing and detailed exploratory data analysis (EDA) reports with just a few lines of code. It generates a comprehensive HTML report that includes visualizations of the distribution of each feature in a dataset, along with comparisons between different subsets of data, if applicable. Here's how Sweet viz can be beneficial for visualizations and has contributed to our project:

- i. Automated Analysis:
Sweet viz automates the process of creating a detailed EDA report, saving time and effort compared to manually plotting each feature. It provides a quick overview that can help identify patterns, outliers, and potential issues in the data.
- ii. Feature Visualizations:
It generates visual comparisons for all features in the dataset, displaying metrics such as count, mean, median, and mode. The visualizations include histograms, bar charts, and pie charts, which make it easy to understand the distribution and relationships of the data.
- iii. Comparative Analysis:
If analyzing two datasets or subsets (such as before and after a certain event), Sweetviz can compare them side by side to highlight key differences. This is useful for understanding the impact of specific factors on the dataset.
- iv. Correlation Analysis:
The library includes a correlation analysis that visually represents how features correlate with each other, which is essential for identifying relationships within the data.
- v. Association Graphs:
For categorical data, Sweetviz can create association graphs that show the proportion of each category, helping to quickly understand the composition of categorical features.
- vi. Missing Data:
Sweet viz reports on missing values within the dataset, providing insights into the completeness of the data, which can inform data cleaning and preprocessing strategies.
- vii. Efficient Sharing:
Since the report is in HTML format, it can be easily shared and accessed by stakeholders, making it a valuable tool for collaborative projects.

For our project, Sweet viz will enable us to conduct a thorough initial analysis of the employment dataset. It will help us to visually present and interpret the data's characteristics, trends, and potential

anomalies. This contributes to a deeper understanding of the employment landscape in New York State and supports informed decision-making for subsequent analyses, reporting, and presentations.

Conclusion:

- Service sector dominance:
We observed a surge in service-oriented sectors like "Private Service Providing," shaping a service-driven economy.
- Regional dynamics:
While New York City reigned supreme, hubs like Nassau-Suffolk and Buffalo-Niagara Falls showcased promising growth. Rural areas, however, faced limited opportunities.
- Shifting trends:
We witnessed the decline of manufacturing sectors like "Engine Manufacturing," suggesting economic transitions. Similarly, resource extraction industries saw a gradual retreat.
- Data-driven insights:
By visualizing these trends, we gained valuable insights that can inform decision-making for future growth and resilience.

References:

- [Data.gov](#) – For dataset
- [Current Employment Statistics](#) – For understanding the base data and how to get the inferences.