

In this homework, I use highcharts package to create different line charts. Since I notice the dataset is time series data, so I choose line chart to show the trend and fluctuation of the sales data. The peaks and falls in the chart indicating the rise and falls of revenue generated by different product and different regions. I separate the data to two aspect. The first is the product. By dividing the sales data for different products, I am able to identify which product is most popular and resistant to time shift. By checking the overlapping of the sales data, I can identify the pattern in different regions. The other respect is the region, I discovered that the rise and falls are highly overlapping, meaning the patterns are highly correlated in different region. Also, I observe the data from the perspective of periodicity, and found that monthly sales have an apparent periodicity while daily sales are not that obvious. I use some aggregation function like sum and average to replace the overall statistics. So we can get rid of excessive information that distract us and generating more general rules behind the data. However, if I can regulate the line charts by assigning alike colors to different products in same region, the chart would be easier to interpret, and might lead to more disclosure of trends and information.