MIS 0855 – Data Science (Section 005) – Fall 2016
In-Class Exercise (Week 6) – Advanced Data Visualization with Tableau

Objective: Learn how to use Tableau's advanced data visualization tools

Learning Outcomes:

- Use advanced visualization toolkits in Tableau for more effective presentation of data analyses

- Understand the purpose and effective use of the following tools – dual line chart, box plot, bubble chart, bullet graph, and histogram

Part 0. Understand the Dataset

In this exercise, we will use the crime statistics in large U.S. cities for 2009-2014. This dataset is obtained from the Uniform Crime Reports (UCR) published by the FBI (https://ucr.fbi.gov/crime-in-the-u.s/2014/crime-in-the-u.s.-2014). This file includes the number of crime occurrences per population (crime rates) and the number of police officers killed or assaulted in the line of duty.

1) Download "Crimes 1999-2014.xlsx" and save it to your computer. Remember where you saved it!

2) Open the file, and browse the data.

3) Take a look at Data Dictionary tab, and understand what each crime data is for.

Part 1. Dual Line Chart

1) Start Tableau.

2) Click on "Microsoft Excel" under "In a file."

3) Navigate to the location where your data file ("Crimes 1999-2014.xlsx") is stored and select it.

4) Drag the "Crimes 1999-2014" sheet to the workspace.

Workbook

Crimes 1999-2014.xlsx

Crimes in 2009-2014

Sheets

Enter sheet name

⊞ Crimes in 2009-2014
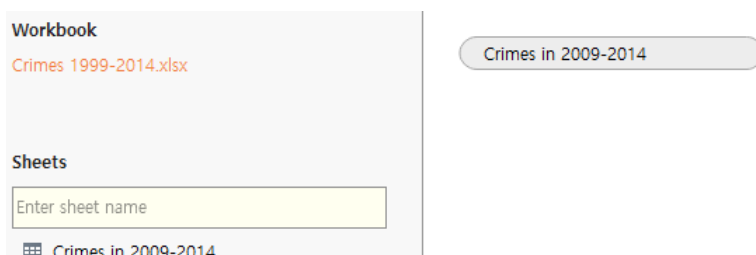
4) Click "Sheet 1" to Go to Worksheet.

5) For us to draw a line chart, there must be at least one date-type dimension. If you look at Dimensions panel, Year dimension is designated as a text field. Right-click Year and select "Change Data Type" > "Date."
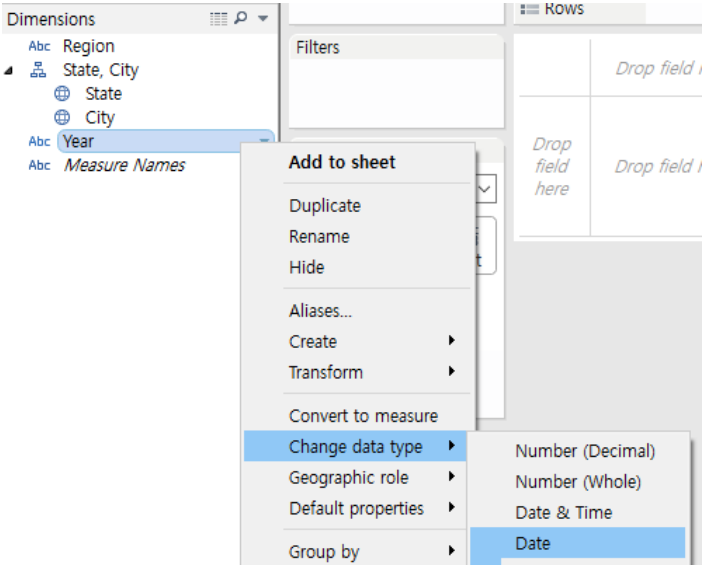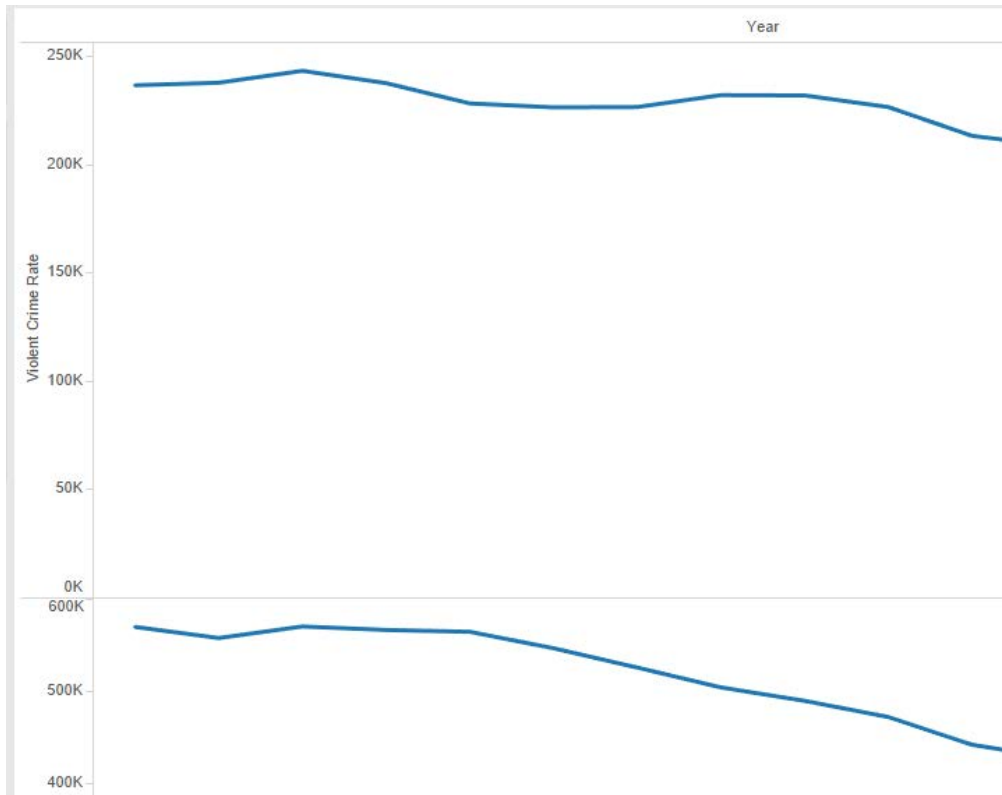

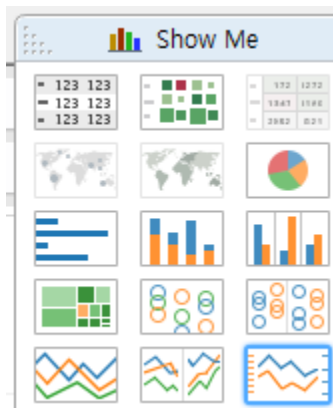You will see a date ▭ icon next to Year.

6) Drag Year from Dimensions to Columns, and drag "Violent Crime Rate" and "Property Crime Rate" from Measures to Rows.
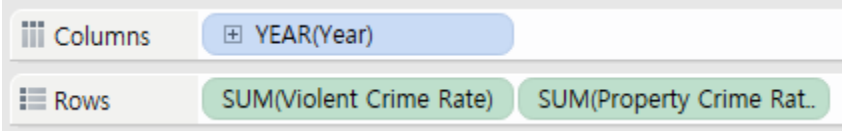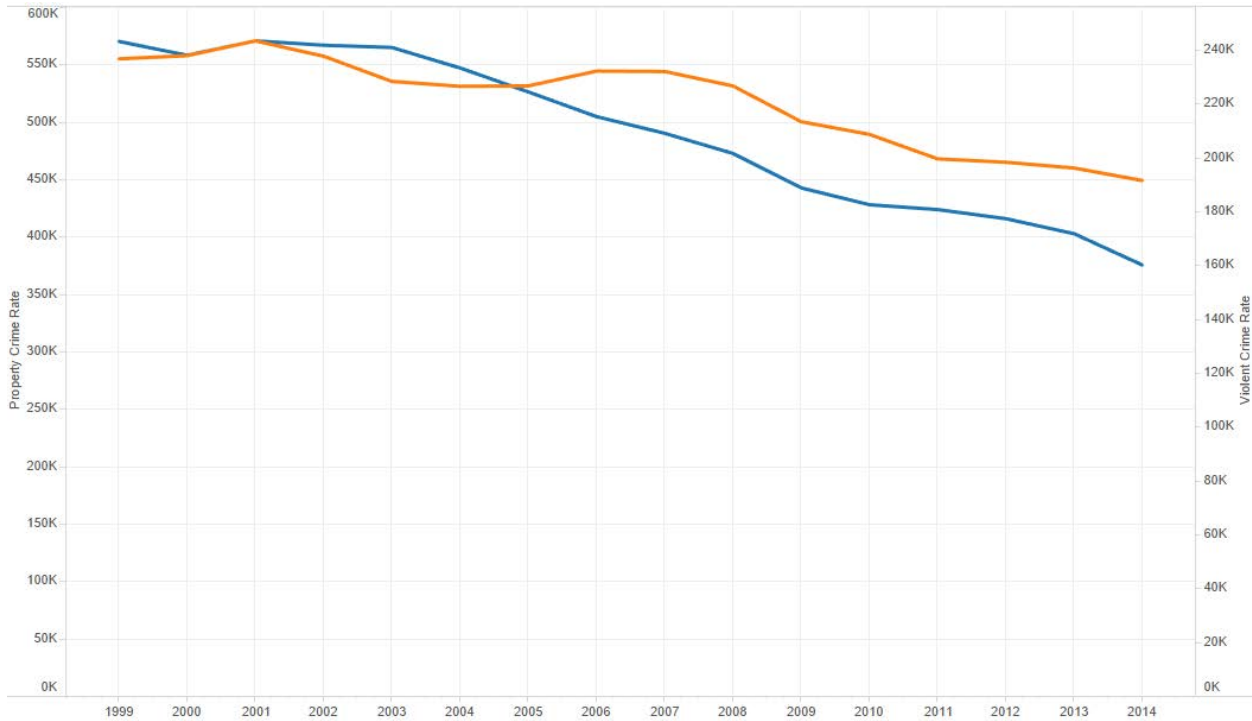

You will see two separate line charts – the top one for the violent crime rate and the bottom one for the property rate.
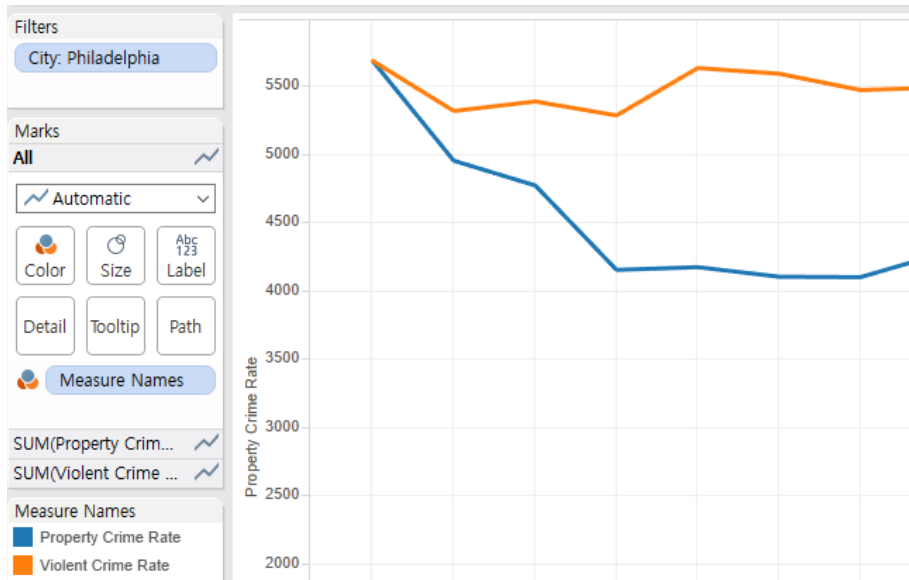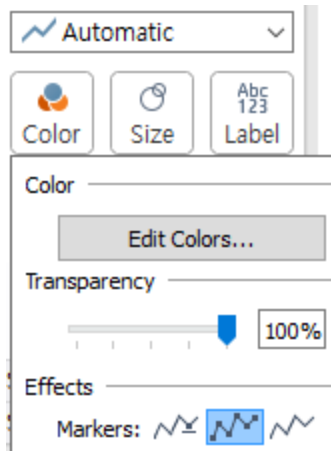
7) Click "Dual Line Chart" on Show Me tab.

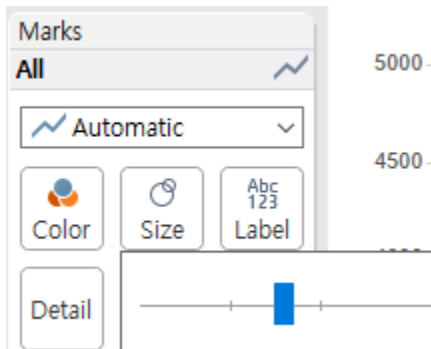You will see that the two charts are combined into one.

8) We'd like to know the crime trends in the city where we live – Philadelphia. Drag City from Dimensions to Filters and check Philadelphia.
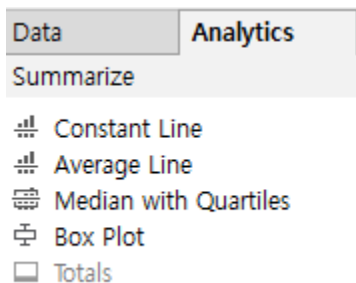

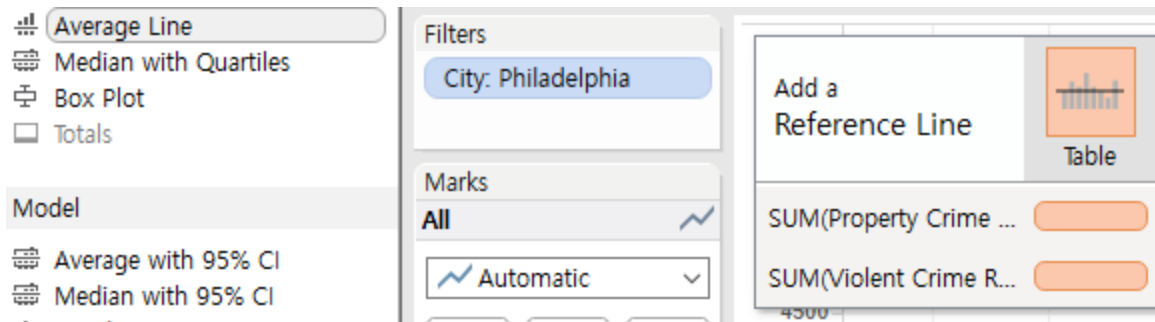9) On Marks panel, click Color and select a dotted line.

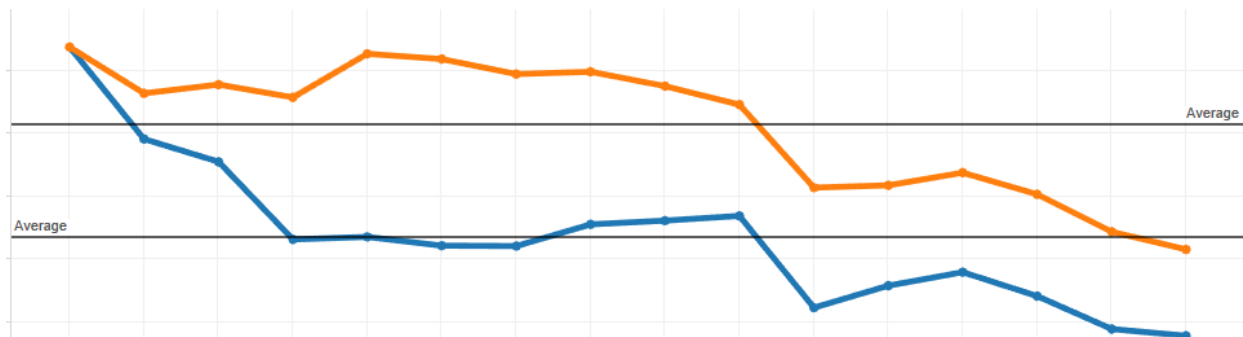10) Right next to Color, click Size, and you can make the lines thicker.

11) Click Analytics tab next to Data.

12) Drag Average Line from Analytics to "Table."

Now you'll see that the chart now displays the average crime rates in 1999-2014, which demonstrates that Philadelphia has become safer to live in recent years.
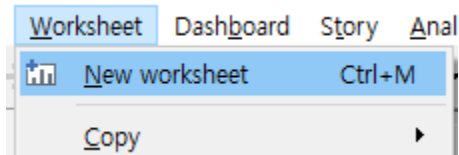

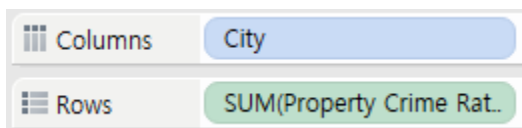13) Rename "Sheet 1" to "Dual Line Charts" and save Tableau.
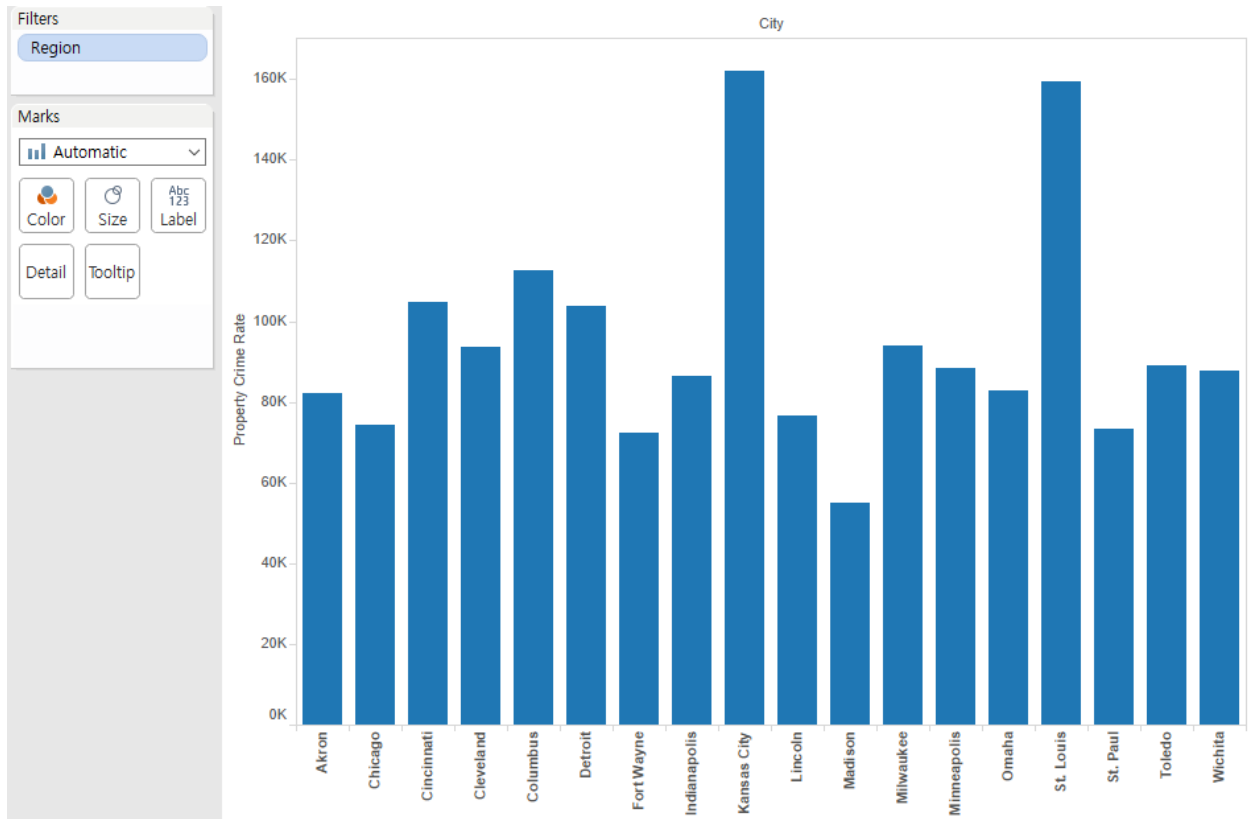

Part 2. Box Plot
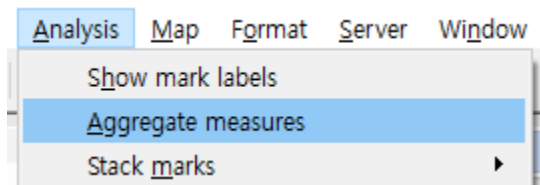
1) Create a new worksheet.

2) Drag City from Dimensions to Columns, and drag Property Crime Rate from Measures to Rows.
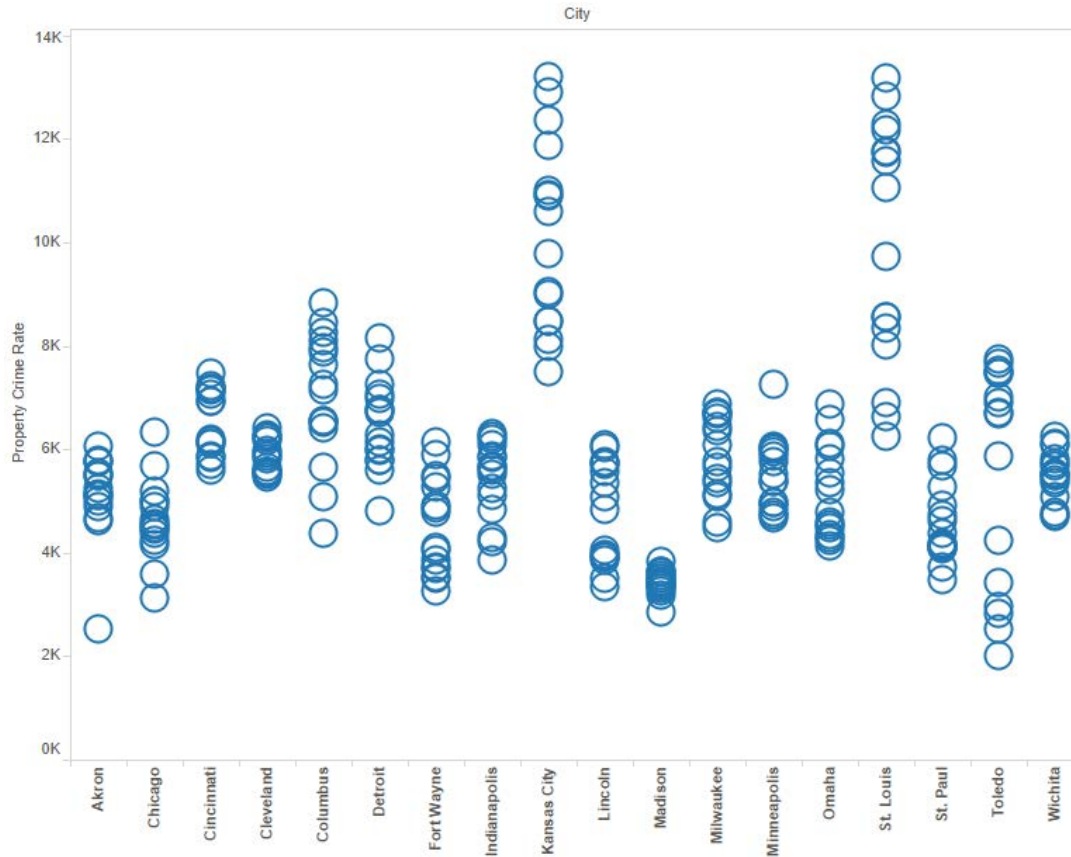
3) Let's focus on the cities in Midwest. Drag Region from Dimensions to Filter. Select East Midwest and West Midwest.

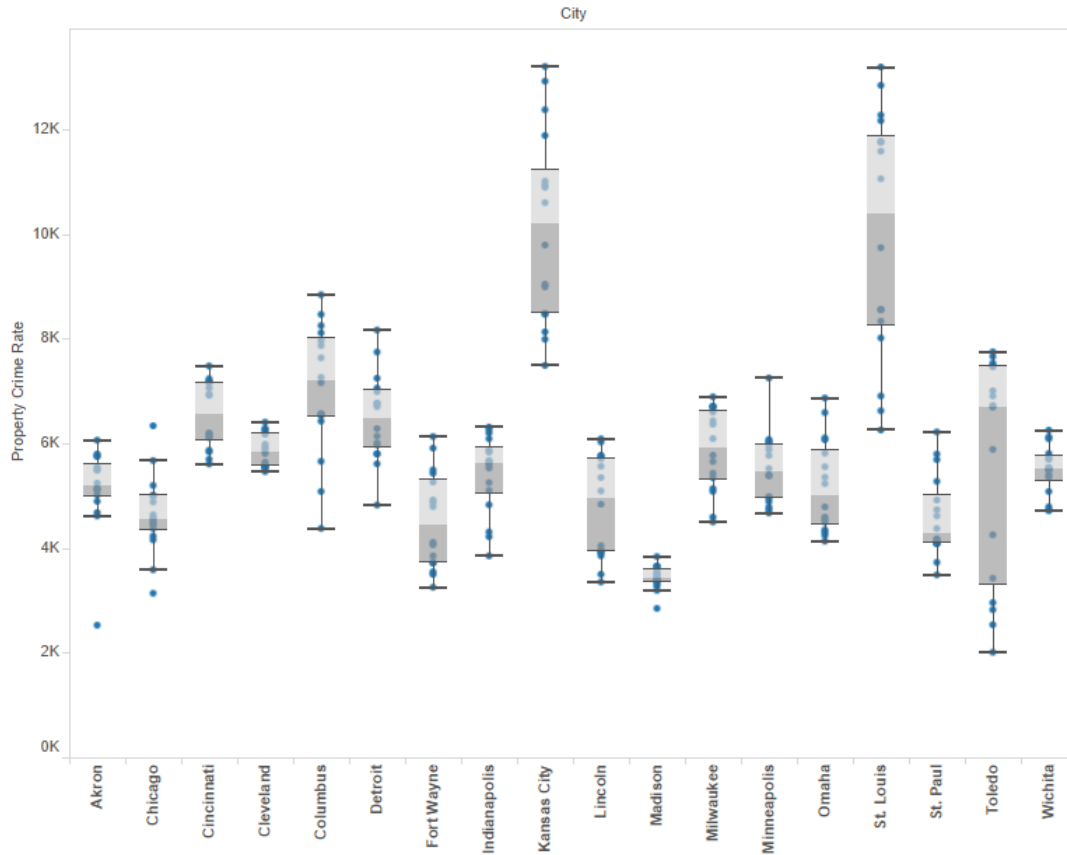4) On the top menu, click Analysis and uncheck Aggregate Measures.

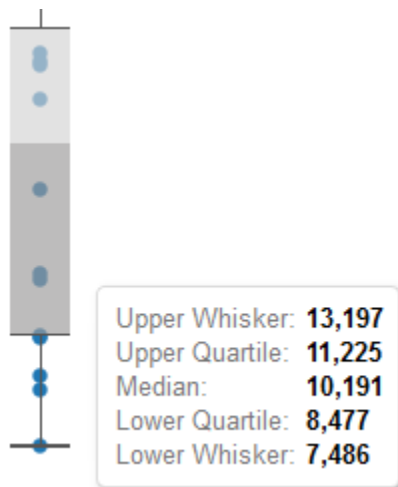You'll see the property crime rate in each year of 1999-2014 appears as a circle.

5) On Show Me tab, click "box-and-whisker plot."

For **box-and-whisker plot**
try

6) Move the mouse over a box, which shows the range of property crime rates in each city in 1999-2014 (minimum, maximum, 25%, 50% and 75% quartiles).

| Upper Whisker: | 13,197 |
Upper Quartile:	11,225
Median:	10,191
Lower Quartile:	8,477
Lower Whisker:	7,486

7) Rename "Sheet 2" to "Box Plot" and save.

Part 3. Bubble Chart

1) Create a new worksheet.

2) Drag Violent Crime Rate from Measures to Columns, and drag Property Crime Rate to Rows.

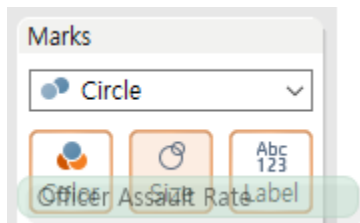|	
Columns	SUM(Violent Crime Rate)
Rows	SUM(Property Crime Rat..

3) On the top menu, click Analysis and uncheck Aggregate Measures, which shows a scatter plot of violent and property crime rates. As expected, the two types of crimes are highly correlated.
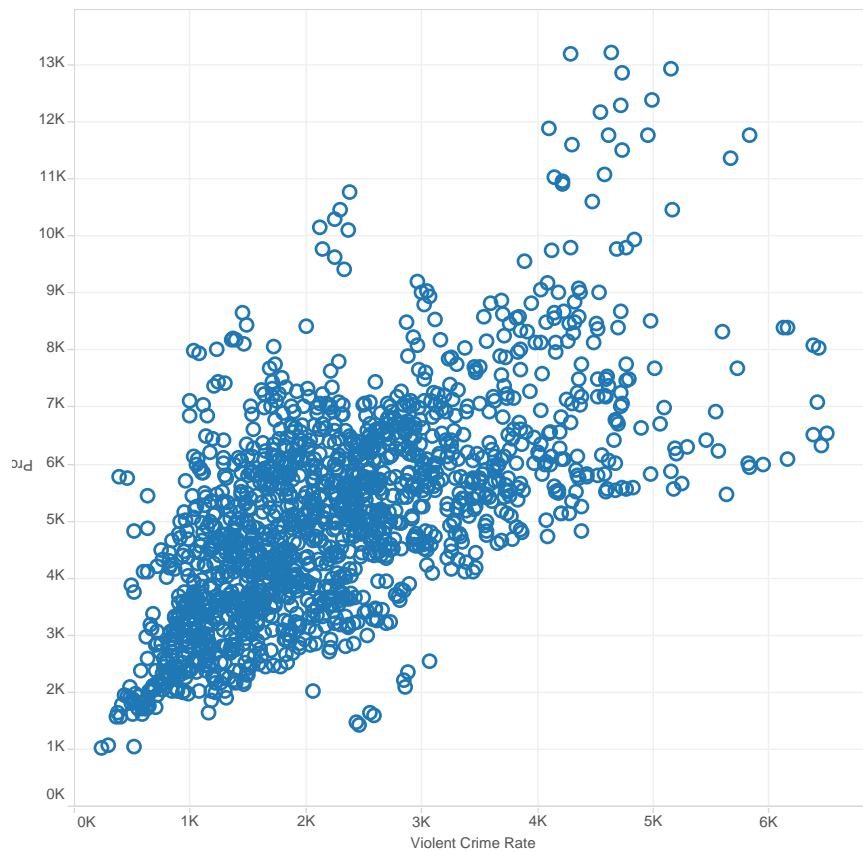
4) On Marks panel, change Automatic to Circle.
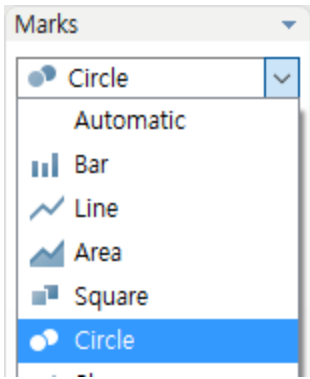
5) Drag Officer Assault Rate to Size on Marks panel.

Note that Officer Assault Rate is the number of police officers killed or assaulted in the line of duty. This bubble chart now shows that police officers are in more danger (as shown in bigger circles) in cities with higher crime rates.

6) Drag Region from Dimensions from Color on Marks panel.
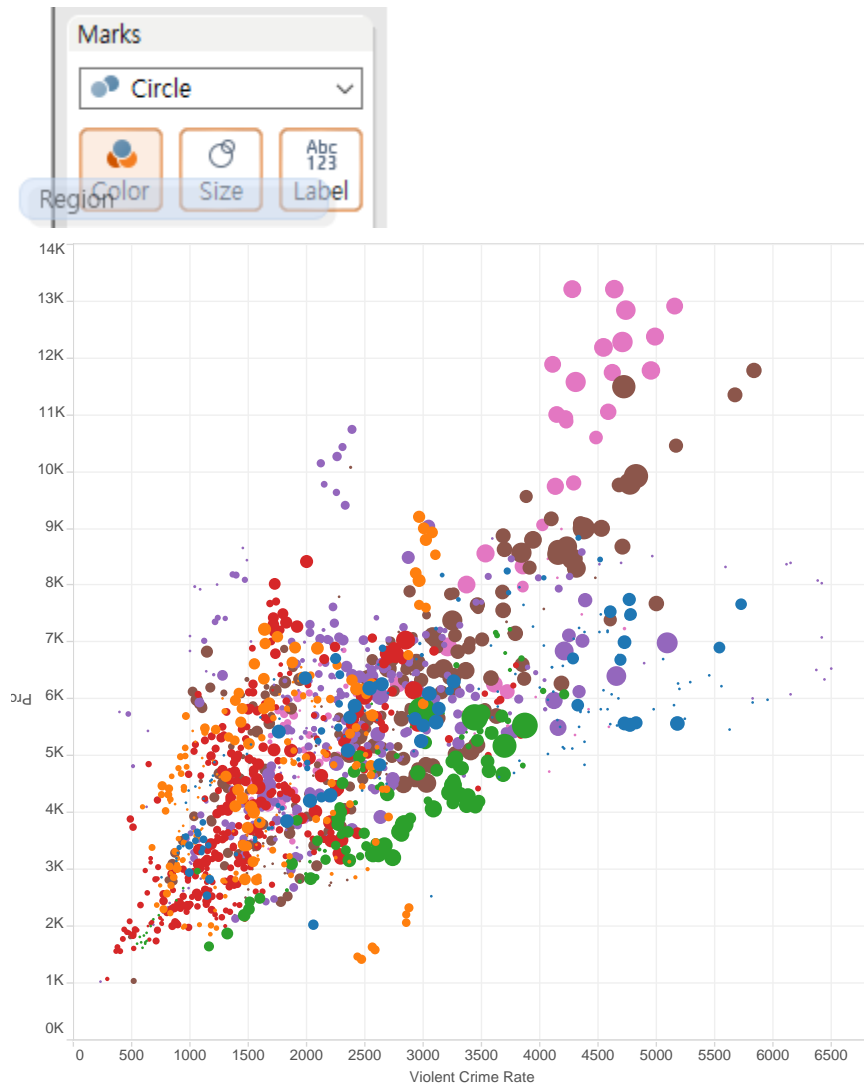
7) While the plot might look pretty, we can make this more informative and visually appealing by using fewer colors. (It is less distracting to viewers if you use as few different colors as possible.) Right-click Region panel, and select Edit colors.
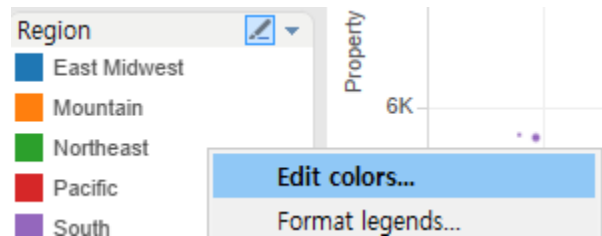
8) Let's assign colors to Regions as follows.

| Northeast	Dark blue
South Atlantic	Light blue
Pacific	Dark red
Mountain	Light red
South	Dark green
East and West Midwest	Light green
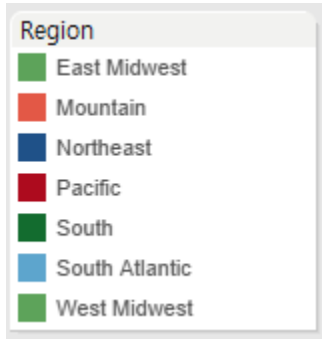
Now this one looks easier to understand (Blue is East Coast, Red is West Coast, and Green is in-between.) It is now clearer that cities in West Midwest (in light green) and South Atlantic (in light blue) are more crime-prone.
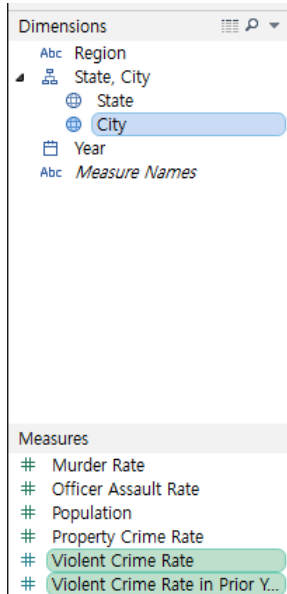
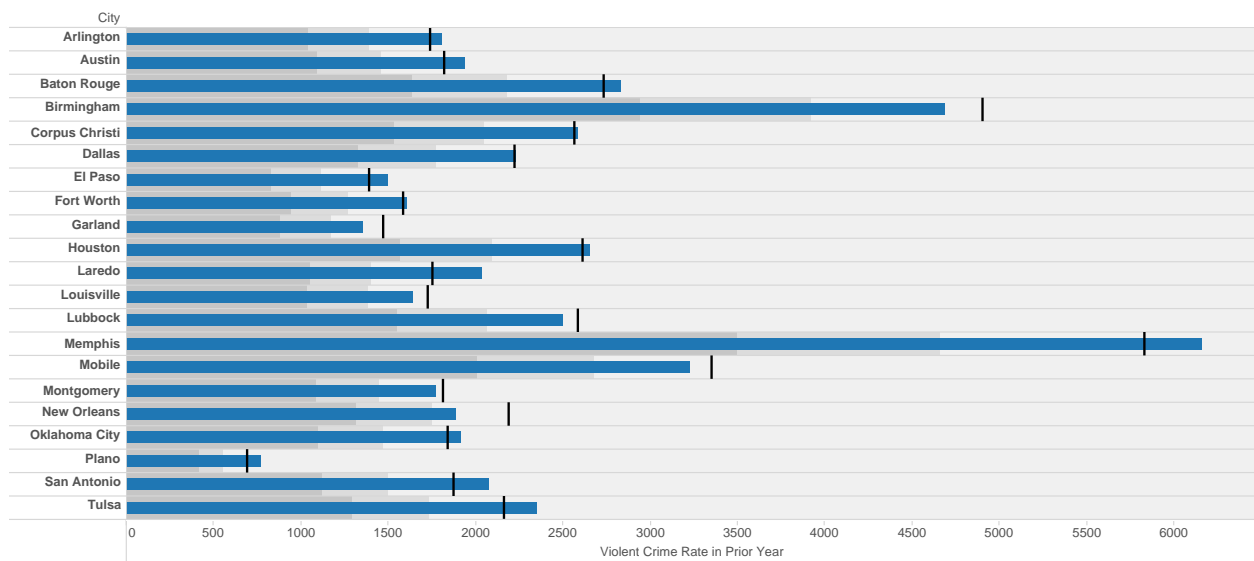7) Rename "Sheet 3" to "Bubble Chart" and save.

Part 4. Bullet Graph

1) Create a new worksheet.

2) Drag Region from Dimensions to Filters. Select South.

3) Drag Year from Dimensions to Filters. Double-click "Years" and select 2014.
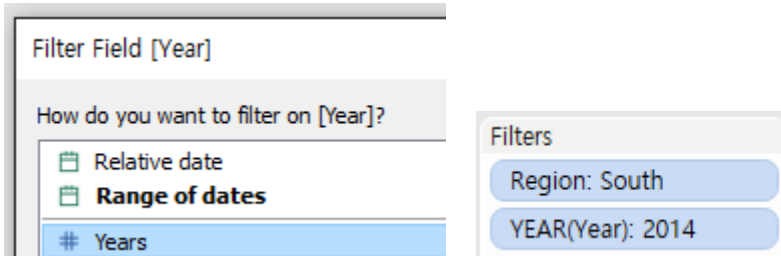
4) With pressing Ctrl key, click City, Violent Crime Rate, and Violet Crime Rate in Prior Year.

5) On Show me tab, click Bullet Graphs.

This graph shows the violent crime rates in a prior year (2013) as bars and the crime rate in a current year (2014) as horizontal line as a reference. It is not informative. We'd rather like a graph that shows the 2014 crime rate as a bar and the 2013 crime rate as a reference.

6) Right-click the horizontal axis and select "Swap reference line fields."
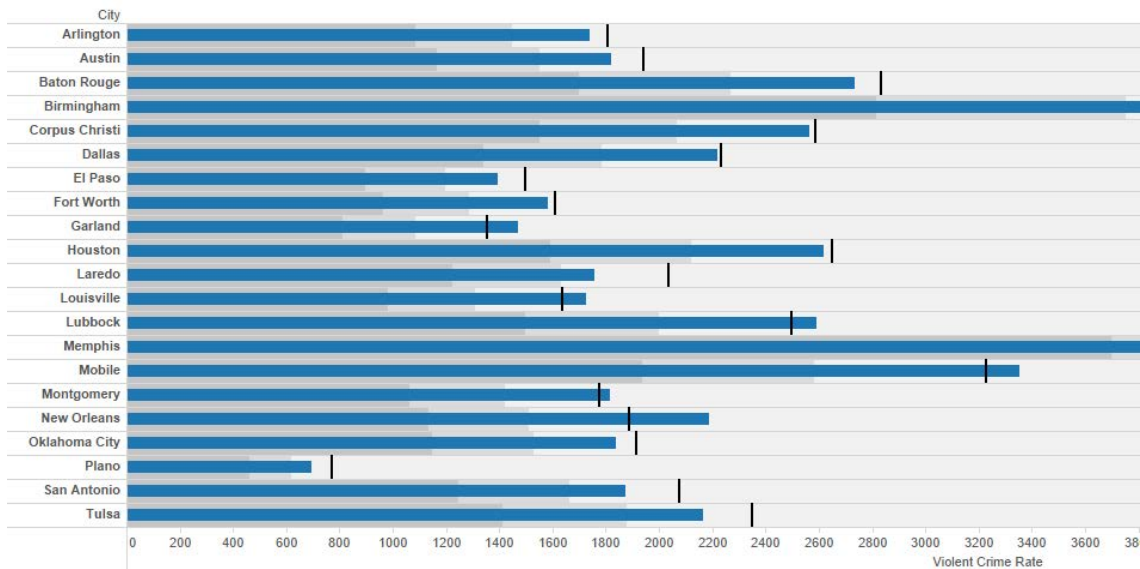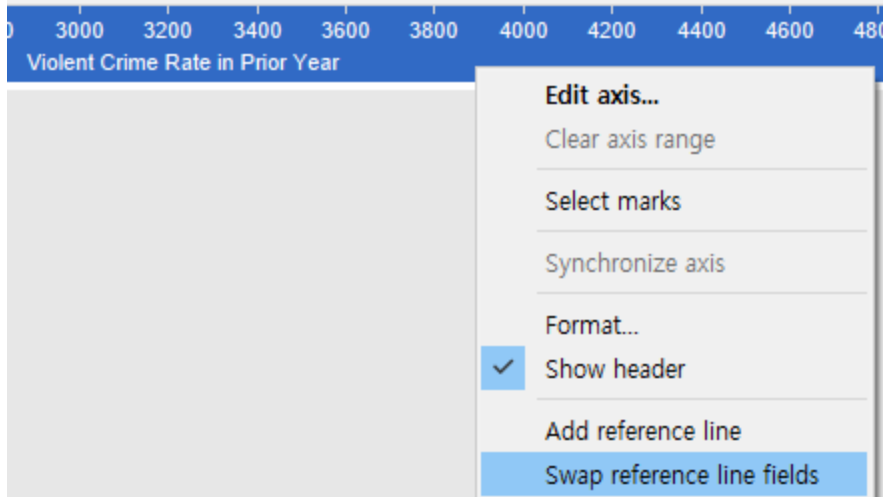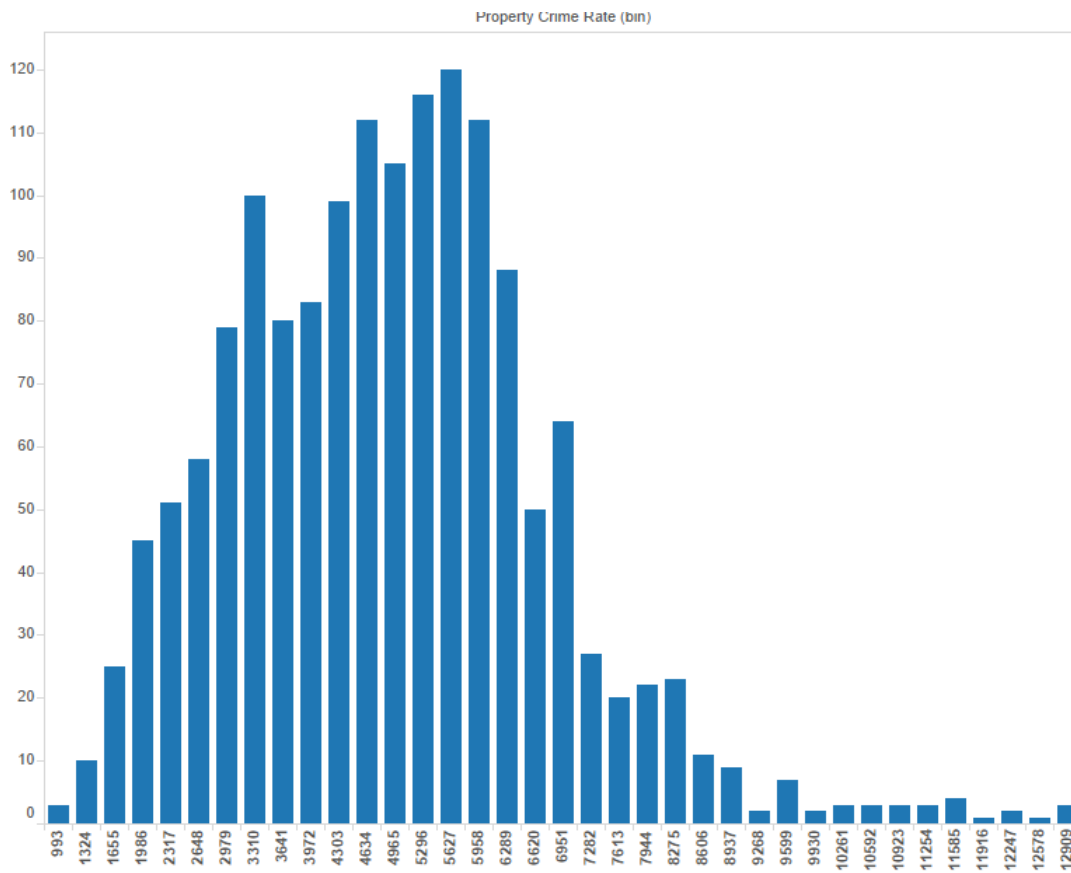
Now we can see how the crime rate has changed from 2013 (line) to 2014 (bar) in each southern city.

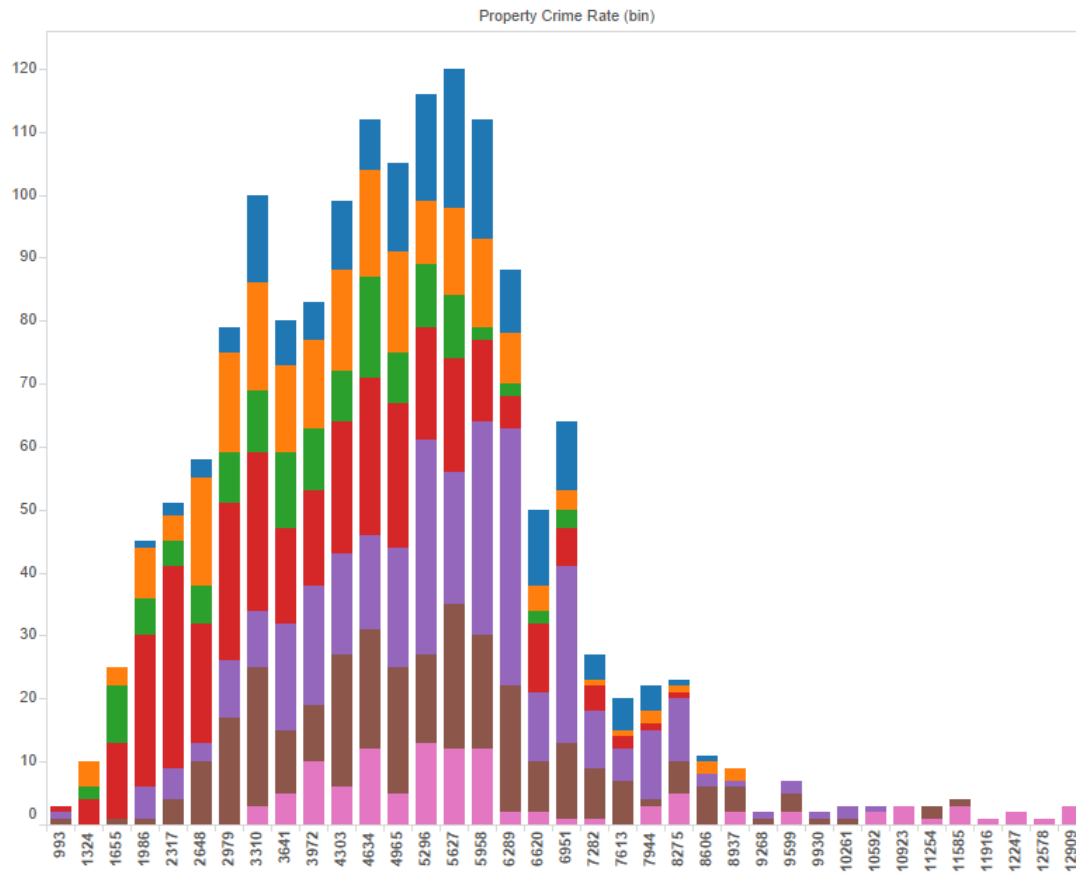7) Rename "Sheet 4" to "Bullet Graph" and save.

Part 4. Histogram

1) Create a new worksheet.

2) Click Property Crime Rate on Measures.

3) On Show Me tab, click Histogram.

4) Drag Region from Dimensions to Color.

Property Crime Rate (bin)

5) Rename "Sheet 5" to "Histogram" and save.