HW 2B, Principles of Statistical Graphs - Nikhil Kotecha

Background

Goal is to get as close to this economist graph as possible (maybe even improve on it)! Here's a link with the orignial: http://www.economist.com/node/21541178

I relied on a harvard tutorial to help me get up to speed with ggplot2:http://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html#orgbe01a81

(I have working knowledge of R from working with data for neural nets but mainly I use python).

Data comes from the linked tutorial with information on country name, Human Development Index, Corruption Perceptions Index, and associated region.

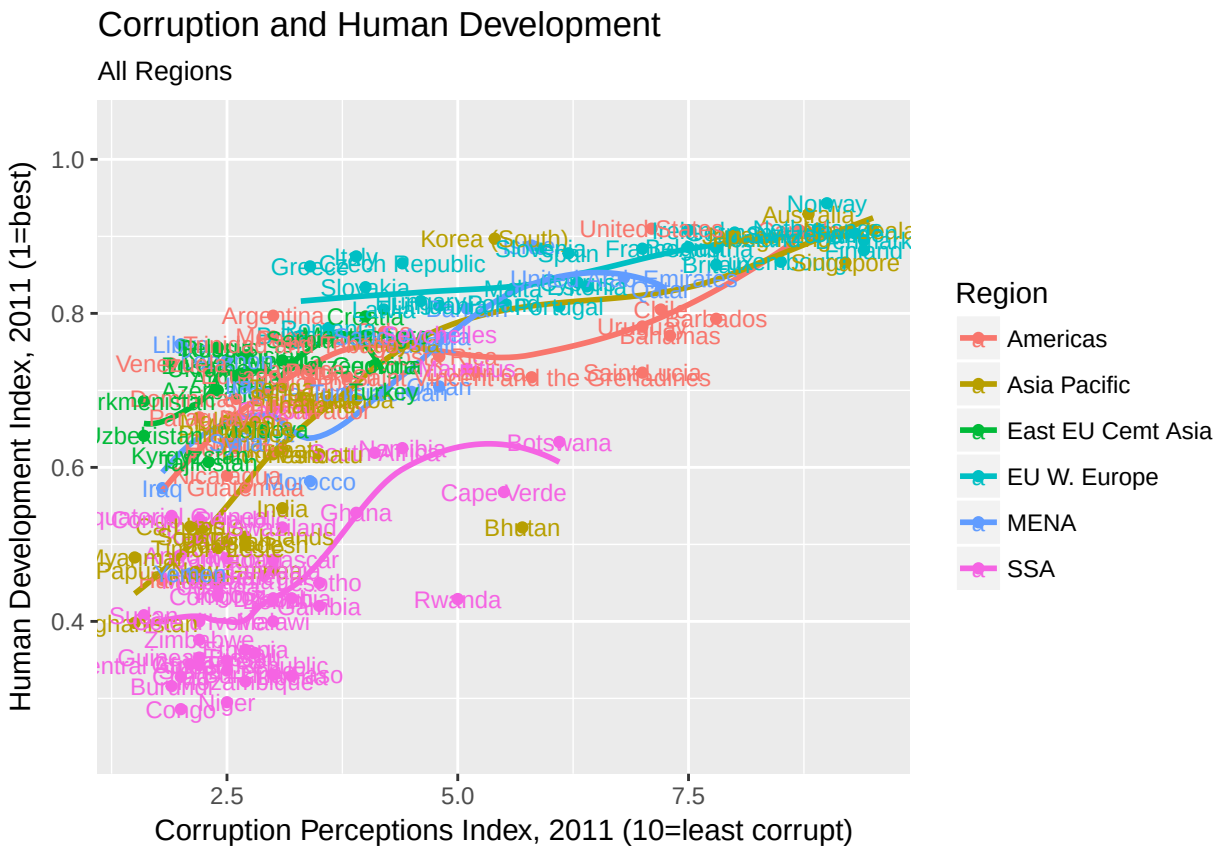
I slowly added information, and found that the most interesting use of space was to show HDI by CPI clustered by region. I initially tried to show all the regions in one graph, but it was too messy. Instead I graphed three regions separately (to stay in the 4 graph limit.) The hypotehtical audience for the plot's are a lay audience interested in new from around the world. My goals in making the graph were to learn how to replicate an economist-esque graph. I was particularly interested in trying a graphic in the vein of a mass media outlet after going through the "Visual Revelations: Improving Data" article in this week's reading. I have tried academic graphics before, but never through R and wanted to bolster my skills here too. In each graph the comparisions are similar: I wanted to show regional differences in perceived corruption and the development of the nation. Initially it was to compare across all regions (the first graph), but later graphs I think more effective because they compare within region.

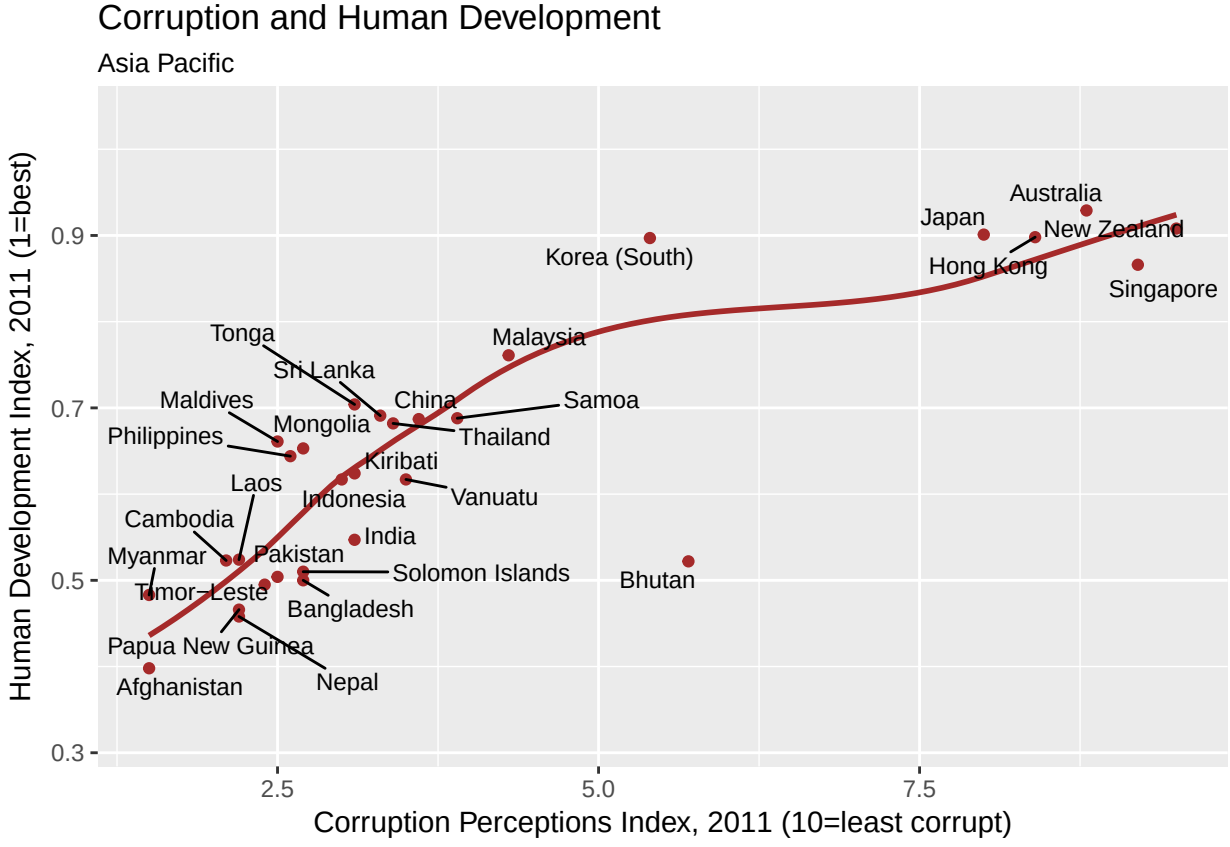
Code Below:

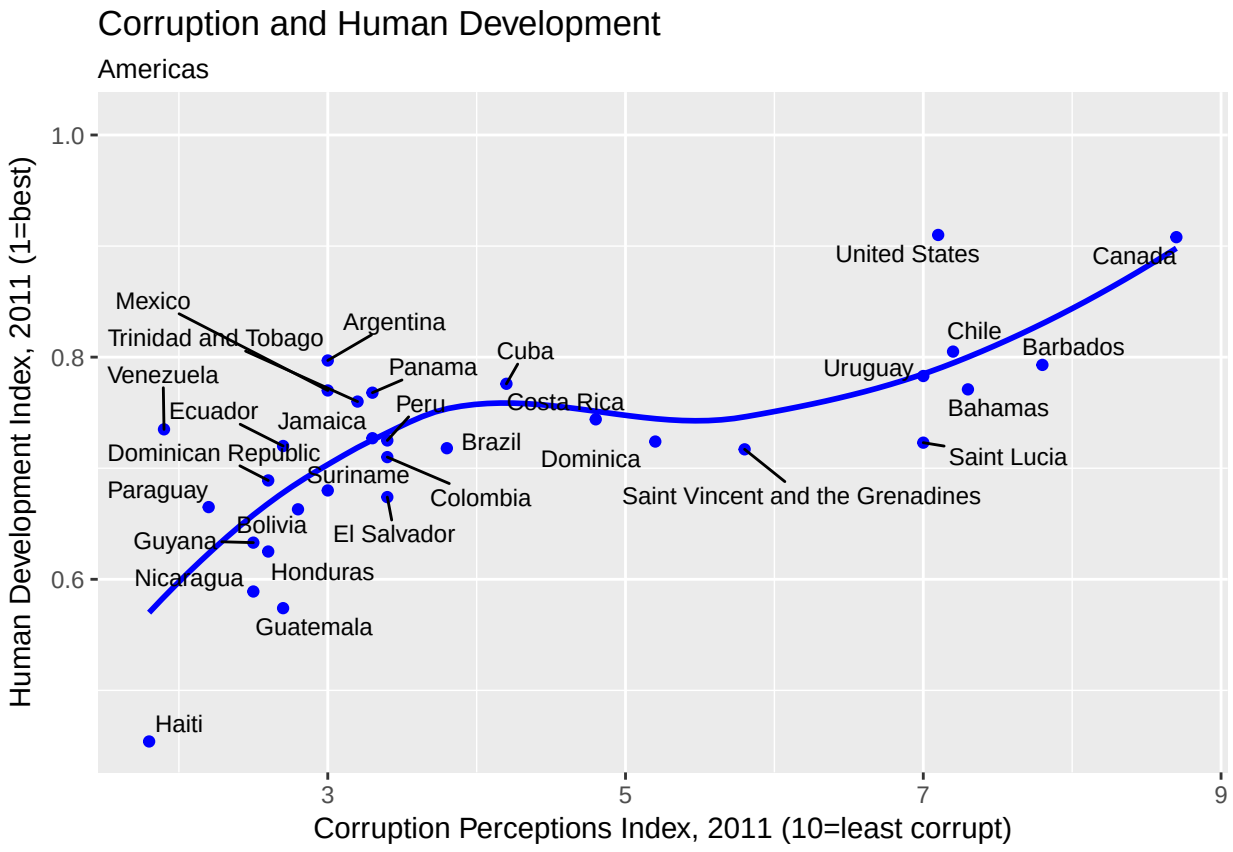
```
library(ggplot2)
library(ggrepel)

EconomistData <- read.csv(file="~/Desktop/EconomistData.csv",head=TRUE,sep=",")

p1 <- ggplot(EconomistData, aes(x = CPI, y = HDI, color=Region)) + geom_point()

p2 <- p1 + geom_smooth(method = loess, fill =NA)

p3 <- p2 + labs(x = "Corruption Perceptions Index, 2011 (10=least corrupt)", y ="Human Development Inde

p4 <- p3 + geom_text(aes(label=Country), size = 3)

#p5 <-p4 + facet_wrap(~Region, ncol = 10)
#p5 was super hard to read, so broke into separate graphs

Asia_Pacific <-
  ggplot(subset(EconomistData, Region %in% c("Asia Pacific")),
  aes(x = CPI,
  y = HDI,
  label=Region)) +
  geom_point(color="brown") +
  geom_smooth(method = loess, fill = NA, color="brown") +
  labs(x = "Corruption Perceptions Index, 2011 (10=least corrupt)", y =
  "Human Development Index, 2011 (1=best)", title = "Corruption and Human Development", subtitle = "Asi
```

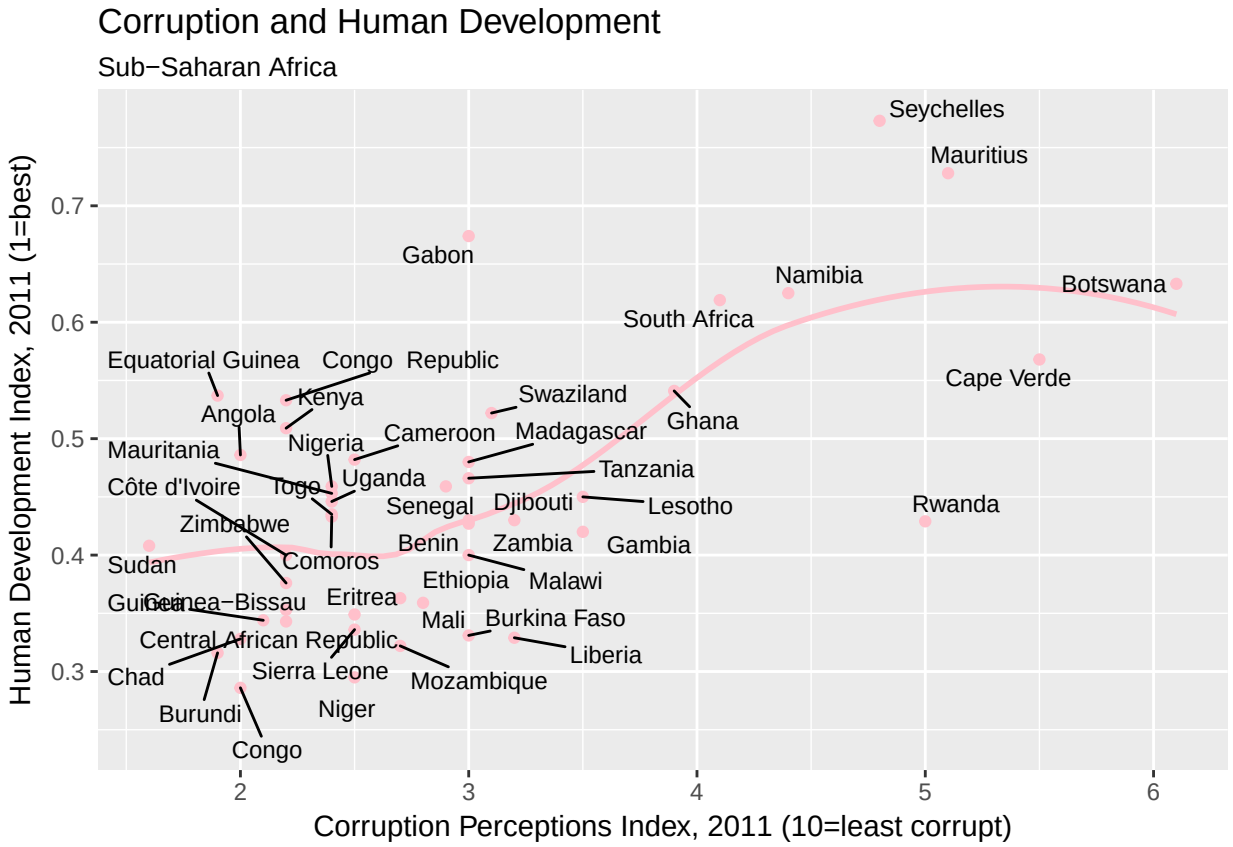
```
  geom_text_repel(aes(label=Country), size = 3)

Americas <- ggplot(subset(EconomistData, Region %in% c("Americas")),
       aes(x=CPI,
           y=HDI,
           label=Region))+
  geom_point(color="blue")+
  geom_smooth(method = loess, fill =NA, color="blue")+
  labs(x = "Corruption Perceptions Index, 2011 (10=least corrupt)", y ="Human Development Index, 2011 (
  geom_text_repel(aes(label=Country), size = 3)

SSA <- ggplot(subset(EconomistData, Region %in% c("SSA")),
       aes(x=CPI,
           y=HDI,
           label=Region))+
  geom_point(color="pink")+
  geom_smooth(method = loess, fill =NA, color="pink")+
  labs(x = "Corruption Perceptions Index, 2011 (10=least corrupt)", y ="Human Development Index, 2011 (
  geom_text_repel(aes(label=Country), size = 3)
```

Plots


Corruption and Human Development
All Regions
Human Development Index, 2011 (1=best)
1.0
0.8
0.6
0.4
2.5
5.0
7.5
Corruption Perceptions Index, 2011 (10=least corrupt)
Region
Americas
Asia Pacific
East EU Cemt Asia
EU W. Europe
MENA
SSA


Corruption and Human Development
Asia Pacific
Human Development Index, 2011 (1=best)
0.9
0.7
0.5
0.3
2.5
5.0
7.5
Corruption Perceptions Index, 2011 (10=least corrupt)
Australia
Japan
New Zealand
Hong Kong
Singapore
Korea (South)
Tonga
Malaysia
Sri Lanka
Maldives
China
Samoa
Mongolia
Philippines
Thailand
Kiribati
Laos
Vanuatu
Indonesia
Cambodia
India
Myanmar
Pakistan
Solomon Islands
Bhutan
Timor-Leste
Bangladesh
Papua New Guinea
Nepal
Afghanistan


Corruption and Human Development
Americas
Human Development Index, 2011 (1=best)
1.0
0.8
0.6
Mexico
Trinidad and Tobago
Venezuela
Ecuador
Dominican Republic
Paraguay
Guyana
Nicaragua
Haiti
Argentina
Panama
Peru
Jamaica
Suriname
Bolivia
Honduras
Guatemala
Cuba
Brazil
Colombia
El Salvador
Costa Rica
Dominica
Saint Vincent and the Grenadines
United States
Canada
Chile
Barbados
Uruguay
Bahamas
Saint Lucia
3
5
7
9
Corruption Perceptions Index, 2011 (10=least corrupt)


Corruption and Human Development
Sub−Saharan Africa
Human Development Index, 2011 (1=best)
0.7
0.6
0.5
0.4
0.3
Seychelles
Mauritius
Gabon
Namibia
Botswana
South Africa
Cape Verde
Equatorial Guinea
Congo Republic
Kenya
Angola
Swaziland
Ghana
Cameroon
Madagascar
Nigeria
Mauritania
Uganda
Tanzania
Côte d'Ivoire
Togo
Senegal
Djibouti
Lesotho
Rwanda
Zimbabwe
Benin
Zambia
Gambia
Sudan
Comoros
Ethiopia
Malawi
Guinea
Guinea−Bissau
Eritrea
Mali
Burkina Faso
Central African Republic
Liberia
Chad
Sierra Leone
Mozambique
Burundi
Niger
Congo
2
3
4
5
6
Corruption Perceptions Index, 2011 (10=least corrupt)