Data Visualization – ggplot2

Grace He

October 29, 2017

Introduction

Data analysis has gradually became more and more important and has played an increasingly important role in various fields such as manufacturing, entertainment, and technology, just to name a few. And after all the data preparation and analysis, we want to present the result in the most effective way. In a lot of cases, our audience do not have the professional background knowledge to understand the data very well; to help others to quickly grasp the information and build a good big picture of the data, we need visual aids such as plots and graphs to enforce the communication. This post focuses on an extremely useful tool – ggplot2 – a plotting system for R that you need to create good graphics. Compared to base graphics, ggplot2 uses a different system for adding plot elemnts and is simpler for complex graphics. The major contents of the posts include some plot basics with ggplot2, scales, themes, and interesting visualizations other than the basic plots.

Dataset Background

The dataset we will be using in this post consists of HDI (Human Development Index), CPI (Corruption Perception Index), and HDI rankings of countries from different regions around the world in 2011. Higher HDI means the country is more developed, and the CPI is on a scale from 0 (highly corrupt) to 10 (very clean). This dataset is orginally used by *The Economist* magazine. The region names in this dataset are abbreviated, and for your convenience, here is a list of the abbreviations and the regions they represent:

- Americas: Americas region
- Asia Pacific: Asia & Ocenia
- SSA: Sub-Sharan Africa
- MENA: Middle East & North Africa
- East EU Cemt Asia: Central & Eastern Europe
- EU W. Europe: Western Europe

Plot Basics

To create a new plot with ggplot2, we need to call ggplot(), and supply it with data(dataframe) and aesthetic mappings with the aes() function. Then we can add different layers, coordinates, facets, and scales with a "+" sign. "Aesthetic" in ggplots means something we can see; some exemples include position, color, fill, shape of points, size, and linetype. To add a title to the plot to make it more appropriate and complete, we can use the function ggtitle(). And before using ggplot2, we need to first load the package.

```
library(ggplot2)
```

Geometric Objects

Geometic objects determine what kinds of plot we want to create – they are the actual marks that appear on our plot. A plot must have at least one geom, but we can have more to create more layers to the plot. Some common examples include:

- geom_point: Points, for scatter plots, dot plots, etc
- geom_abline, geom_hline, geom_vline: Reference lines (horizontal, vertical, and diagonal)
- geom_bar: Bar charts
- geom_boxplot: boxplots
- geom_label, geom_text: texts that appear on the plot

Before creating any plot, let's start with the dataset, and I will name the dataframe "development".

```
development <- read.csv("data/economistData.csv")
development
```

```
##       X                         Country HDI.Rank   HDI CPI
## 1     1                     Afghanistan      172 0.398 1.5
## 2     2                         Albania       70 0.739 3.1
## 3     3                         Algeria       96 0.698 2.9
## 4     4                          Angola      148 0.486 2.0
## 5     5                       Argentina       45 0.797 3.0
## 6     6                         Armenia       86 0.716 2.6
## 7     7                       Australia        2 0.929 8.8
## 8     8                         Austria       19 0.885 7.8
## 9     9                      Azerbaijan       91 0.700 2.4
## 10   10                         Bahamas       53 0.771 7.3
## 11   11                         Bahrain       42 0.806 5.1
## 12   12                      Bangladesh      146 0.500 2.7
## 13   13                        Barbados       47 0.793 7.8
## 14   14                         Belarus       65 0.756 2.4
## 15   15                         Belgium       18 0.886 7.5
## 16   16                           Benin      167 0.427 3.0
## 17   17                          Bhutan      141 0.522 5.7
## 18   18                         Bolivia      108 0.663 2.8
## 19   19           Bosnia and Herzegovina       74 0.733 3.2
## 20   20                        Botswana      118 0.633 6.1
## 21   21                          Brazil       84 0.718 3.8
## 22   22                         Britain       28 0.863 7.8
## 23   23                        Bulgaria       55 0.771 3.3
## 24   24                    Burkina Faso      181 0.331 3.0
## 25   25                         Burundi      185 0.316 1.9
```

```
## 26   26                       Cambodia      139 0.523 2.1
## 27   27                       Cameroon      150 0.482 2.5
## 28   28                         Canada        6 0.908 8.7
## 29   29                     Cape Verde      133 0.568 5.5
## 30   30        Central African Republic      179 0.343 2.2
## 31   31                           Chad      183 0.328 2.0
## 32   32                          Chile       44 0.805 7.2
## 33   33                          China      101 0.687 3.6
## 34   34                       Colombia       87 0.710 3.4
## 35   35                        Comoros      163 0.433 2.4
## 36   36                          Congo      187 0.286 2.0
## 37   37                 Congo  Republic      137 0.533 2.2
## 38   38                     Costa Rica       69 0.744 4.8
## 39   39                  Côte d'Ivoire      170 0.400 2.2
## 40   40                        Croatia       46 0.796 4.0
## 41   41                           Cuba       51 0.776 4.2
## 42   42                         Cyprus       31 0.840 6.3
## 43   43                Czech Republic       27 0.865 4.4
## 44   44                        Denmark       16 0.895 9.4
## 45   45                       Djibouti      165 0.430 3.0
## 46   46                       Dominica       81 0.724 5.2
## 47   47              Dominican Republic       98 0.689 2.6
## 48   48                        Ecuador       83 0.720 2.7
## 49   49                          Egypt      113 0.644 2.9
## 50   50                    El Salvador      105 0.674 3.4
## 51   51                 Equatorial Guinea      136 0.537 1.9
## 52   52                        Eritrea      177 0.349 2.5
## 53   53                        Estonia       34 0.835 6.4
## 54   54                       Ethiopia      174 0.363 2.7
## 55   55                        Finland       22 0.882 9.4
## 56   56                         France       20 0.884 7.0
## 57   57                          Gabon      106 0.674 3.0
## 58   58                         Gambia      168 0.420 3.5
## 59   59                        Georgia       75 0.733 4.1
## 60   60                        Germany        9 0.905 8.0
## 61   61                          Ghana      135 0.541 3.9
## 62   62                         Greece       29 0.861 3.4
## 63   63                      Guatemala      131 0.574 2.7
## 64   64                         Guinea      178 0.344 2.1
## 65   65                  Guinea-Bissau      176 0.353 2.2
## 66   66                         Guyana      117 0.633 2.5
## 67   67                          Haiti      158 0.454 1.8
## 68   68                       Honduras      121 0.625 2.6
## 69   69                      Hong Kong       13 0.898 8.4
## 70   70                        Hungary       38 0.816 4.6
## 71   71                        Iceland       14 0.898 8.3
## 72   72                          India      134 0.547 3.1
## 73   73                      Indonesia      124 0.617 3.0
## 74   74                           Iran       88 0.707 2.7
## 75   75                           Iraq      132 0.573 1.8
## 76   76                        Ireland        7 0.908 7.5
## 77   77                         Israel       17 0.888 5.8
## 78   78                          Italy       24 0.874 3.9
## 79   79                        Jamaica       79 0.727 3.3
## 80   80                          Japan       12 0.901 8.0
## 81   81                         Jordan       95 0.698 4.5
## 82   82                     Kazakhstan       68 0.745 2.7
## 83   83                          Kenya      143 0.509 2.2
## 84   84                       Kiribati      122 0.624 3.1
## 85   85                  Korea (South)       15 0.897 5.4
## 86   86                         Kuwait       63 0.760 4.6
## 87   87                      Kyrgyzstan      126 0.615 2.1
## 88   88                           Laos      138 0.524 2.2
## 89   89                         Latvia       43 0.805 4.2
## 90   90                        Lebanon       71 0.739 2.5
## 91   91                        Lesotho      160 0.450 3.5
## 92   92                        Liberia      182 0.329 3.2
## 93   93                          Libya       64 0.760 2.0
## 94   94                      Lithuania       40 0.810 4.8
## 95   95                     Luxembourg       25 0.867 8.5
## 96   96                     Madagascar      151 0.480 3.0
## 97   97                         Malawi      171 0.400 3.0
## 98   98                       Malaysia       61 0.761 4.3
## 99   99                       Maldives      109 0.661 2.5
## 100 100                           Mali      175 0.359 2.8
## 101 101                          Malta       36 0.832 5.6
## 102 102                     Mauritania      159 0.453 2.4
## 103 103                      Mauritius       77 0.728 5.1
## 104 104                         Mexico       57 0.770 3.0
## 105 105                        Moldova      111 0.649 2.9
## 106 106                       Mongolia      110 0.653 2.7
## 107 107                     Montenegro       54 0.771 4.0
## 108 108                        Morocco      130 0.582 3.4
## 109 109                     Mozambique      184 0.322 2.7
## 110 110                        Myanmar      149 0.483 1.5
```

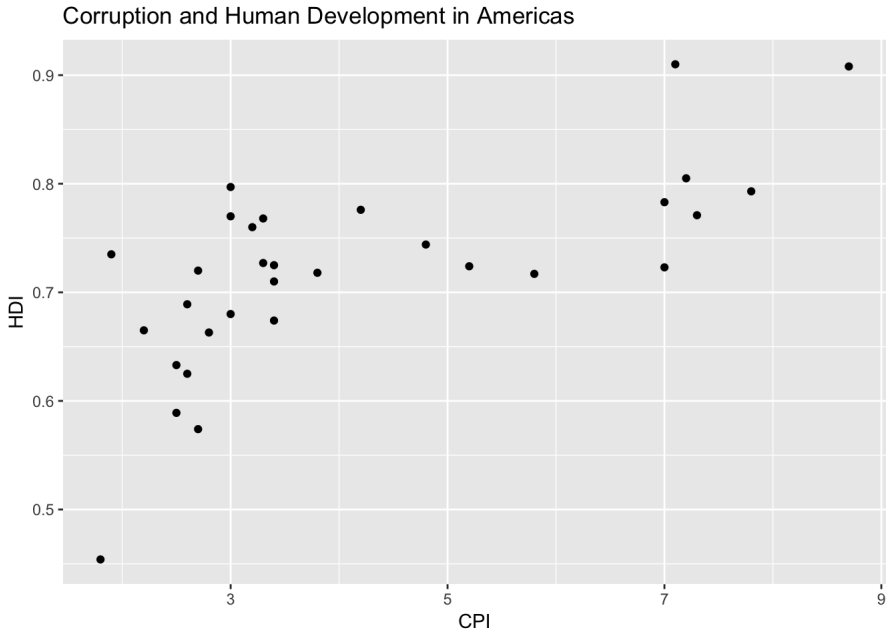
```
## 111 111                          Namibia      120 0.625 4.4
## 112 112                            Nepal      157 0.458 2.2
## 113 113                      Netherlands        3 0.910 8.9
## 114 114                      New Zealand        5 0.908 9.5
## 115 115                        Nicaragua      129 0.589 2.5
## 116 116                            Niger      186 0.295 2.5
## 117 117                          Nigeria      156 0.459 2.4
## 118 118                           Norway        1 0.943 9.0
## 119 119                             Oman       89 0.705 4.8
## 120 120                         Pakistan      145 0.504 2.5
## 121 121                           Panama       58 0.768 3.3
## 122 122                 Papua New Guinea      153 0.466 2.2
## 123 123                         Paraguay      107 0.665 2.2
## 124 124                             Peru       80 0.725 3.4
## 125 125                      Philippines      112 0.644 2.6
## 126 126                           Poland       39 0.813 5.5
## 127 127                         Portugal       41 0.809 6.1
## 128 128                            Qatar       37 0.831 7.2
## 129 129                          Romania       50 0.781 3.6
## 130 130                           Russia       66 0.755 2.4
## 131 131                           Rwanda      166 0.429 5.0
## 132 132                      Saint Lucia       82 0.723 7.0
## 133 133 Saint Vincent and the Grenadines       85 0.717 5.8
## 134 134                            Samoa       99 0.688 3.9
## 135 135                     Saudi Arabia       56 0.770 4.4
## 136 136                          Senegal      155 0.459 2.9
## 137 137                           Serbia       59 0.766 3.3
## 138 138                       Seychelles       52 0.773 4.8
## 139 139                     Sierra Leone      180 0.336 2.5
## 140 140                        Singapore       26 0.866 9.2
## 141 141                         Slovakia       35 0.834 4.0
## 142 142                         Slovenia       21 0.884 5.9
## 143 143                  Solomon Islands      142 0.510 2.7
## 144 144                     South Africa      123 0.619 4.1
## 145 145                            Spain       23 0.878 6.2
## 146 146                        Sri Lanka       97 0.691 3.3
## 147 147                            Sudan      169 0.408 1.6
## 148 148                         Suriname      104 0.680 3.0
## 149 149                        Swaziland      140 0.522 3.1
## 150 150                           Sweden       10 0.904 9.3
## 151 151                      Switzerland       11 0.903 8.8
## 152 152                            Syria      119 0.632 2.6
## 153 153                       Tajikistan      127 0.607 2.3
## 154 154                         Tanzania      152 0.466 3.0
## 155 155                         Thailand      103 0.682 3.4
## 156 156                      Timor-Leste      147 0.495 2.4
## 157 157                             Togo      162 0.435 2.4
## 158 158                            Tonga       90 0.704 3.1
## 159 159             Trinidad and Tobago       62 0.760 3.2
## 160 160                          Tunisia       94 0.698 3.8
## 161 161                           Turkey       92 0.699 4.2
## 162 162                     Turkmenistan      102 0.686 1.6
## 163 163                           Uganda      161 0.446 2.4
## 164 164                          Ukraine       76 0.729 2.3
## 165 165            United Arab Emirates       30 0.846 6.8
## 166 166                   United States        4 0.910 7.1
## 167 167                          Uruguay       48 0.783 7.0
## 168 168                       Uzbekistan      115 0.641 1.6
## 169 169                         Vanuatu      125 0.617 3.5
## 170 170                       Venezuela       73 0.735 1.9
## 171 171                           Yemen      154 0.462 2.1
## 172 172                          Zambia      164 0.430 3.2
## 173 173                        Zimbabwe      173 0.376 2.2
##               Region
## 1        Asia Pacific
## 2   East EU Cemt Asia
## 3               MENA
## 4                SSA
## 5           Americas
## 6   East EU Cemt Asia
## 7        Asia Pacific
## 8        EU W. Europe
## 9   East EU Cemt Asia
## 10          Americas
## 11              MENA
## 12       Asia Pacific
## 13          Americas
## 14  East EU Cemt Asia
## 15       EU W. Europe
## 16               SSA
## 17       Asia Pacific
## 18          Americas
## 19  East EU Cemt Asia
## 20               SSA
## 21          Americas
## 22       EU W. Europe
```

```
## 22       EU W. Europe
## 23       EU W. Europe
## 24               SSA
## 25               SSA
## 26       Asia Pacific
## 27               SSA
## 28           Americas
## 29               SSA
## 30               SSA
## 31               SSA
## 32           Americas
## 33       Asia Pacific
## 34           Americas
## 35               SSA
## 36               SSA
## 37               SSA
## 38           Americas
## 39               SSA
## 40  East EU Cemt Asia
## 41           Americas
## 42       EU W. Europe
## 43       EU W. Europe
## 44       EU W. Europe
## 45               SSA
## 46           Americas
## 47           Americas
## 48           Americas
## 49               MENA
## 50           Americas
## 51               SSA
## 52               SSA
## 53       EU W. Europe
## 54               SSA
## 55       EU W. Europe
## 56       EU W. Europe
## 57               SSA
## 58               SSA
## 59  East EU Cemt Asia
## 60       EU W. Europe
## 61               SSA
## 62       EU W. Europe
## 63           Americas
## 64               SSA
## 65               SSA
## 66           Americas
## 67           Americas
## 68           Americas
## 69       Asia Pacific
## 70       EU W. Europe
## 71       EU W. Europe
## 72       Asia Pacific
## 73       Asia Pacific
## 74               MENA
## 75               MENA
## 76       EU W. Europe
## 77               MENA
## 78       EU W. Europe
## 79           Americas
## 80       Asia Pacific
## 81               MENA
## 82  East EU Cemt Asia
## 83               SSA
## 84       Asia Pacific
## 85       Asia Pacific
## 86               MENA
## 87  East EU Cemt Asia
## 88       Asia Pacific
## 89       EU W. Europe
## 90               MENA
## 91               SSA
## 92               SSA
## 93               MENA
## 94       EU W. Europe
## 95       EU W. Europe
## 96               SSA
## 97               SSA
## 98       Asia Pacific
## 99       Asia Pacific
## 100              SSA
## 101      EU W. Europe
## 102              SSA
## 103              SSA
## 104          Americas
## 105 East EU Cemt Asia
## 106      Asia Pacific
## 107 East EU Cemt Asia
```

```
## 108              MENA
## 109               SSA
## 110      Asia Pacific
## 111               SSA
## 112      Asia Pacific
## 113      EU W. Europe
## 114      Asia Pacific
## 115          Americas
## 116               SSA
## 117               SSA
## 118      EU W. Europe
## 119              MENA
## 120      Asia Pacific
## 121          Americas
## 122      Asia Pacific
## 123          Americas
## 124          Americas
## 125      Asia Pacific
## 126      EU W. Europe
## 127      EU W. Europe
## 128              MENA
## 129      EU W. Europe
## 130 East EU Cemt Asia
## 131               SSA
## 132          Americas
## 133          Americas
## 134      Asia Pacific
## 135              MENA
## 136               SSA
## 137 East EU Cemt Asia
## 138               SSA
## 139               SSA
## 140      Asia Pacific
## 141      EU W. Europe
## 142      EU W. Europe
## 143      Asia Pacific
## 144               SSA
## 145      EU W. Europe
## 146      Asia Pacific
## 147               SSA
## 148          Americas
## 149               SSA
## 150      EU W. Europe
## 151      EU W. Europe
## 152              MENA
## 153 East EU Cemt Asia
## 154               SSA
## 155      Asia Pacific
## 156      Asia Pacific
## 157               SSA
## 158      Asia Pacific
## 159          Americas
## 160              MENA
## 161 East EU Cemt Asia
## 162 East EU Cemt Asia
## 163               SSA
## 164 East EU Cemt Asia
## 165              MENA
## 166          Americas
## 167          Americas
## 168 East EU Cemt Asia
## 169      Asia Pacific
## 170          Americas
## 171              MENA
## 172               SSA
## 173               SSA
```

In the following example, we will cretae a scatterplot of HDI(Human Development Index, 2011) and CDI(Corruption Perceptions Index, 2011) of countries in the region Americas to see the relationship between corruption and human development.

```
americas <- development[development$Region == "Americas",] # We just want data of countries in the region Americas
ggplot(americas, aes(x = CPI, y = HDI)) + geom_point() + ggtitle("Corruption and Human Development in Americas")
```

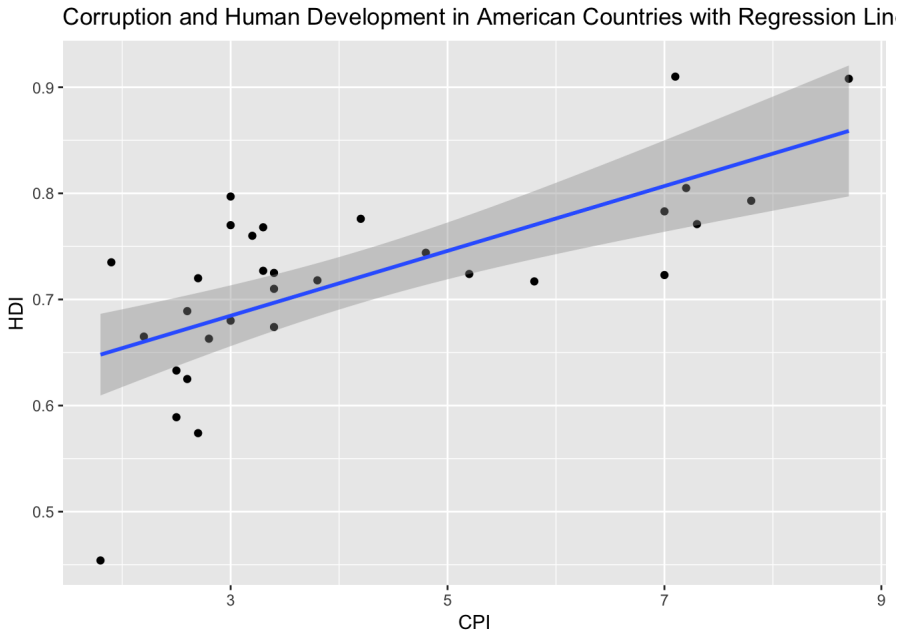


Adding Lines

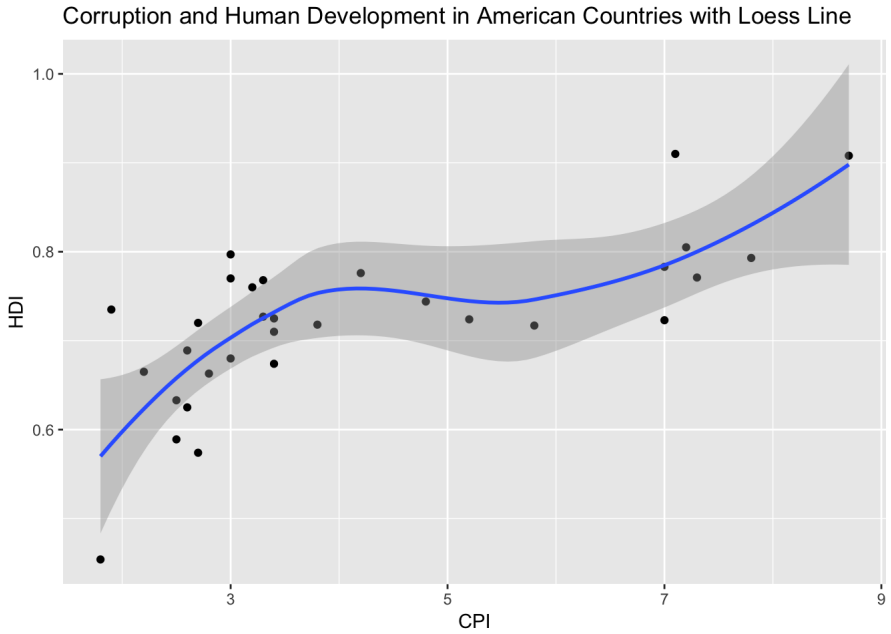
To make our plots more interesting and also to help ourselves to see the patterns of data more easily, we can add lines to our plot with the function geom_smooth. The argument “method” specifies which function or smoothing method that we want to use for the data. For example, if we want to draw a regression line, then we would use “method = lm”, and similarly, we would use “method = loess” to add a loess line. The argument “se” (logical) controls whether to display the confidence interval around the line. If you don’t want the shaded confidence region, then you would specify: “se = FALSE”.

To demonstrate, we will first add a regressgion and then a loess line to the plot that we just created above.

```
ggplot(americas, aes(x = CPI, y = HDI)) + geom_point() + geom_smooth(method = lm) +
  ggtitle("Corruption and Human Development in American Countries with Regression Line")
```




```
ggplot(americas, aes(x = CPI, y = HDI)) + geom_point() + geom_smooth(method = loess) +
  ggtitle("Corruption and Human Development in American Countries with Loess Line")
```

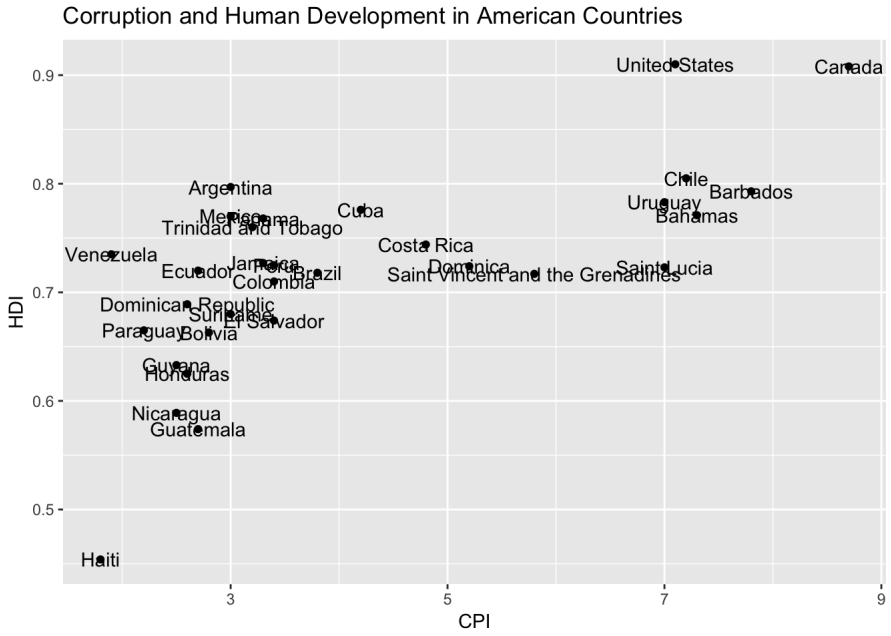


From the plots, we indeed can observe that there seem to be a positive correlation between less corruption (higher CPI means less corruption) and human development in the countries.

Adding Text

We can also add labels and texts to the plot to make it clearer and more comprehensive. There are two functions that we can use for this purpose: geom_text() and geom_label(). The difference between the two is that geom_text adds text directly to the plot, whereas geom_label adds a rectangle behind the text. To continue with the plot that we created in previous section, we will add the country names to the dots on the plot.

```
ggplot(americas, aes(x = CPI, y = HDI)) + geom_point() + geom_text( aes(label = Country)) +
  ggtitle("Corruption and Human Development in American Countries")
```



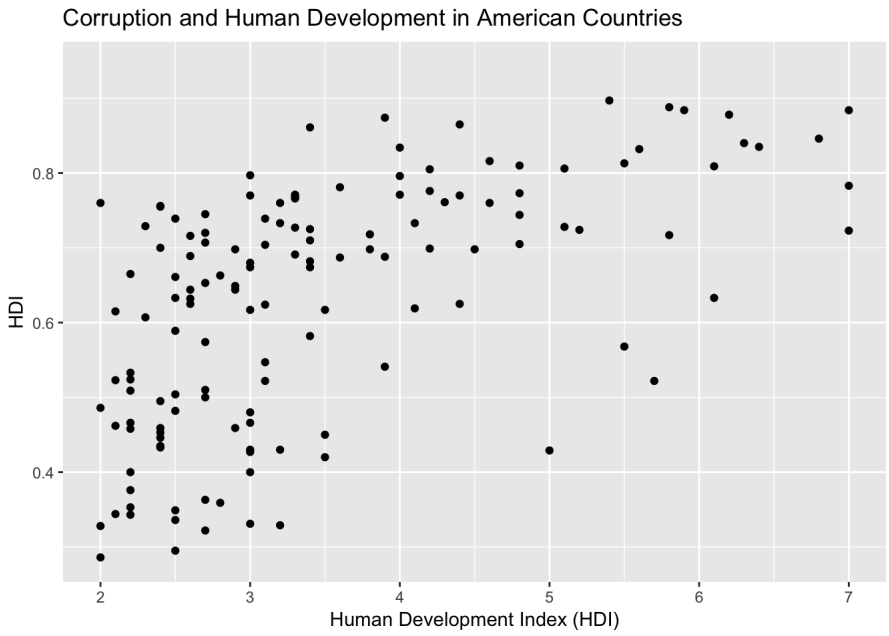
Scales

Up to now, we have been focusing on controlling what is plotted, and in this section, we are going to talk about how to control the visual details of the elements on the plot using scales. With scales, we can restruct the plot to a particular range of variables, set axis names, and change the color scale of a plot. Scales follows scale__ naming scheme. For example, control of the x and y values for continuous varaibles is done with the functions scale-x_continuous and scale_y_continuous. Arguments used by these two functions include “name,” “breaks,” “labels,” “limits,” and “trans”.

If we want to relabel the x-aixs and change range of CPI (limit CPI values from 2 to 7), we will do as the following:

```
correlation <- ggplot(development, aes(x = CPI, y = HDI)) + geom_point() +
  ggtitle("Corruption and Human Development in American Countries") #This is the first basic scatterplot we create
d previously
correlation + scale_x_continuous(name = "Human Development Index (HDI)", limits = c(2, 7))
```

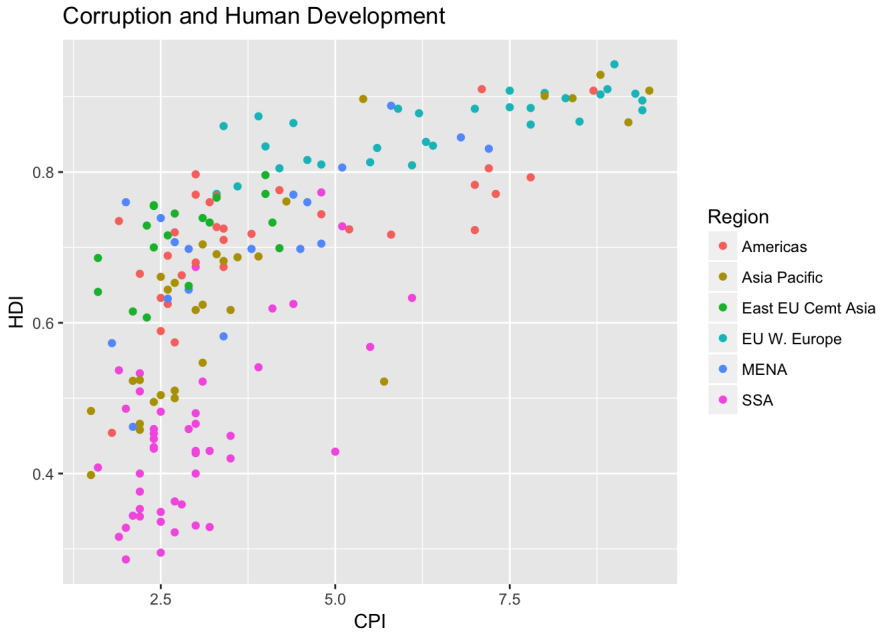
```
## Warning: Removed 34 rows containing missing values (geom_point).
```



Color Scales and Legends

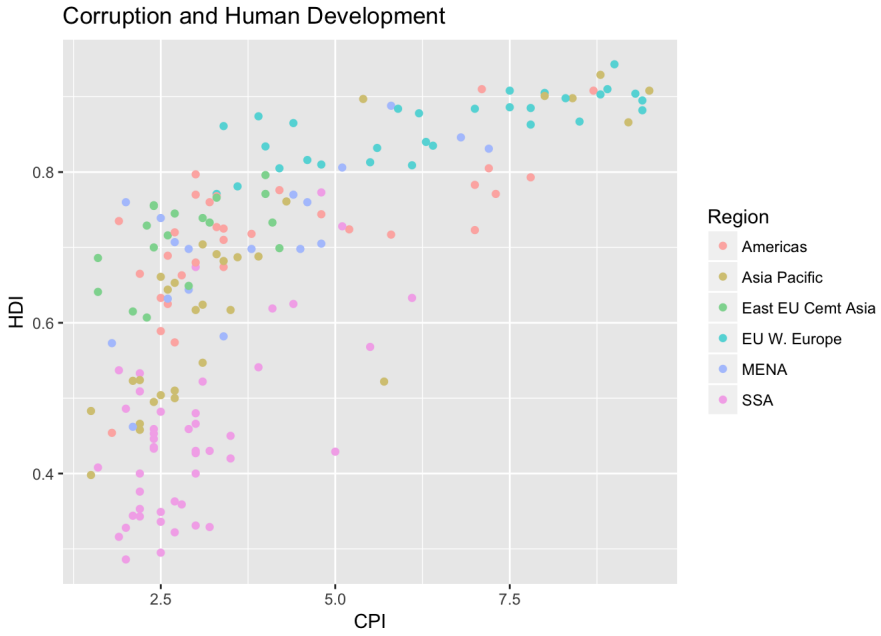
Aside from changing various argument for x and y axis, we can also make the plot more interesting and informative by using functions such as scale_color_discrete or scale_color_hue to change parameters of the color aesthetic mapped, including lengend name and labels, limits, and hue values. A few major paramters are h = range of hues in [0, 360], c = chroma (intensity of color), and l = luminance (lightness) in [0, 100]. We can also use other set color palettes such as ColorBrewer which works better for people with common types of color blindness. Let's try with an example!

```
hdicpi <- ggplot(development, aes(x = CPI, y = HDI, color = Region)) + geom_point() + ggtitle("Corruption and Huma
n Development")
hdicpi
```



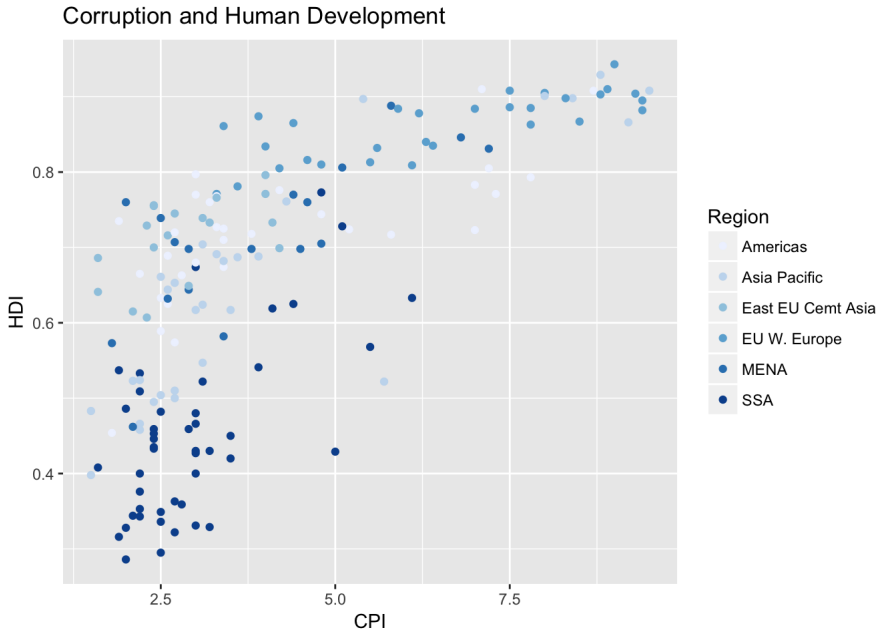
Right now the graph above uses the default categorical scale which picks colors that are evenly spaced around the colour wheel. We can easily modify the colors:

```
hdicpi + scale_color_hue(l = 80, c = 50)
```

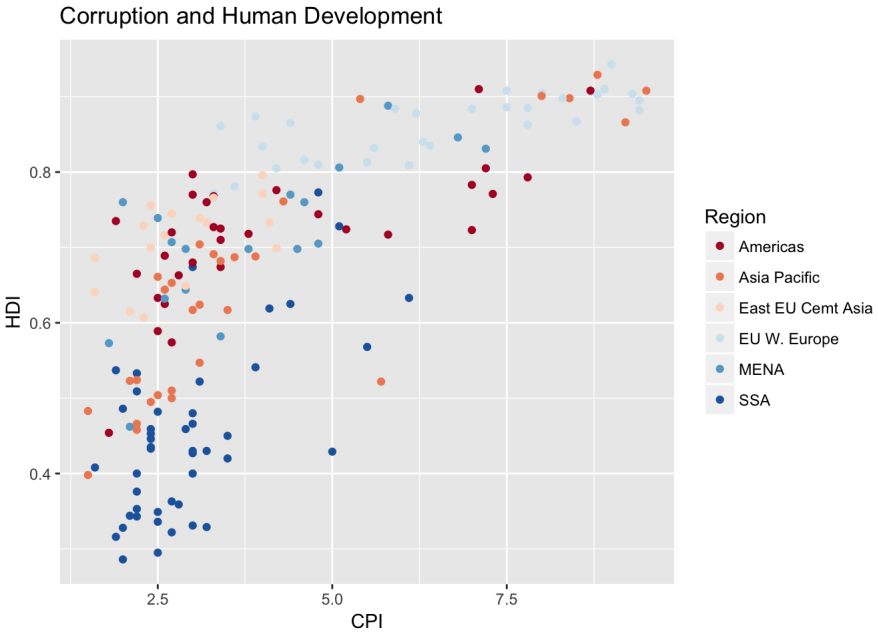


We can also use the ColorBrewer set instead, which is tuned to work better for people with common typles of color blindeness:

```
hdicpi + scale_color_brewer()
```




```
hdicpi + scale_color_brewer(palette = "RdBu")
```



If you are interested, you can find the complete list of all ColorBrewer palattes online at http://colorbrewer2.org/.

Themes

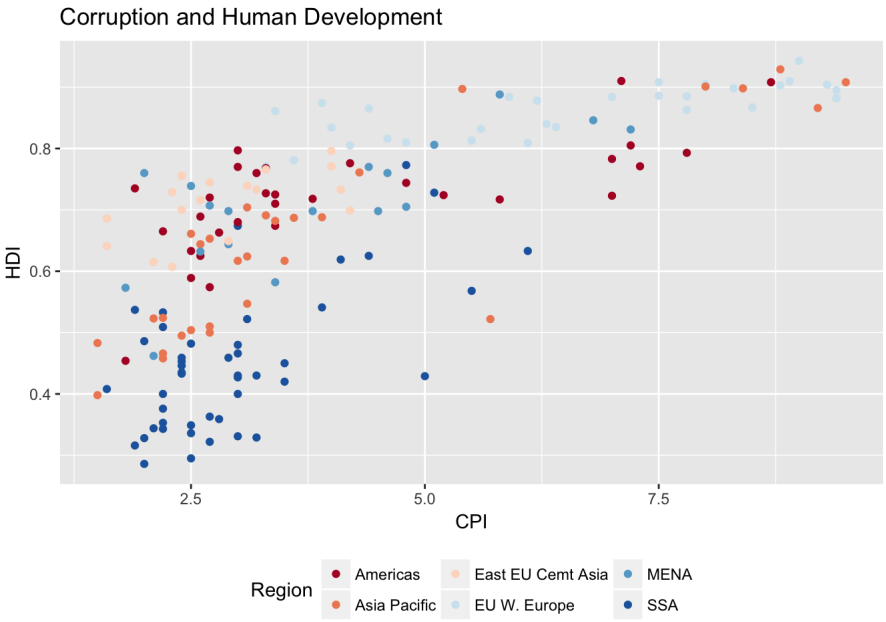
To adjust precisely how we want different parts of the plot to look we want to use the theme command. It can be a little confusing in the beginning to know right away when to use scle and when to use theme, but as you gain experience and start to become more familiar with the functions, it will be clearer for you where to use them. But a general rule is that unlike scales, theme only takes care of the non-data plot elements such as plot background, color, font and size; it cannot add words or change ranges of variables. There are eight themes in ggplot2 by default, and here is a list of them:

- them_bw(): white background with grid lines
- theme_light(): light axes and grid lines
- theme_classic(): axes without grid lines
- theme_linedraw(): black axes and grid lines
- theme_dark(): dark plot backrgound
- theme_minimal(): no background
- theme_gray(): grey background; default theme
- theme_void(): no axes or gridlines, only geoms are shown

Though there are only eight defaut themes, but we can override specific theme elements using the function theme(). There are so many parameters to this function, and in this post we will just show you a few as exemples.

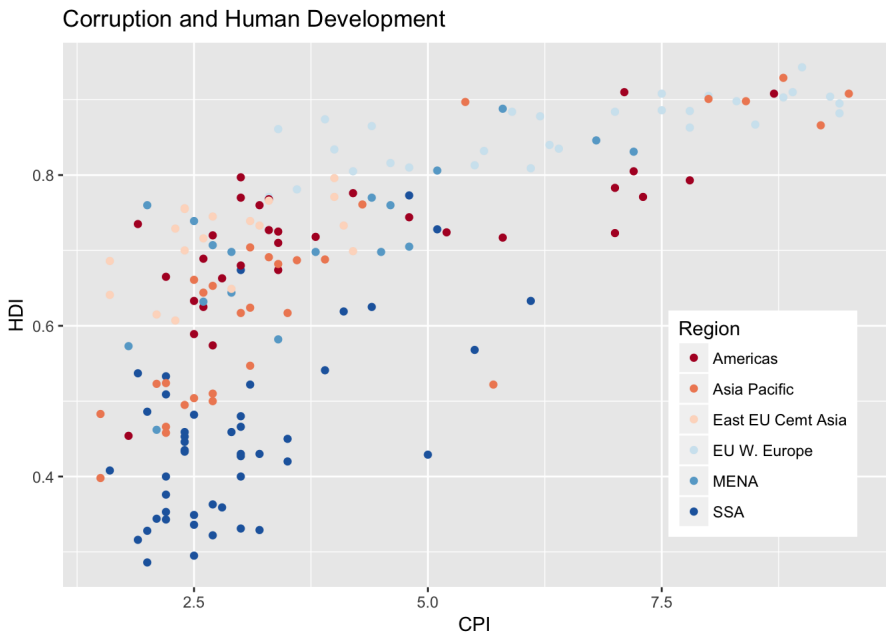
If we want to change the legend position of the plot in the previous section, we can do the following:

```
hdicpi <- ggplot(development, aes(x = CPI, y = HDI, color = Region)) + geom_point() + ggtitle("Corruption and Huma
n Development") + scale_color_brewer(palette = "RdBu")
hdicpi + theme(legend.position = "bottom")
```



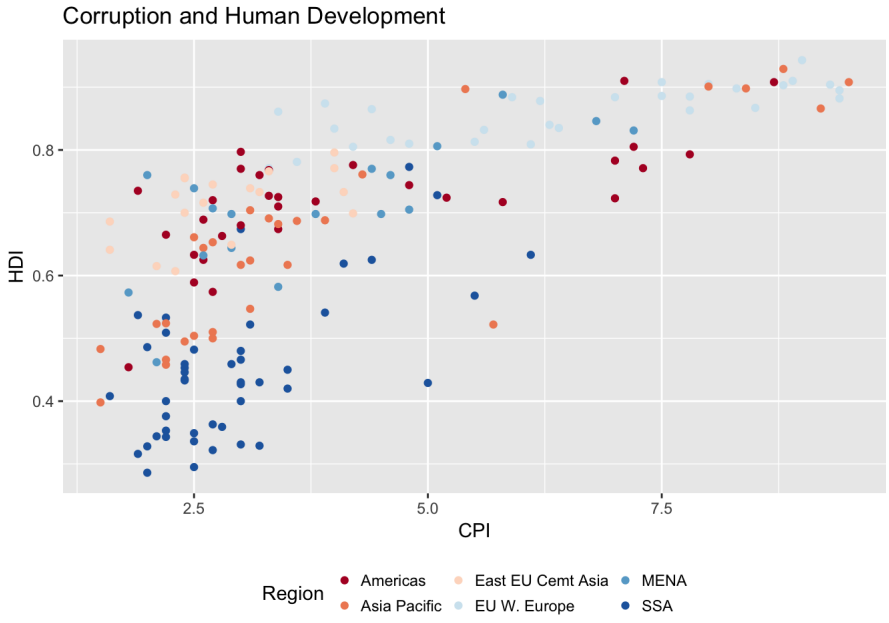
We can also move the legend to the interior of the plot:

```
hdicpi + theme(legend.position = c(0.85, 0.3))
```



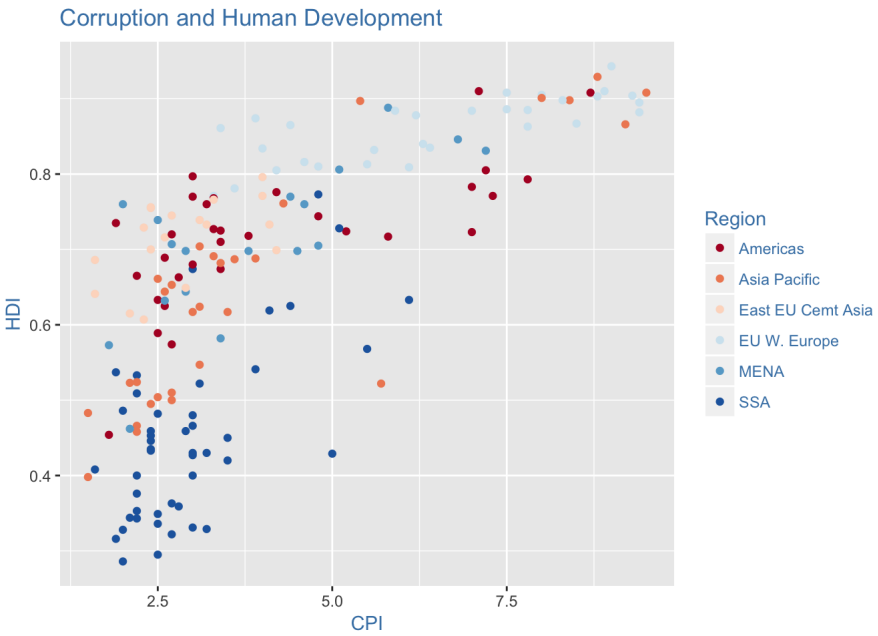
We can also change the size and fill of legend key:

```
hdicpi + theme(legend.position = "bottom", legend.key.size = unit(5, "mm"), legend.key = element_rect(fill = "white"))
```



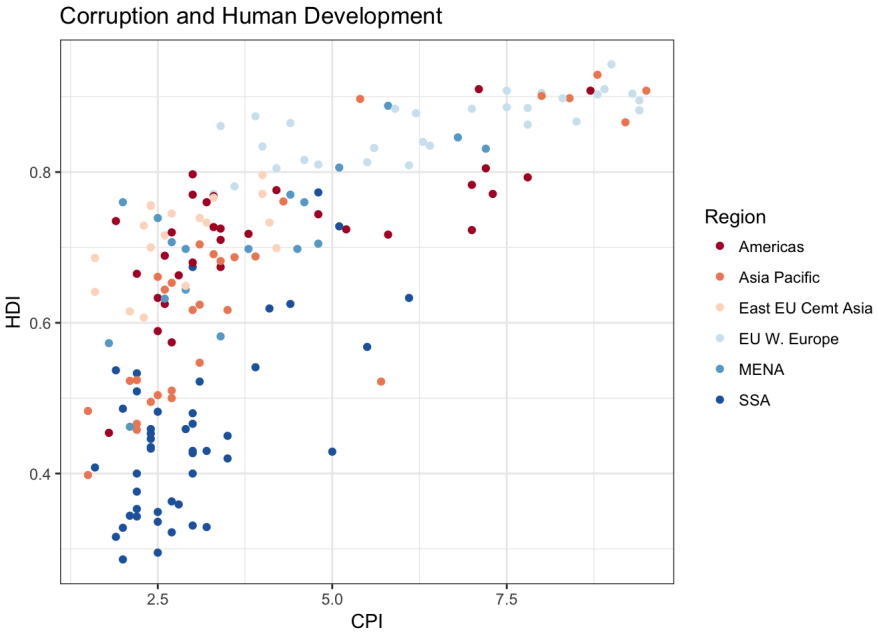
We can also change the text color of our axis labels:

```
hdicpi + theme(text = element_text(color = "Steelblue"))
```



And if we don't like the gray plot background and want to change it to white, we can change the default theme to theme_bw():

```
hdicpi + theme(text = element_text(color = "Steelblue")) + theme_bw()
```



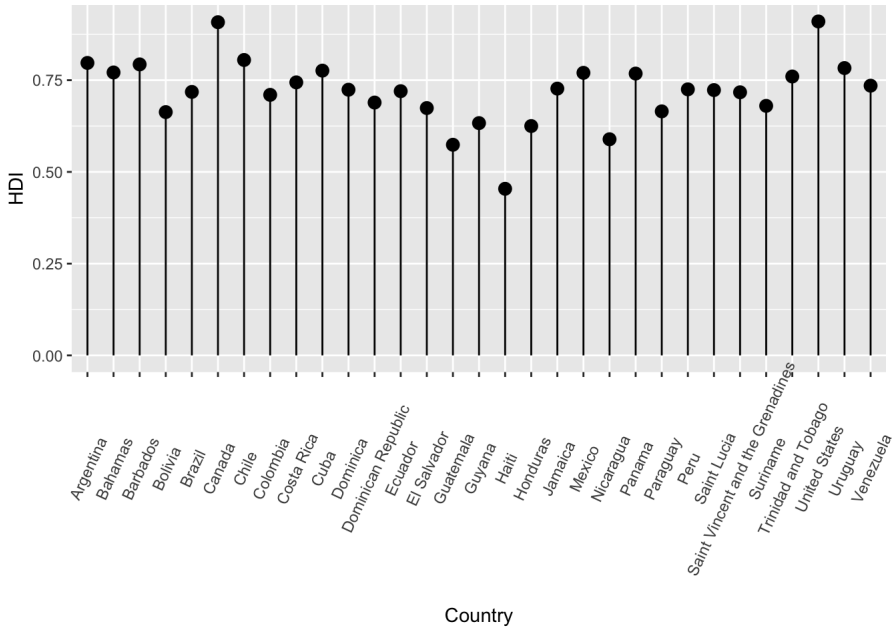
Other ggplot2 Visualizations

Other than the common plots that we use frequently such as scattplots and bar graphs, there are so many other interesting, nice-looking charts ggplot2 can create that also convey the right information effectively. In this final section of the post, we will explore some not very common yet still useful visualizations.

Lollipop Chart

Many of us are familiar with bar charts, and lollipop charts conveys the same information as in bar charts, but instead of having the traidtional thick bars, it has "bars" in the shape of lollipops, which make the chart looks more modern. Lollipop chart is done by adding **geom_point()** and **geom_segment()** that work together to give the shape of lollipop. In the following example, we are creating a lollipop chart ranking the HDI of countries in the regions Americas.

```
ggplot(americas, aes(x = Country, y = HDI)) +
  geom_point(size = 3) +
  geom_segment(aes(x = Country,
                   xend = Country,
                   y = 0,
                   yend = HDI)) +  # In this step we just created the "bars" of the lollipops
  theme(axis.text.x = element_text(angle=65, vjust=0.6))
```



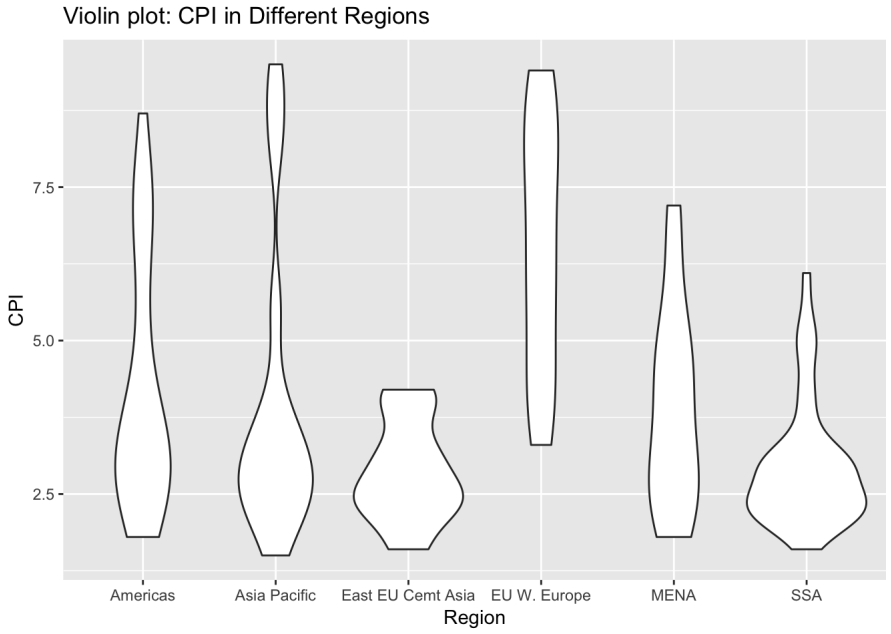

```
  ggtitle("Lollipop Chart: HDI Ranking in America by Country")
```

```
## $title
## [1] "Lollipop Chart: HDI Ranking in America by Country"
##
## $subtitle
## NULL
##
## attr(,"class")
## [1] "labels"
```

Violin Plot

Violin plot is a great tool to visualize the distribution; its function is similiar to that of a boxplot. While boxplot focuses more on summary statistics such as mean, median, and interquartile ranges, the violin plot shows the full distribution of the data. The function to create violin plot is **geom_violin()**. In the following example, we will use violin plot to help us examine the distribution of CPI in different regions.

```
cpi_region <- ggplot(development, aes(Region, CPI))
cpi_region + geom_violin() +
  labs(title="Violin plot: CPI in Different Regions",
       x="Region",
       y="CPI")
```

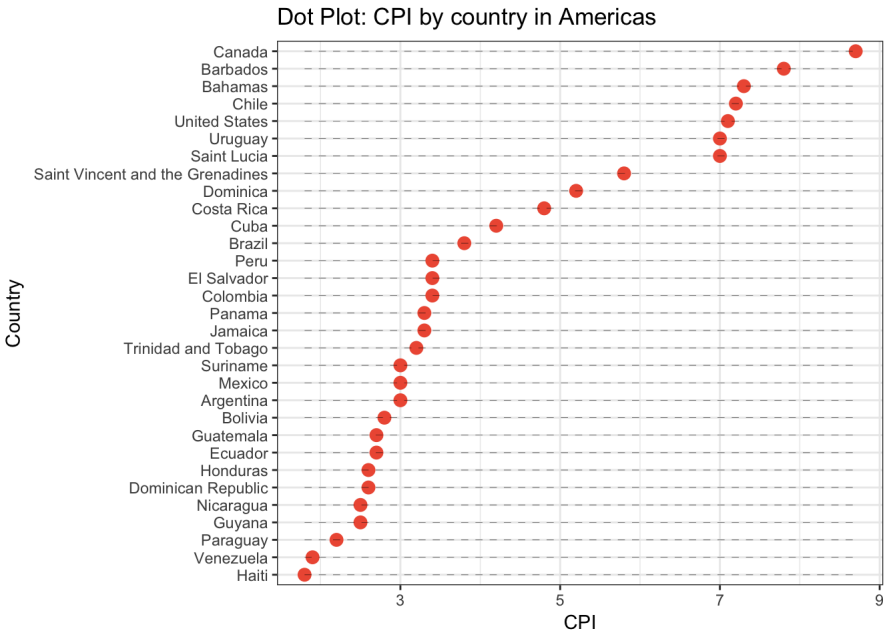


Dot Plot

Dot plot in many situations serves the same purpose as ordered bar chart, which we have studied and practiced in our past homework, but dot plot emphasizes more on the ranking with respect to actual values and how far apart are the values. If you are tired of ordered bar charts, dot plots is another great choice. One inconvenience with dot plot is that we need to manually draw the segnments for each x-value ourselves. In the

following example, we will use dot plot to show the CPI ranking of Countries in the region Americas.

```
ggplot(americas, aes(x = reorder(Country, CPI), y = CPI)) +  # we want to first order the CPI values for our ranki
ng
  geom_point(col = "tomato2", size = 3) +
  geom_segment(aes(x = Country,
                   xend = Country,
                   y = min(CPI),
                   yend = max(CPI)),
               linetype = "dashed",
               size = 0.1) + # Draw dashed lines
   coord_flip() + labs( x = "Country", y = "CPI") +
  ggtitle("Dot Plot: CPI by country in Americas") + coord_flip() + theme_bw()
```



From the above graph, we can observe that in 2011, Canada is the least corrupt country in Americas, and Haiti has the most corruption.

Conclusion (Message)

In this post we have explored a lot about ggplot2, from constructing basic scatterplot to more complex visualizations. From all the examples and the plots that we have constructed, I hope it becomes clear for you that when we are doing data analysis, graphs are extremely useful in terms of presenting the information clearly and effectively. Rankings, patterns, and correlation relationships are easier to examine with a nice graph. And to construct a beautiful and informative plot, ggplot2 is a great tool to use. It adds the components of the graph layer by layer, and it allows us to play with our visual details with a great amount of creativity (as shown in the theme section of this post). Other than the basic plots, ggplot2 also offers a variety of more modern visualizations such as lollipop plot and violin plot that give us more freedom when choosing the best presentation of our data. There is a lot more to explore about ggplot2 beyond this post, and I hope this post inspired you about how ggplot2 can help us to analyze and present data in a very interesting and effective way.

Reference

1. The “graphics for communication chapter” in *R for Data Science*: http://r4ds.had.co.nz/graphics-for-communication.html#annotations
2. Harvard ggplot2 Workshop: http://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html#introduction
3. Function references at the website http://ggplot2.tidyverse.org/reference/index.html
4. *Cookbook for R*: http://www.cookbook-r.com/Graphs/
5. Scales and themes in ggplot2: https://www3.nd.edu/~steve/computing_with_data/12_Scales_themes/scales_themes.html
6. Top 50 ggplot2 Visualizations: http://r-statistics.co/Top50-Ggplot2-Visualizations-MasterList-R-Code.html#Scatterplot%20With%20Encircling
7. *The Economist* dataset used in this tutorial: http://databeauty.com/data/EconomistData.csv