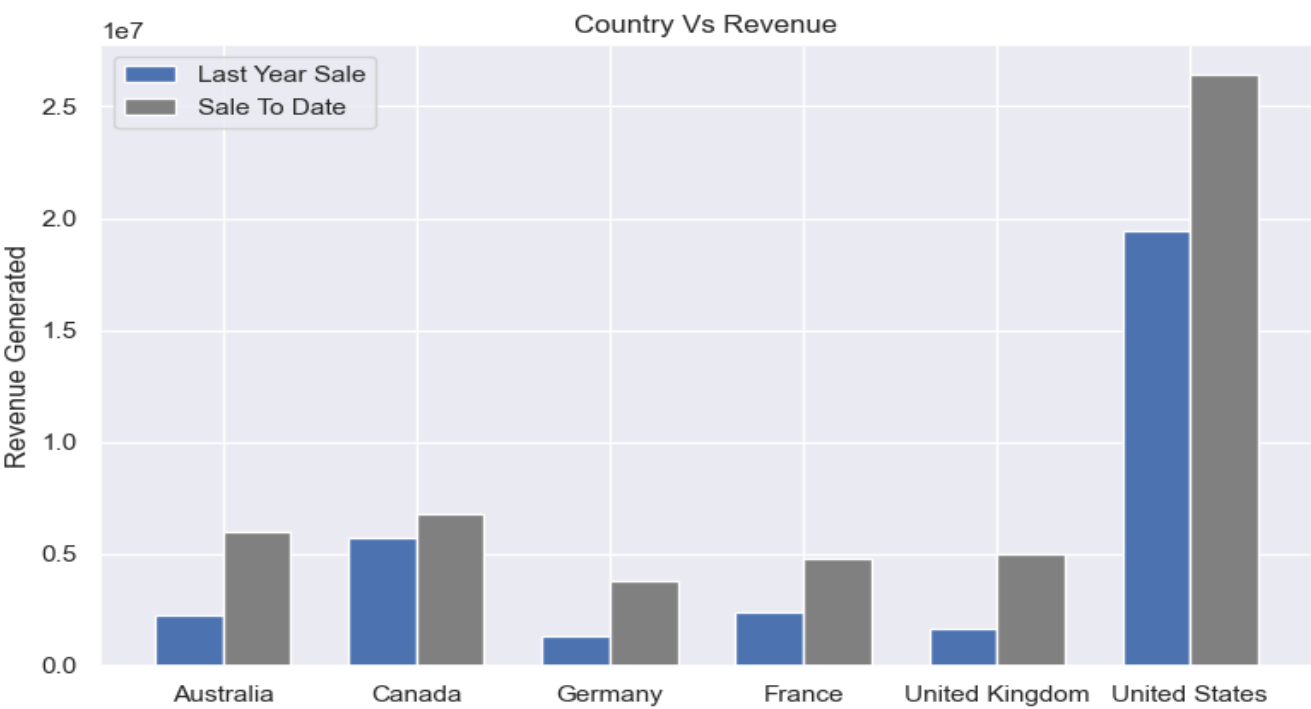
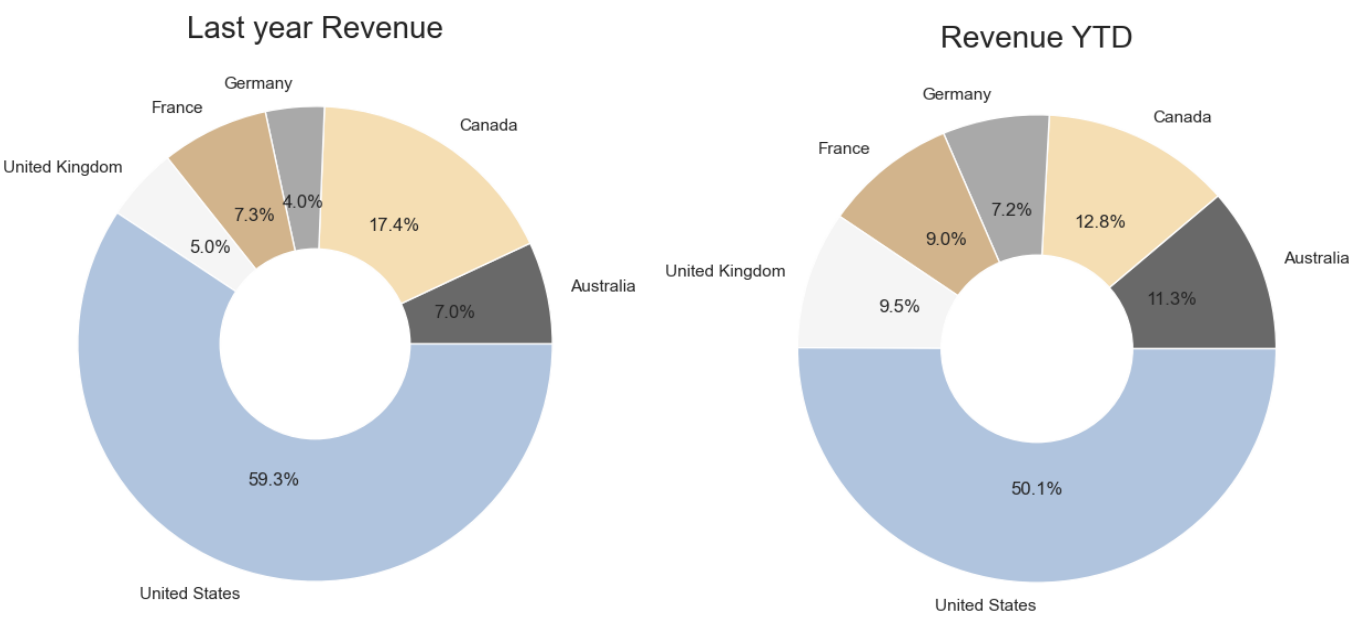


What is the relationship between Country and Revenue?

From the data given in AdventureWorks database about the revenue generated by each country from sales last year and year to date, from the visualization, it can be interpreted that the US contributes the largest part of the revenue generated by each country as the total of this year and last year's sales.



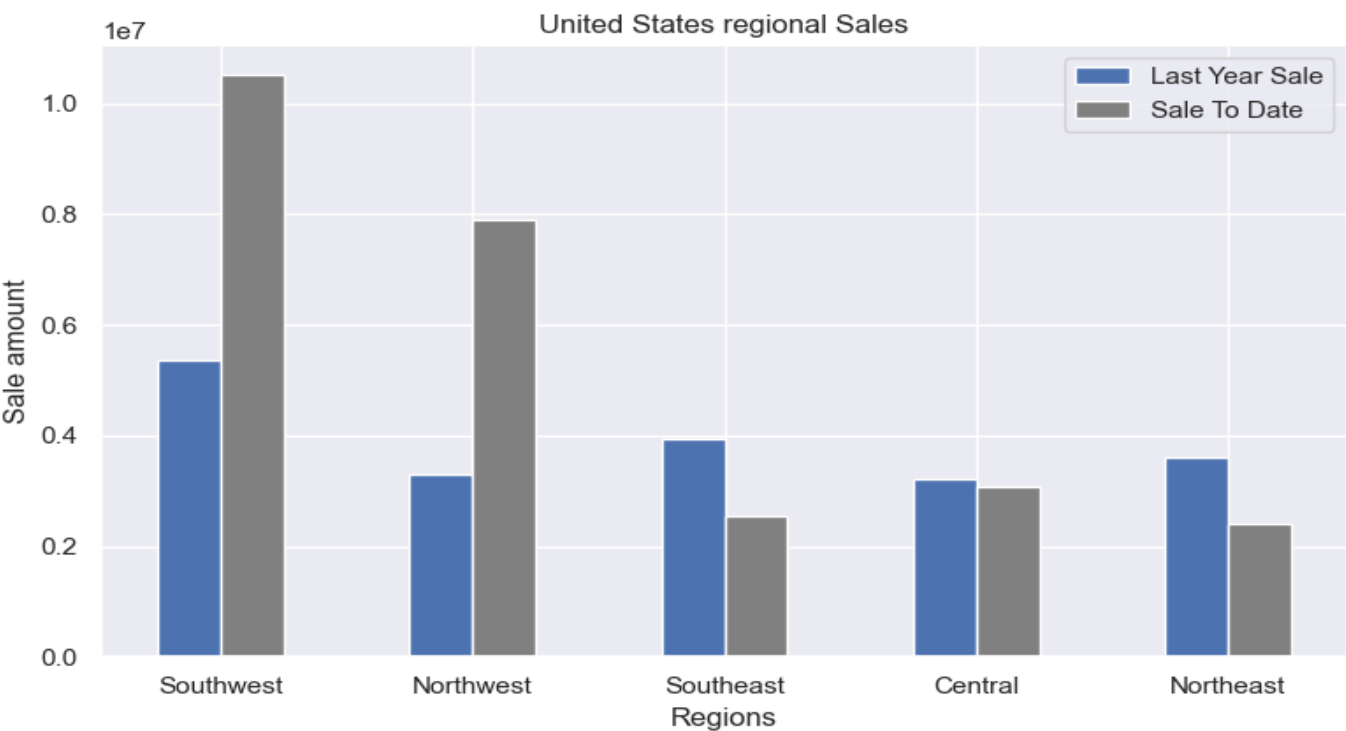
Although overall YTD total revenue in other regions increased as compared to last year's revenue i.e., 40.7% last year's revenue increased by 9.2%, reaching 49.9% YTD revenue, yet the US contributes more than 50% in the revenue.



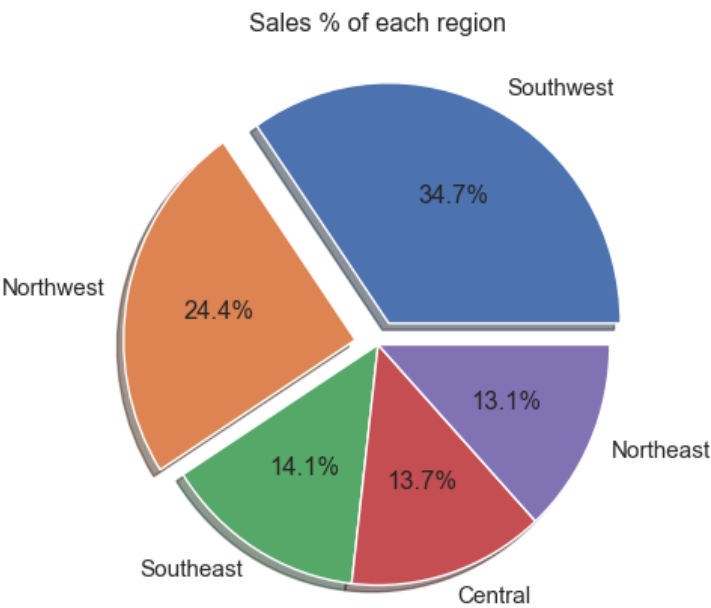
The US has several regions for sales than other countries which is contributing to higher revenue than other countries. Recommendations could be to expand the market in Europe to reach more countries. Plan attractive sales strategies to improve sales in the low revenue generating area.

What are the regional sales in the best-performing country?

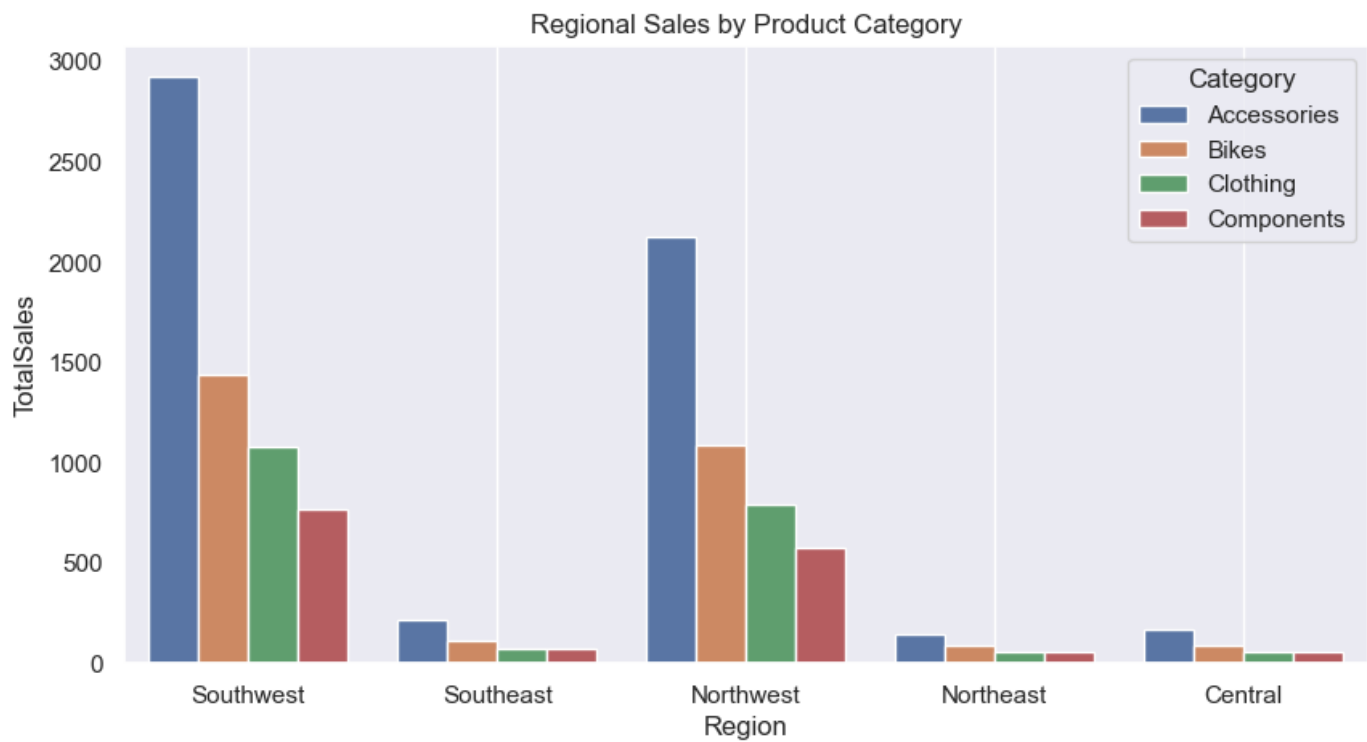
The analysis above shows that the US is the best-performing country in terms of total sales of last year and year to date among all the countries. The US has several regions of sales like Southcoast, Northcoast and central. we can see from the bar chart that the southwest and northwest are the biggest contributors to US sales.



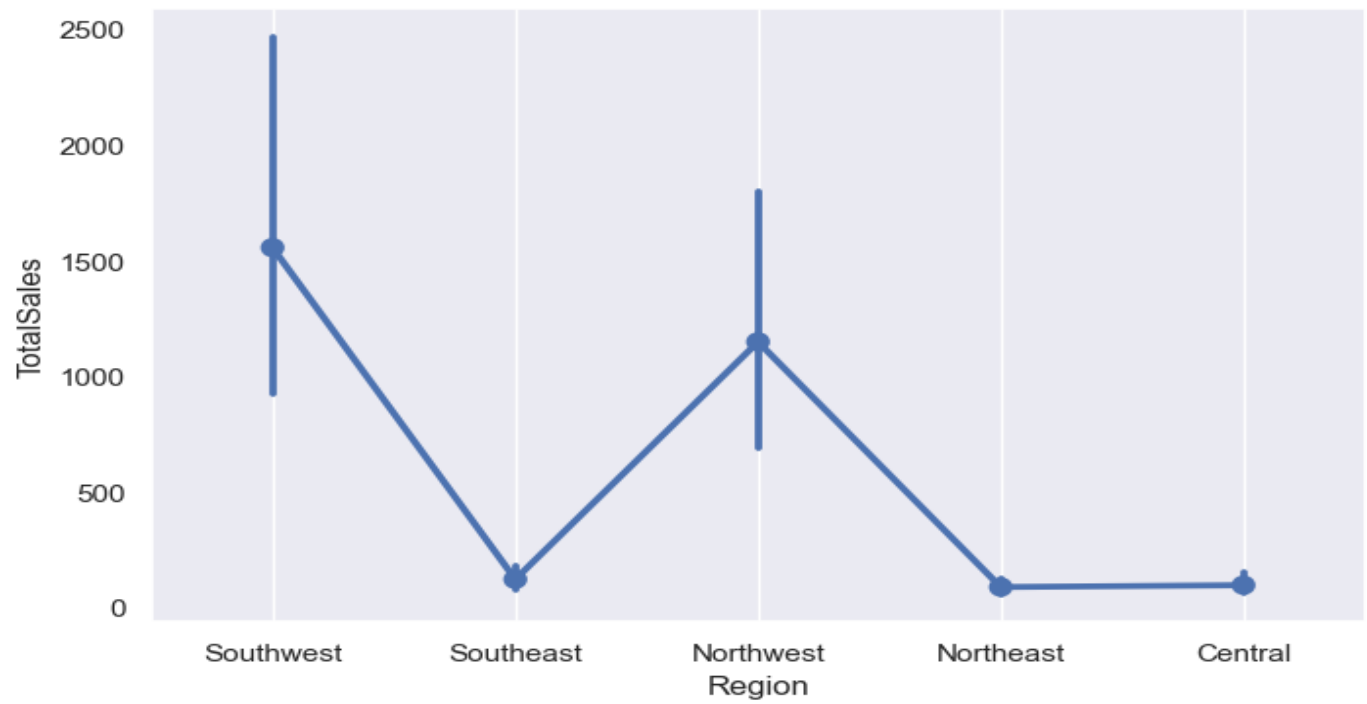
also from the pie chart, we can see that south and northwest combined contribute approx. 60 sales. Southwest has 34.7% and Northwest h 24.4% of contribution in sales. While all other combined sales of 40.9%.



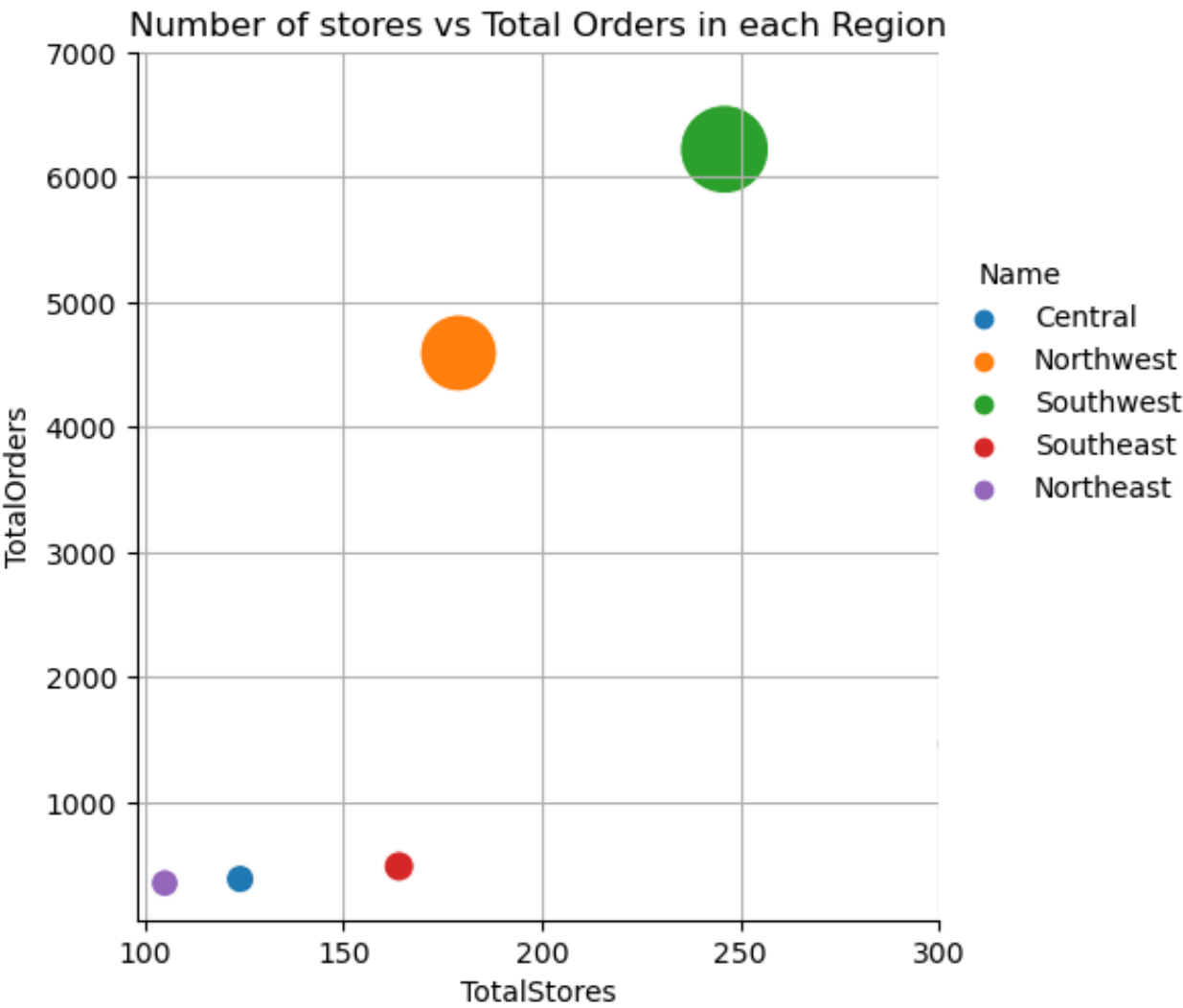
for further investigation to know why some regions have higher sales than others. I processed the data further to get information about each region's best-performing product and total orders for each region, I plotted a bar graph to visualize how the sale of each product's category is contributing to each region's sales. We can see that the category accessories is the best-selling product in both the southwest and northwest.



Using the same data, I plotted a point plot to see the average sales of each region and the average also gives us a clear picture of each region's sales.



To see why two regions are performing well while others are not doing good, I processed the data further to find the total number of orders and the total number of stores in each region, I plotted the scatter plot to do the comparison, which shows that regions with the higher number of total stores have a higher sale but also southeast's sales is very much less than its total number of stores.

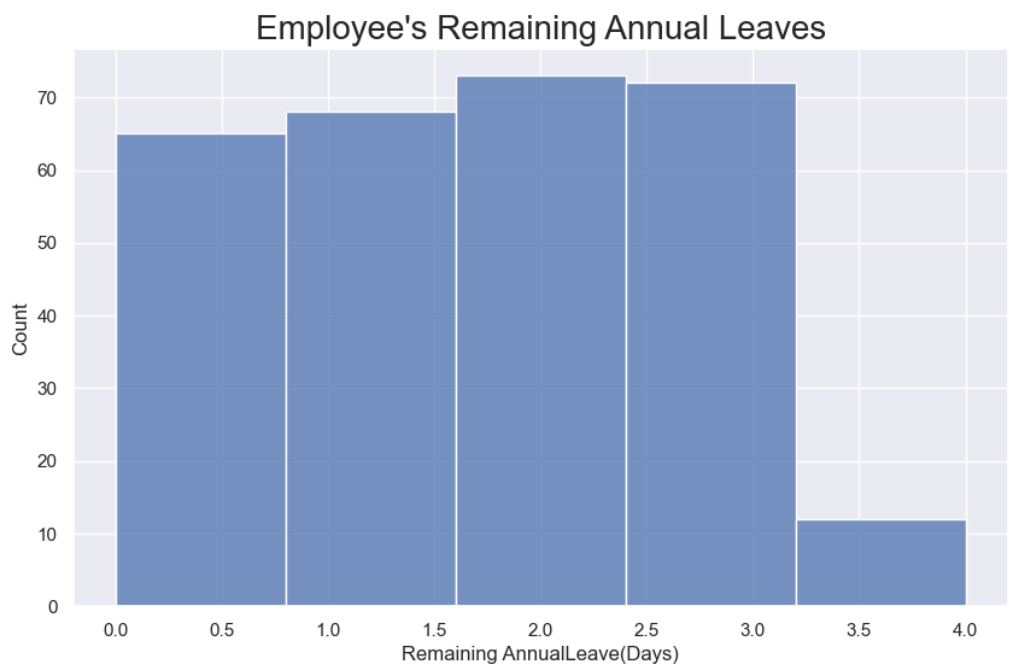


From the above analysis, we can conclude that

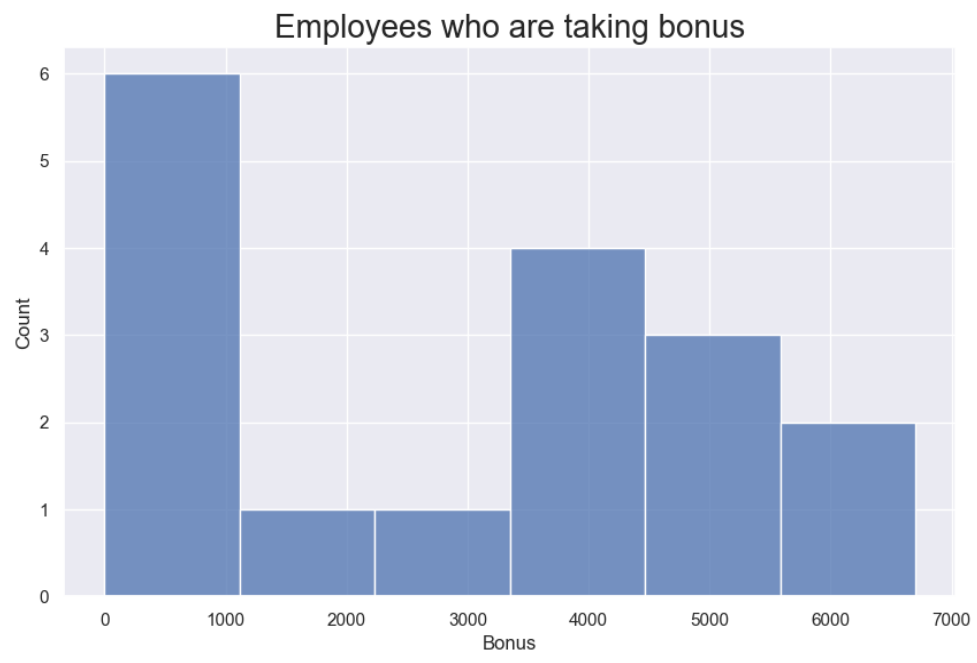
- For those regions, not performing well, the number of stores needs to be increased
- Other regions like Central, Northeast, and Southeast need to follow the sales strategy of Northwest which has much higher sales as compared to its total number of stores

What is the relationship between annual leave taken and bonus?

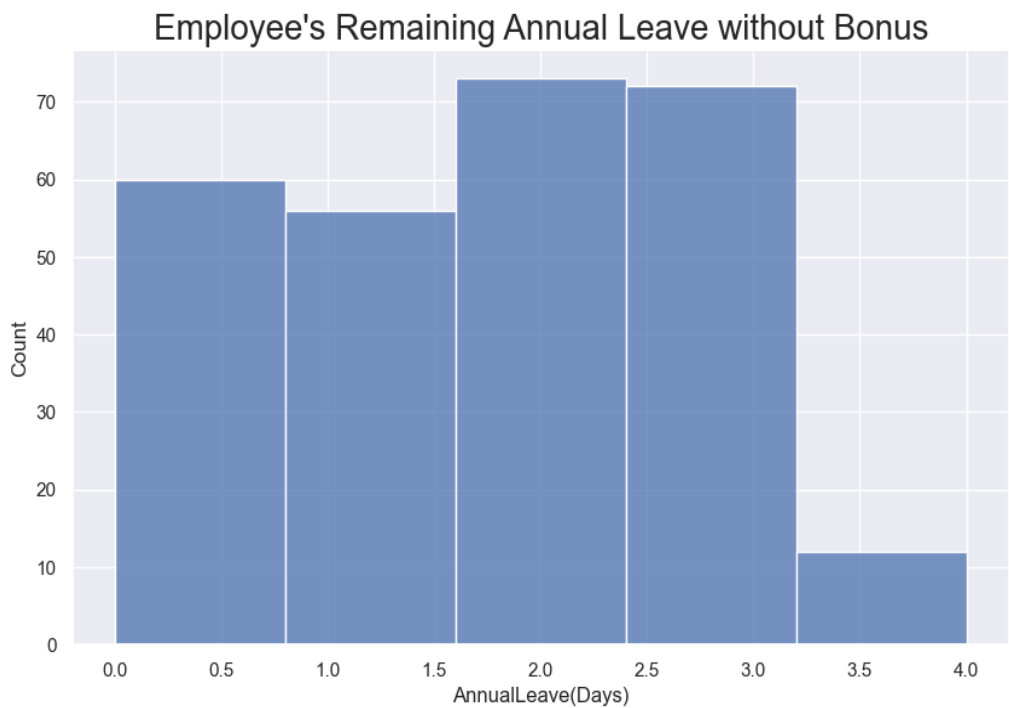
To find out the relationship between annual leave taken and bonus I first converted the vacation hours to days, in the data given, vacation hours given are the remaining leave for each employee. I created a view named Annual Leave which includes the remaining vacation days and bonus of each employee, but the data shows that more employees are not getting bonuses as compared to those who are getting bonuses, to visualize that I used a histogram to see the distribution of remaining vacation days of each employee, it shows almost every employee has taken the leave and few days are left for each employee.



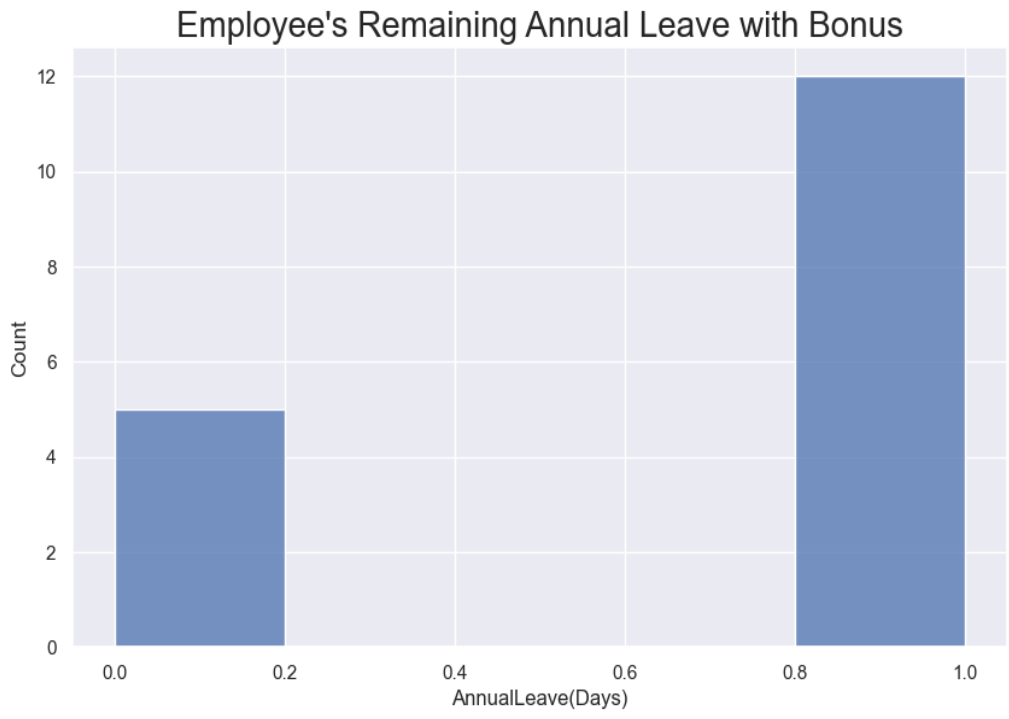
The distribution of bonuses shows there is a very small number of employees who are given bonuses. If we compare the height of distribution, we can see the difference in the number of employees who have taken annual leaves and who are given bonuses.



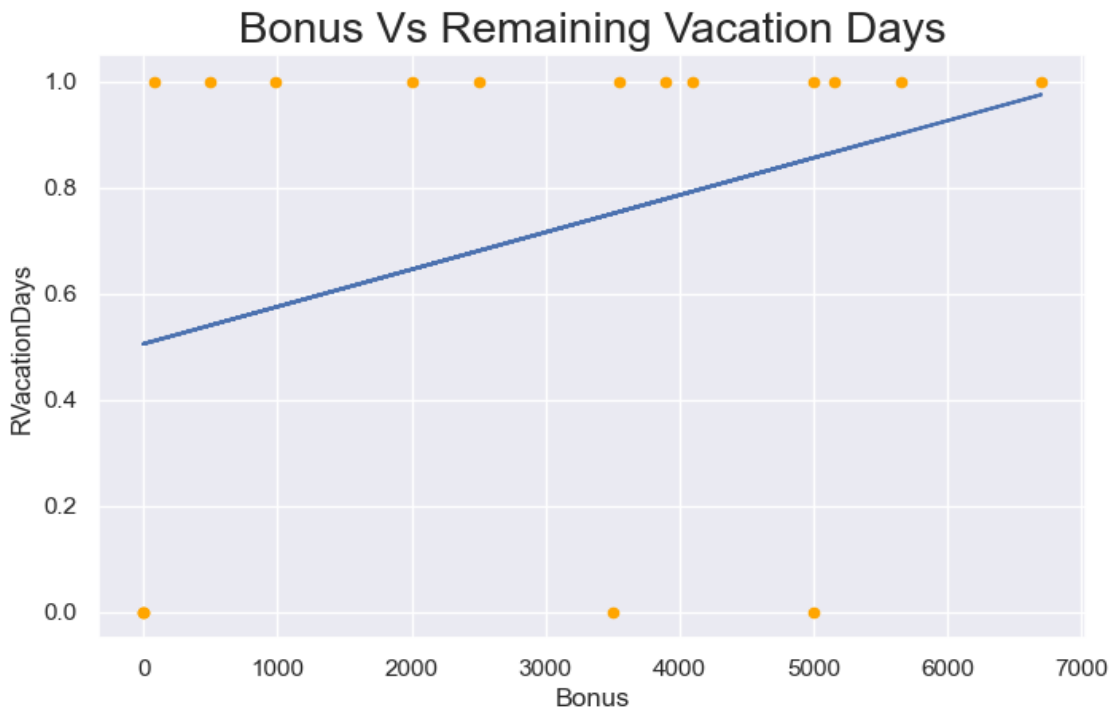
To further analyze the relationship, I used a histogram to see the distribution of employees taking annual leaves with and without bonuses, using the same table from the database, the distribution shows the employees who do not take bonuses about half of the employees have less than 4 days of leaves left.



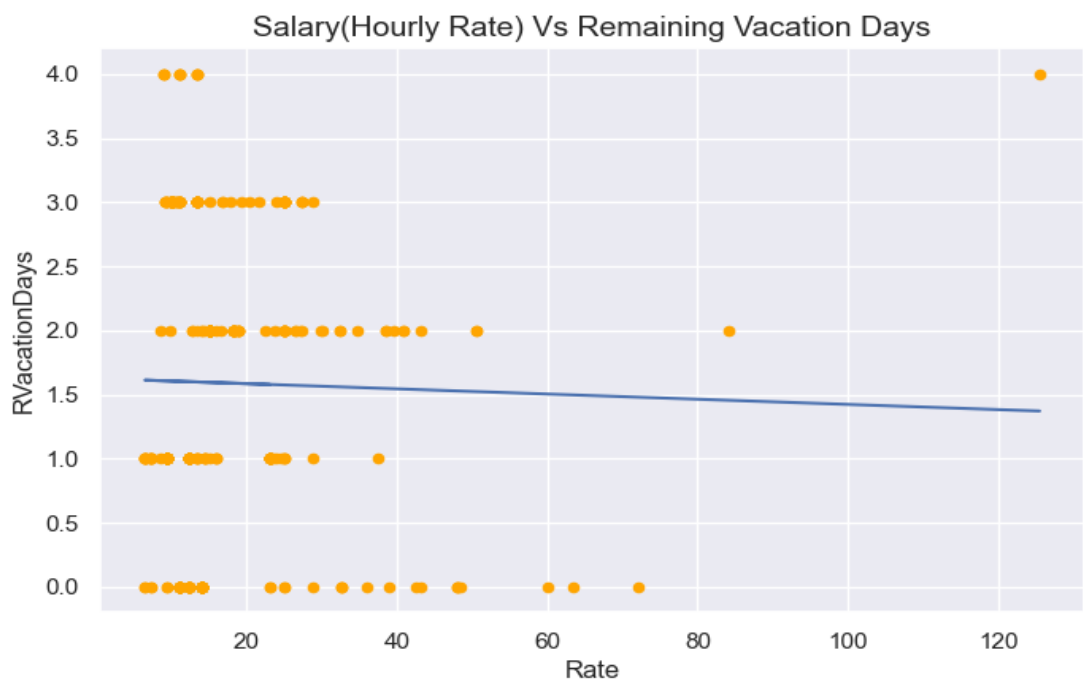
For those employees who take bonuses, very few employees have less than a day of leave left. As the number of employees who do not take bonuses is much higher than those who take bonuses, the relationship between them cannot be found with certainty.



To find out the relationship between bonuses and annual leave, I used a scatter plot to find out the correlation between them, which shows a very weak positive relationship (corr = 0.34) between these two variables i.e., if the bonus increases the remaining annual leave also increases (slightly) means the annual leave taken will decrease,



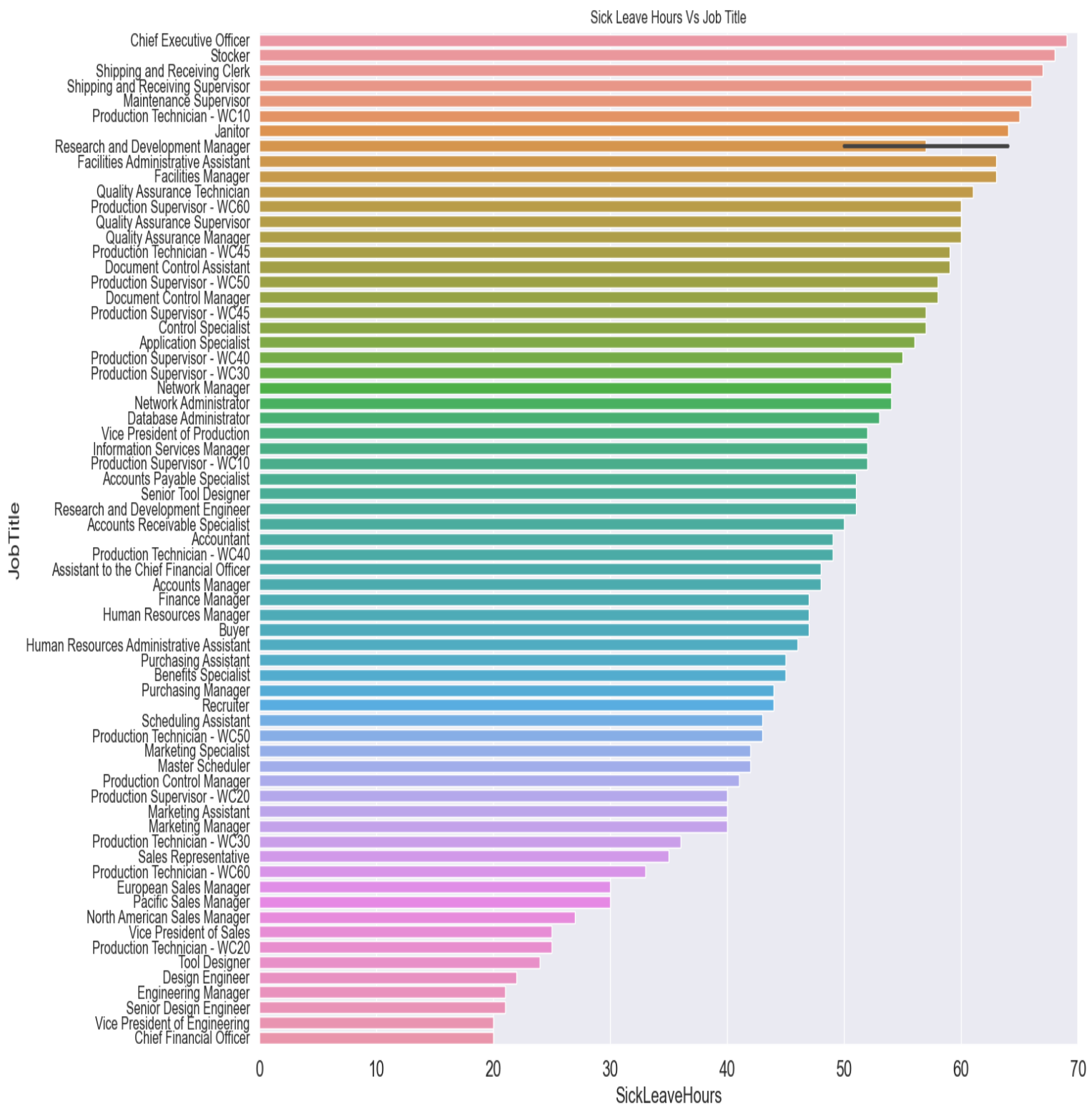
to see if there is any confounder variable affecting the annual leave, I used a scatter plot to see the correlation between salary (hourly rate) and annual leave, it shows a weak negative relationship (corr=-0.02) means remaining leave decreases (i.e., annual leave taken increases) as hourly rate increases, that means Annual leave is not much affected by bonus but by confounder variable (slightly).



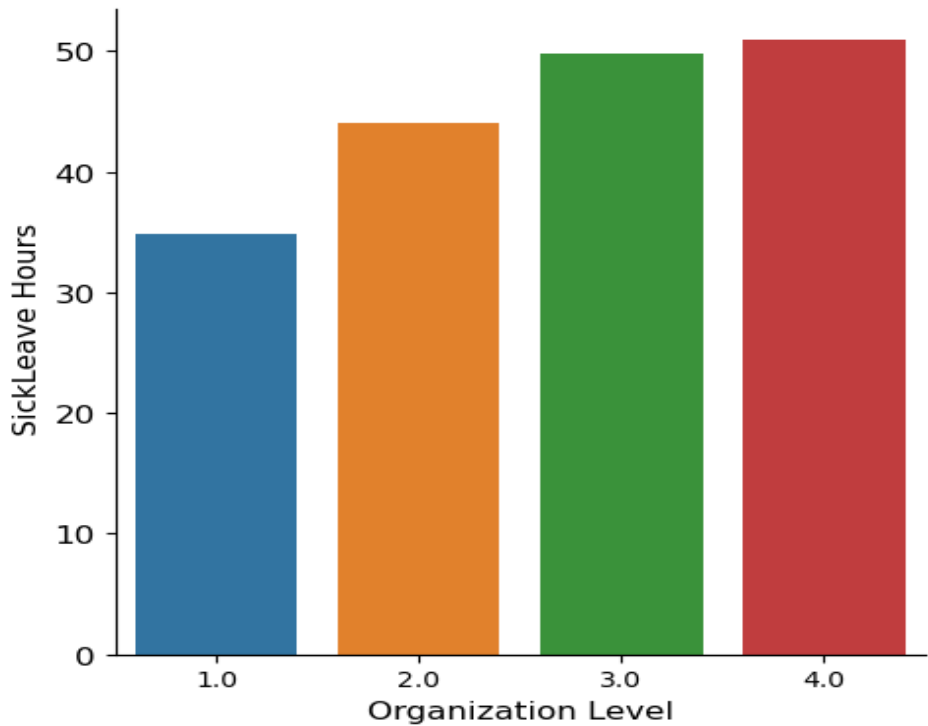
To conclude, it can be said that by the data given, there is a very weak relationship between annual leave taken and bonus.

What is the relationship between sick leave and Job Title?

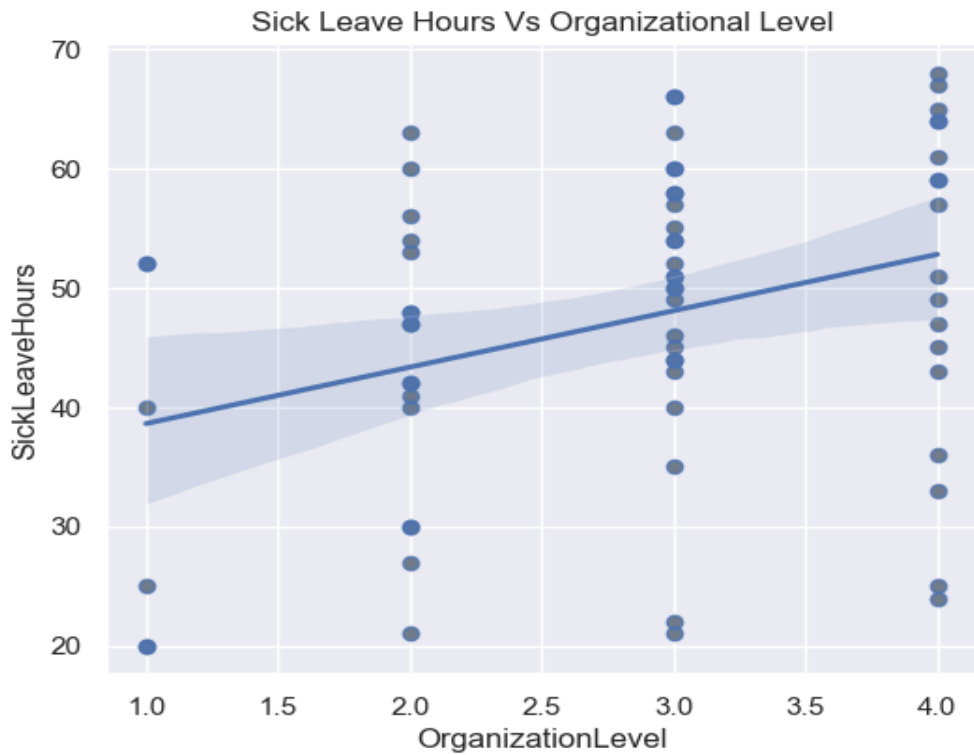
To find out the relationship between sick leave and job title, I used the bar plot to see the relation between job title and sick leave, we can see after sorting, the graph follows a pattern from maximum to minimum sick leave taken, CEO has taken the highest average sick leave, and CFO has taken the minimum average sick leave.



To further investigate the relationship, I plotted a bar graph with Organization level and sick leave hours, It can be seen that there is a clear relationship between Organization level and sick leave hours, higher Organization levels have fewer sick leave hours. (Note: Low organization level number=higher organization level)

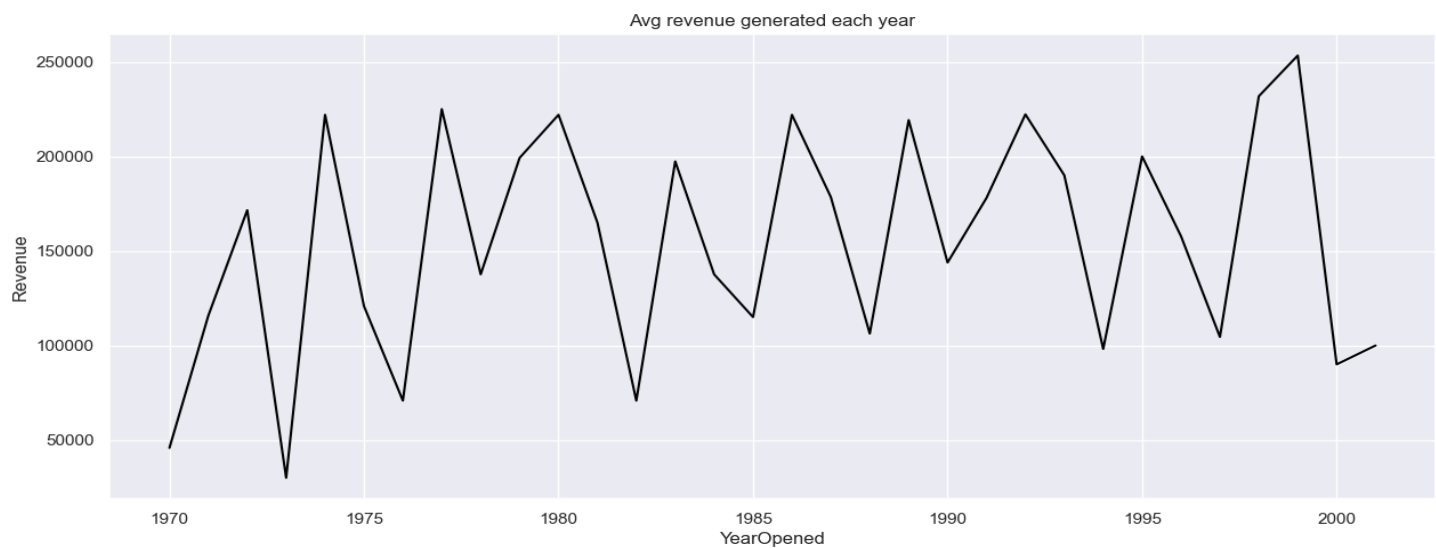


The scatter plot below shows a weak positive correlation i.e., **0.33**, which indicates that as the number of organization level increases the sick leave hours also increases.



What is the relationship between store trading duration and revenue?

To find out the relationship between store trading duration and revenue I first plotted a line graph to see the average revenue generated by each store opened in a specific year, the graph below shows no pattern in an increase of revenue each year, the maximum revenue was generated in 1999 while the lowest was generated in 1973.

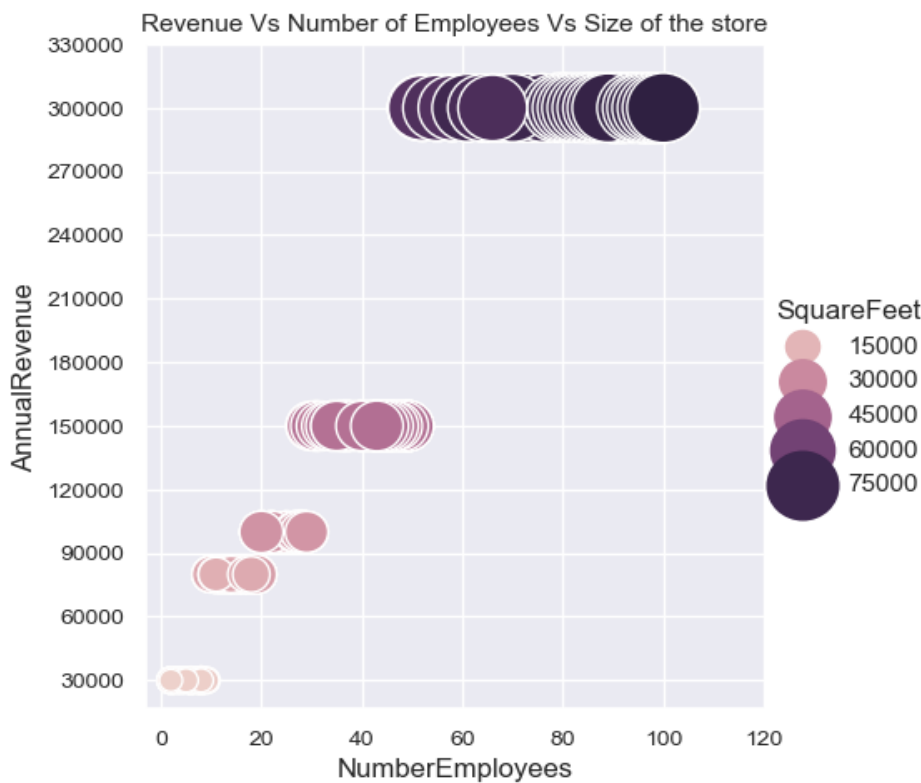


The scatter plot below shows a weak negative correlation between the duration of a store opening and revenue i.e., **-0.22**, which indicates that as the duration of a store opening increases the revenue generated decreases

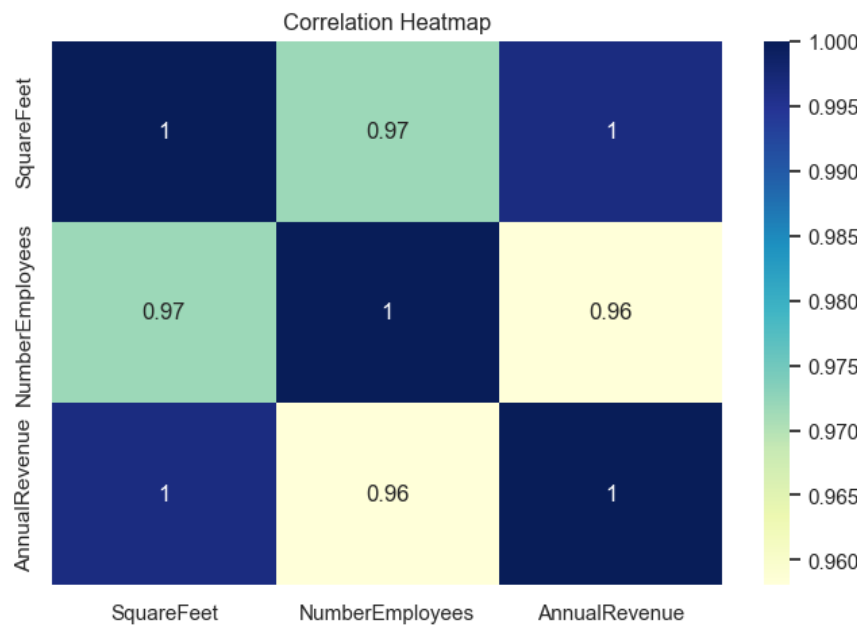


What is the relationship between the size of the stores, the number of employees, and revenue?

To find out the relationship between the size of the stores, the number of employees, and the revenue I used a scatter plot to visualize the relationship among these three variables, it can be seen that as the number of employees in a store increases and also when the size of the store is greater the revenue generated by these stores is also greater.



The heatmap below shows a very strong correlation among all these three variables with the lowest being 0.96 and the highest being the perfect 1.



Conclusion:

- The US has several regions for sales than other countries which is contributing to higher revenue than other countries. Recommendations could be to expand the market in Europe to reach more countries. Plan attractive sales strategies to improve sales in the low-revenue-generating area
- For those regions in the US, not performing well, the number of stores needs to be increased
- Other regions like Central, Northeast, and Southeast need to follow the sales strategy of Northwest which has much higher sales as compared to its total number of stores
- There is a weak positive correlation between organization level and sick leave, which indicates that as the number of organization level increases the sick leave hours also increases. There must be some factors affecting the lower organization level employees' health making them take more sick leave
- There is a weak negative correlation between store opening duration and revenue, which indicates that older stores generated less revenue than newer stores
- There is a very strong correlation between the number of employees, the size of the store, and revenue, which indicates that if a store size is greater and has a greater number of employees, it generates more revenue than other stores.