

Visualization

Shrinath

May 29, 2018

```
library(ggplot2)
```

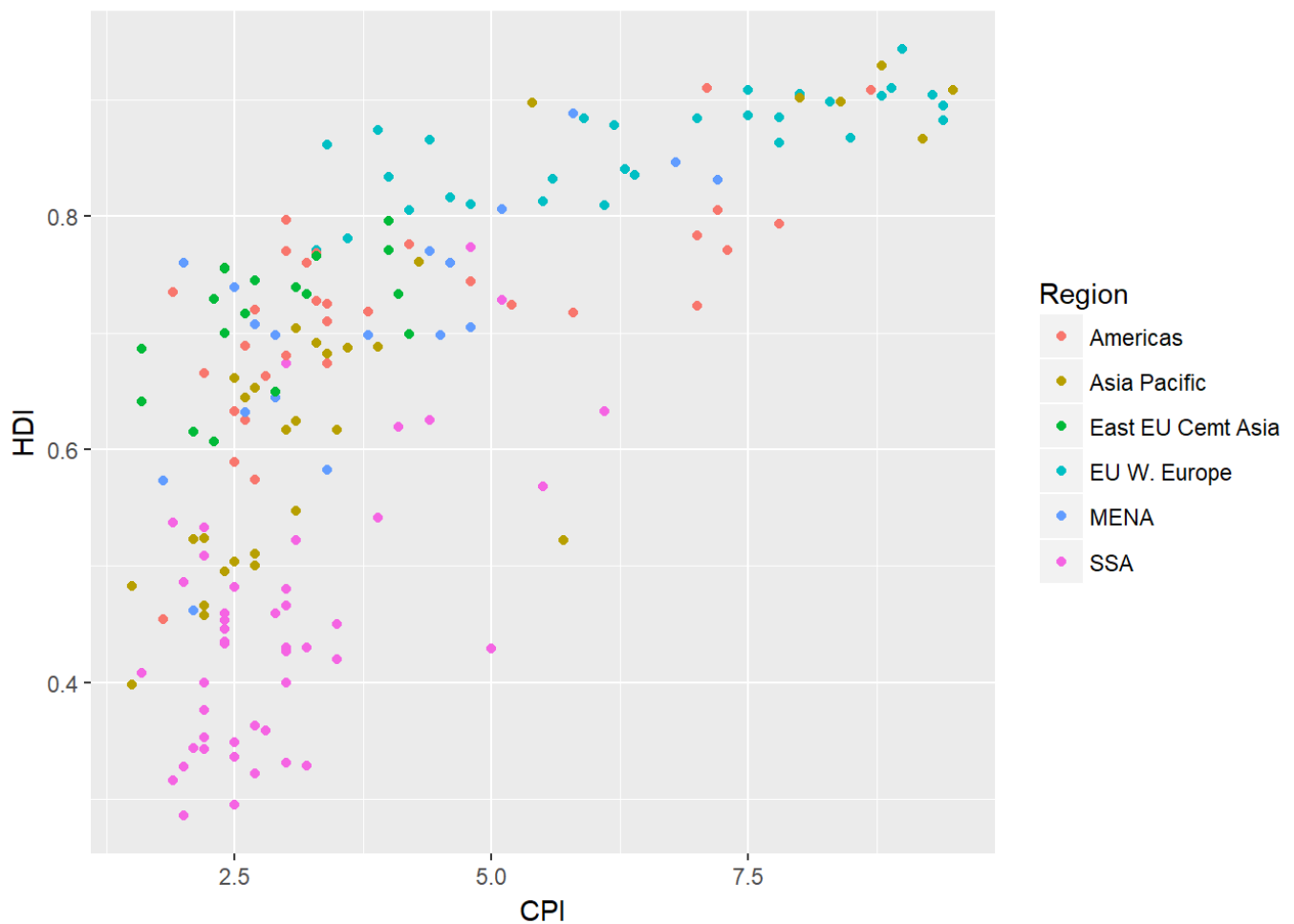
```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
library(data.table)
df <- fread('Economist_Assignment_Data.csv',drop=1)
```

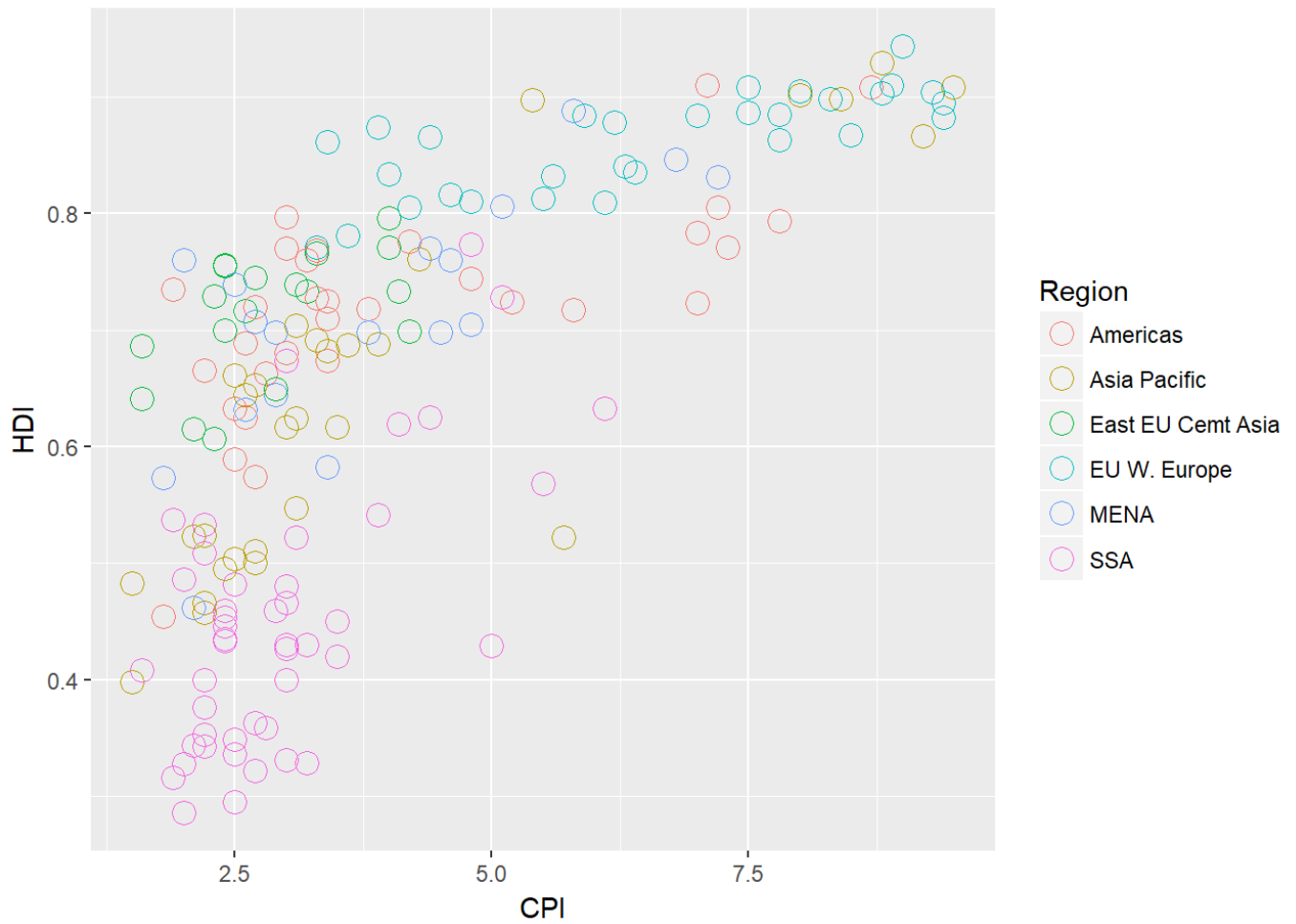
```
head(df)
```

```
##      Country HDI.Rank  HDI CPI      Region
## 1: Afghanistan    172 0.398 1.5 Asia Pacific
## 2:  Albania        70 0.739 3.1 East EU Cemt Asia
## 3:  Algeria        96 0.698 2.9      MENA
## 4:  Angola       148 0.486 2.0      SSA
## 5: Argentina       45 0.797 3.0      Americas
## 6:  Armenia        86 0.716 2.6 East EU Cemt Asia
```

```
ggplot(df,aes(x=CPI,y=HDI,color=Region)) + geom_point()
```

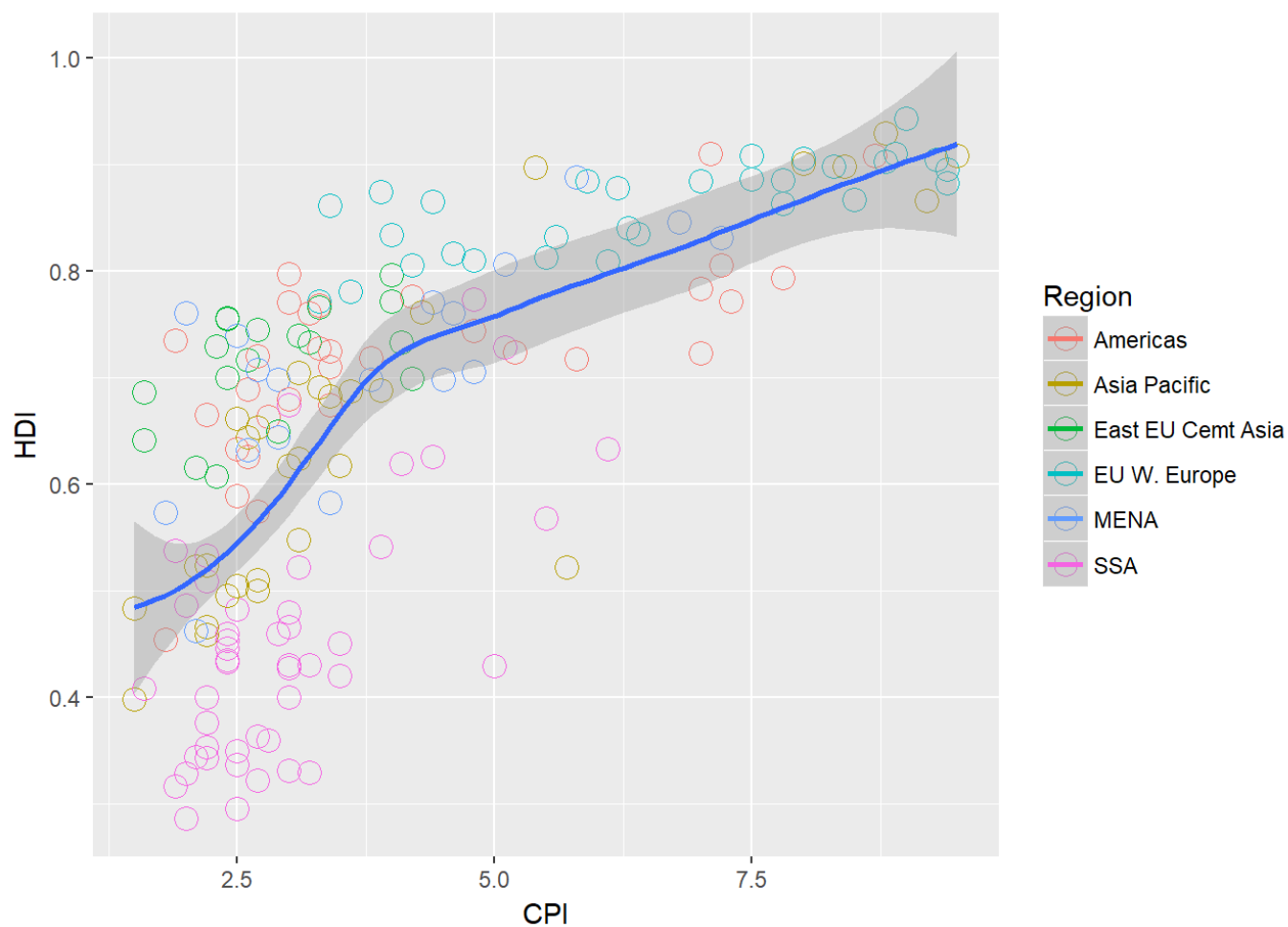


```
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1)
```



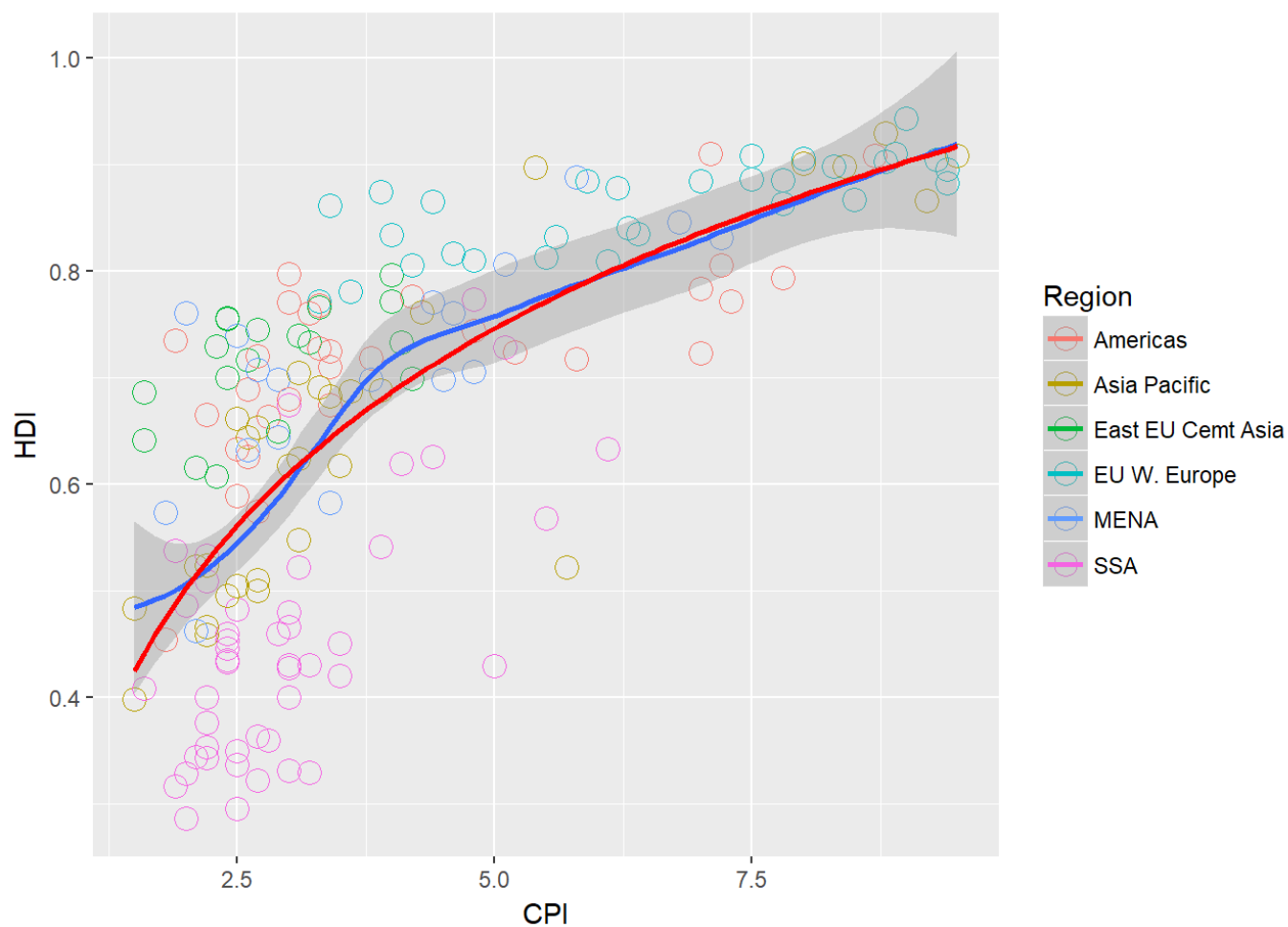
```
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1) + geom_smooth(aes(group=1))
```

```
## `geom_smooth()` using method = 'loess'
```



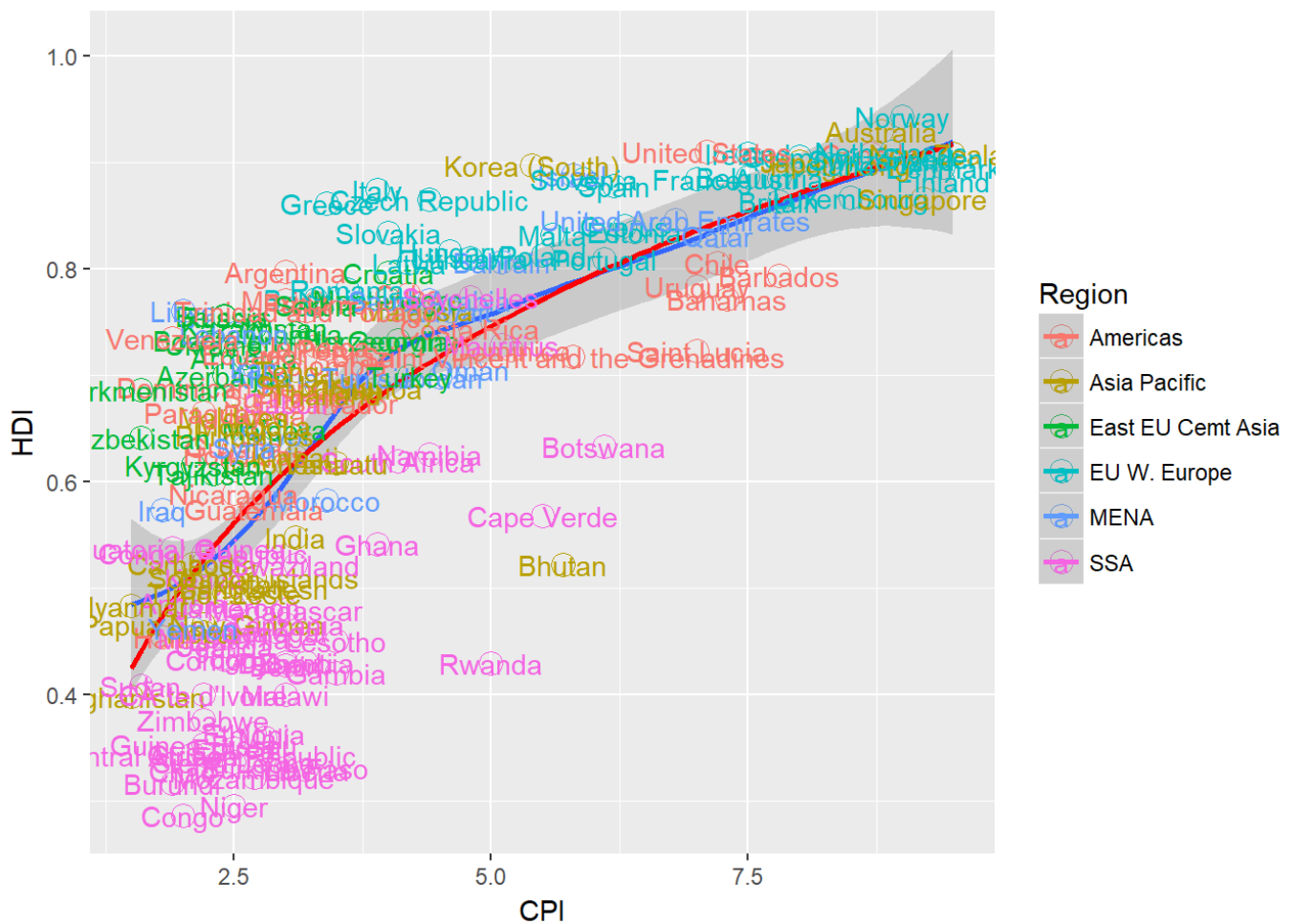
```
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1) + geom_smooth(aes(group=1)) + geom_smooth(aes(group=1), method='lm', formula = y~log(x), se=FALSE, color='red')
```

```
## `geom_smooth()` using method = 'loess'
```



```
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1) + geom_smooth(aes(group=1)) + geom_smooth(aes(group=1), method='lm', formula = y~log(x), se=FALSE, color='red') + geom_text(aes(label=Country))
```

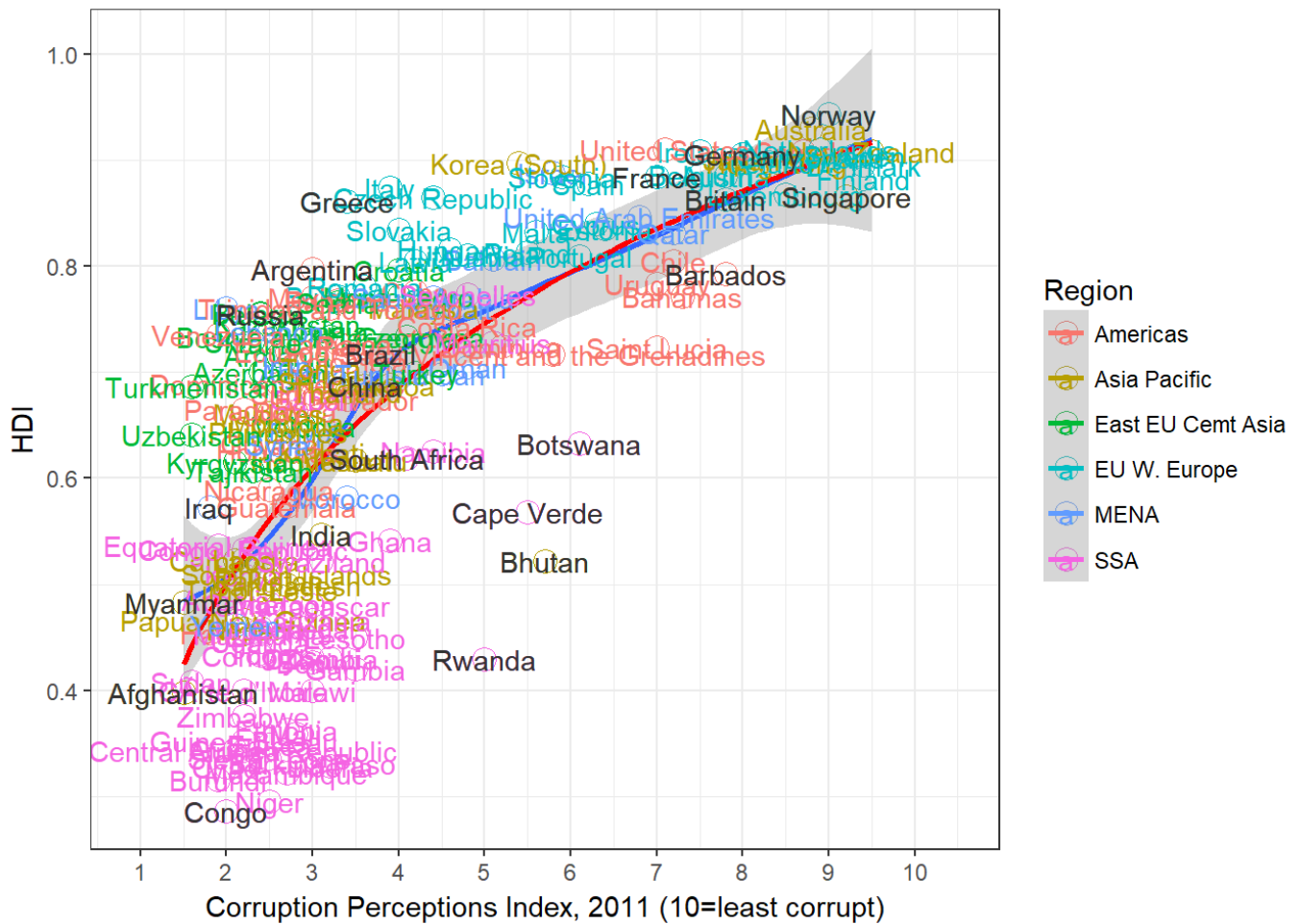
```
## `geom_smooth()` using method = 'loess'
```



```
pointsToLabel <- c("Russia", "Venezuela", "Iraq", "Myanmar", "Sudan",
  "Afghanistan", "Congo", "Greece", "Argentina", "Brazil",
  "India", "Italy", "China", "South Africa", "Spain",
  "Botswana", "Cape Verde", "Bhutan", "Rwanda", "France",
  "United States", "Germany", "Britain", "Barbados", "Norway", "Japan",
  "New Zealand", "Singapore")

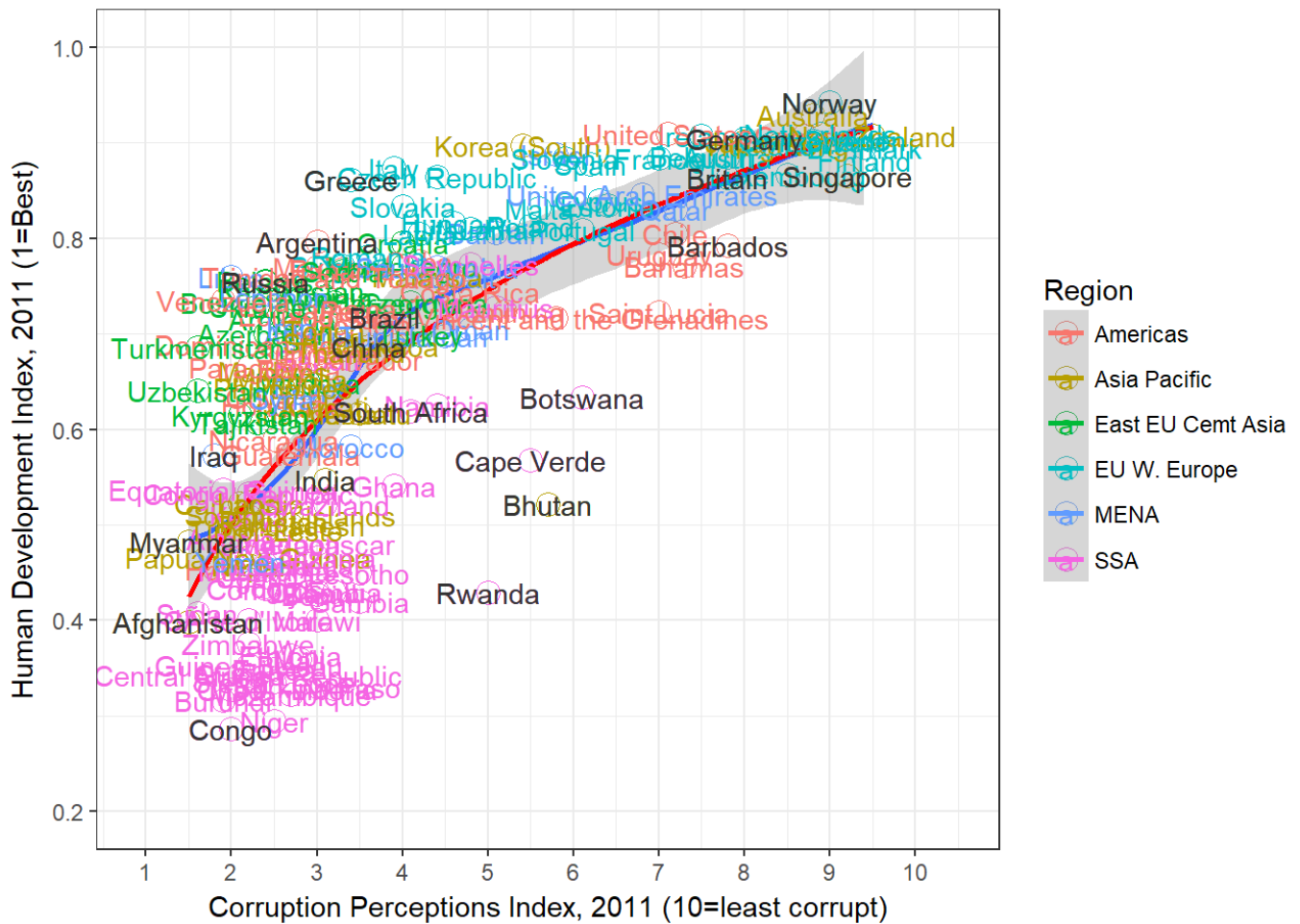
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1) + geom_smooth(aes(group=1)) +
  geom_smooth(aes(group=1), method = 'lm', formula = y~log(x), se=FALSE, color='red') +
  geom_text(aes(label=Country)) + geom_text(aes(label = Country), color = "gray20",
  data = subset(df, Country %in% pointsToLabel), check_overlap = TRUE) + theme_bw()
```

```
## `geom_smooth()` using method = 'loess'
```

```
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1) + geom_smooth(aes(group=1)) +
  geom_smooth(aes(group=1), method = 'lm', formula = y~log(x), se=FALSE, color='red') +
  geom_text(aes(label=Country)) + geom_text(aes(label = Country), color = "gray20",
  data = subset(df, Country %in% pointsToLabel), check_overlap = TRUE) +
  theme_bw() +
  scale_x_continuous(name = "Corruption Perceptions Index, 2011 (10=least corrupt)",
    limits = c(.9, 10.5), breaks=1:10) +
  scale_y_continuous(name = "Human Development Index, 2011 (1=Best)", limits = c(0.2, 1.0))
```

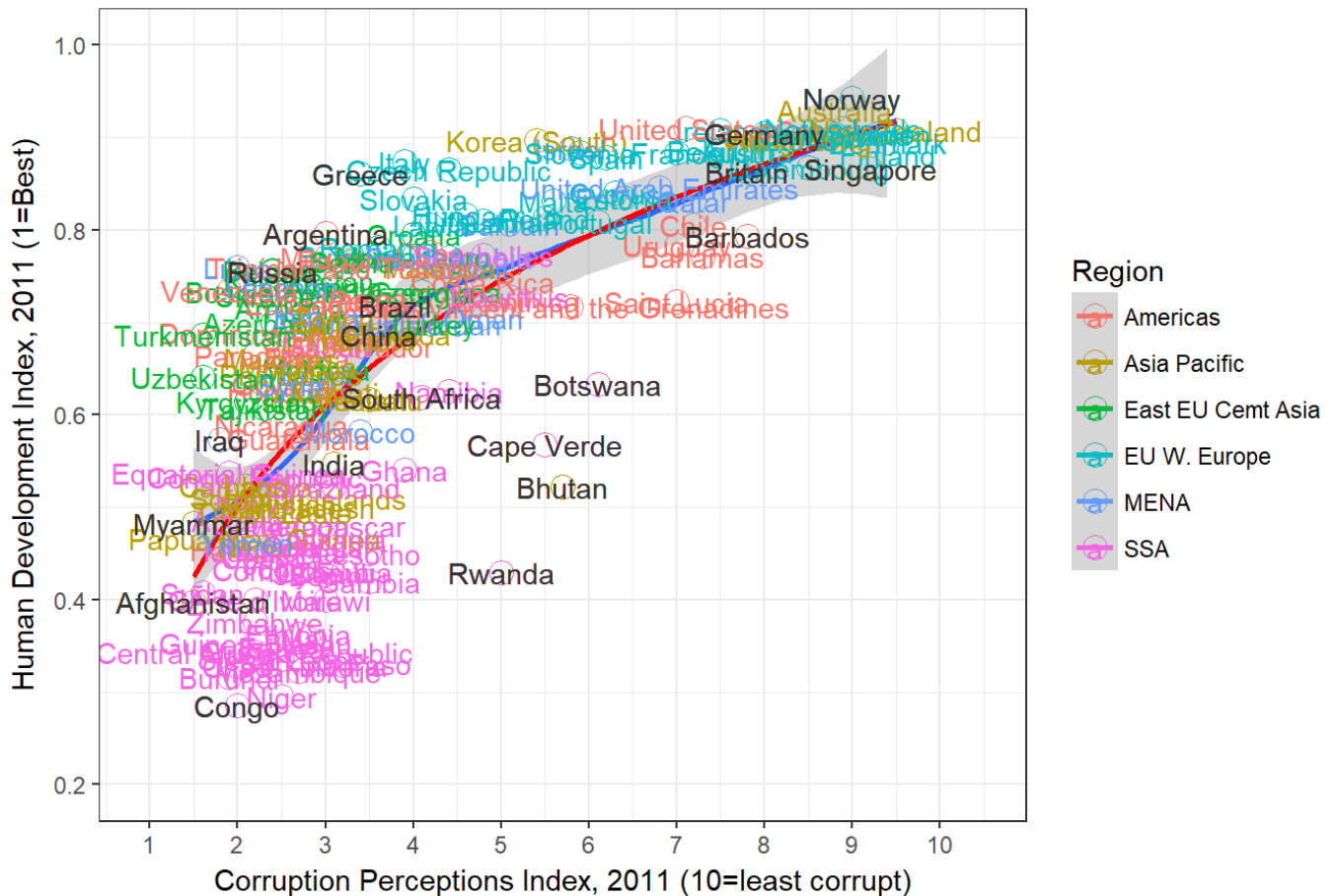
```
## `geom_smooth()` using method = 'loess'
```



```
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1) + geom_smooth(aes(group=1)) +
  geom_smooth(aes(group=1), method='lm', formula = y~log(x), se=FALSE, color='red') +
  geom_text(aes(label=Country)) + geom_text(aes(label = Country), color = "gray20",
  data = subset(df, Country %in% pointsToLabel), check_overlap = TRUE) +
  theme_bw() +
  scale_x_continuous(name = "Corruption Perceptions Index, 2011 (10=least corrupt)",
    limits = c(.9, 10.5), breaks=1:10) +
  scale_y_continuous(name = "Human Development Index, 2011 (1=Best)",
    limits = c(0.2, 1.0)) +
  ggtitle("Corruption and Human development")
```

```
## `geom_smooth()` using method = 'loess'
```


Corruption and Human development



library (ggthemes)

```
ggplot(df, aes(x=CPI, y=HDI, color=Region)) + geom_point(size=4, shape=1) + geom_smooth(aes(group=1)) + geom_smooth(aes(group=1), method = 'lm', formula=y~log(x), se=FALSE, color = 'red') + geom_text(aes(label=Country)) + geom_text(aes(label = Country), color = "gray20", data = subset(df, Country %in% pointsToLabel), check_overlap = TRUE) + theme_bw() + scale_x_continuous(name = "Corruption Perceptions Index, 2011 (10=least corrupt)", limits = c(.9, 10.5), breaks=1:10) + scale_y_continuous(name = "Human Development Index, 2011 (1=Best)", limits = c(0.2, 1.0)) + ggtitle("Corruption and Human development") + theme_economist_white()
```

```
## `geom_smooth()` using method = 'loess'
```

Corruption and Human development

