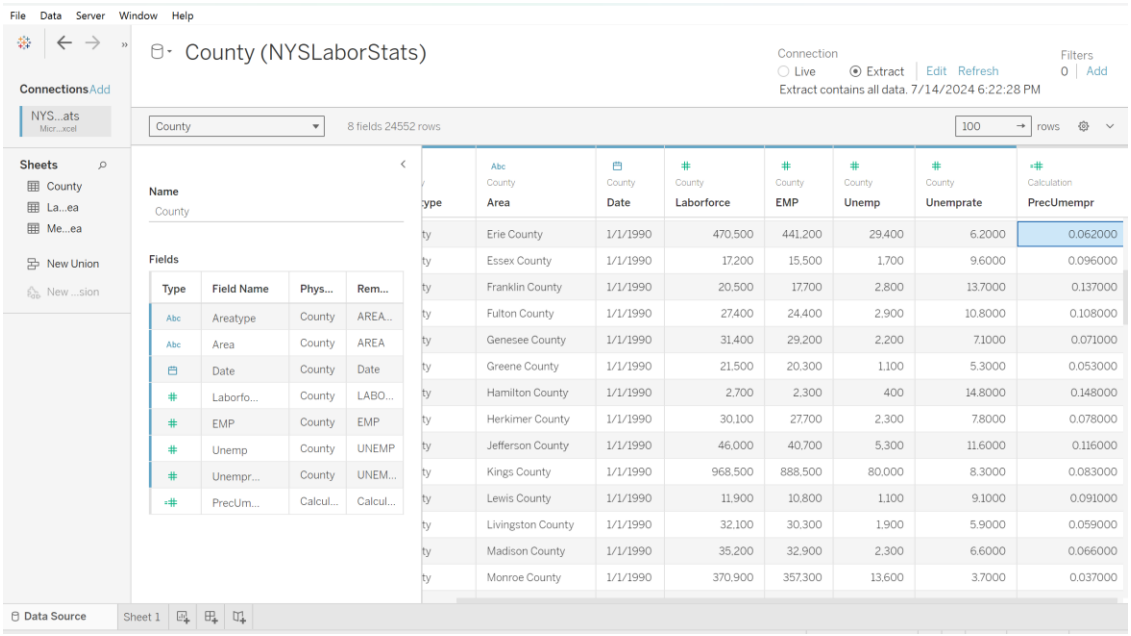
Homework Assignment 1 Report

6. Scroll down the data preview table and take a screenshot of the table when you find “Erie County” on “1/1/1990”. Be sure to include the PercUnempr value in this screenshot.



8. What is the resulting output? Why might this be bad for your data analysis?

Answer: The resulting output is 0 because the values in the ‘PercUnempr’ field are all less than 1. This is due to the fact that we divided the original Unemprate values by 100 to convert them into a percentage format. As a result, when these decimal values are rounded to whole numbers, they all round down to 0. This means that the representation of the unemployment rate will be inaccurate and misleading, as it fails to reflect the actual percentage values we intended to capture.

9. We’ll learn how to make these charts more aesthetically pleasing in a bit, but for now, hover over the different points on the chart. On average, which month had the highest employment numbers across the dataset?

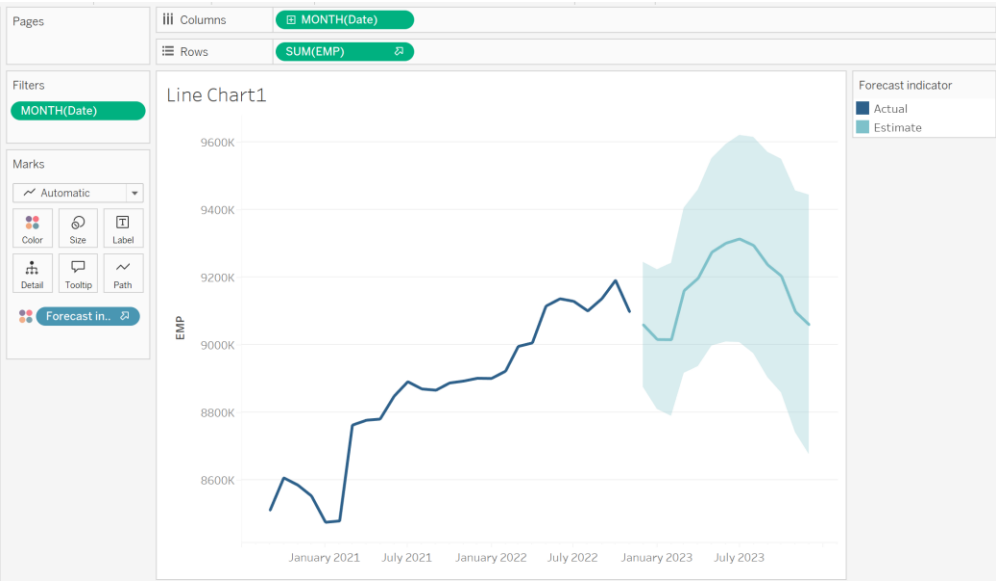
Answer: On average, July month had the highest employment numbers of 142685 across the dataset.

22. You might be thinking “That forecast is too wide of a forecast!” and you would be right! Forecasts are directly driven by past data. If our past data has a high amount of variation, the forecasts for the future will be less predictable. What is happening in our current chart that you feel is contributing to the increase in “uncertainty” in our forecast? Set your filter to happen after that event (September 2020). Does doing that increase the “certainty” of your forecast? Paste a screenshot of your forecast.

Answer: The forecast line is very wide in our current chart, which means that the forecast is very uncertain. This is likely due to the fact that the past data has a high amount of variation. There is a high variation in data from February 2020 to August 2020 which is the cause of uncertainty in the forecast data.

After filtering the chart by keeping the starting date as "9/1/2020" to only show data from September 2020 onwards, it is observed that the forecast line is now much narrower, which means that the forecast is more certain. This is because the data from September 2020 onwards has less variation than the data from before that date.

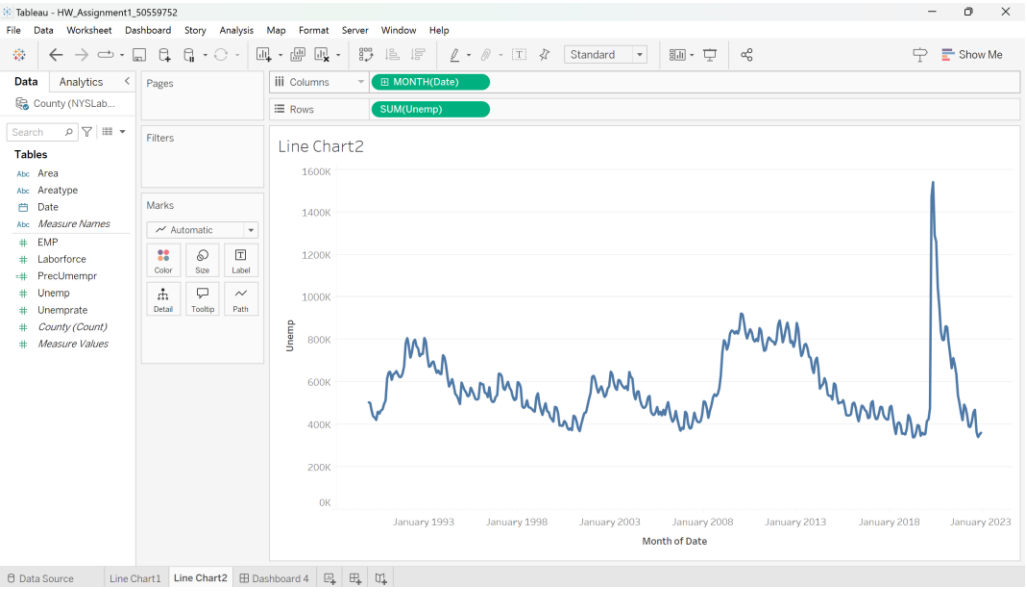
Here is a screenshot of the forecast after you set the filter to September 2020 (changed the color from blue to teal):



Trying it on your own

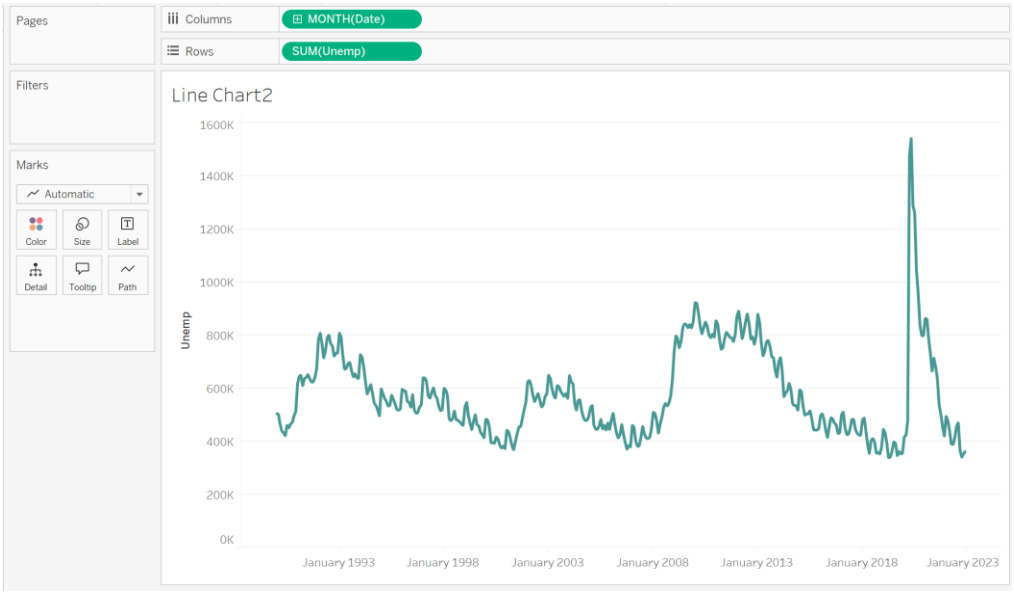
1. Now that you have the basics down, create a second sheet with a line chart using the “Unemp” field. HINT: Don’t forget to change the date hierarchy to match the time values from the previous section Paste a picture of the resulting line chart.

Answer: Here’s the screenshot of new line chart,



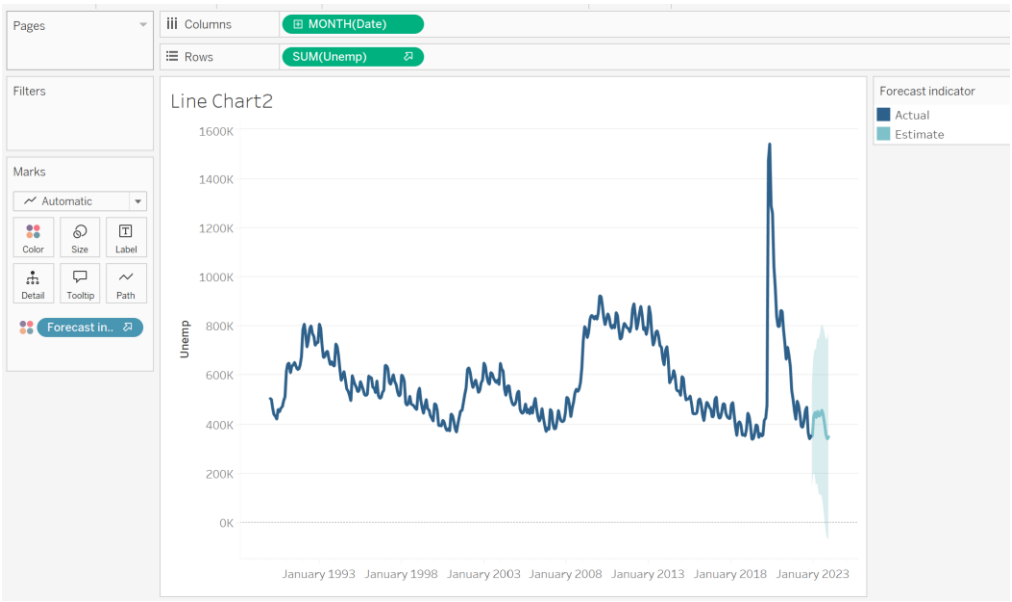
2. Change the color of the line chart to something other than blue. Paste a picture of the new line chart in your report.

Answer: Changed the color to teal.



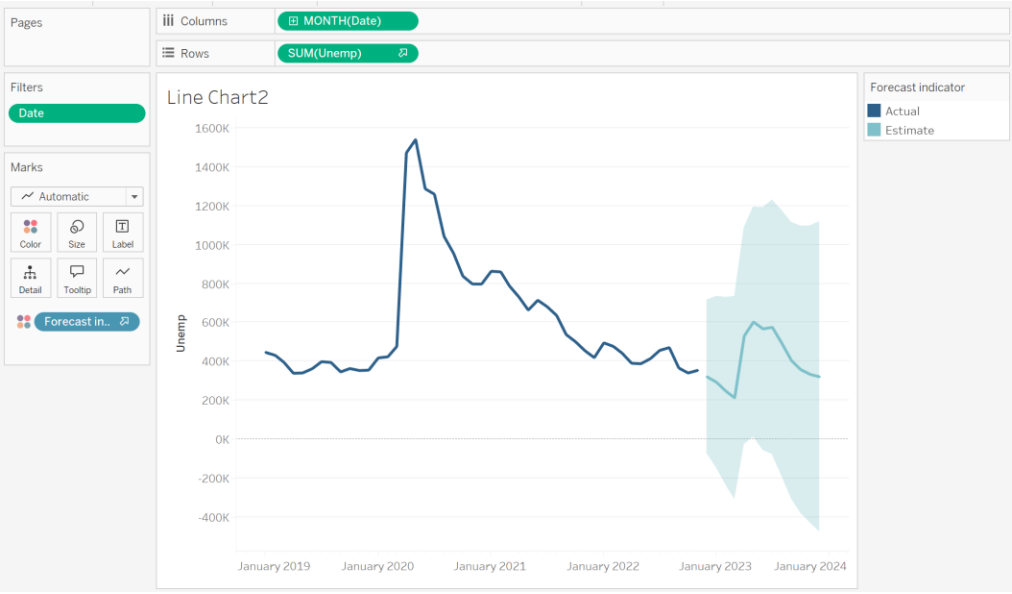
3. Create a forecast, setting the model to custom and the season to additive. Paste a picture of the line chart in your report.

Answer:



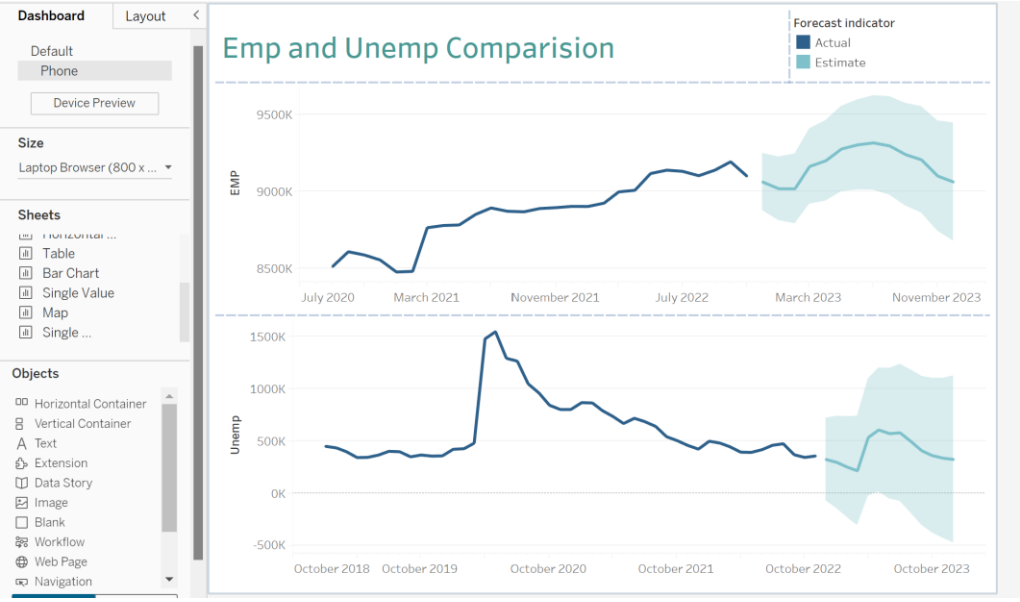
4. Change your line chart to filter for a Relative date showing items since 2019.

Answer:



5. Create a Dashboard and add your two sheets to it in a configuration of your choosing. Paste a screenshot of this dashboard in your report. Tableau has a dashboard button on the bottom row of options, which allows you to combine multiple charts into one "dashboard" view.

Answer:



6. Write 2-4 sentences in your report detailing any insights you can draw from these two line charts.

Answer: The two line charts show the number of employees (EMP) and the number of unemployed (UNEMP) over time. The EMP chart shows a general upward trend, with a few dips in 2020 and 2022. The UNEMP chart shows a general downward trend, with a few spikes in 2020. The two charts seem to be inversely related, which means that when the number of employees increases, the number of unemployed decreases, and vice versa. This suggests that there is a correlation between the two variables. Similarly, forecast line that the number of employees is expected to continue increasing in the future (in July 2023) and unemployed is expected to continue decreasing in the future (in March 2023).

Creating a dashboard

Your client wants to assess the viability of the job market across New York state, particularly considering the Erie County region as compared to statewide statistics. The client wants to understand what the data for 2016-2022 looked like for Erie County and what the 2022-2024 looks like, including any forecasted pitfalls in the employment statistics. The client wants more than just a simple "number of employed" and would like you to provide some insights into unemployment rates across various counties and where Erie county has ranked over this time period.

Q. Write 1 page (double-spaced) describing your dashboard (tell your story!)

Storytelling through Data Visualization: New York State Job Market Analysis (2016-2024) – focused on Erie County

This dashboard offers valuable insights into New York's job market dynamics, focusing on Erie County in comparison to the statewide trends. The employed population, unemployment rates, and Erie County's ranking among other counties are all explored through the following elements:

1. **Line Chart with Forecast (Top Left):**
    - **Gestalt Principles Used:** In the initial chart, variation of values is not clearly visible so, realigned axis by using independent axis range for rows (**Proximity**). Also, to distinguish Statewide and Erie County employment (**Similarity**) assigned different color palettes (blue for statewide, orange for Erie County). Additionally, it lines have smooth upward trend (**Continuity**).

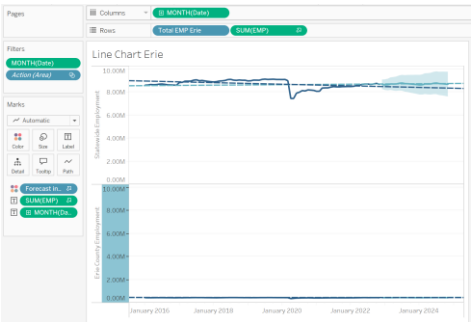


Figure 1: Initial Line Chart

Figure 2: Final Line Chart

    - **Insights:** This chart reveals a positive trend in employment across New York State (blue line) from 2016 to 2024 (with forecast from 2022 to 2024). The orange line shows a similar pattern for Erie County, with employment rising steadily. The forecasted growth suggests a continuation of this positive trend for both regions. But there is an uncertainty in the forecasted employment statistics because the past data has a high amount of variation in data from January 2020 to July 2021.

2. **Horizontal Bar Chart (Top Right):**

    - **Gestalt Principles Used:** Assigned blue-teal and orange color palette for consistent comparison similar as line chart (**Similarity**). Also, added Reference line for average Umemprate rank to help in grouping Counties based on Unemployment rate (**Connectivity**).

    - **Insights:** This chart delves deeper into unemployment rates. Here, we see Erie County's unemployment rate (orange) has been ranked as 25 across all counties and this ranking is slightly less than the average line throughout the period 2016-2022. This contextualizes the unemployment rate data by revealing Erie County's relative position among other counties.

3. **Rank Table (Bottom Left):**

- **Gestalt Principles Used:** Updated the table by removing the unwanted borders, header for year (**Closure**) and formatted the values to match all other charts (**Similarity**). Added additional field "Emp%" using calculated field to provide details on employment percentage throughout the period (**Connectivity**).

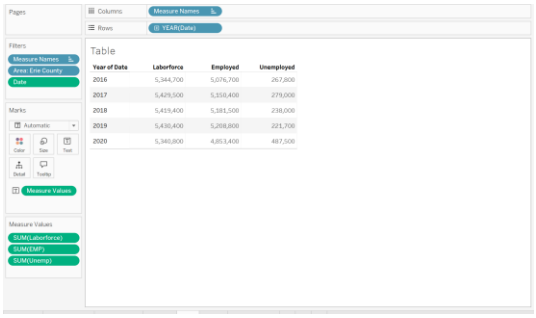


Figure 3: Initial table

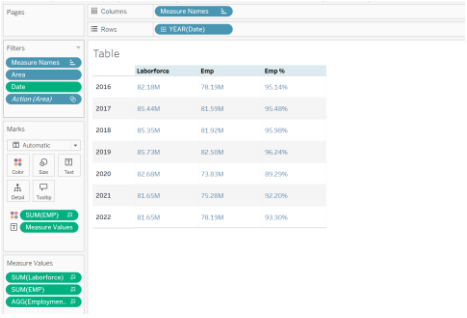


Figure 4: Final table

- **Insights:** This table sheds light on Erie County's employment for each year between 2016 and 2022. It is observed that Erie County had highest employment rate of 95.98% in 2018 and seen drop in 2020.

4. **Bar graph** (Bottom Middle):

   - **Gestalt Principles Used:** Assigned same color palette to forecasted and actual data for both Employed and Unemployed (**Similarity**).

   - **Insights:** This chart offers insights about the comparison of number of employed and unemployed people in Erie county from 2016 to 2022 and forecasted values from 2022 to 2024. It is observed that employment and unemployment are inversely correlated except for the year 2020.

5. **Interactive Map** (Bottom Right):

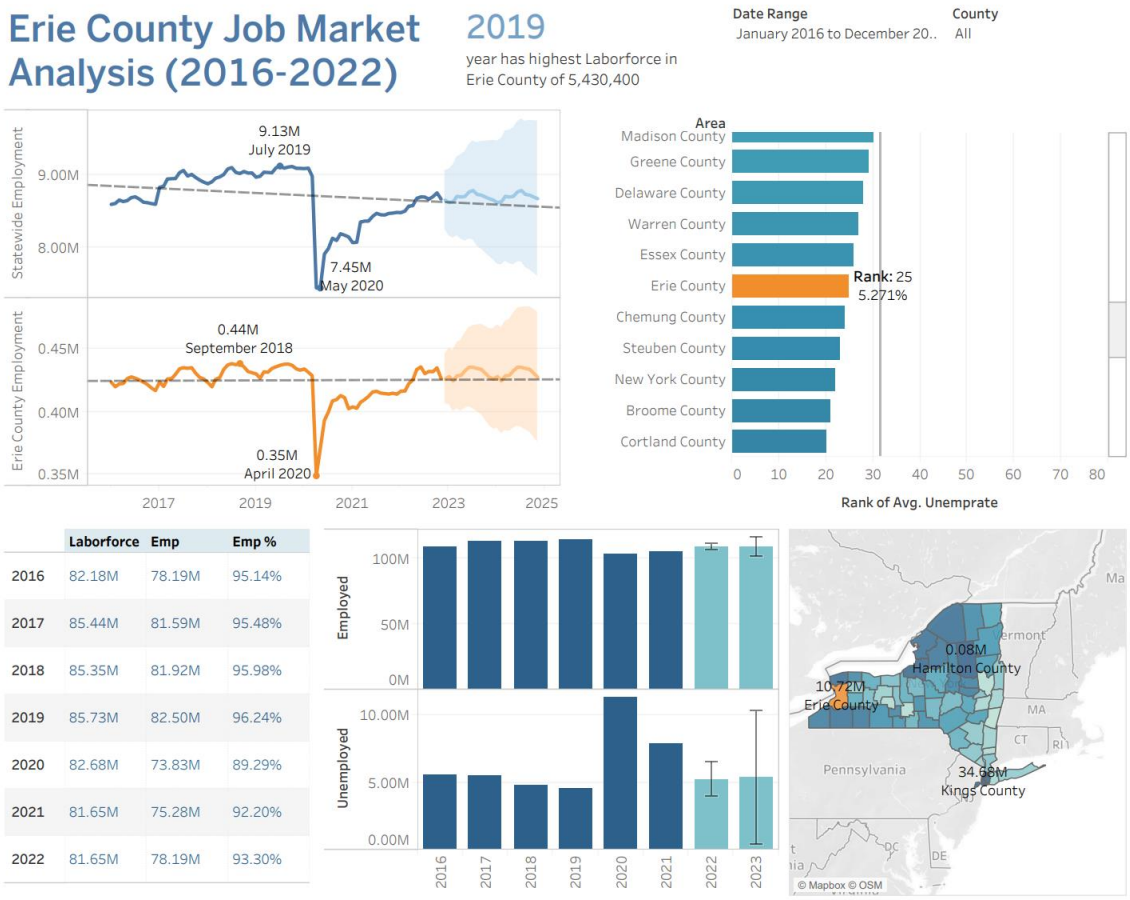
   - **Gestalt Principles Used:** Assigned blue-teal and orange color palette for consistent comparison similar as line chart and horizontal bar chart (**Similarity**). **Proximity** (Erie County in relation to surrounding counties)

   - **Insights:** This map offers a geographically insightful view of the number of unemployed across New York State from 2016 to 2022. It is seen that Kings County has a high number of unemployed people in that time period.

The dashboard adheres to Gestalt principles for clear communication. A consistent color scheme (blue for statewide, orange for Erie County) promotes **similarity**. The interactive map allows for user exploration, fostering a sense of **closure**. **Overall, this data visualization effectively communicates the strengths and weaknesses of New York's job market, with a particular focus on Erie County.**

Q. Paste a screenshot of your dashboard in your final report. You can also use the built in "Export to PDF" option in Tableau.

Here's the screenshot (Tableau print to PDF) of the final dashboard:



- Added filters to select particular date range (by default it'll select data from January 2016 to December 2022) and county in New York state in the top right corner.
- Created single value element to display the year (in period 2016 to 2022) that has highest Laborforce in Erie County.
- Added a filter action as below screenshot. So, we can check the charts, map and table for a selected county and find the insights by comparing it with Erie County (by default the dashboard will show all details for Erie County).

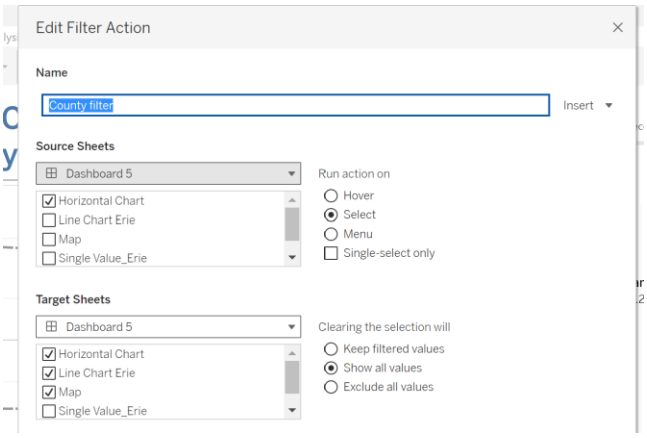


Figure 1: Filter action

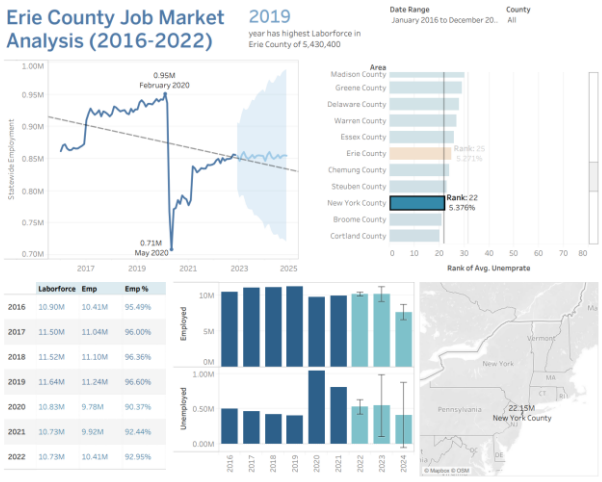


Figure 2: New York County selected