Static to live: Combining Stata with Google Charts API

Stata Conference
Chicago 2016

Belen Chavez
William Matsuoka
July 28, 2016
Overview

• Motivation
• Introduction
• Examples
• Future features
• Questions
• Contact
Motivation

- Presenting data through graphics that are transparent
- Reporting through interactive web-based graphics
- Combining Google Charts API with Stata (screenshots in next slide) https://developers.google.com/chart/
  - Full suite of graphs
- Standardizing syntax for ease of use
Introduction

• Introducing gcharts – beta version
  • Modeled after twoway command
  • Comprehensive library
  • No need for external software
  • No need to learn JavaScript and/or HTML
  • No need to parse arrays
  • Easy to automate and update graphics
The process
Translating this:

```javascript
var formatternum1 = new google.visualization.NumberFormat({
  pattern: "$$#,##0.##"
});
var formatternum2 = new google.visualization.NumberFormat({
  pattern: "$$#";
});
var formatternum3 = new google.visualization.NumberFormat({
  pattern: "%#"
});
formatternum1.format(data, 1);
formatternum2.format(data, 2);
formatternum3.format(data, 3);
var options = {
  alignment: 'center',
  position: 'bottom',
};
series: {
  0: {
    pointsVisible: false, pointSize: 8,
    color: 'green', lineWidth: 4,
  },
  1: {
    pointShape: 'star', pointSize: 10,
    color: 'blue', lineWidth: 3,
  }
},
title: 'SP500',
titleTextStyle: {
  color: 'blue',
},
chartArea: {
  backgroundColor: {
    strokeWeight: 6,
    fill: '#B2c8e9',
    stroke: '#c10534'
  },
},
gchart area open high close date, msymbol(none star triangle) ///
msize(medium 10 vsmall) ///
legend(alignment(center) position(bottom)) ///
lwidth(medthin thin medthick) ///
title("SP500", color(blue)) ///
nformat(open "$$,###.##" high "$$,###" close "%%") ///
mcolor(green blue red) ///
orientation(Vertical) ytitle("date", color(green) font(calibri) italic) ///
xlabel(#10, format($#,###)) yscale(reverse) ylabel(#20, format(yy-d-MMM)) ///
xline(1350, lcolor(green)) ///
selectmode(multiple) curve(none function function) ///
crosshair(both, scolor(orange) hcolor(red)) ///
plotregion(fcolor(eltblue) lwidth(medthick) lcolor(cranberry)) ///
replace graphregion(fcolor(gray) lcolor(black) lwidth(vvthick)) step
```
Google Chart Suite
Google Chart Gallery

https://developers.google.com/chart/interactive/docs/gallery
gcharts (beta)
Available **gchart** types in this release

- Scatter
- Line
- Column/Bar
- Area
- Pie/Donut
- Bubble
- Org Chart
- Sankey
- Treemap
- Gauge
- Gantt
- Calendar
Stata-like graph options
Stata-like options

• Line options:
  • lwidth(), lcolor(), lpattern()

• Scatter options:
  • msymbol(), mcolor(), msize()

• Label options:
  • xlabel(), ylabel(), xtitle(), ytitle(), xscale(), yscale(), xline() yline()

• Legend options

• Title options

• Plot and graph region options
Google chart options
Google chart-specific options

We tried making all Google chart configuration options as “Stata” like as possible. Some of them include:

- `crosshair()`
- `selectmode()`
- `trendlines()`
- `mfcolor()`
- `opac()`
Google charts examples
Example – table chart

https://developers.google.com/chart/interactive/docs/gallery/table

Let’s say we want to create an HTML table. Using `gcharts` and `sysuse auto`, we type:

```
<table>
<thead>
<tr>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>gchart table price mpg (lab) foreign (bool) foreign (val) make foreign rep78 (bool) rep78 in 1/10</td>
</tr>
</tbody>
</table>
```

Notice how it looks a lot like collapse
Example – table chart

Resulting in:

<table>
<thead>
<tr>
<th>Price</th>
<th>Mileage (mpg)</th>
<th>Car type</th>
<th>Make and Model</th>
<th>Repair Record 1978</th>
<th>Repair Record 1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>4,099</td>
<td>22</td>
<td>Domestic</td>
<td>AMC Concord</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4,749</td>
<td>17</td>
<td>Domestic</td>
<td>AMC Pacer</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>3,799</td>
<td>22</td>
<td>Domestic</td>
<td>AMC Spirit</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4,816</td>
<td>20</td>
<td>Domestic</td>
<td>Buick Century</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>7,827</td>
<td>15</td>
<td>Domestic</td>
<td>Buick Electra</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>5,788</td>
<td>18</td>
<td>Domestic</td>
<td>Buick LeSabre</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4,453</td>
<td>26</td>
<td>Domestic</td>
<td>Buick Opel</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5,189</td>
<td>20</td>
<td>Domestic</td>
<td>Buick Regal</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>10,372</td>
<td>16</td>
<td>Domestic</td>
<td>Buick Riviera</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>4,082</td>
<td>19</td>
<td>Domestic</td>
<td>Buick Skylark</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
Let’s say we want to look at the flow of customers to and from different rates. Sankey Diagrams are useful for visualizing flows of data. Our example data looks like:

To create a Sankey diagram using `gcharts` we type:

```
gchart sankey, from(rate_from) to(rate_to) weight(custs)
```
Example – Sankey diagram

Resulting in:
Example – calendar chart

https://developers.google.com/chart/interactive/docs/gallery/calendar

Suppose we want to look at monthly or weekly trends in a year. Calendar charts are useful for that. Let’s look at the daily average temperature at SeaTac airport in 2015.

Using gcharts this looks like:

<table>
<thead>
<tr>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>gchart calendar temp date</code></td>
</tr>
</tbody>
</table>

.tagName
Example – calendar chart

Resulting in the following default chart:
Example – calendar chart

Let’s add in a title and a few other options:

```
gchart calendar temp date, title("SeaTac Average Temperature") gsize(20) ///
gc(opac(0.75)) mfcOLOR(blue lawngreen) ///
yrlabel(color(blue) bold) outline(color(orange)) ///
molab(color(black))
```

This gives us:

![Calendar chart](image-url)
Stata-like charts
Example – line chart

https://developers.google.com/chart/interactive/docs/gallery/linechart

In this example we’ll be using S&P500 (sysuse sp500)

Using **gcharts** we type:

<table>
<thead>
<tr>
<th>Command</th>
</tr>
</thead>
<tbody>
<tr>
<td>gchart line low high date</td>
</tr>
</tbody>
</table>

Resulting in:
Example – line chart
Example – bar chart

Stata bar charts are known as “vertical bar” charts or “column charts” in Google Charts

https://developers.google.com/chart/interactive/docs/gallery/columnchart

The options are similar as line charts. Let’s continue working with S&P500 data.

Using gcharts, we type:

```
Command
```
gchart column change date, title("Closing Price Change")
```

Which gives us the following graph:
Example – bar chart

![Closing Price Change Chart]

-100  -50   0    50    100

17Sep2001 change: -53.77
Example – scatter chart

https://developers.google.com/chart/interactive/docs/gallery/scatterchart

Scatter is also very similar. Let’s use the auto dataset to create a scatter plot of price versus MPG

Using `gcharts` we type:

```
gchart scatter price mpg, msymb(star) mcolor(blue) msize(vhuge) ///
title("Price vs MPG", color(red)) xtitle("MPG") ytitle("Price")
```

Which gives us:
Example – scatter chart
Example – pie chart

https://developers.google.com/chart/interactive/docs/gallery/piechart

In this example we’ll make a pie chart of car prices by type

```
Command

gchart pie price, over(foreign)
```

This gives us:
For more examples

• A portfolio of examples with additional options is available at: www.belenchavez.com/stata
Future **gchart** types (available soon)

- Timeline
- Intervals
- Histogram
- Candlestick/Waterfall
- Diff Charts
Future **gchart** features (available soon)

- Multiple axis charts (secondary axis options)
  - Improved integration with combo charts
- Bar chart – bar widths
- Making charts by-able
Questions?
Contact Information

William Matsuoka
- William.matsuoka@gmail.com
- www.wmatsuoka.com/stata*
- Twitter: @WilliamMatsuoka

Belen Chavez
- belen@belenchavez.com
- www.belenchavez.com/stata*
- Twitter: @_belenchavez

* Not affiliated with Stata, we just like it. A lot.
Additional Examples
Example – combo chart

https://developers.google.com/chart/interactive/docs/gallery/combochart

We can combine different chart types into one. The possible types of series are bars, line, area, scatter, and stepped area.

Using gcharts, let’s combine line, scatter and a bar chart together.

gchart (scatter low high date, msymbol(star circle) mcolor(blue orange)) ///
  (line close date, lpattern(dash) lcolor(red)) (bar open date)
Example – combo chart

- low
- high
- close
- open

*high
May 18, 2001, 1,292.06*
Let’s add a few more options to our table chart example:

```
gchart table price mpg (lab) foreign (bool) foreign (val) make foreign rep78 (bool) rep78 in 1/10, html format(arrow(1, base(4000))) bar(2, base(20) barl) color(7, addrange(3,4,"orange", "yellow") addrange(1,3,"red", "blue")) style(header(fcolor(red)) oddrow(fcolor(green)) hoverrow(fcolor(lilac) bold italic ))
```

This gives us the following informative yet overwhelming table:
## Example – table chart

<table>
<thead>
<tr>
<th>Price</th>
<th>Mileage (mpg)</th>
<th>Car type</th>
<th>Make and Model</th>
<th>Car type</th>
<th>Repair Record 1978</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4,099</td>
<td>17</td>
<td>Domestic</td>
<td>AMC Concord</td>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>$4,749</td>
<td>17</td>
<td>Domestic</td>
<td>AMC Pacer</td>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>$3,799</td>
<td>22</td>
<td>Domestic</td>
<td>AMC Spirit</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>$4,816</td>
<td>20</td>
<td>Domestic</td>
<td>Buick Century</td>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>$7,827</td>
<td>15</td>
<td>Domestic</td>
<td>Buick Electra</td>
<td>x</td>
<td>4</td>
</tr>
<tr>
<td>$5,788</td>
<td>18</td>
<td>Domestic</td>
<td>Buick LeSabre</td>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>$4,453</td>
<td>26</td>
<td>Domestic</td>
<td>Buick Opel</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>$5,189</td>
<td>20</td>
<td>Domestic</td>
<td>Buick Regal</td>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>$10,372</td>
<td>16</td>
<td>Domestic</td>
<td>Buick Riviera</td>
<td>x</td>
<td>3</td>
</tr>
<tr>
<td>$4,082</td>
<td>19</td>
<td>Domestic</td>
<td>Buick Skylark</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>
Example – sankey diagram

Adding in a few more features to our sankey example:

\[
gchart sankey, from(rate_from) to(rate_to) weight(custs) node(nodepadding(40))\n\hline
labelpadding(12) width(10) interactivity
\hline
label(font("Times-Roman") size(12) italic bold)) \hline
\hline
link(colormode(gradient) color(fcolor("#efd")) \hline
\hline
colors(blue red purple brown)) \hline
\hline
tooltip(text(bold italic font("Times-Roman") size(10) color(blue)))
\hline
\hline
\]

This results in the following visualization:
Example – sankey diagram
Example – pie chart

Adding a few more options:

gchart pie price, width(100%) height(200 px) name(my_graph) ///
tooltip(price, ignore ishtml text(percent) textstyle(color(red) ///
font(Calibri)) trigger(selection)) ///
legend(textstyle(color(red)) position(labeled) alignment(end)) ///
over(foreign) piehole(.1)

We get:
Example – line chart

Let’s add a title and customize the line graph a bit:

```
sysuse sp500, clear
gchart line low high date, title("S&P500", color(blue) bold) ///
     lcolor(green purple) lwidth(thick medthick) lpattern(dash_dot solid)
```

This gives us the following output:
Example – line chart

![S&P500 chart showing data from April 2001 to October 2001 with a spike on 16th July 2001 reaching a high of 1,219.63.]