

# The explanation of Three Charts

## Data Introduction

In this project, the data used to do visualization was the PM 2.5 value of Shanghai, which was download from the website of United States Embassy Shanghai. The date range of the data is from Dec, 2011 to Jul, 2016. The unit of PM 2.5 is  $\mu\text{g}/\text{m}^3$ .

The Chinese Air Quality Problem catches the crowd's eyes in recent years. To solve the problem, improve air quality and give the blue sky back to national folks, the Chinese government took lots of measure and promulgate many policies in each cities. Therefore, the PM 2.5 data of Shanghai was chosen to check if the policies worked well to control the concentration of PM 2.5 in Shanghai. The data analysis and testing part, which used T-test and ANOVA to check, were included in another more complete report. The following charts were the part of data exhibition. Their target was used to simply reflect data distribution.

The next, let's begin the visualization part.

### For Stacked Bar Chart

This plot was drawn by year and quarter. Each bar was consisted of at least 1 month or at most 3 months. The numbers mean the monthly average PM2.5 concentration of Shanghai. The first quarter and fourth quarter emitted more than the second and third quarter.

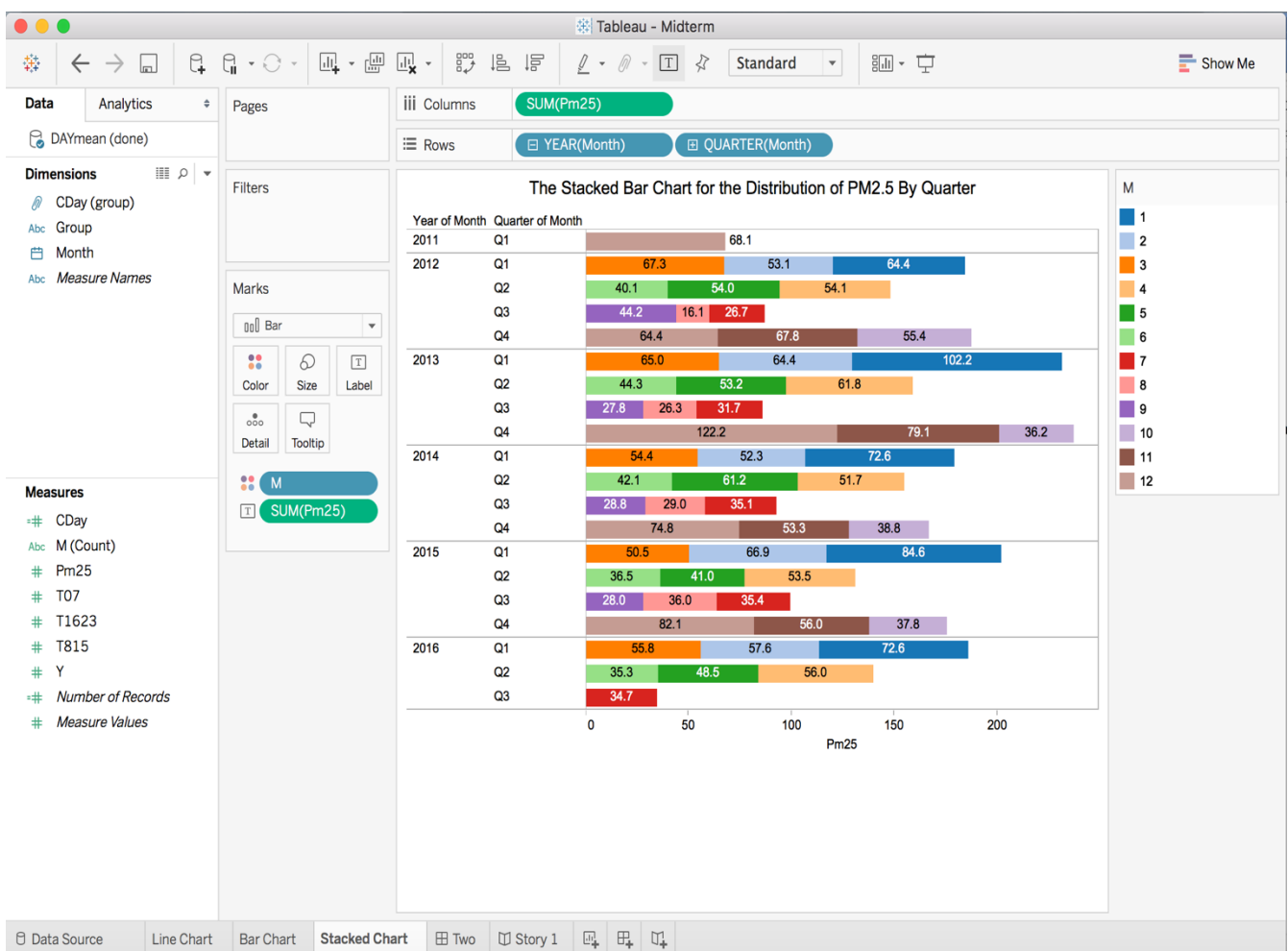
The concentration of December, 2013 is 122.2 ug/m<sup>3</sup>, which was the biggest one among the dataset.

That of January, 2013 is 102.2 ug/m<sup>3</sup>, which was the second biggest one among the dataset.

That of August, 2012 is 16.1  $\mu\text{g}/\text{m}^3$ , which was the smallest one among the dataset.

### The step to draw:

1. Put the variable Pm25 in the columns, put the variable Month in the rows, choose stacked bar chart.
2. Put variable M in the Color, put variable Pm25 in the label.



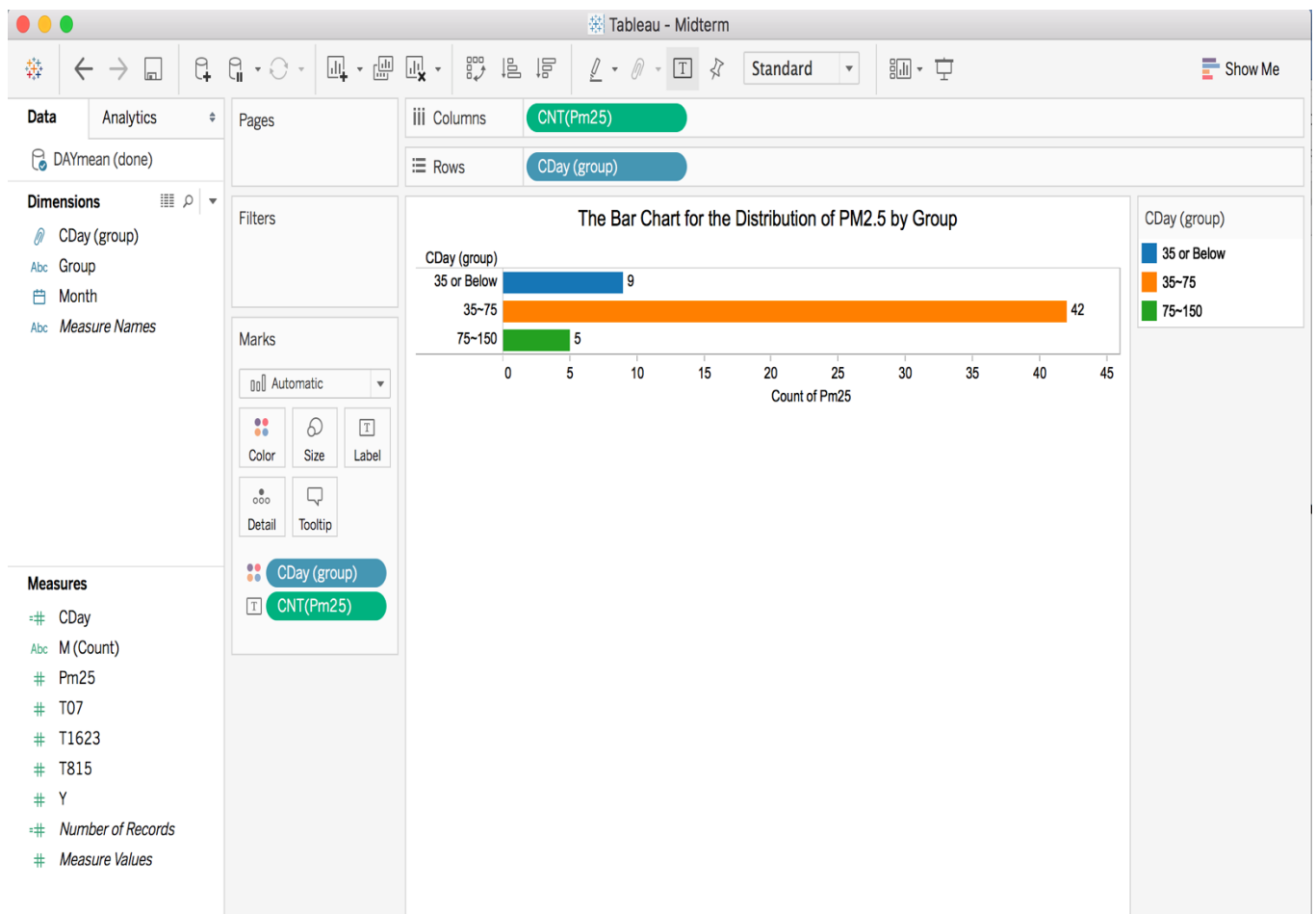
## For Bar Chart

This plot was drawn by different concentration group from year 2011 to 2016.

The day number of 35 ug/m3 or Below was 9; that of 35 ~ 75 ug/m3 was 42; that of 75 ~ 150 ug/m3 was 5. In shanghai from 2011 to 2016, the concentration of PM2.5 of most months was 35 to 75 ug/m3. It's percentage was 75.0%

## The step to draw:

1. Made a group variable CDay according to the variable Pm25, which had three groups, 35 or Below, 35 ~ 75 and 75 ~ 150.
2. Put the variable Pm25 in the Column, change it to "count", put the variable CDay in the Rows.
3. Put the variable Pm25 in the Color, put the variable Pm25 in the Label, change Pm25 to "count".



## For Line Chart

This plot was drawn by 3 time-periods, Hour 0 to 7, Hour 8 to 15 and Hour 16 to 23, from 2011 to 2016. The Blue Area was for Hour 0 to 7; the Red Area was for Hour 8 to 15; the Green Area was for Hour 16 to 23. The three area had similar attribution. So it means that the PM2.5's day average number can be strong enough to depict the whole day's concentration of PM2.5.

## The step to draw:

1. Put the variable Month in the Columns, click it to MONTH.
2. Put the variable T07, T815 and T1623 in the Row2.
3. Change the color of three part to Blue, Red and Green.

