

## **Homework / Assignment 5**

### **Integrating Warehouses with BI Tools**

MGS 6577LEC F1S: Cloud Data Warehousing & Data Engn (19774 Fall24)

**Submitted By**

**Anjali Pandey**

# Table of Contents

## Contents

Table of Contents .....	2
Executive Summary for Power BI and Oracle Cloud Integration Assignment .....	3
Key Activities: .....	3
1. PowerBI Installation and Data Connection: .....	3
2. Data Modeling: .....	3
3. Data Preparation: .....	3
4. Visualization: .....	3
5. Dashboard Creation: .....	4
6. Insights: .....	4
Appendix A .....	5
1. Click the button and paste a screenshot into your report. Answer the following questions: Did autodetect find any relationships between the data? Why do you think this was the case? .....	5
2. Paste a screenshot of the resulting screen in your report. Along with it, answer the following question: What was the TrueEmployed value for Saturday, February 1, 2014? .....	6
3. Now that you have the basics down, create a second line chart using the “Unemployment Stats” worksheet. It should leverage the following fields: .....	7
▪ Reflecting Week Ended .....	7
▪ Insured Unemployment Rate .....	7
4. Change the color of the line chart to something other than blue. Paste a picture of the new line chart in your report? .....	7
5. Create a forecast, setting the forecast length to 26 and the seasonality to 52 (this matches the data seasonality). Paste a picture of the line chart in your report. ....	8
6. Change your line chart to filter for a Relative date showing items in the last 3 calendar years. Paste a picture of the full page (with both line charts) in your report. ....	8
7. Open the “Model” workspace from the left panel. You should see all of the imported tables and relationships. If you do not, use the autodetect function to add relationships between the tables (hint: review the earlier sections of the assignment if you don’t know how). Take a screenshot of the resulting relationship diagram and attach it to your report. ....	9
DashBoard .....	10
References .....	11

# Executive Summary for Power BI and Oracle Cloud Integration Assignment

This assignment main goal is to integrate Power BI to connect, visualize, and analyse the data from warehouse integrated along with Oracle Cloud. The main objective of this assignment was to establish a secure connection between PowerBI and Oracle Autonomous Database(ADB) for seamless access to warehouse data. There was various key task involved in installing Oracle Data Access Components and Oracle Wallet for database connection. PowerBI was the main tool to visualize and utilized to import, model and prepare employment statistics data for visualization. This project aligns with the following key activities and outcomes:

## Key Activities:

### 1. PowerBI Installation and Data Connection:

- PowerBI was installed and configured to visualize the data.
- The previous Oracle Wallet directory was used again to establish a strong and secure connection with Oracle Autonomous Database (ADB). There are steps involved in modifying system environmental variables and running the install.bat script for Oracle Data Access Components (ODAC).

### 2. Data Modeling:

- Various facts and Dimension Table were imported into PowerBI from the warehouse established already.
- Relationships between the tables were managed and maintained, with the “Manage Relationship” tool which was utilized for autodetection and manual adjustments wherever required.

### 3. Data Preparation:

- There was a calculation included in this such as:

*TrueEmployed = 'Employment Numbers'[Employed] \* 1000*

Implementation of this was done to standardize and expand the dataset for visualization. This was the addressed issues like abbreviations in the employed numbers.

### 4. Visualization:

- Many different visualizations were created, including:
  - Line chart denoting employment trends with forecast.
  - Second line chart denoted unemployment trends by leveraging “Reflecting Week Ended” and “Insured Unemployment Rate” fields were adjusted for date hierarchy.
  - Forecasts were adjusted and customized by setting seasonality and length parameters, and filters were applied to focus the main and the relevant data (e.g., last three calendar years).

## 5. Dashboard Creation:

- A one page of dashboard was created, which incorporated data visualization, and the insights derived from datasets. The dashboard was specifically for forecast visualization and clear story narrative for business decision-making.

## 6. Insights:

- The TrueEmployed Data formula helped in revealing employment patterns, highlighting the trends such as significant dips during the 2020 due to the pandemic.
- Unemployment trends also showed seasonal variations and provided a detailed and comprehensive overview of decision-making.

## Appendix A

1. Click the button and paste a screenshot into your report. Answer the following questions: Did autodetect find any relationships between the data? Why do you think this was the case?

Ans – Using the mutual State field that joins employment and unemployment data, there is the connection of "Employment Numbers" and "Unemployment Stats" thus stemming. This action permits a wider analysis in terms of statistics that are linked through employment figures and unemployment data. One of the possible reasons why the Autodetect might fail to identify the fact that the relationship is many-to-many with the data or the data sets are not consistent. Through the establishment of such relations, the results of employment trends and their correlation with unemployment statistics are precise and correct.

The screenshot displays the Microsoft Power BI Desktop interface. On the left, the 'Data Model' view shows a diagram with tables: 'Unemployment Stats', 'Employment Numbers', 'FACT SALES', 'DIM DATE', and 'DIM CUSTOMER'. The 'Unemployment Stats' and 'Employment Numbers' tables are connected by a relationship line. The 'Edit relationship' dialog box is open on the right, showing the relationship between 'Employment Numbers' (From table) and 'Unemployment Stats' (To table). The 'State' column in 'Employment Numbers' is linked to the 'State' column in 'Unemployment Stats'. The cardinality is set to 'Many to many (\*\*)'. The 'Cross-filter direction' is set to 'Both'. The 'Make this relationship active' checkbox is checked. A warning message at the bottom states: 'This relationship has cardinality Many-Many. This should only be used if it is expected that neither column (State and State) contains unique values, and that the significantly different behavior of Many-Many relationships is understood. [Learn more](#)'.

**From table: Employment Numbers**

date	Employed	State	TrueEmployed	Year
01 June 2021	8717	NY	8717000	2021
01 May 2021	8628	NY	8628000	2021
01 April 2021	8661	NY	8661000	2021

**To table: Unemployment Stats**

Continued CL...	Covered Empl...	Filed week en...	Initial Claims	Reflecting We...	State	Unemploy
264725	8466102	04 January 20...	52096	28 December ...	NY	3.13
279382	8493299	11 January 20...	34077	04 January 20...	NY	3.29
255766	8493299	18 January 20...	28584	11 January 20...	NY	3.01

**Cardinality:** Many to many (\*\*) **Cross-filter direction:** Both

☒ Make this relationship active ☐ Apply security filter in both directions

☐ Assume referential integrity

**Warning:** This relationship has cardinality Many-Many. This should only be used if it is expected that neither column (State and State) contains unique values, and that the significantly different behavior of Many-Many relationships is understood. [Learn more](#)

**Buttons:** Save, Cancel

2. Paste a screenshot of the resulting screen in your report. Along with it, answer the following question: What was the TrueEmployed value for Saturday, February 1, 2014?

Ans -

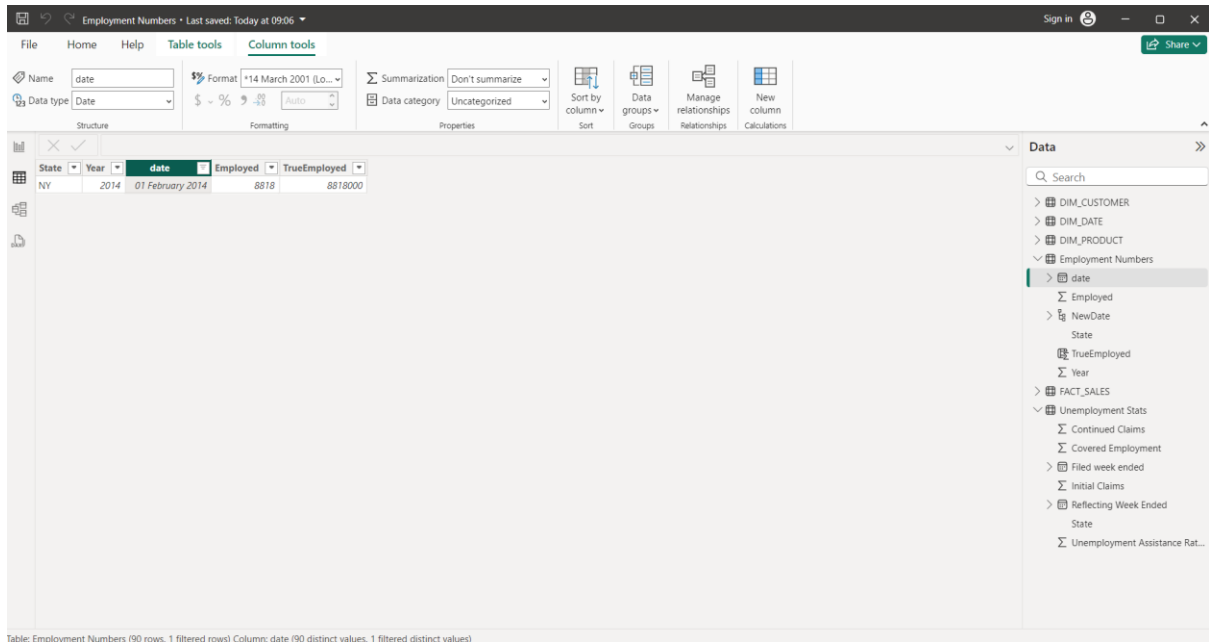


Table: Employment Numbers (90 rows, 1 filtered rows) Column: date (90 distinct values, 1 filtered distinct values)

Based on the formula  $\text{TrueEmployed} = \text{'Employment Numbers'}[Employed] * 1000$  and the data shown in the screenshot, the value of Employed for Saturday, February 1, 2014, is 8818. By applying the formula:

$$\text{TrueEmployed} = \text{'Employment Numbers'}[Employed] * 1000$$

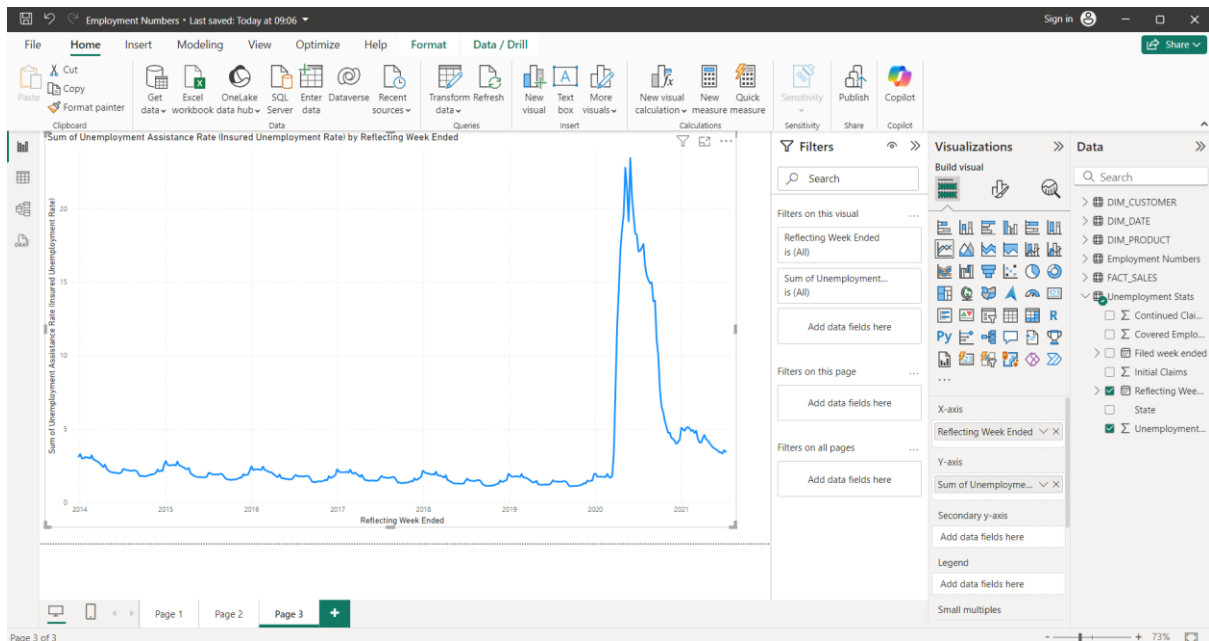
$$\text{TrueEmployed} = 8818 \times 1000 = 8,818,000$$

Thus, the TrueEmployed value for Saturday, February 1, 2014, is 8,818,000.

3. Now that you have the basics down, create a second line chart using the “Unemployment Stats” worksheet. It should leverage the following fields:

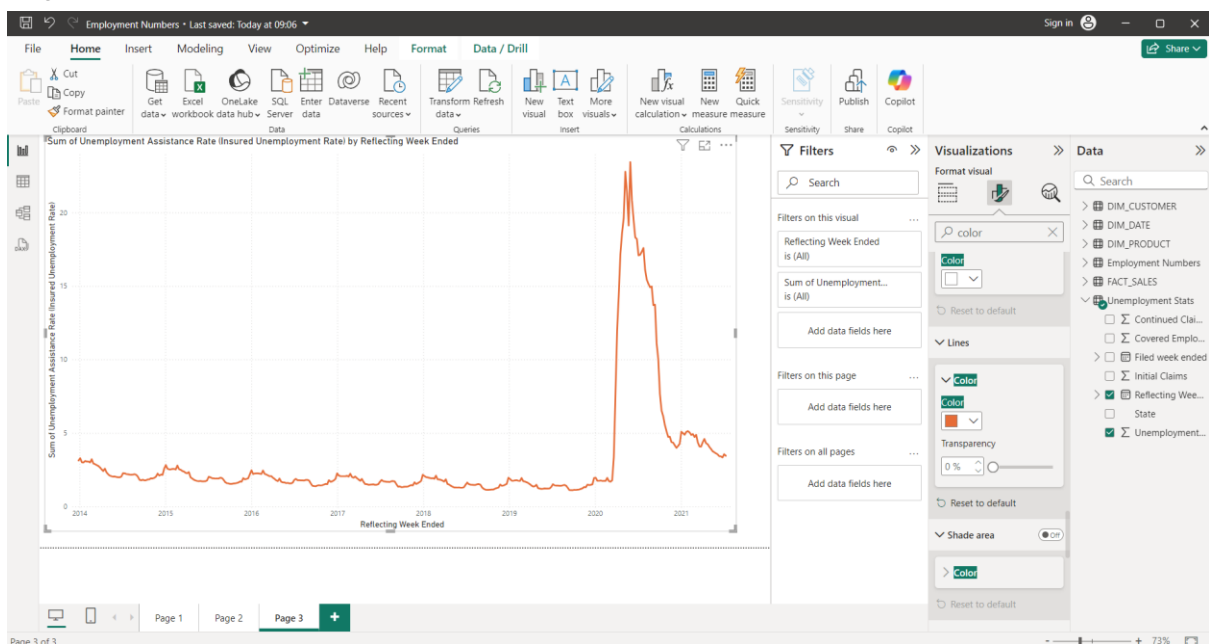
- Reflecting Week Ended
- Insured Unemployment Rate

Ans -



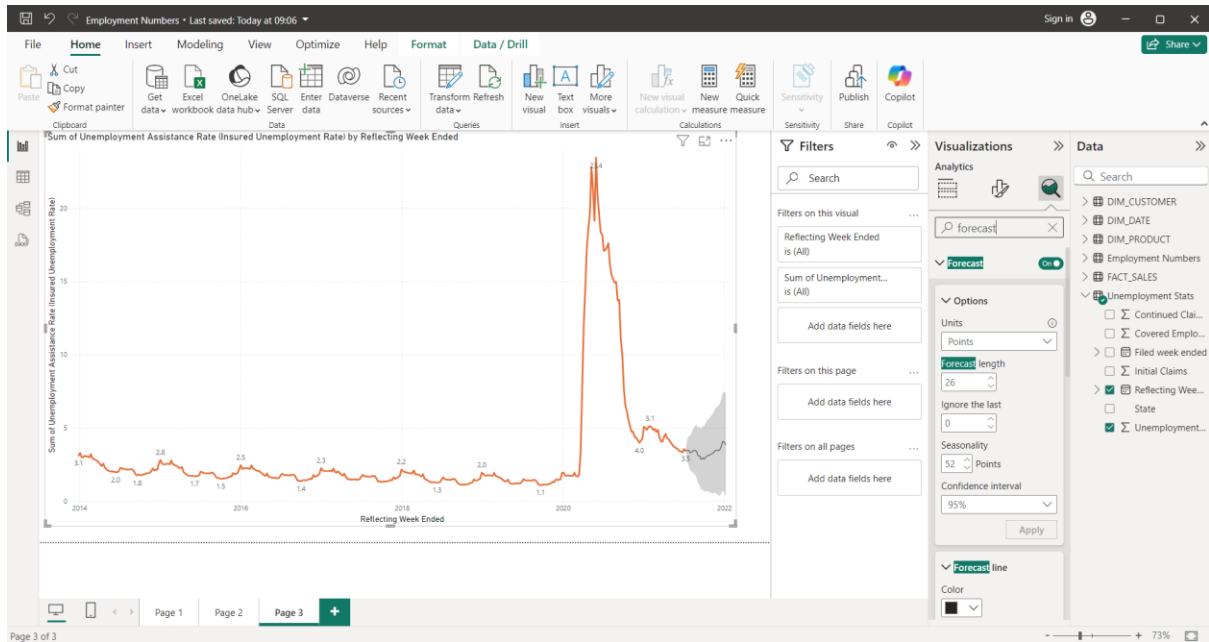
4. Change the color of the line chart to something other than blue. Paste a picture of the new line chart in your report?

Ans -



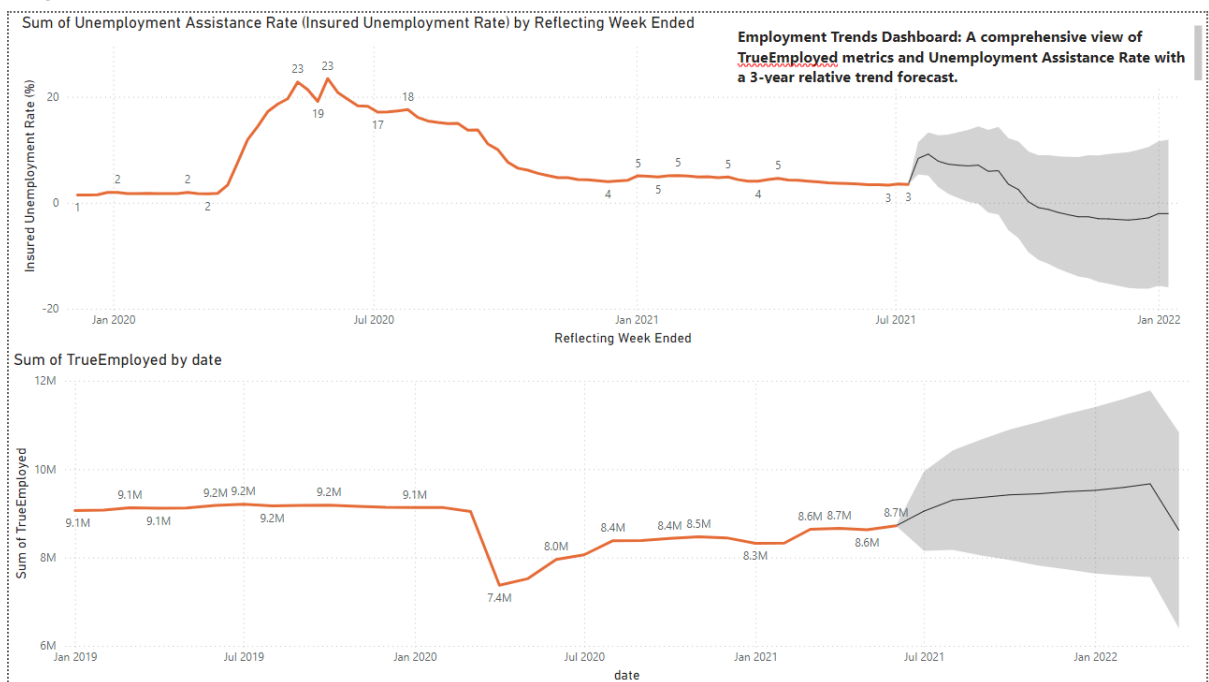
5. Create a forecast, setting the forecast length to 26 and the seasonality to 52 (this matches the data seasonality). Paste a picture of the line chart in your report.

Ans -



6. Change your line chart to filter for a Relative date showing items in the last 3 calendar years. Paste a picture of the full page (with both line charts) in your report.

Ans -

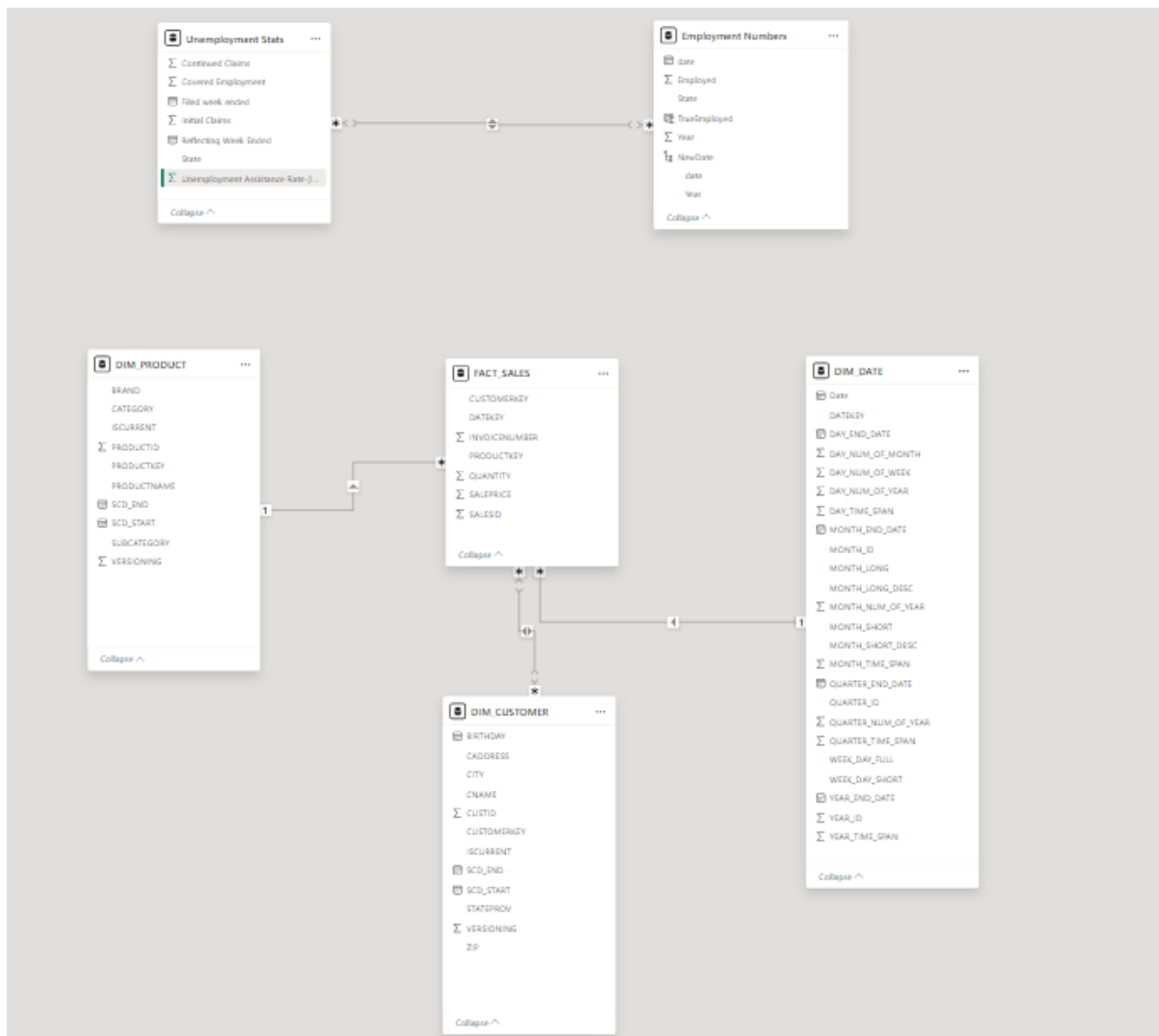




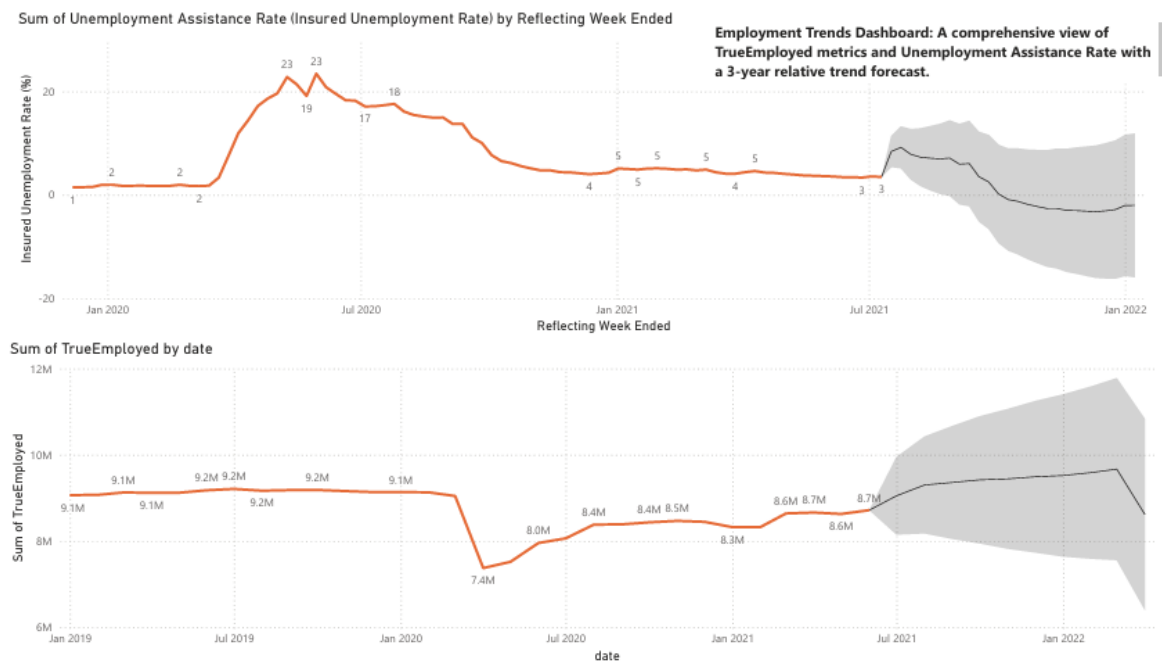
The charts offers a clear overview of employment and unemployment trends over the past three years. The TrueEmployed chart marks a period of decline at the time of serious economic disruptions and then the economic recovery steps in. The chart of the unemployment assistance rate shows a strong increase at the time of the peak of youths running away from work, however, steadying over time. All these conclusions emphasize the flexible aspect of labor market fluctuations and recovery phases.

7. Open the “Model” workspace from the left panel. You should see all of the imported tables and relationships. If you do not, use the autodetect function to add relationships between the tables (hint: review the earlier sections of the assignment if you don’t know how). Take a screenshot of the resulting relationship diagram and attach it to your report.

Ans –



## Dashboard



This dashboard gives a comprehensive reviews of employment and unemployment trends over the past three years, the main metrics of which are including the Unemployment Assistance Rate and the TrueEmployed numbers. It concentrates on the core indicators, presenting examples of changes such as the advancement of unemployment in 2020 and the subsequent recovery stages. The addition of forecasted trends provides valuable insights to businesses and policymakers about how these trends might develop in the future, thus, they can plan for potential problems and opportunities. By keeping data in a neat, objective, and understandable form, the dashboard enables such a complicated labor market data to impair stakeholders to make effective, strategic decisions.

## References

1. ChatGPT for paraphrasing the executive summary.