As a precursor to connecting our Oracle Data Warehouse and analyzing the business data, we performed a hands-on exercise to better grasp the concepts of calculated fields, forecasts, and gleaning insights from the data using data visualizations on the business intelligence layer. We are using Tableau Software for this last leg of the Data Warehouse Lifecycle.

The first part of the assignment was to connect to and analyze an Excel workbook that contained New York State Employment Statistics data for the time period spanning January 2014 to May 2021. Prior to the analysis we had to hammer out certain details to ensure the accuracy of the analysis. These involved creating calculated fields to convert the truncated *Employed* numbers back to 1000s and another calculated field to correctly read the *Insured Unemployment Rate.* We then proceeded to create the line graph for analysis of the data. While we have this all this data for describing what happened during the past seven years, the real benefit of historical data of this kind is when we can draw meaningful forecasts from there to base certain business decisions off of. We achieved this by creating a forecast using Tableau’s in-built capabilities.

After connecting both data sources Employment numbers and Unemployment stats, Tableau automatically formed the relation between them using the ‘state’ field.

At first, the forecast was a straight line (flat) that does not provide meaningful insights. The reason this happened was due to a massive anomaly in the dataset – the New York State economy lost more than 1.7 Million private sector jobs in April 2020 due to the direct effects of the COVID-19 pandemic ([NYS Department of Labor](https://dol.ny.gov/system/files/documents/2021/03/press-release-1-april-2020-final_2.pdf)). The forecast was impacted by this as Tableau assesses a seasonal cycle for the data based on the time series indicated in the dataset. Since the graph was using *Month* as the basis for drawing the trendline, the sudden plunge in the employment stats disrupted the automatic detection of a seasonal cycle ([Tableau Software – Seasonality](https://help.tableau.com/current/pro/desktop/en-us/forecast_how_it_works.htm#seasonality)). Post this, there was a rectification reflected in the graph after we manually changed the *Automatic Forecast* to *Custom* and set the *Seasonality* to *Additive*. This final change produced a meaningful data visualization that can be used for further discussions and decisions based on the data.

The second and more prominent part of this assignment was to connect our Data Warehouse to Tableau and then analyzing the data to generate business insights. To begin with, we had to install Oracle Instant Client and setup our Oracle Home environment variable to set up our device for the appropriate settings. Furthermore, to successfully authenticate, we downloaded the Oracle Wallet files from the Data Warehouse and placed them in the requisite location on the local machine.

After performing the initial setup, next phase is to connect the Data Warehouse to Tableau using a connection string available on the Oracle Data Warehouse environment in addition to the user access credentials for the Oracle Data Warehouse. Once the data connection is successfully established, the penultimate step is to add the facts and dimensions to Tableau’s workspace where Tableau automatically finds the connecting fields between the tables and sets up those connections for us. The final and most valuable part of this exercise was creating meaningful data visualizations that support an inference of business insights.

**Screenshot of the formula used for Unemployment Stats:**

A screenshot of a computer

Description automatically generated

**Line Chart of Unemployment Stats:**

Chart

Description automatically generated

**Line Chart of Unemployment Stats after changing colour:**

Chart, line chart

Description automatically generated

**Line Chart of Unemployment Stats with Forecast:**

Chart

Description automatically generated

**Dashboard of Employment numbers and Unemployment Stats:**

Chart

Description automatically generated with medium confidence

Unemployment rate and employment numbers show strong correlation in this chart. The post covid-19 pandemic recovery shows promising growth with unemployment numbers going down at a meager but steady rate with the employment numbers showing some peaks and troughs throughout the recovery period (2021-2023).

**Screenshot relationship diagram:**

Graphical user interface

Description automatically generated

**Business Insights for the Oracle DW Dataset:**

Graphical user interface, chart, application

Description automatically generated

**Pie chart for Total Sales per Customer:**

At a glance, the amount of sales for male customers (Dominic Sellitto, James Bond, and Jeep Jeeperson) is higher than those for female customers (Sally Sallerson & Jennifer Lopez). From this, we can infer that we must try to market more to the female customer base.

**Line graph for Sales in Buffalo & Rochester:**

The sales in Rochester are higher than in Buffalo. However Buffalo appears to have a steady forecast of sales with lesser troughs in the graph. From this, we could arrive at a decision of increasing the sales in Buffalo and solidifying the customer base in Rochester so as to have a steady flow of orders throughout the year.