

# DATA VISUALIZATION USING TABLEAU

## ASSIGNMENT 2

SEPTEMBER 26, 2022.

INDIANA UNIVERSITY  
(BLOOMINGTON)

Authored by: VARSHA R  
IU ID: 2000751388  
Sem: Fall 2022



---

# Introduction

Data visualization is achieved using Tableau, a software package focusing on business intelligence (BI). The word tableau refers to a graphic representation or description. In this assignment I have implemented the concepts of marking, grouping and categorizing. Let us discuss each concept in depth in this report.

## Understanding the dataset

Quarterly census of employment and wages annual data: Beginning 2000 - [Dataset](#)

The **Quarterly Census of Employment and Wages (QCEW)** program publishes a quarterly count of employment and wages reported by employers covering more than 95 percent of U.S. jobs, available at the county, MSA, state and national levels by industry. New York State's Employers' Unemployment Insurance Law (UI Law) requires employers to submit employment and wage data on a quarterly basis (QCEW). This program is coordinated with the U.S. At the State, regional, and county levels, QCEW data cover approximately 97 percent of New York's nonfarm employment, offering the most comprehensive employment and wage data available. In general, the term "covered" employment refers to both private-sector employees and employees of state, county, and municipal governments. For purposes of the program, federal employees are covered under separate laws. A few agricultural workers, railroad workers, private household workers, students, and unpaid family workers are not covered by UI. As with CES data, QCEW data reflect jobs by place of employment; therefore, if an individual holds two jobs, he or she will be counted twice. As QCEW measures only employment covered by unemployment insurance laws, its totals will differ from CES employment totals due to employee categories excluded by UI.

For this assignment, we will be using the following data from the dataset:

- Total wage – The total wage earned by the person (in trillions).
- Annual average salary – The annual average salary of that region (in thousands).
- Average employment – Average employment rate in that region (in millions).
- Year – The year the data was collected (2000 – 2016).
- Area – The name of the region.
- Area type – The type of area the region is identified as.

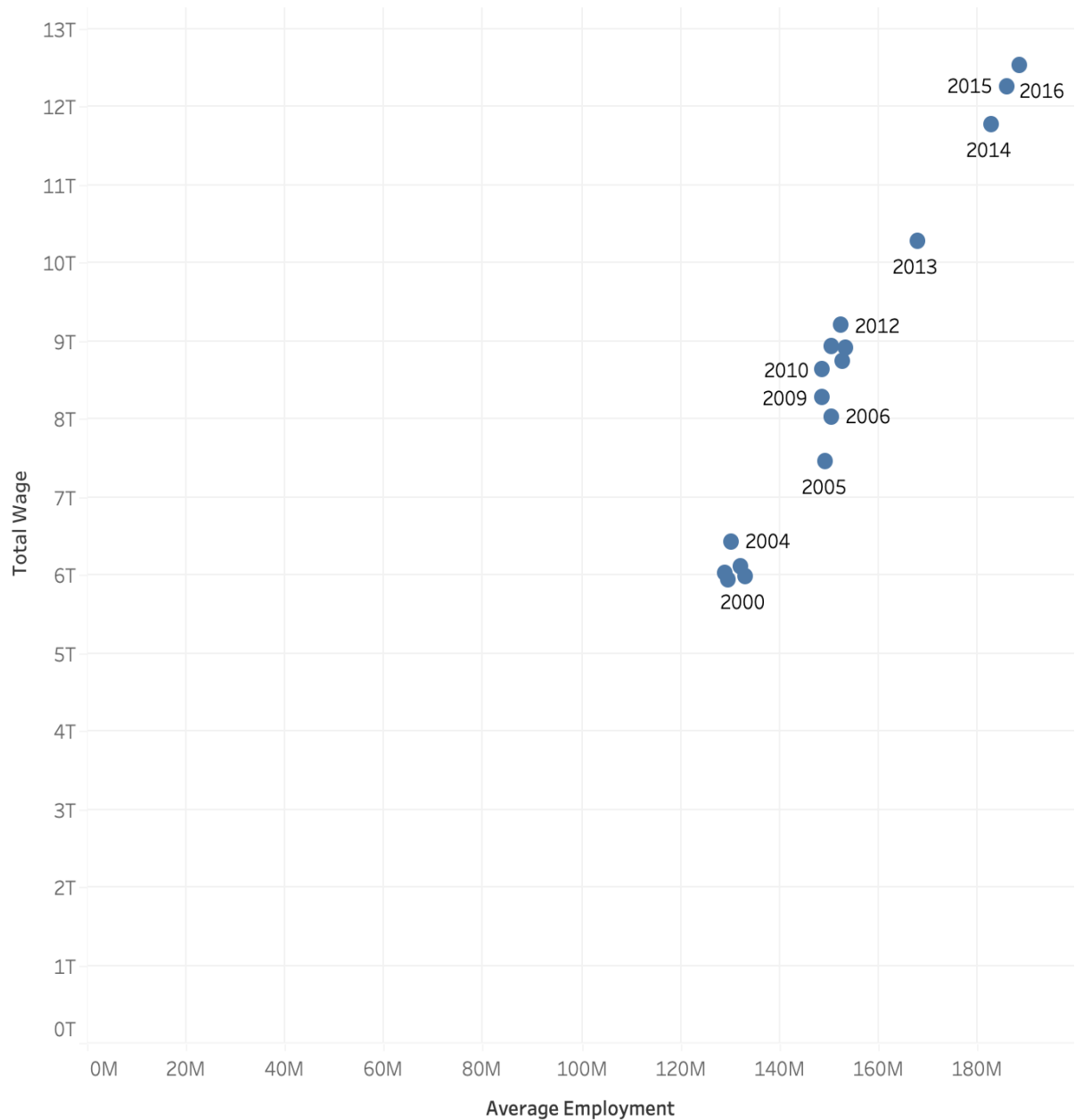
---

# Visualizations

## 1) Q1

### ○ Plot

Q1



Sum of Average Employment vs. sum of Total Wage. The marks are labeled by Year.



## ○ Plot analysis

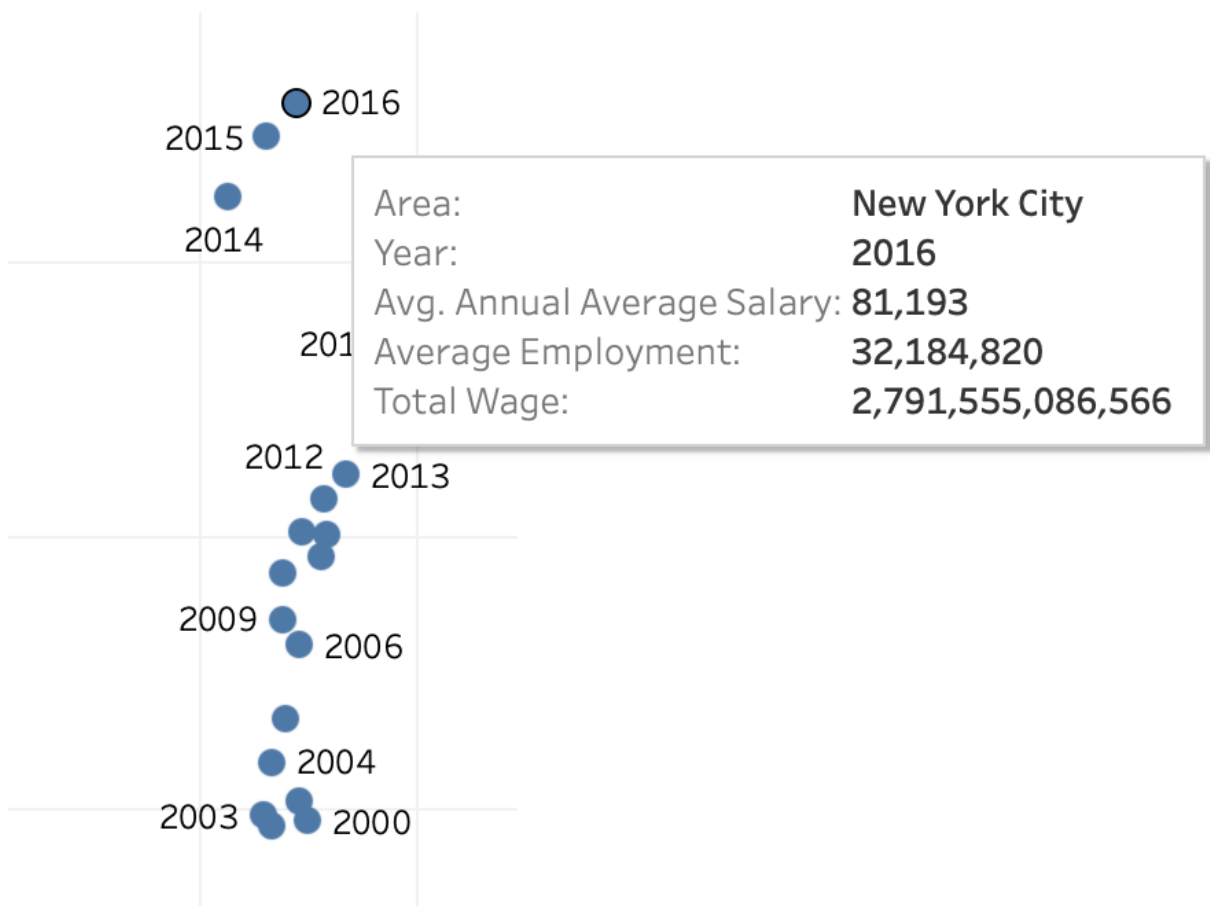
Attributes used for this plot: Average employment, Total wage

Columns: SUM(Average employment)

Rows: SUM(Total wage)

Details: AVG(Annual Average Salary), Area

As we have plotted average employment vs total wage, we can see that we get an incremental line as year increases. This shows that the total wage and employment rate is increasing as years pass by. After using marks to insert the average annual salary and area in the details section, we get the following hovering data for 2016 which is the latest year of data available in the dataset:

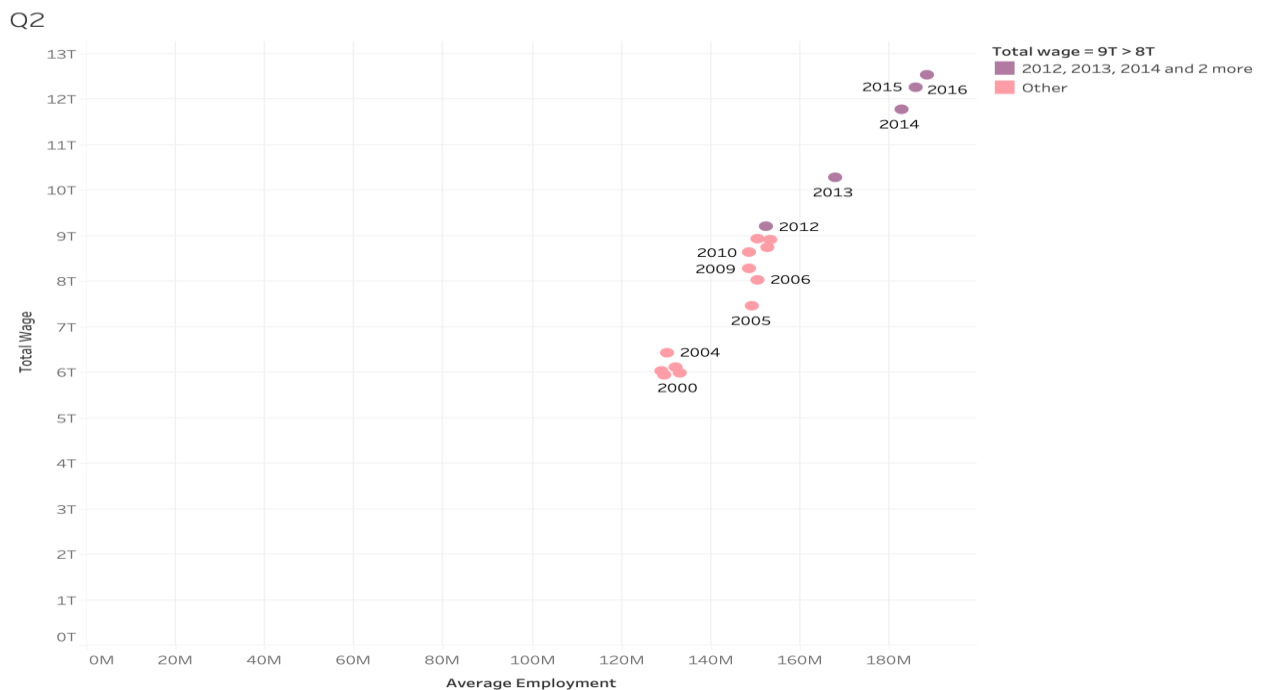


## ○ Inferences drawn from the visualizations

- The employment rate and total wage increases over the years.
- The highest total wage is from the New York City area which has approx. 32M employees.
- The average annual salary is \$81,193 for this area.

## 2) Q2

### ○ Plot



Sum of Average Employment vs. sum of Total Wage. Color shows details about Total wage = 9T > 8T. The marks are labeled by Year.

### ○ Plot analysis

Attributes used for this plot: Average employment, Total wage

Columns: SUM(Average employment)

Rows: SUM(Total wage)

Details: AVG(Annual Average Salary)

Colors: Year

Let us use the concept of groups to create a visualization using similar data attributes as that of the previous visualization. But in this one we are not going to use the area attribute in the details section. Here we are grouping all the years that has a total wage of over 9 trillion together. Since we have grouped them together, it becomes easy to differentiate the years more clearly.

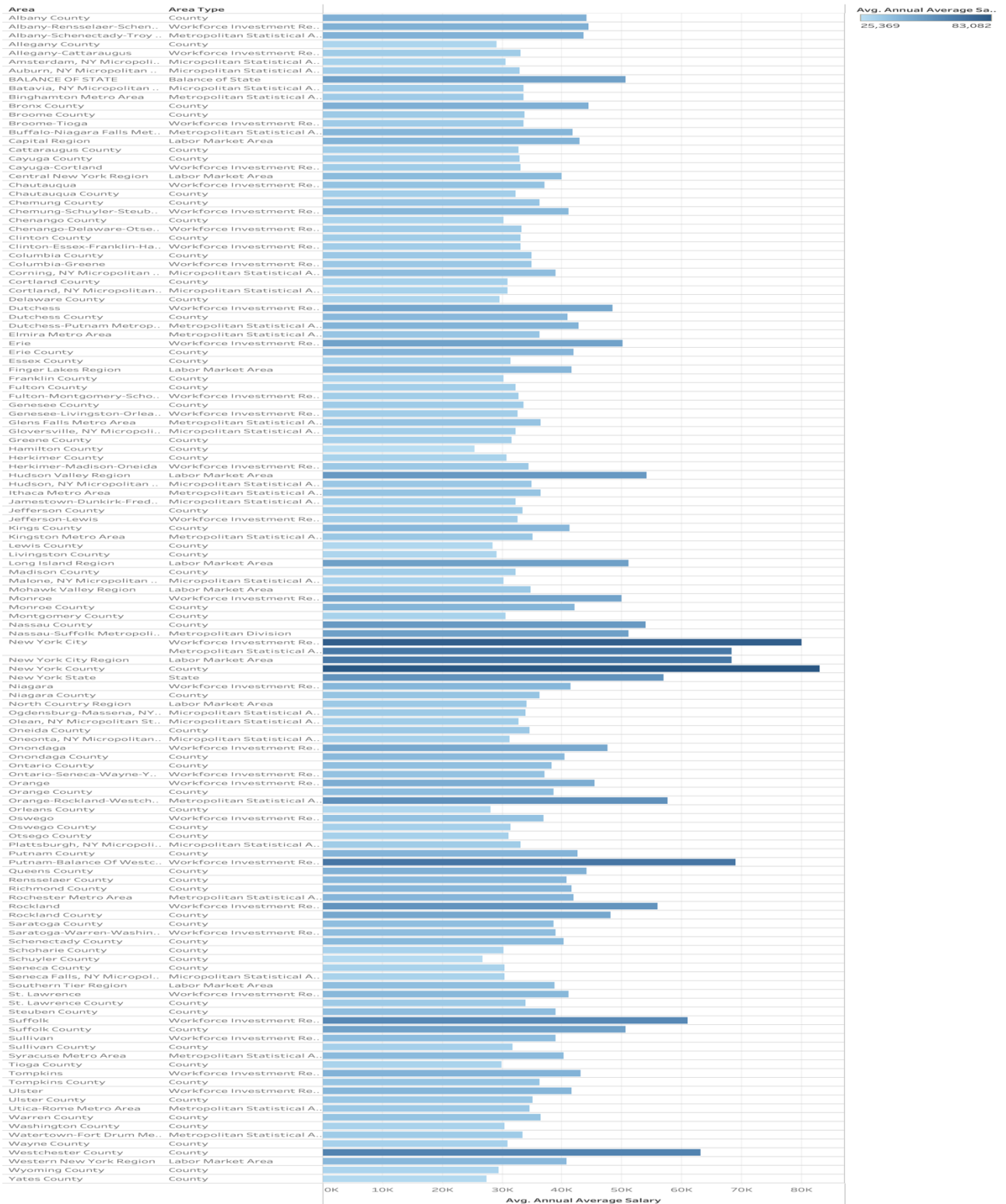
### ○ Inferences drawn from the visualizations

- The total wage is over 9 trillion from the year 2012.
- Grouping the years from 2011 to 2016 together allows us to highlight and distinguish them from the other years easily.

### 3) Q3 and Q4

#### Plot

Q3 and Q4



---

## ○ Plot analysis

Attributes used for this plot: Annual average salary, Area, Area type

Columns: AVG(Annual average salary)

Rows: Category – Area, Sub-category – Area type

Color: AVG(Annual average salary)

As we implement categories and hierarchy into the visualization for this dataset, we can get more insights about each region and how much is the average annual salary in these regions. We can use this to derive insights about which area and the area type has the highest or is the top 3 for average annual salary. We can use AVG(Annual average salary) in colors to highlight the ones with high average salary values.

Here we have used the area data as the main category and the area type as its sub-category. Also, since they both are related and can be prioritized one after another, we can create a hierarchy containing area and area type below area. Using these dimensions, we can visualize and derive inferences accordingly. Using these concepts allows us to make visualizing easier and more efficient.

## ○ Inferences drawn from the visualizations

- The highest annual average salary is shown to be in the New York County area which is a county area type with over \$83k.
- The second is followed by the New York City which has the area type defined as workforce investment region with over \$80k.
- Most of the areas that has a high annual average salary value is in New York.

## References

<https://data.amerigeoss.org/nl/dataset/quarterly-census-of-employment-and-wages-annual-data-beginning-2000>