

About the dataset:

The dataset provided to us namely 'employment-and-wages-annual-data-beginning-2000' is a quarterly census data that encompasses information about employees and their wages at the various area types such as state, county, labor market area, metropolitan statistical area and workforce investment region. The dataset has 9 features namely Area Type, Area, NAICS (code given to various work), NAICS title (Name of the industry/domain), Year, Establishment (Number of organizations in the domain), Average Employment, Total Wages and Annual Average salary. Out of the 9 given features 3 features Area Type, Area and NAICS are categorical and the rest of the 6 features are numerical.

Fig 1 - quarterly-census-of-employment-and-wages-annual-data-beginning-2000

Area Type	Area	NAICS	NAICS Title	Year	Establishments	Average Employment	Total Wage	Annual Average Salary
State	New York State	11	Agriculture, Forestry, Fishing and Hunting	2016	2600	25849	872749893.00	33763.00
State	New York State	111	Crop Production	2016	1128	12379	377915334.00	30529.00
State	New York State	112	Animal Production	2016	833	10204	366444605.00	35912.00
State	New York State	113	Forestry and Logging	2016	196	644	24957438.00	38754.00
State	New York State	114	Fishing, Hunting and Trapping	2016	42	103	3515772.00	34134.00
State	New York State	115	Agriculture & Forestry Support Activity	2016	402	2519	99916744.00	39665.00
State	New York State	21	Mining	2016	366	4409	281253726.00	63791.00
State	New York State	211	Oil and Gas Extraction	2016	32	326	22490945.00	69991.00
State	New York State	212	Mining (except Oil and Gas)	2016	262	3749	233708577.00	62339.00
State	New York State	213	Support Activities for Mining	2016	72	333	25054204.00	75238.00
State	New York State	22	Utilities	2016	589	37390	4279229991.00	114449.00
State	New York State	221	Utilities	2016	589	37390	4279229991.00	114449.00
State	New York State	23	Construction	2016	48002	374247	26290108729.00	70248.00
State	New York State	236	Construction of Buildings	2016	16962	98074	6753128950.00	68857.00
State	New York State	237	Heavy and Civil Engineering Construction	2016	1532	29986	2973768680.00	99172.00
State	New York State	238	Specialty Trade Contractors	2016	29508	246186	16563211099.00	67279.00
State	New York State	31	Manufacturing	2016	17009	447472	28777984475.00	64312.00
State	New York State	311	Food Manufacturing	2016	2370	55795	2465883744.00	44195.00
State	New York State	312	Beverage & Tobacco Product Manufacturing	2016	432	8945	552341861.00	61749.00
State	New York State	313	Textile Mills	2016	215	3497	209676183.00	59959.00
State	New York State	314	Textile Product Mills	2016	326	3735	185473748.00	49658.00
State	New York State	315	Apparel Manufacturing	2016	1126	16263	870897174.00	53551.00
State	New York State	316	Leather and Allied Product Manufacturing	2016	107	1603	89261168.00	55684.00
State	New York State	321	Wood Product Manufacturing	2016	492	8336	362363388.00	43470.00
State	New York State	322	Paper Manufacturing	2016	276	15892	949559268.00	59751.00
State	New York State	323	Printing and Related Support Activities	2016	1619	20964	1062127605.00	50664.00
State	New York State	324	Petroleum & Coal Products Manufacturing	2016	111	1824	126969553.00	69611.00
State	New York State	325	Chemical Manufacturing	2016	610	37423	3049878840.00	81497.00
State	New York State	326	Plastics & Rubber Products Manufacturing	2016	432	21207	1119452714.00	52787.00
State	New York State	327	Nonmetallic Mineral Product Mfg	2016	683	15512	928555304.00	59860.00
State	New York State	331	Primary Metal Manufacturing	2016	191	10096	651768159.00	64557.00
State	New York State	332	Fabricated Metal Product Manufacturing	2016	2268	50046	2856719138.00	57082.00
State	New York State	333	Machinery Manufacturing	2016	1007	39688	2925598665.00	73715.00
State	New York State	334	Computer and Electronic Product Mfg	2016	950	58040	5620384376.00	96836.00
State	New York State	335	Electrical Equipment and Appliances	2016	358	13472	873142847.00	64812.00
State	New York State	336	Transportation Equipment Manufacturing	2016	348	19794	1414316480.00	71488.00
State	New York State	337	Furniture and Related Product Mfg	2016	1017	13026	633661222.00	48646.00
State	New York State	339	Miscellaneous Manufacturing	2016	2071	32326	1829953038.00	56609.00
State	New York State	42	Wholesale Trade	2016	34906	336448	27861205784.00	82810.00
State	New York State	423	Merchant Wholesalers, Durable Goods	2016	14703	147413	11593008167.00	78643.00
State	New York State	424	Merchant Wholesalers, Nondurable Goods	2016	11948	146280	11624135385.00	79465.00
State	New York State	425	Electronic Markets and Agents/Brokers	2016	8255	42755	4644062232.00	108620.00
State	New York State	44	Retail Trade	2016	74685	938267	31814913634.00	33908.00

Fig. 1 Screenshot of dataset

Use of dataset:

This dataset can be used to study detailed statistics about various industries (also can be used to get an idea about which industry employs maximum or minimum depending on their location i.e., area type), their employment depending. Also, this dataset can be used to visualize annual salary in an industry based on their location. Plus, the dataset is also helpful in getting a general idea about the total wages and number of establishments in a particular job domain.

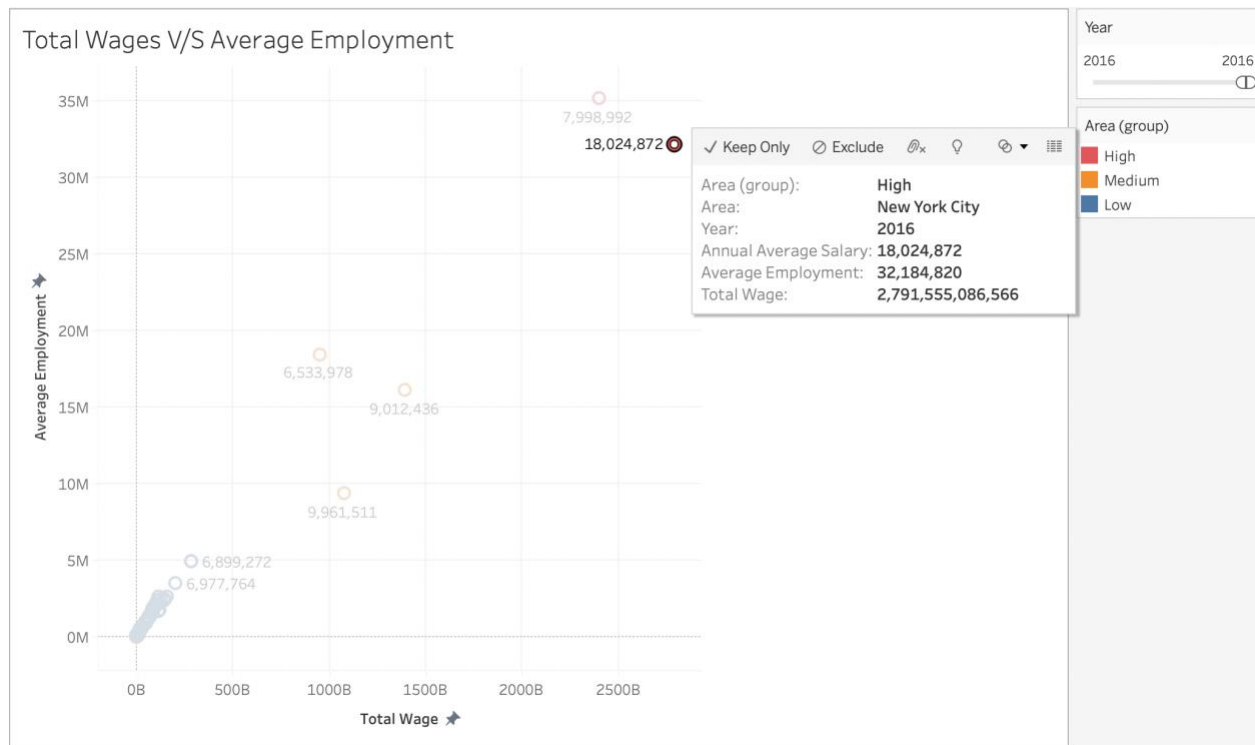


Fig. 2 Total Wages V/S Average Employment

Inference from visualization:

In this scatterplot we created 3 categories: High, Medium and Low depending on their position in the scatter plot (Wages vs Average Employment)

This graph shows us which areas have high average employment and total wages (New York City and New York State), areas which have medium average employment and total wages (New York County, Balance of state and New York City Region) and areas with low average employment and total wages (Long Island Region, Hudson Valley region, etc.)

Filter has been applied to the years measure to display the data of the year 2016 and on total wages as it consisted of 7 null values which needed to be taken care of.

The aim of creating this sheet was to create a scatterplot to visualize the "Total Wage" versus "Average Employment".

Further, in the plot, shows that New York City (area) having the average employment of 32,184,820 has the Highest Total Wage in 2016.

Also, as required by the question the marks display the Annual Average Salary in that area i.e., 18,024,872.

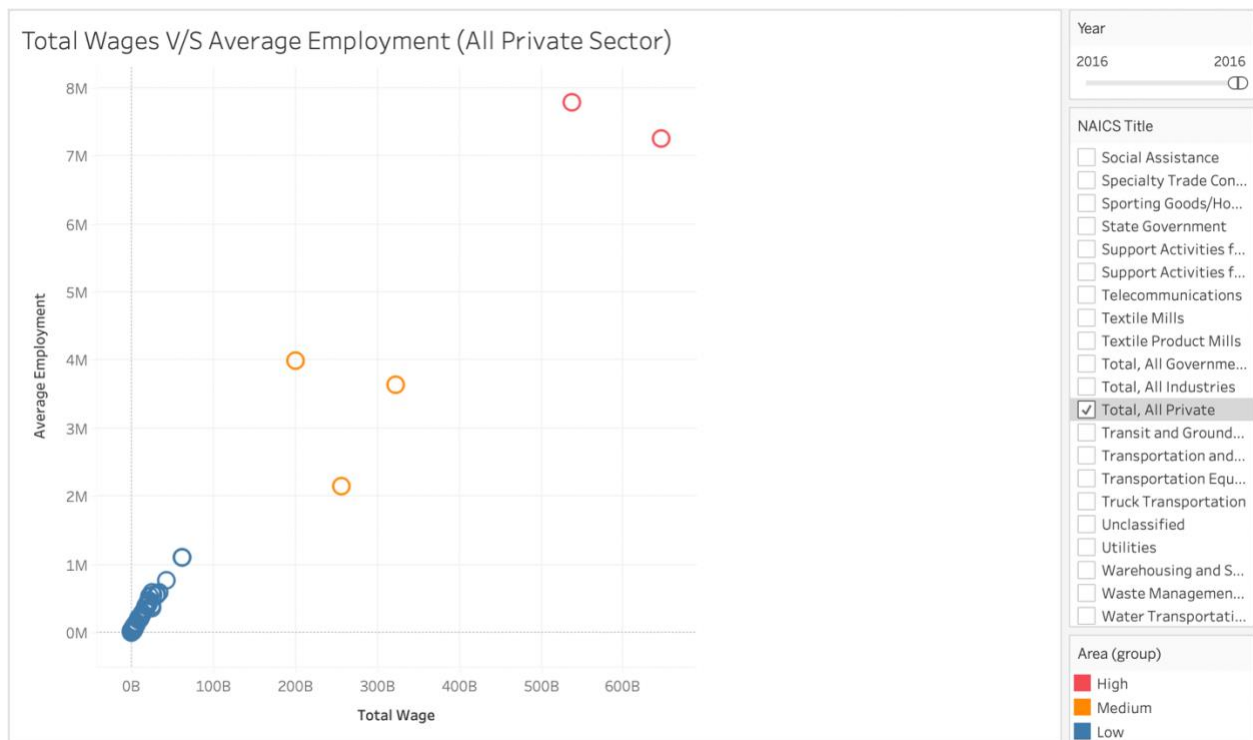


Fig. 3 Total Wages V/S Average Employment (All Private Sector)

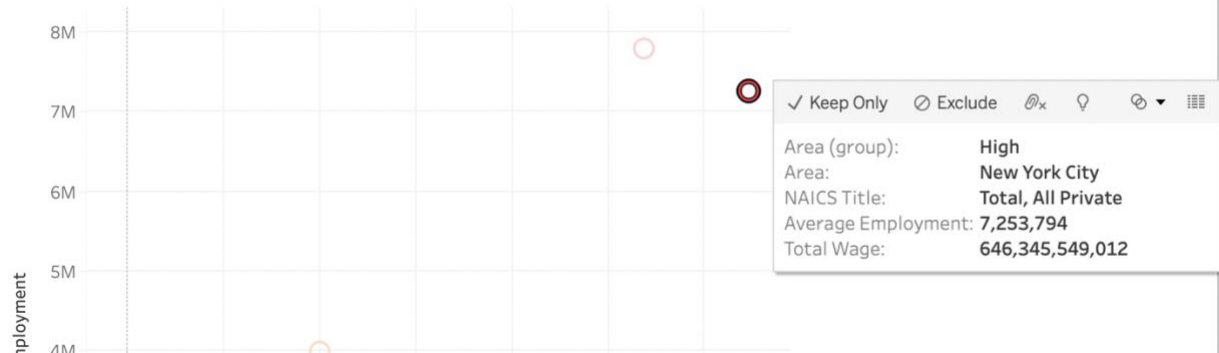
Inference from visualization:

The above scatter plot was plotted to find out why there is a difference between the employment and total wages of New York City and New York State.

To analyze that we further divided the category Area into 2 parts: One containing NAICS title (All private: Sheet 2) and another NAICS title (All Government) to evaluate if the employment by these two sector is due to which there is a gap in the average employment and total wages of both the regions.



### Total Wages V/S Average Employment (All Private Sector)



The comparison between the Private sector yields the following result:

Region	Average Employment	Total Wages
New York City	7,253,794	646,354,549,012
New York State	7,781,866	537,782,552,480

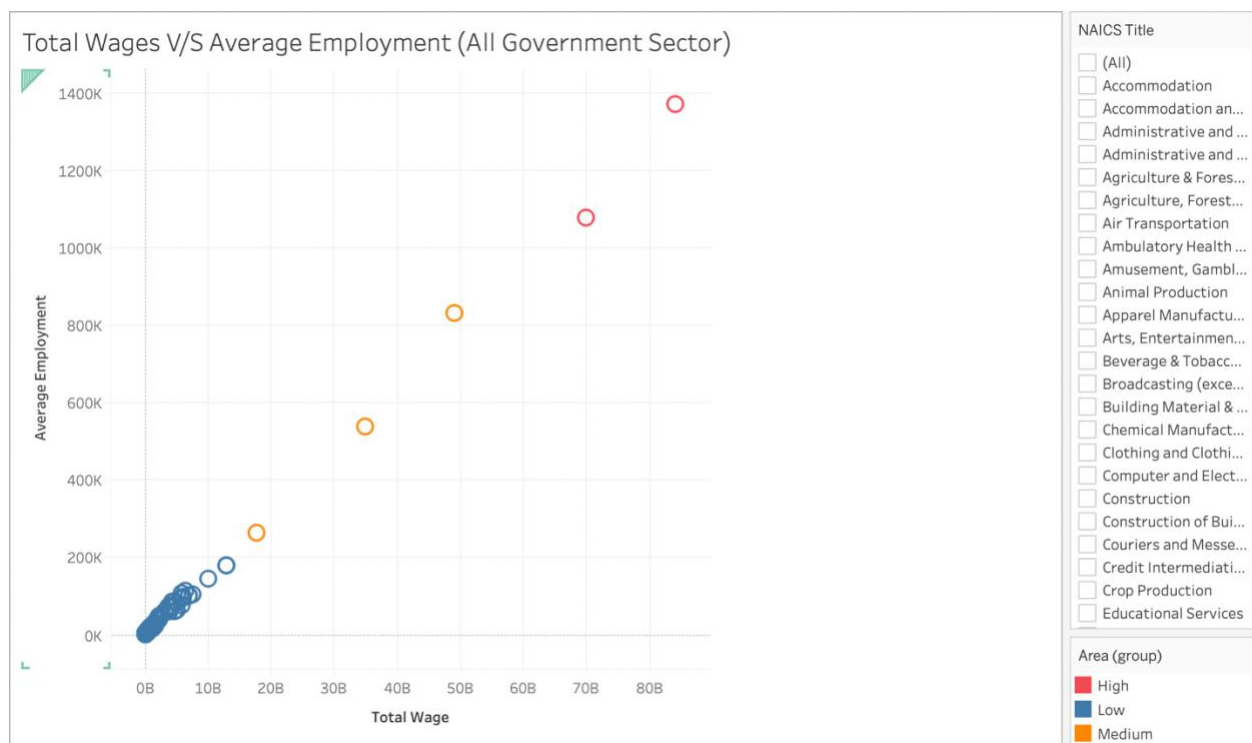


Fig. 4 Total Wages V/S Average Employment (All Government Sector)

Inference from visualization:

Having further analyzed the private sector we move forward and check whether the employment and wages of the government sector are the reason for disparity for the two regions.



The comparison between the government sector of the two regions show us that the average employment by the New York State is higher but the total wages of the New York City is much more. This difference is due to the total wages earned by the private sector of New York City is very much greater than that of New York State.

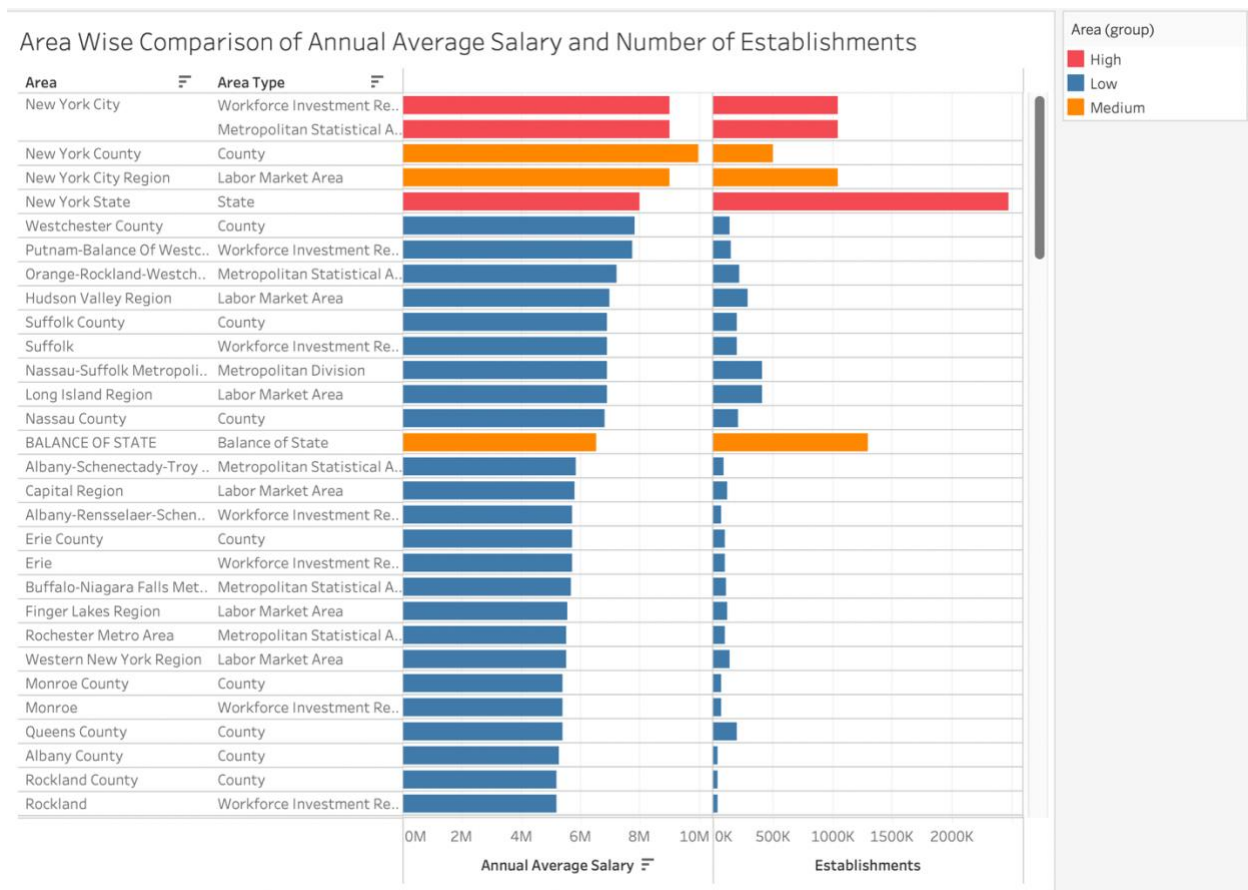


Fig. 5 Comparison of Annual Average Salary and Establishments (Area Wise)

Inference from the visualization:

This bar plot helps us infer why the average employment in New York State is more.

The reason being the number of establishments present in the region is more. This trend follows for the whole dataset as the regions which have a higher number of establishments generate more employment.

But that does not necessarily mean that the annual average salary of the region is more which can be clearly seen in the graph above.

New York City has a significantly smaller number of establishments then New York State but the average annual salary for New York City is comparatively greater. The same inference can also be made about New York County and New York City Region.

