Exploratory Data Visualizations

Joan Wang

October 15, 2017

```
library(tidyverse)
```

```
## Loading tidyverse: ggplot2
## Loading tidyverse: tibble
## Loading tidyverse: tidyr
## Loading tidyverse: readr
## Loading tidyverse: purrr
## Loading tidyverse: dplyr
```

```
## Conflicts with tidy packages ----------------------------------------------
```

```
## filter(): dplyr, stats
## lag():    dplyr, stats
```

```
countries <- read_csv('../data/Countries_merged.csv')
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   X1 = col_integer(),
##   id = col_integer(),
##   name_DOL = col_character(),
##   num_territories = col_integer(),
##   region = col_character(),
##   region_id = col_integer(),
##   name_WB = col_character(),
##   `Country Name` = col_character(),
##   name_HTI = col_character()
## )
```

```
## See spec(...) for full column specifications.
```

```
assessments <- read_csv('../data/Assessments_cleaned.csv')
```

```
## Parsed with column specification:
## cols(
##   X1 = col_integer(),
##   `Unnamed: 0` = col_integer(),
##   advancement_level = col_character(),
##   advancement_level_id = col_integer(),
##   country = col_character(),
##   country_id = col_integer(),
##   description = col_character(),
##   id = col_integer(),
##   year = col_integer()
## )
```

```
country_goods <- read_csv('../data/CountryGoods.csv')
```

```
## Parsed with column specification:
## cols(
```

```
##   X1 = col_integer(),
##   assessment_id = col_integer(),
##   child_labor = col_character(),
##   country = col_character(),
##   country_id = col_integer(),
##   forced_child_labor = col_character(),
##   forced_labor = col_character(),
##   good = col_character(),
##   good_id = col_integer(),
##   id = col_integer(),
##   regionname = col_character(),
##   year = col_integer()
## )
```

```
statistics <- read_csv('../data/Statistics_cleaned.csv')
```

```
## Parsed with column specification:
## cols(
##   X1 = col_integer(),
##   `Unnamed: 0` = col_integer(),
##   age_range_of_children_attending_school = col_character(),
##   age_range_of_children_working_and_studying = col_character(),
##   age_range_of_working_children = col_character(),
##   assessment_id = col_integer(),
##   country = col_character(),
##   country_id = col_integer(),
##   id = col_integer(),
##   percent_of_children_attending_school = col_double(),
##   percent_of_children_working_and_studying = col_double(),
##   percent_of_working_children = col_double(),
##   percent_of_working_children_agriculture = col_double(),
##   percent_of_working_children_industry = col_double(),
##   percent_of_working_children_services = col_double(),
##   population_of_working_children = col_double(),
##   primary_completion_rate = col_double(),
##   year = col_integer()
## )
```

Bubble Chart

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
stats_2016 <- filter(statistics, year == 2016)
```

```
## Warning: package 'bindrcpp' was built under R version 3.3.3
```

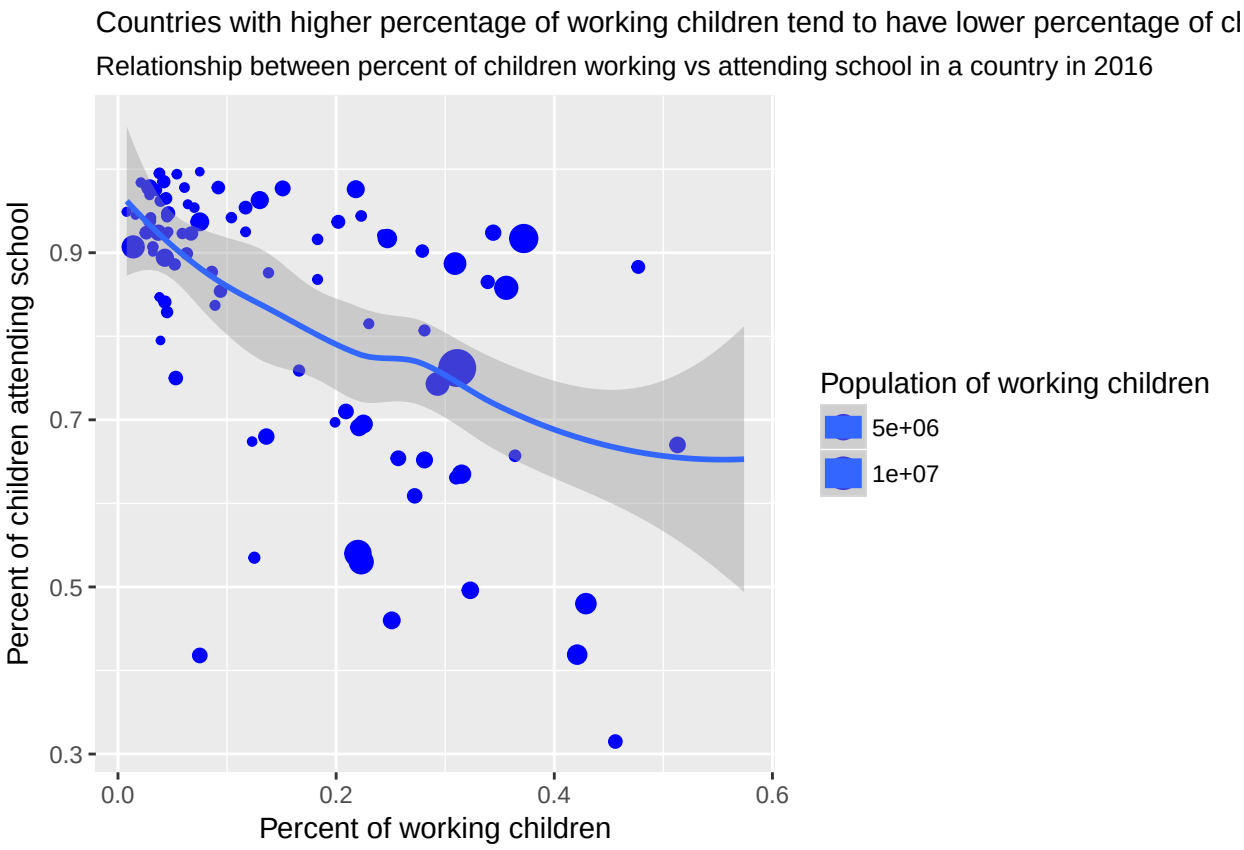
```
ggplot(data=stats_2016, aes(x=percent_of_working_children, y=percent_of_children_attending_school, size=
  geom_point(na.rm=TRUE, color="Blue") +
  geom_smooth(model = lm) +
  labs(title = "Countries with higher percentage of working children tend to have lower percentage of ch
       subtitle = "Relationship between percent of children working vs attending school in a country in
```

```
      x = "Percent of working children",
      y = "Percent of children attending school",
      size = "Population of working children") +
  theme(plot.title = element_text(size=11))
```

```
## Warning: Ignoring unknown parameters: model
```

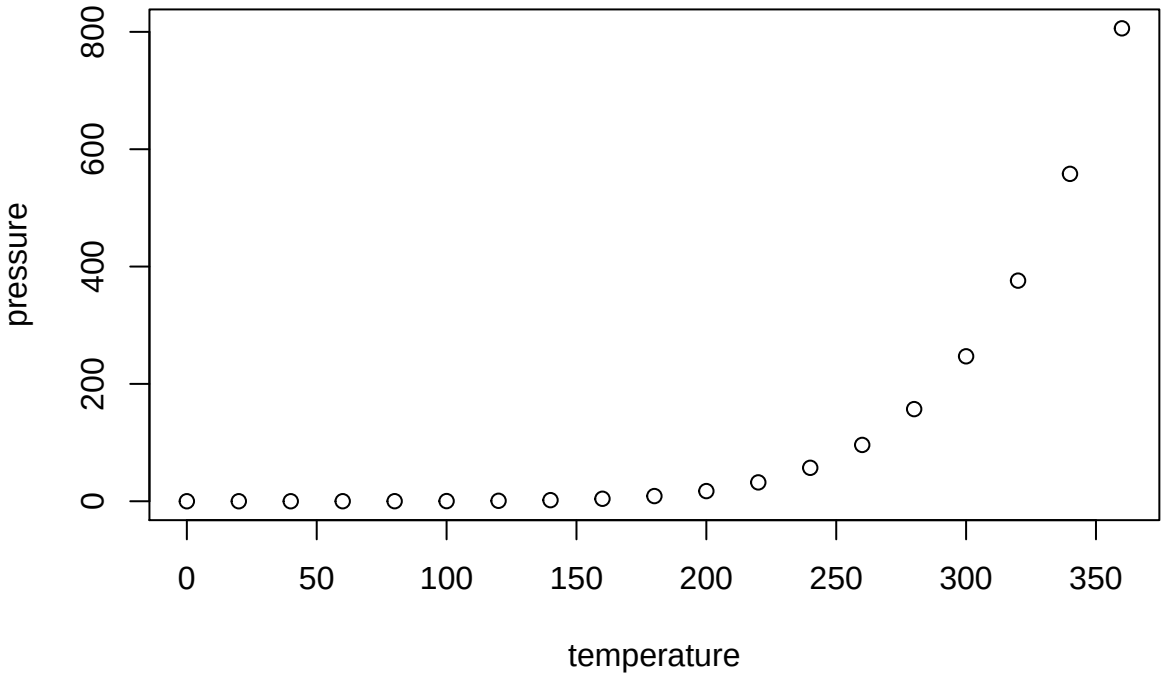
```
## `geom_smooth()` using method = 'loess'
```

```
## Warning: Removed 38 rows containing non-finite values (stat_smooth).
```



Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.