Lesson 6

Xiaoni

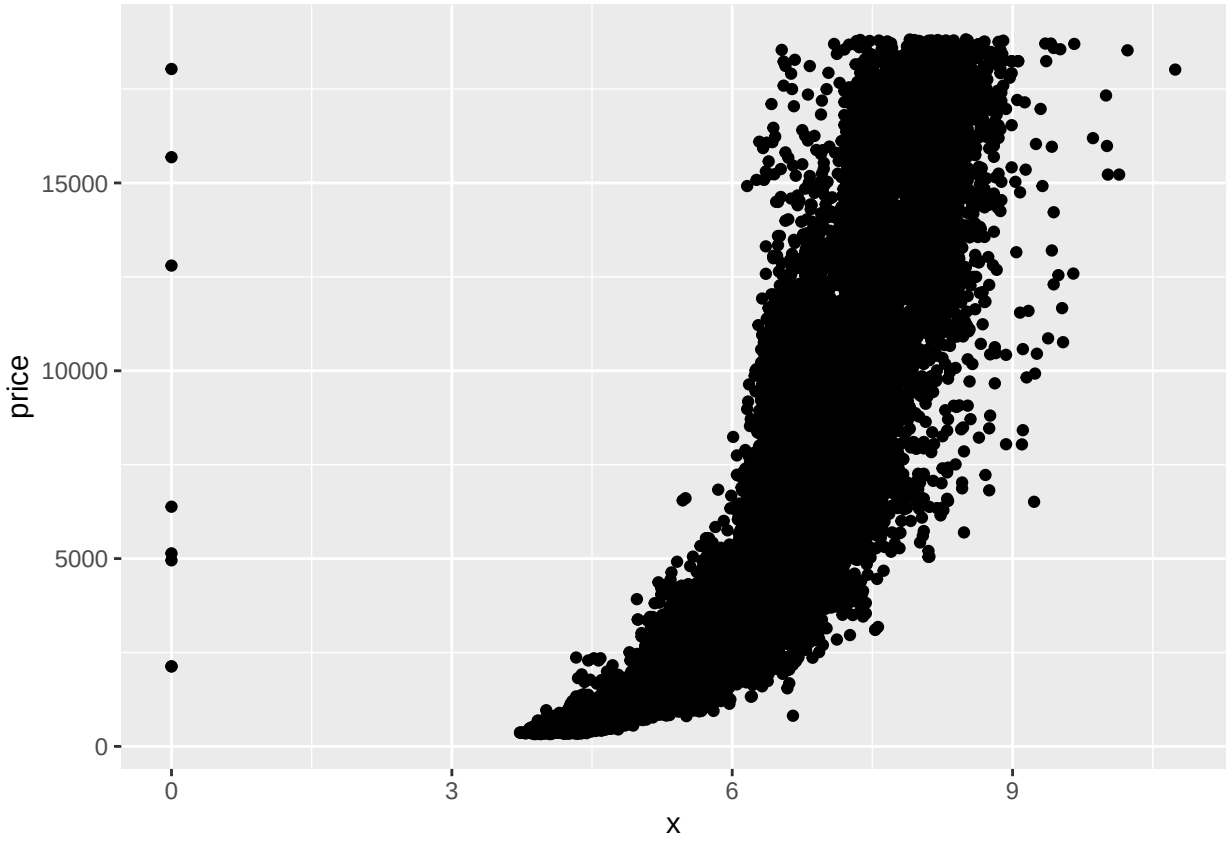
4/1/2020

Quiz 1.price vs. x

```
# In this problem set, you'll continue
# to explore the diamonds data set.

# Your first task is to create a
# scatterplot of price vs x.
# using the ggplot syntax.

# This assignment is not graded and
# will be marked as correct when you submit.

# ENTER YOUR CODE BELOW THIS LINE
# ================================================
library(ggplot2)
data(diamonds)
names(diamonds)
```

```
##  [1] "carat"   "cut"     "color"   "clarity" "depth"   "table"   "price"
##  [8] "x"       "y"       "z"
```

```
ggplot(aes(x = x, y = price), data = diamonds) +
  geom_point()
```

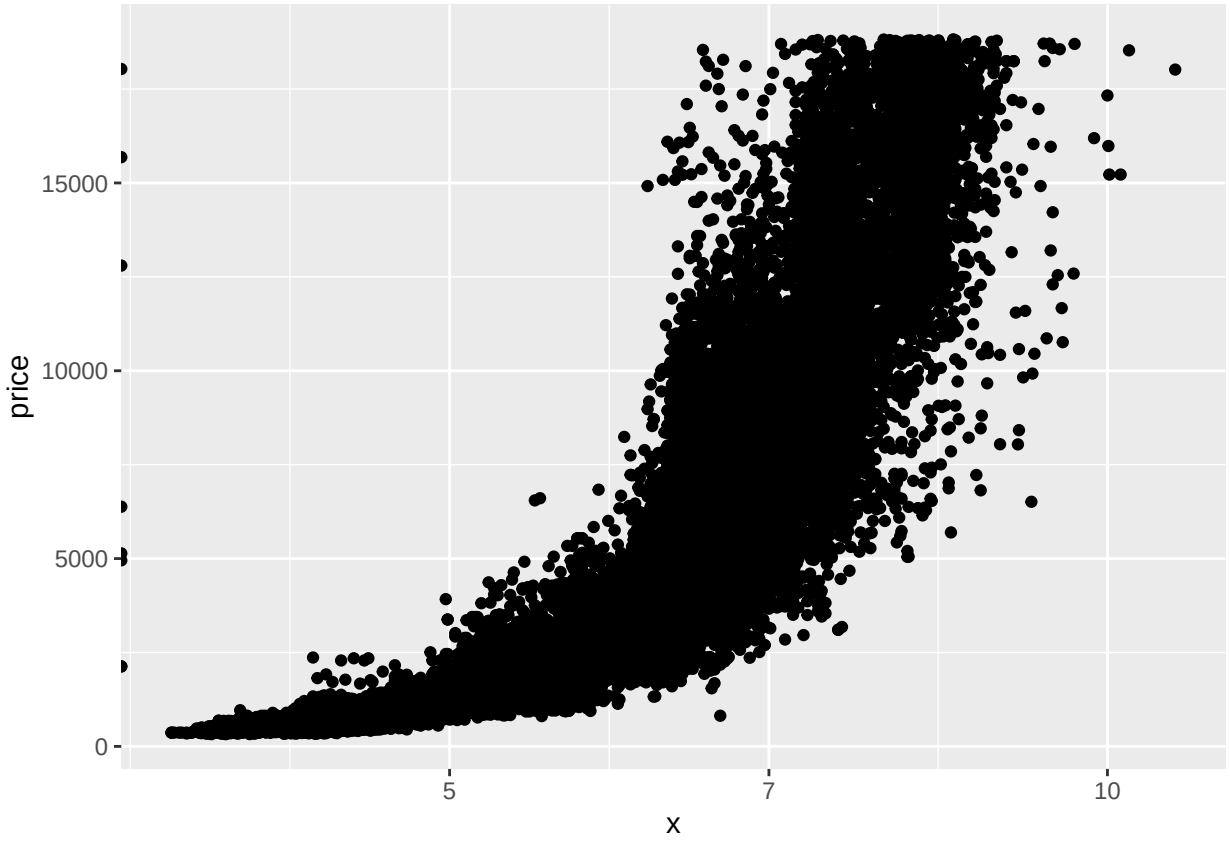



```
ggsave('PriceX1.png')
```

```
## Saving 6.5 x 4.5 in image
```

```
ggplot(aes(x = x, y = price), data = diamonds) +
  geom_point() +
  scale_x_log10() +

ggsave('PriceX2.png')
```

```
## Saving 6.5 x 4.5 in image

## Warning: Transformation introduced infinite values in continuous x-axis

## Warning: Transformation introduced infinite values in continuous x-axis
```



Quiz 2. Findings - price vs. x

What are your observations about the scatterlot of price vs. x? Response: It looks like that there are some outliers and an exponential relationship between price and x.

Quiz 3. Correlations

1.What is the correlation between price and x? Response:0.88 2.What is the correlation between price and y? Response:0.87 3. What is the correlation between price and z? Response:0.86

```
with(diamonds, cor(price, x))
```

```
## [1] 0.8844352
```

```
with(diamonds, cor(price, y))
```

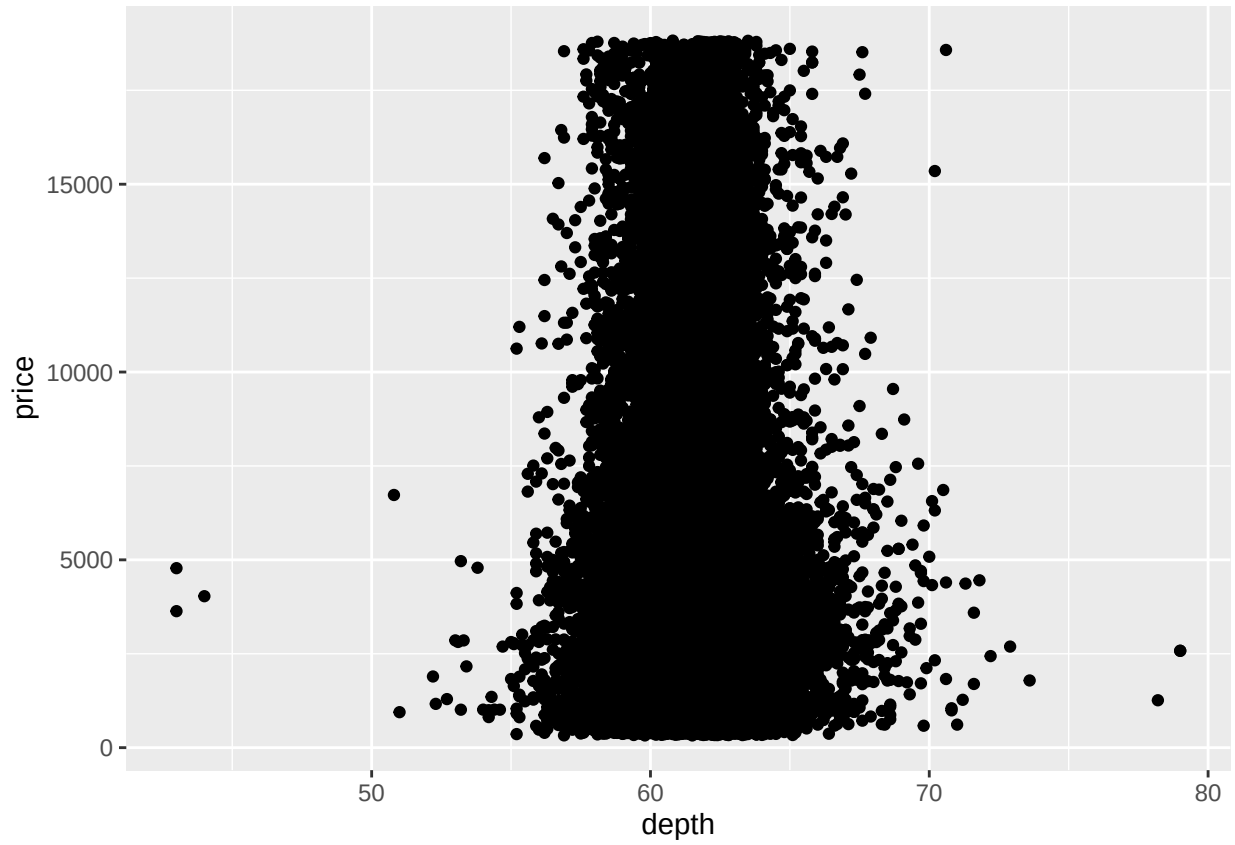
```
## [1] 0.8654209
```

```
with(diamonds, cor(price, z))
```

```
## [1] 0.8612494
```

#Quiz 4. price vs. depth

```
# Create a simple scatter plot of price vs depth.

# This assignment is not graded and
# will be marked as correct when you submit.

# ENTER YOUR CODE BELOW THIS LINE
#================================================
ggplot(aes(x = depth, y = price), data = diamonds) +
  geom_point()
```



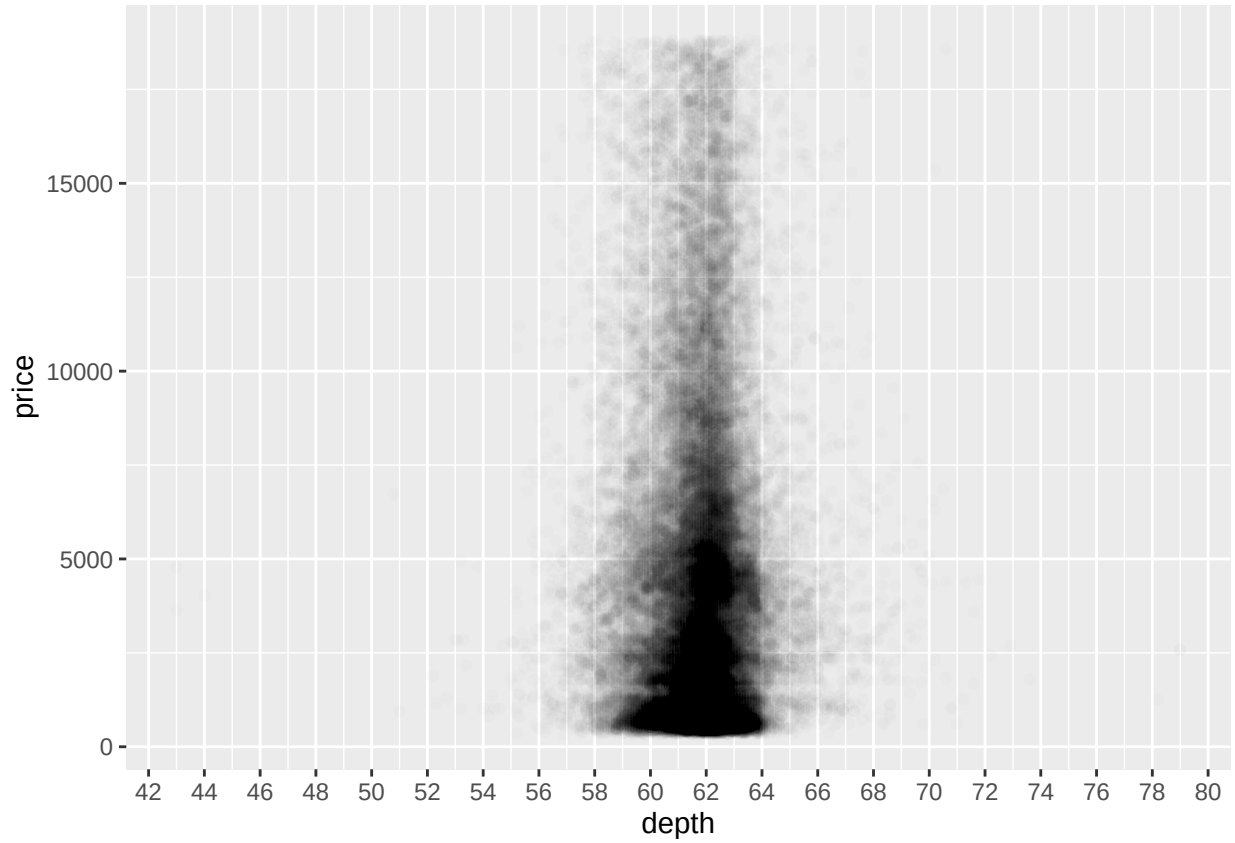

```
ggsave('Pricedepth.png')
```

```
## Saving 6.5 x 4.5 in image
```

#Quiz 5. Adjustments - price vs. depth

```
# Change the code to make the transparency of the
# points to be 1/100 of what they are now and mark
# the x-axis every 2 units. See the instructor notes
# for two hints.

# This assignment is not graded and
```

```
# will be marked as correct when you submit.

# ALTER THE CODE BELOW THIS LINE
#==========================================
ggplot(data = diamonds, aes(x = depth, y = price)) +
  geom_point(alpha = 1/100) +
  scale_x_continuous(breaks = seq(0, 18830, 2))
```




```
ggsave('PricedepthAdj.png')
```

```
## Saving 6.5 x 4.5 in image
```

Quiz 6. Typical Depth Range

Base on the sactterplot of depth vs. price,most diamonds are between what values of depth? Reponse:58-64

```
summary(diamonds$depth)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   43.00   61.00   61.80   61.75   62.50   79.00
```

Quiz 7. Correlation - price and depth

1. What's the correlation between depth vs. price? Response: -0.01
2. Base on the correation coefficient would you use depth to preict the price of diamonds? Response: No
3. Why? Response: R value is around 0.

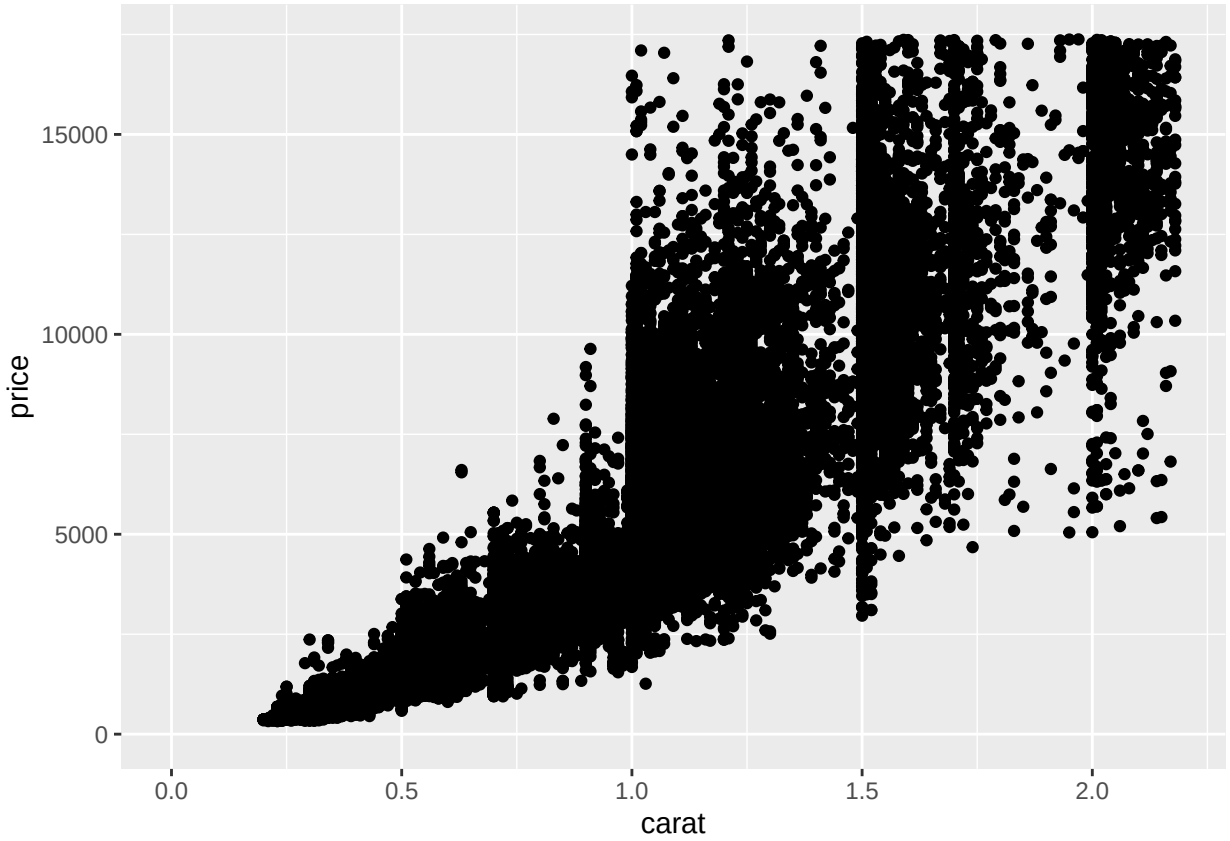
```
with(diamonds, cor.test(depth, price, method = 'pearson'))
```

```
##
##  Pearson's product-moment correlation
##
## data:  depth and price
## t = -2.473, df = 53938, p-value = 0.0134
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.019084756 -0.002208537
## sample estimates:
##        cor
## -0.0106474
```

Quiz 8.price vs. carat

```
# Create a scatterplot of price vs carat
# and omit the top 1% of price and carat
# values.

# This assignment is not graded and
# will be marked as correct when you submit.

# ENTER YOUR CODE BELOW THIS LINE
# ======================================
ggplot(aes(x = carat, y = price), data = diamonds) +
  geom_point() +
  xlim(0, quantile(diamonds$carat, 0.99)) +
  ylim(0, quantile(diamonds$price, 0.99))
```

```
## Warning: Removed 926 rows containing missing values (geom_point).
```




```
ggsave('0.95priceCarat.png')
```

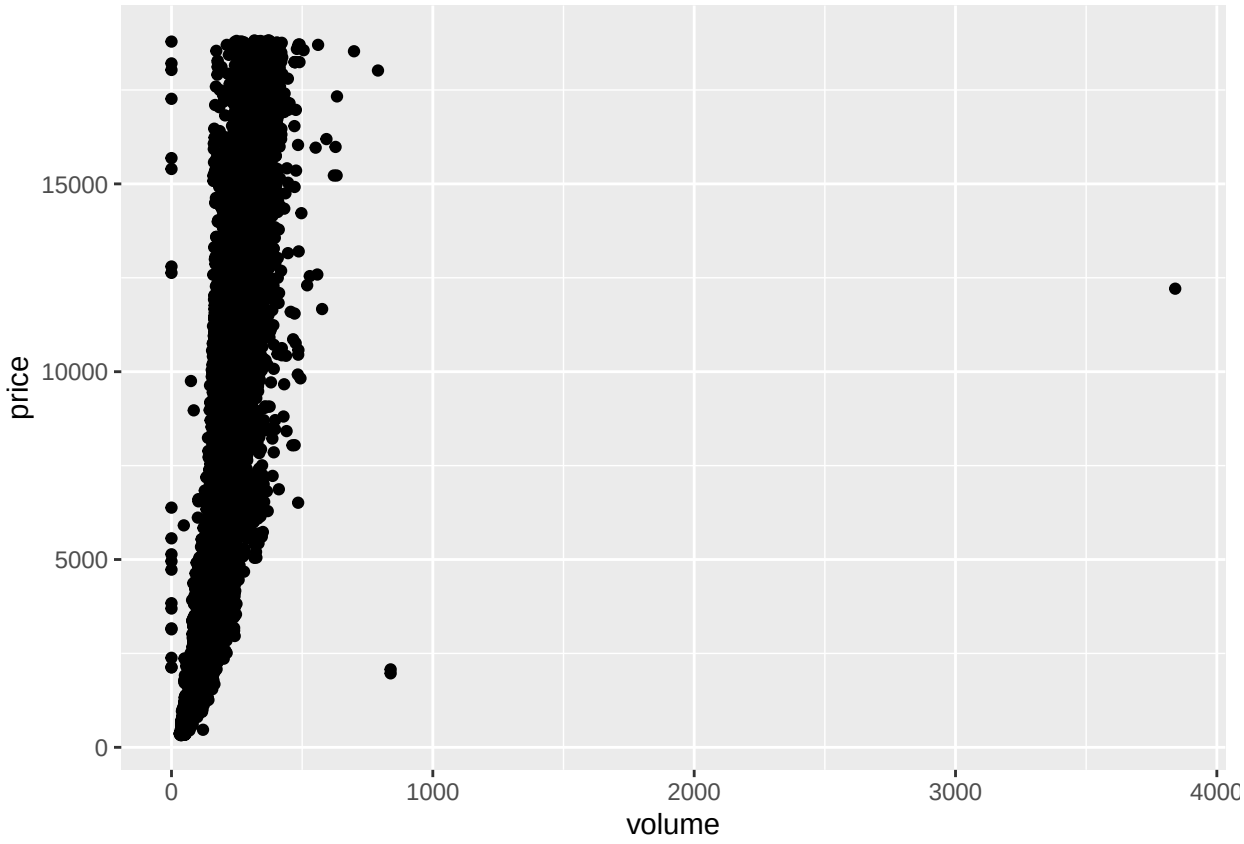
```
## Saving 6.5 x 4.5 in image
```

```
## Warning: Removed 926 rows containing missing values (geom_point).
```

Quiz 9.price vs. volume

```
# Create a scatterplot of price vs. volume (x * y * z).
# This is a very rough approximation for a diamond's volume.

# Create a new variable for volume in the diamonds data frame.
# This will be useful in a later exercise.

# Don't make any adjustments to the plot just yet.

# This assignment is not graded and
# will be marked as correct when you submit.

# ENTER YOUR CODE BELOW THIS LINE
# ================================================================
diamonds$volume = diamonds$x * diamonds$y * diamonds$z
```

```
ggplot(aes(x = volume, y = price), data = diamonds) +
  geom_point()
```

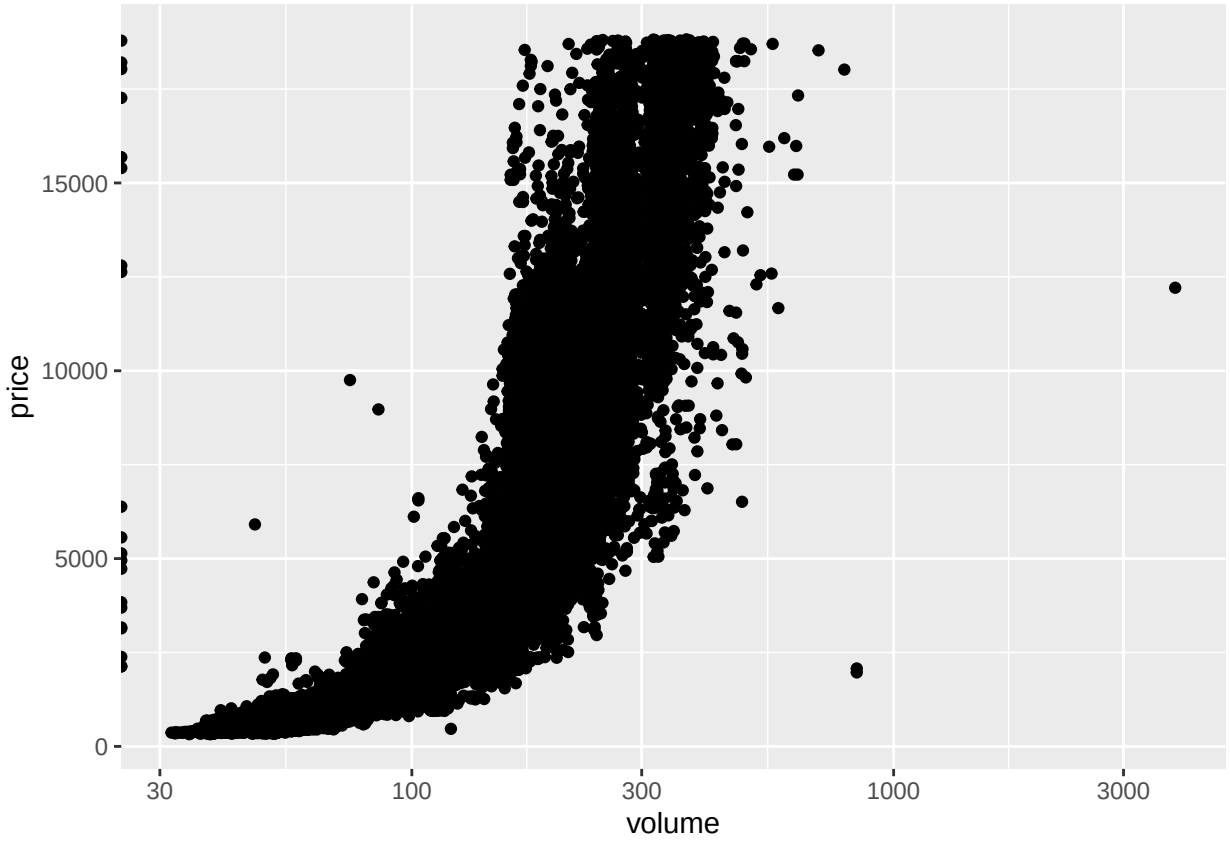



```
ggsave('Pricevolume1.png')
```

```
## Saving 6.5 x 4.5 in image
```

```
ggplot(aes(x = volume, y = price), data = diamonds) +
  geom_point() +
  scale_x_log10()
```

```
## Warning: Transformation introduced infinite values in continuous x-axis
```




```
ggsave('Pricevolume2.png')
```

```
## Saving 6.5 x 4.5 in image
```

```
## Warning: Transformation introduced infinite values in continuous x-axis
```

Quiz 10.Findings - price vs. volume

What are your observations about the scatterlot of price vs. volume scatterplot? Response: It looks like that there are some outliers and an exponential relationship between price and volume.

```
library(plyr)
count(diamonds$volume == 0)
```

```
##       x  freq
## 1 FALSE 53920
## 2  TRUE    20
```

Quiz 11. Correlations on Subsets

What's the correlation of price and volume? Exclude diamonds that have a volume of 0 or that are greater than or equal to 800. Response: 0.9235455

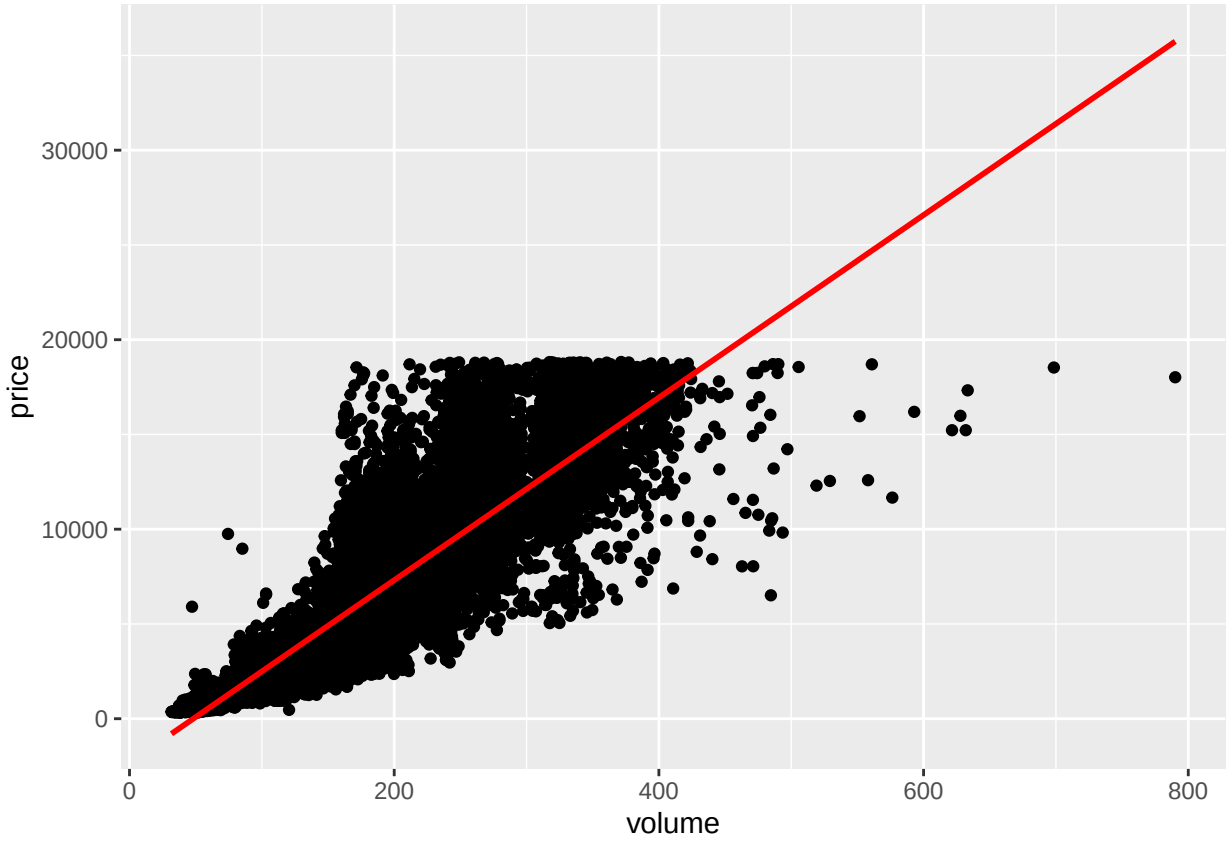
```r
with(subset(diamonds, diamonds$volume < 800 & diamonds$volume > 0),
     cor.test(volume, price))
```

```
##
##  Pearson's product-moment correlation
##
## data:  volume and price
## t = 559.19, df = 53915, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.9222944 0.9247772
## sample estimates:
##       cor
## 0.9235455
```

Quiz 12.Adjustments - price vs. volume

Do you think this would be a useful model to estimate the price of diamonds? Why or why not? Response: Yes. Because cor value is almost 1.

```r
# Subset the data to exclude diamonds with a volume
# greater than or equal to 800. Also, exclude diamonds
# with a volume of 0.

# Adjust the transparency of the points and add a
# linear model to the plot. (See the
# Instructor Notes or look up the documentation of
# geom_smooth() for more details about smoothers.)

# We encourage you to think about this next question and
# to post your thoughts in the discussion section.

# Do you think this would be a useful model to estimate
# the price of diamonds? Why or why not?

# This assignment is not graded and
# will be marked as correct when you submit.

# ENTER YOUR CODE BELOW THIS LINE
# ========================================

ggplot(aes(x = volume, y = price), data = subset(diamonds, diamonds$volume < 800 & diamonds$volume > 0))
  geom_point() +
  stat_smooth(method = "lm", formula = y ~ x, color = 'red',size = 1)
```




```
ggsave('smooth1.png')
```

```
## Saving 6.5 x 4.5 in image
```

#Quiz 13. Mean Price by Clarity

```
# Use the function dplyr package
# to create a new data frame containing
# info on diamonds by clarity.

# Name the data frame diamondsByClarity

# The data frame should contain the following
# variables in this order.

#       (1) mean_price
#       (2) median_price
#       (3) min_price
#       (4) max_price
#       (5) n

# where n is the number of diamonds in each
# level of clarity.

# This assignment WILL BE automatically
```

```r
# graded!

# DO NOT ALTER THE NEXT THREE LINES OF CODE.
# ==========================================================
suppressMessages(library(ggplot2))
suppressMessages(library(dplyr))
#data(diamonds)

# ENTER YOUR CODE BELOW THIS LINE
# ==========================================================
#uninstall package(plyr)
#detach("package:plyr", unload=TRUE)

diamondsByClarity = diamonds %>%
  group_by(clarity) %>%
  summarise(mean_price = mean(price),
            median_price = median(price),
            min_price = min(price),
            max_price = max(price),
            count = n()) %>%
  arrange(clarity)
```

Quiz 14. Bar Charts of Mean Price

```r
# We've created summary data frames with the mean price
# by clarity and color. You can run the code in R to
# verify what data is in the variables diamonds_mp_by_clarity
# and diamonds_mp_by_color.

# Your task is to write additional code to create two bar plots
# on one output image using the grid.arrange() function from the package
# gridExtra.

# This assignment is not graded and
# will be marked as correct when you submit.

# See the Instructor Notes for more info on bar charts
# and for a hint on this task.

# DO NOT DELETE THE LINES OF CODE BELOW
# ================================================================
# data(diamonds)
# library(dplyr)

diamonds_by_clarity <- group_by(diamonds, clarity)
diamonds_mp_by_clarity <- summarise(diamonds_by_clarity, mean_price = mean(price))

diamonds_by_color <- group_by(diamonds, color)
diamonds_mp_by_color <- summarise(diamonds_by_color, mean_price = mean(price))

# ENTER YOUR CODE BELOW THIS LINE
```

```
# ================================================================

library(ggplot2)
head(diamonds_by_clarity)
```

```
## # A tibble: 6 x 11
## # Groups:   clarity [5]
##   carat cut       color clarity depth table price     x     y     z volume
##   <dbl> <ord>     <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>  <dbl>
## 1 0.23  Ideal     E     SI2      61.5    55   326  3.95  3.98  2.43   38.2
## 2 0.21  Premium   E     SI1      59.8    61   326  3.89  3.84  2.31   34.5
## 3 0.23  Good      E     VS1      56.9    65   327  4.05  4.07  2.31   38.1
## 4 0.290 Premium   I     VS2      62.4    58   334  4.2   4.23  2.63   46.7
## 5 0.31  Good      J     SI2      63.3    58   335  4.34  4.35  2.75   51.9
## 6 0.24  Very Good J     VVS2     62.8    57   336  3.94  3.96  2.48   38.7
```

```
head(diamonds_mp_by_clarity)
```

```
## # A tibble: 6 x 2
##   clarity mean_price
##   <ord>        <dbl>
## 1 I1           3924.
## 2 SI2          5063.
## 3 SI1          3996.
## 4 VS2          3925.
## 5 VS1          3839.
## 6 VVS2         3284.
```

```
head(diamonds_by_color)
```

```
## # A tibble: 6 x 11
## # Groups:   color [3]
##   carat cut       color clarity depth table price     x     y     z volume
##   <dbl> <ord>     <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>  <dbl>
## 1 0.23  Ideal     E     SI2      61.5    55   326  3.95  3.98  2.43   38.2
## 2 0.21  Premium   E     SI1      59.8    61   326  3.89  3.84  2.31   34.5
## 3 0.23  Good      E     VS1      56.9    65   327  4.05  4.07  2.31   38.1
## 4 0.290 Premium   I     VS2      62.4    58   334  4.2   4.23  2.63   46.7
## 5 0.31  Good      J     SI2      63.3    58   335  4.34  4.35  2.75   51.9
## 6 0.24  Very Good J     VVS2     62.8    57   336  3.94  3.96  2.48   38.7
```

```
head(diamonds_mp_by_color)
```

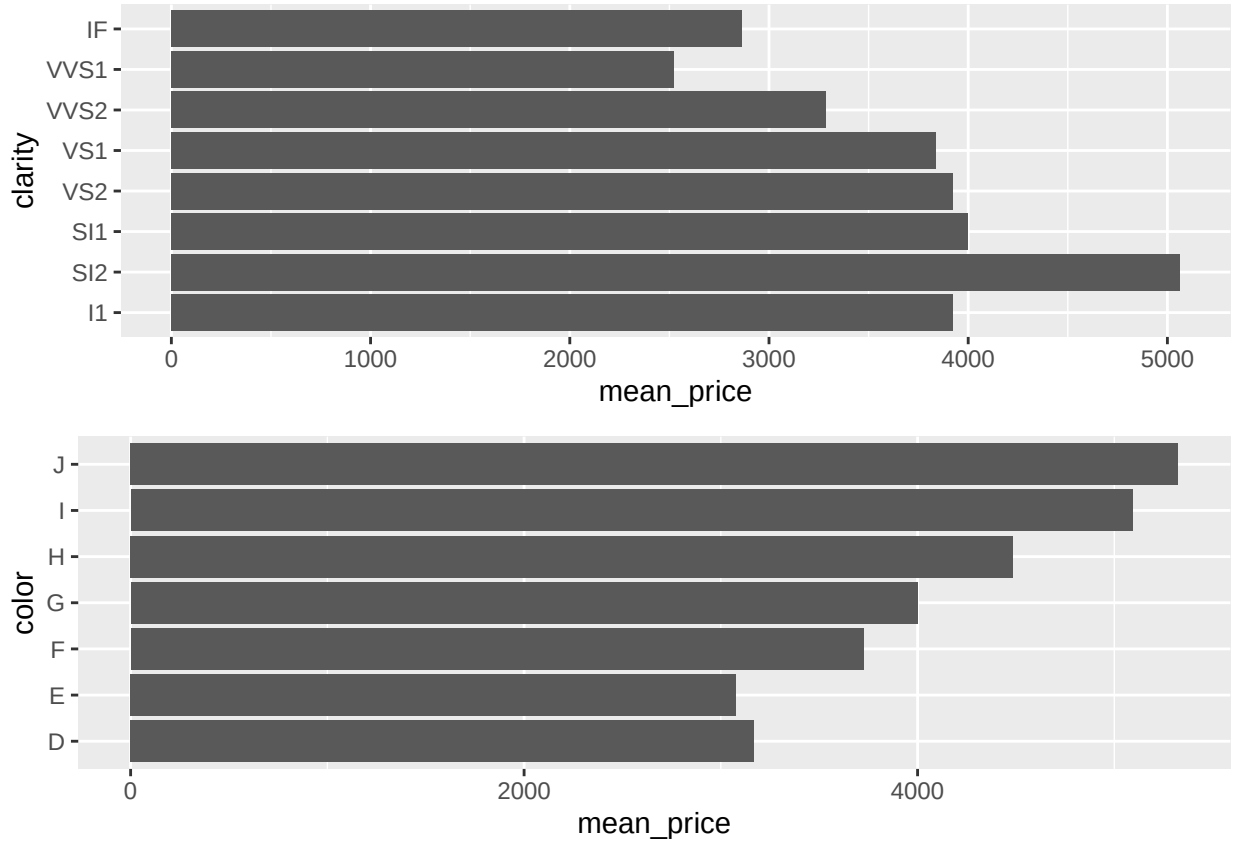
```
## # A tibble: 6 x 2
##   color mean_price
##   <ord>      <dbl>
## 1 D          3170.
## 2 E          3077.
## 3 F          3725.
## 4 G          3999.
## 5 H          4487.
## 6 I          5092.
```

```
pbar1 = ggplot(aes(x = mean_price, y= clarity), data = diamonds_mp_by_clarity) +
  stat_summary(fun = mean, geom = "bar")
pbar2 = ggplot(aes(x = mean_price, y= color), data = diamonds_mp_by_color) +
  stat_summary(fun = mean, geom = "bar")

library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
grid.arrange(pbar1,pbar2, ncol = 1)
```




```
ggsave('bar12.png')
```

```
## Saving 6.5 x 4.5 in image
```

Quiz 15. Trends in Mean Price

What do you notice in each of bar charts for mean price by clarity and mean price by color? Response: 1. According to the difine of clarity, a measurement of how clear the diamond is (I1 (worst), SI2, SI1, VS2,

VS1, VVS2, VVS1, IF (best)).However, mean price tends to decrease as clarity improves.

#Quiz 16.Gapminder Revisited

```
# The Gapminder website contains over 500 data sets with information about
# the world's population. Your task is to continue the investigation you did at the
# end of Problem Set 3 or you can start fresh and choose a different
# data set from Gapminder.

# If you're feeling adventurous or want to try some data munging see if you can
# find a data set or scrape one from the web.

# In your investigation, examine pairs of variable and create 2-5 plots that make
# use of the techniques from Lesson 4.

# You can find a link to the Gapminder website in the Instructor Notes.

# Once you've completed your investigation, create a post in the discussions that includes:
#       1. the variable(s) you investigated, your observations, and any summary statistics
#       2. snippets of code that created the plots
#       3. links to the images of your plots

# Copy and paste all of the code that you used for
# your investigation, and submit it when you are ready.
# ======================================================================
babies = read.csv('children_per_woman_total_fertility.csv')
names(babies)
```

```
##   [1] "country" "X1800"   "X1801"   "X1802"   "X1803"   "X1804"   "X1805"
##   [8] "X1806"   "X1807"   "X1808"   "X1809"   "X1810"   "X1811"   "X1812"
##  [15] "X1813"   "X1814"   "X1815"   "X1816"   "X1817"   "X1818"   "X1819"
##  [22] "X1820"   "X1821"   "X1822"   "X1823"   "X1824"   "X1825"   "X1826"
##  [29] "X1827"   "X1828"   "X1829"   "X1830"   "X1831"   "X1832"   "X1833"
##  [36] "X1834"   "X1835"   "X1836"   "X1837"   "X1838"   "X1839"   "X1840"
##  [43] "X1841"   "X1842"   "X1843"   "X1844"   "X1845"   "X1846"   "X1847"
##  [50] "X1848"   "X1849"   "X1850"   "X1851"   "X1852"   "X1853"   "X1854"
##  [57] "X1855"   "X1856"   "X1857"   "X1858"   "X1859"   "X1860"   "X1861"
##  [64] "X1862"   "X1863"   "X1864"   "X1865"   "X1866"   "X1867"   "X1868"
##  [71] "X1869"   "X1870"   "X1871"   "X1872"   "X1873"   "X1874"   "X1875"
##  [78] "X1876"   "X1877"   "X1878"   "X1879"   "X1880"   "X1881"   "X1882"
##  [85] "X1883"   "X1884"   "X1885"   "X1886"   "X1887"   "X1888"   "X1889"
##  [92] "X1890"   "X1891"   "X1892"   "X1893"   "X1894"   "X1895"   "X1896"
##  [99] "X1897"   "X1898"   "X1899"   "X1900"   "X1901"   "X1902"   "X1903"
## [106] "X1904"   "X1905"   "X1906"   "X1907"   "X1908"   "X1909"   "X1910"
## [113] "X1911"   "X1912"   "X1913"   "X1914"   "X1915"   "X1916"   "X1917"
## [120] "X1918"   "X1919"   "X1920"   "X1921"   "X1922"   "X1923"   "X1924"
## [127] "X1925"   "X1926"   "X1927"   "X1928"   "X1929"   "X1930"   "X1931"
## [134] "X1932"   "X1933"   "X1934"   "X1935"   "X1936"   "X1937"   "X1938"
## [141] "X1939"   "X1940"   "X1941"   "X1942"   "X1943"   "X1944"   "X1945"
## [148] "X1946"   "X1947"   "X1948"   "X1949"   "X1950"   "X1951"   "X1952"
## [155] "X1953"   "X1954"   "X1955"   "X1956"   "X1957"   "X1958"   "X1959"
## [162] "X1960"   "X1961"   "X1962"   "X1963"   "X1964"   "X1965"   "X1966"
## [169] "X1967"   "X1968"   "X1969"   "X1970"   "X1971"   "X1972"   "X1973"
## [176] "X1974"   "X1975"   "X1976"   "X1977"   "X1978"   "X1979"   "X1980"
```

```
## [183] "X1981"   "X1982"   "X1983"   "X1984"   "X1985"   "X1986"   "X1987"
## [190] "X1988"   "X1989"   "X1990"   "X1991"   "X1992"   "X1993"   "X1994"
## [197] "X1995"   "X1996"   "X1997"   "X1998"   "X1999"   "X2000"   "X2001"
## [204] "X2002"   "X2003"   "X2004"   "X2005"   "X2006"   "X2007"   "X2008"
## [211] "X2009"   "X2010"   "X2011"   "X2012"   "X2013"   "X2014"   "X2015"
## [218] "X2016"   "X2017"   "X2018"   "X2019"   "X2020"   "X2021"   "X2022"
## [225] "X2023"   "X2024"   "X2025"   "X2026"   "X2027"   "X2028"   "X2029"
## [232] "X2030"   "X2031"   "X2032"   "X2033"   "X2034"   "X2035"   "X2036"
## [239] "X2037"   "X2038"   "X2039"   "X2040"   "X2041"   "X2042"   "X2043"
## [246] "X2044"   "X2045"   "X2046"   "X2047"   "X2048"   "X2049"   "X2050"
## [253] "X2051"   "X2052"   "X2053"   "X2054"   "X2055"   "X2056"   "X2057"
## [260] "X2058"   "X2059"   "X2060"   "X2061"   "X2062"   "X2063"   "X2064"
## [267] "X2065"   "X2066"   "X2067"   "X2068"   "X2069"   "X2070"   "X2071"
## [274] "X2072"   "X2073"   "X2074"   "X2075"   "X2076"   "X2077"   "X2078"
## [281] "X2079"   "X2080"   "X2081"   "X2082"   "X2083"   "X2084"   "X2085"
## [288] "X2086"   "X2087"   "X2088"   "X2089"   "X2090"   "X2091"   "X2092"
## [295] "X2093"   "X2094"   "X2095"   "X2096"   "X2097"   "X2098"   "X2099"
## [302] "X2100"
```

```
head(babies,8)
```

```
##               country X1800 X1801 X1802 X1803 X1804 X1805 X1806 X1807 X1808
## 1         Afghanistan  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2             Albania  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3             Algeria  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4              Angola  6.93  6.93  6.93  6.93  6.93  6.93  6.93  6.94  6.94
## 5 Antigua and Barbuda  5.00  5.00  4.99  4.99  4.99  4.98  4.98  4.97  4.97
## 6           Argentina  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80
## 7             Armenia  7.80  7.80  7.81  7.81  7.81  7.82  7.82  7.82  7.83
## 8           Australia  6.50  6.48  6.46  6.44  6.42  6.40  6.38  6.36  6.34
##   X1809 X1810 X1811 X1812 X1813 X1814 X1815 X1816 X1817 X1818 X1819 X1820 X1821
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  6.94  6.94  6.94  6.94  6.94  6.94  6.94  6.94  6.94  6.94  6.94  6.94  6.95
## 5  4.97  4.96  4.96  4.96  4.95  4.95  4.94  4.94  4.94  4.93  4.93  4.93  4.92
## 6  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80
## 7  7.83  7.83  7.83  7.84  7.84  7.84  7.85  7.85  7.85  7.86  7.86  7.86  7.87
## 8  6.32  6.30  6.28  6.26  6.24  6.22  6.20  6.18  6.16  6.14  6.12  6.10  6.08
##   X1822 X1823 X1824 X1825 X1826 X1827 X1828 X1829 X1830 X1831 X1832 X1833 X1834
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  6.95  6.95  6.95  6.95  6.95  6.95  6.95  6.95  6.95  6.95  6.95  6.95  6.95
## 5  4.92  4.92  4.91  4.91  4.90  4.90  4.90  4.89  4.89  4.89  4.88  4.88  4.87
## 6  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80
## 7  7.87  7.87  7.88  7.88  7.88  7.89  7.89  7.89  7.89  7.90  7.90  7.90  7.91
## 8  6.06  6.04  6.02  6.00  5.96  5.92  5.87  5.83  5.79  5.75  5.70  5.66  5.62
##   X1835 X1836 X1837 X1838 X1839 X1840 X1841 X1842 X1843 X1844 X1845 X1846 X1847
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  6.96  6.96  6.96  6.96  6.96  6.96  6.96  6.96  6.96  6.96  6.96  6.96  6.96
## 5  4.87  4.87  4.86  4.86  4.86  4.85  4.85  4.84  4.84  4.84  4.83  4.83  4.83
```

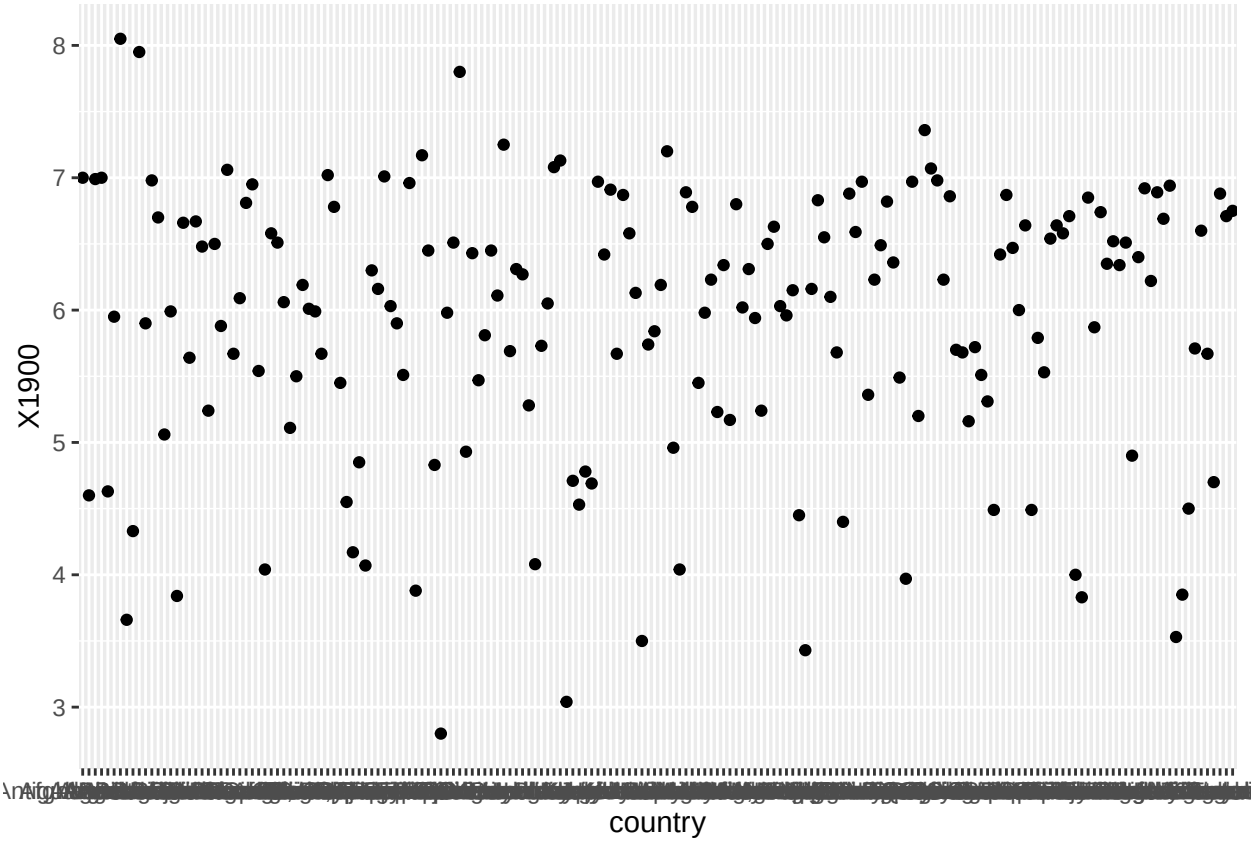
```
## 6  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80
## 7  7.91  7.91  7.92  7.92  7.92  7.93  7.93  7.93  7.94  7.94  7.94  7.95  7.95
## 8  5.58  5.54  5.49  5.45  5.41  5.37  5.32  5.28  5.24  5.20  5.16  5.11  5.07
##    X1848 X1849 X1850 X1851 X1852 X1853 X1854 X1855 X1856 X1857 X1858 X1859 X1860
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  6.96  6.97  6.97  6.97  6.97  6.97  6.97  6.97  6.97  6.97  6.97  6.97  6.97
## 5  4.82  4.82  4.82  4.81  4.81  4.80  4.80  4.80  4.79  4.79  4.79  4.78  4.78
## 6  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80
## 7  7.95  7.95  7.96  7.96  7.96  7.97  7.97  7.97  7.98  7.98  7.98  7.99  7.99
## 8  5.03  4.99  4.94  5.01  4.70  5.03  4.86  5.32  5.19  5.63  5.71  5.75  5.71
##    X1861 X1862 X1863 X1864 X1865 X1866 X1867 X1868 X1869 X1870 X1871 X1872 X1873
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  6.97  6.97  6.98  6.98  6.98  6.98  6.98  6.98  6.98  6.98  6.98  6.98  6.98
## 5  4.77  4.77  4.77  4.76  4.76  4.76  4.75  4.75  4.75  4.74  4.74  4.73  4.73
## 6  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.80  6.73  6.65  6.58  6.58
## 7  7.99  8.00  8.00  8.00  8.01  8.01  8.01  8.01  8.02  8.02  8.02  8.03  8.03
## 8  5.67  5.80  5.59  5.75  5.64  5.33  5.41  5.43  5.19  5.19  5.09  4.97  5.01
##    X1874 X1875 X1876 X1877 X1878 X1879 X1880 X1881 X1882 X1883 X1884 X1885 X1886
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  6.98  6.98  6.98  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 5  4.73  4.72  4.72  4.72  4.71  4.71  4.70  4.70  4.70  4.69  4.69  4.69  4.68
## 6  6.57  6.57  6.57  6.57  6.56  6.56  6.56  6.56  6.55  6.47  6.39  6.30  6.22
## 7  8.03  8.04  8.04  8.04  8.05  8.05  8.05  8.06  8.06  8.06  8.07  8.07  8.07
## 8  4.93  4.81  4.81  4.69  4.74  4.80  4.73  4.73  4.62  4.66  4.77  4.78  4.74
##    X1887 X1888 X1889 X1890 X1891 X1892 X1893 X1894 X1895 X1896 X1897 X1898 X1899
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  6.99  6.99  6.99  6.99  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 5  4.68  4.67  4.67  4.67  4.66  4.66  4.66  4.65  4.65  4.65  4.64  4.64  4.63
## 6  6.14  6.09  6.04  5.99  5.94  5.90  6.26  6.63  7.00  6.48  5.96  5.96  5.95
## 7  8.07  8.08  8.08  8.08  8.09  8.09  8.09  8.10  8.10  8.09  8.08  8.07  8.06
## 8  4.77  4.76  4.65  4.69  4.62  4.52  4.40  4.13  4.07  3.81  3.78  3.64  3.66
##    X1900 X1901 X1902 X1903 X1904 X1905 X1906 X1907 X1908 X1909 X1910 X1911 X1912
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  7.00  7.00  7.00  7.00  7.00  7.01  7.01  7.01  7.01  7.01  7.01  7.01  7.01
## 5  4.63  4.63  4.62  4.62  4.62  4.61  4.61  4.60  4.60  4.60  4.59  4.59  4.59
## 6  5.95  5.94  5.94  5.88  5.82  5.76  5.70  5.64  5.56  5.49  5.41  5.33  5.25
## 7  8.05  8.04  8.03  8.02  8.01  8.00  7.99  7.98  7.97  7.96  7.95  7.94  7.93
## 8  3.66  3.64  3.58  3.39  3.54  3.51  3.35  3.35  3.35  3.35  3.35  3.51  3.51
##    X1913 X1914 X1915 X1916 X1917 X1918 X1919 X1920 X1921 X1922 X1923 X1924 X1925
## 1  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00  7.00
## 2  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60  4.60
## 3  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99  6.99
## 4  7.01  7.01  7.01  7.01  7.01  7.01  7.02  7.02  7.02  7.02  7.02  7.02  7.02
## 5  4.58  4.58  4.58  4.57  4.57  4.56  4.56  4.56  4.55  4.55  4.55  4.54  4.54
```

```
## 6  5.28  5.30  5.23  5.15  5.08  5.01  4.94  4.86  4.79  4.72  4.65  4.57  4.50
## 7  7.92  7.91  7.90  7.89  7.88  7.87  7.86  7.85  7.84  7.83  7.82  7.81  7.80
## 8  3.51  3.51  3.51  3.07  3.07  3.07  3.07  3.07  3.12  3.11  3.02  2.97  2.95
##    X1926 X1927 X1928 X1929 X1930 X1931 X1932 X1933 X1934 X1935 X1936 X1937 X1938
## 1  7.03  7.05  7.08  7.11  7.14  7.16  7.19  7.21  7.24  7.26  7.29  7.31  7.33
## 2  4.42  4.25  4.07  3.89  3.72  3.54  3.36  3.57  3.78  3.98  4.19  4.40  4.61
## 3  7.02  7.05  7.07  7.10  7.12  7.15  7.18  7.20  7.22  7.25  7.27  7.29  7.31
## 4  7.02  7.01  7.01  7.00  7.00  7.00  7.00  6.99  6.99  6.99  7.00  7.00  7.00
## 5  4.53  4.53  4.53  4.52  4.52  4.52  4.51  4.51  4.51  4.50  4.50  4.49  4.49
## 6  4.43  4.36  4.26  4.17  4.07  3.98  3.89  3.80  3.71  3.62  3.53  3.44  3.44
## 7  7.65  7.50  7.34  7.19  7.04  6.89  6.74  6.58  6.43  6.28  6.13  5.98  5.82
## 8  2.85  2.80  2.77  2.64  2.58  2.36  2.19  2.17  2.11  2.12  2.18  2.21  2.21
##    X1939 X1940 X1941 X1942 X1943 X1944 X1945 X1946 X1947 X1948 X1949 X1950 X1951
## 1  7.36  7.38  7.40  7.42  7.44  7.46  7.48  7.50  7.52  7.54  7.56  7.57  7.56
## 2  3.75  4.21  3.77  4.43  4.61  4.79  4.96  5.14  5.33  5.51  5.69  5.87  5.97
## 3  7.33  7.35  7.37  7.38  7.40  7.41  7.43  7.44  7.45  7.47  7.48  7.49  7.47
## 4  7.00  7.01  7.02  7.02  7.03  7.04  7.05  7.06  7.07  7.08  7.10  7.11  7.13
## 5  4.49  4.48  4.48  4.48  4.47  4.47  4.46  4.46  4.46  4.45  4.45  4.45  4.46
## 6  3.43  3.43  3.42  3.42  3.37  3.33  3.29  3.24  3.20  3.19  3.18  3.16  3.16
## 7  5.67  5.52  4.56  3.34  1.82  1.74  2.39  3.58  4.10  3.63  4.05  4.53  4.51
## 8  2.22  2.26  2.36  2.38  2.57  2.63  2.74  2.99  3.08  2.99  2.99  3.07  3.06
##    X1952 X1953 X1954 X1955 X1956 X1957 X1958 X1959 X1960 X1961 X1962 X1963 X1964
## 1  7.55  7.54  7.53  7.52  7.51  7.49  7.48  7.46  7.45  7.45  7.45  7.45  7.45
## 2  6.13  6.27  6.37  6.45  6.50  6.53  6.54  6.53  6.49  6.40  6.28  6.13  5.96
## 3  7.44  7.43  7.41  7.41  7.41  7.43  7.45  7.48  7.52  7.57  7.61  7.65  7.67
## 4  7.15  7.17  7.20  7.23  7.27  7.31  7.36  7.42  7.48  7.52  7.56  7.59  7.61
## 5  4.50  4.52  4.53  4.53  4.53  4.51  4.49  4.46  4.43  4.39  4.34  4.30  4.25
## 6  3.16  3.15  3.15  3.14  3.14  3.13  3.12  3.12  3.11  3.10  3.09  3.08  3.07
## 7  4.49  4.48  4.48  4.49  4.51  4.53  4.55  4.56  4.55  4.51  4.44  4.32  4.16
## 8  3.18  3.19  3.20  3.28  3.33  3.42  3.42  3.44  3.44  3.52  3.41  3.32  3.16
##    X1965 X1966 X1967 X1968 X1969 X1970 X1971 X1972 X1973 X1974 X1975 X1976 X1977
## 1  7.45  7.45  7.45  7.45  7.45  7.45  7.45  7.45  7.45  7.45  7.45  7.45  7.45
## 2  5.77  5.58  5.39  5.22  5.06  4.91  4.78  4.64  4.51  4.37  4.24  4.10  3.97
## 3  7.68  7.68  7.67  7.67  7.66  7.64  7.62  7.60  7.56  7.51  7.43  7.34  7.23
## 4  7.62  7.62  7.61  7.61  7.60  7.60  7.60  7.61  7.61  7.61  7.62  7.61  7.59
## 5  4.19  4.13  4.04  3.94  3.82  3.68  3.52  3.33  3.12  2.91  2.70  2.51  2.36
## 6  3.06  3.05  3.05  3.05  3.06  3.07  3.10  3.15  3.20  3.26  3.32  3.37  3.39
## 7  3.98  3.79  3.61  3.45  3.31  3.21  3.12  3.03  2.94  2.85  2.75  2.65  2.56
## 8  3.00  2.92  2.87  2.88  2.86  2.81  2.85  2.68  2.48  2.33  2.18  2.09  2.03
##    X1978 X1979 X1980 X1981 X1982 X1983 X1984 X1985 X1986 X1987 X1988 X1989 X1990
## 1  7.45  7.45  7.45  7.45  7.45  7.45  7.46  7.46  7.46  7.46  7.46  7.46  7.47
## 2  3.84  3.73  3.62  3.53  3.45  3.38  3.32  3.27  3.22  3.16  3.11  3.05  2.98
## 3  7.11  6.96  6.79  6.62  6.43  6.24  6.04  5.84  5.63  5.41  5.19  4.96  4.73
## 4  7.57  7.54  7.50  7.47  7.44  7.41  7.39  7.38  7.37  7.35  7.32  7.29  7.25
## 5  2.24  2.16  2.12  2.10  2.10  2.10  2.10  2.09  2.08  2.07  2.06  2.06  2.06
## 6  3.39  3.37  3.33  3.27  3.21  3.16  3.12  3.09  3.07  3.06  3.04  3.02  2.99
## 7  2.48  2.42  2.39  2.37  2.38  2.41  2.44  2.48  2.52  2.55  2.57  2.57  2.54
## 8  1.97  1.92  1.90  1.92  1.91  1.91  1.86  1.90  1.87  1.86  1.85  1.85  1.88
##    X1991 X1992 X1993 X1994 X1995 X1996 X1997 X1998 X1999 X2000 X2001 X2002 X2003
## 1  7.48  7.50  7.54  7.57  7.61  7.63  7.64  7.62  7.57  7.49  7.39  7.27  7.14
## 2  2.91  2.83  2.75  2.67  2.59  2.51  2.42  2.33  2.25  2.16  2.07  1.98  1.90
## 3  4.48  4.22  3.96  3.70  3.45  3.21  2.99  2.80  2.64  2.51  2.44  2.40  2.41
## 4  7.19  7.13  7.06  6.99  6.92  6.85  6.79  6.73  6.68  6.64  6.60  6.57  6.54
## 5  2.07  2.10  2.13  2.17  2.21  2.25  2.28  2.30  2.32  2.32  2.31  2.29  2.27
```

```
## 6  2.95  2.91  2.87  2.82  2.76  2.71  2.66  2.62  2.59  2.56  2.54  2.52  2.50
## 7  2.48  2.40  2.30  2.18  1.80  1.80  1.70  1.50  1.40  1.30  1.20  1.20  1.40
## 8  1.86  1.87  1.85  1.84  1.82  1.81  1.79  1.77  1.76  1.76  1.76  1.78  1.79
##    X2004 X2005 X2006 X2007 X2008 X2009 X2010 X2011 X2012 X2013 X2014 X2015 X2016
## 1  6.99  6.83  6.65  6.46  6.25  6.04  5.82  5.60  5.38  5.17  4.98  4.80  4.64
## 2  1.82  1.75  1.70  1.67  1.65  1.65  1.65  1.67  1.69  1.70  1.71  1.71  1.71
## 3  2.44  2.50  2.58  2.66  2.75  2.83  2.89  2.93  2.94  2.92  2.89  2.84  2.78
## 4  6.50  6.47  6.42  6.37  6.31  6.24  6.16  6.08  6.00  5.92  5.84  5.77  5.69
## 5  2.25  2.22  2.20  2.18  2.16  2.15  2.13  2.12  2.10  2.09  2.08  2.06  2.05
## 6  2.47  2.45  2.43  2.41  2.40  2.38  2.37  2.36  2.35  2.34  2.32  2.31  2.29
## 7  1.40  1.40  1.30  1.40  1.40  1.60  1.55  1.50  1.73  1.71  1.69  1.66  1.63
## 8  1.82  1.86  1.89  1.92  1.94  1.94  1.93  1.92  1.90  1.89  1.87  1.86  1.85
##    X2017 X2018 X2019 X2020 X2021 X2022 X2023 X2024 X2025 X2026 X2027 X2028 X2029
## 1  4.48  4.33  4.18  4.04  3.90  3.77  3.65  3.53  3.43  3.32  3.23  3.14  3.06
## 2  1.71  1.71  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70
## 3  2.71  2.64  2.59  2.54  2.49  2.46  2.42  2.39  2.36  2.33  2.30  2.28  2.25
## 4  5.62  5.55  5.48  5.41  5.34  5.28  5.21  5.14  5.07  5.01  4.94  4.88  4.81
## 5  2.04  2.03  2.02  2.01  2.00  1.99  1.98  1.97  1.96  1.95  1.95  1.94  1.93
## 6  2.28  2.26  2.25  2.23  2.22  2.20  2.19  2.18  2.16  2.15  2.14  2.12  2.11
## 7  1.60  1.60  1.59  1.59  1.58  1.58  1.58  1.58  1.58  1.58  1.59  1.59  1.59
## 8  1.84  1.83  1.82  1.81  1.81  1.80  1.80  1.80  1.79  1.79  1.79  1.78  1.78
##    X2030 X2031 X2032 X2033 X2034 X2035 X2036 X2037 X2038 X2039 X2040 X2041 X2042
## 1  2.98  2.91  2.85  2.79  2.73  2.68  2.63  2.58  2.53  2.49  2.45  2.41  2.37
## 2  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70  1.70
## 3  2.23  2.21  2.19  2.17  2.15  2.13  2.11  2.10  2.08  2.07  2.06  2.04  2.03
## 4  4.75  4.69  4.63  4.57  4.51  4.45  4.39  4.34  4.28  4.23  4.17  4.12  4.07
## 5  1.92  1.92  1.91  1.90  1.89  1.89  1.88  1.88  1.87  1.87  1.86  1.86  1.85
## 6  2.10  2.09  2.07  2.06  2.05  2.04  2.03  2.02  2.01  2.00  1.99  1.98  1.97
## 7  1.60  1.60  1.60  1.61  1.61  1.61  1.62  1.62  1.62  1.63  1.63  1.64  1.64
## 8  1.78  1.77  1.77  1.77  1.77  1.77  1.77  1.77  1.76  1.76  1.76  1.76  1.76
##    X2043 X2044 X2045 X2046 X2047 X2048 X2049 X2050 X2051 X2052 X2053 X2054 X2055
## 1  2.34  2.30  2.27  2.24  2.21  2.18  2.15  2.13  2.10  2.08  2.05  2.03  2.01
## 2  1.71  1.71  1.71  1.71  1.71  1.71  1.71  1.72  1.72  1.72  1.72  1.72  1.73
## 3  2.02  2.01  2.00  1.99  1.98  1.97  1.97  1.96  1.95  1.95  1.94  1.93  1.93
## 4  4.02  3.97  3.92  3.87  3.82  3.77  3.73  3.69  3.64  3.60  3.56  3.52  3.48
## 5  1.85  1.85  1.84  1.84  1.84  1.84  1.83  1.83  1.83  1.83  1.83  1.82  1.82
## 6  1.96  1.96  1.95  1.94  1.94  1.93  1.92  1.92  1.91  1.90  1.90  1.89  1.89
## 7  1.64  1.65  1.65  1.65  1.66  1.66  1.67  1.67  1.67  1.68  1.68  1.68  1.69
## 8  1.76  1.76  1.76  1.76  1.76  1.76  1.76  1.76  1.76  1.76  1.76  1.76  1.76
##    X2056 X2057 X2058 X2059 X2060 X2061 X2062 X2063 X2064 X2065 X2066 X2067 X2068
## 1  1.99  1.97  1.95  1.94  1.92  1.91  1.89  1.88  1.87  1.85  1.84  1.83  1.82
## 2  1.73  1.73  1.73  1.73  1.73  1.74  1.74  1.74  1.74  1.74  1.75  1.75  1.75
## 3  1.92  1.92  1.91  1.91  1.91  1.90  1.90  1.90  1.89  1.89  1.89  1.89  1.88
## 4  3.44  3.40  3.37  3.33  3.30  3.26  3.23  3.20  3.17  3.14  3.11  3.08  3.05
## 5  1.82  1.82  1.82  1.82  1.82  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81
## 6  1.88  1.88  1.88  1.87  1.87  1.87  1.86  1.86  1.86  1.85  1.85  1.85  1.85
## 7  1.69  1.69  1.69  1.70  1.70  1.70  1.71  1.71  1.71  1.72  1.72  1.72  1.73
## 8  1.77  1.77  1.77  1.77  1.77  1.77  1.77  1.77  1.77  1.77  1.78  1.78  1.78
##    X2069 X2070 X2071 X2072 X2073 X2074 X2075 X2076 X2077 X2078 X2079 X2080 X2081
## 1  1.81  1.81  1.80  1.79  1.79  1.78  1.77  1.77  1.76  1.76  1.76  1.75  1.75
## 2  1.75  1.75  1.75  1.75  1.76  1.76  1.76  1.76  1.76  1.76  1.77  1.77  1.77
## 3  1.88  1.88  1.88  1.88  1.87  1.87  1.87  1.87  1.87  1.87  1.86  1.86  1.86
## 4  3.02  2.99  2.97  2.94  2.91  2.89  2.86  2.84  2.82  2.79  2.77  2.75  2.73
## 5  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81
```

```
## 6  1.85  1.84  1.84  1.84  1.84  1.84  1.84  1.84  1.83  1.83  1.83  1.83  1.83
## 7  1.73  1.73  1.73  1.74  1.74  1.74  1.74  1.74  1.75  1.75  1.75  1.75  1.75
## 8  1.78  1.78  1.78  1.78  1.78  1.78  1.78  1.79  1.79  1.79  1.79  1.79  1.79
##   X2082 X2083 X2084 X2085 X2086 X2087 X2088 X2089 X2090 X2091 X2092 X2093 X2094
## 1  1.75  1.75  1.75  1.75  1.74  1.74  1.74  1.74  1.74  1.74  1.74  1.74  1.74
## 2  1.77  1.77  1.77  1.77  1.78  1.78  1.78  1.78  1.78  1.78  1.78  1.78  1.79
## 3  1.86  1.86  1.86  1.86  1.86  1.86  1.86  1.86  1.86  1.86  1.86  1.86  1.86
## 4  2.71  2.69  2.67  2.65  2.63  2.61  2.59  2.57  2.55  2.54  2.52  2.50  2.48
## 5  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81  1.81
## 6  1.83  1.83  1.83  1.83  1.83  1.82  1.82  1.82  1.82  1.82  1.82  1.82  1.82
## 7  1.76  1.76  1.76  1.76  1.76  1.77  1.77  1.77  1.77  1.77  1.77  1.77  1.77
## 8  1.79  1.79  1.79  1.79  1.79  1.79  1.80  1.80  1.80  1.80  1.80  1.80  1.80
##   X2095 X2096 X2097 X2098 X2099 X2100
## 1  1.74  1.74  1.74  1.74  1.74  1.74
## 2  1.79  1.79  1.79  1.79  1.79  1.79
## 3  1.86  1.86  1.86  1.86  1.86  1.86
## 4  2.47  2.45  2.43  2.42  2.40  2.40
## 5  1.81  1.81  1.81  1.82  1.82  1.82
## 6  1.82  1.82  1.82  1.82  1.82  1.82
## 7  1.78  1.78  1.78  1.78  1.78  1.78
## 8  1.80  1.80  1.80  1.80  1.81  1.81
```

```
library(ggplot2)
ggplot(aes(x = country, y = X1900), data = babies) +
  geom_point()
```



```
ggsave('point1.png')
```

```
## Saving 6.5 x 4.5 in image
```

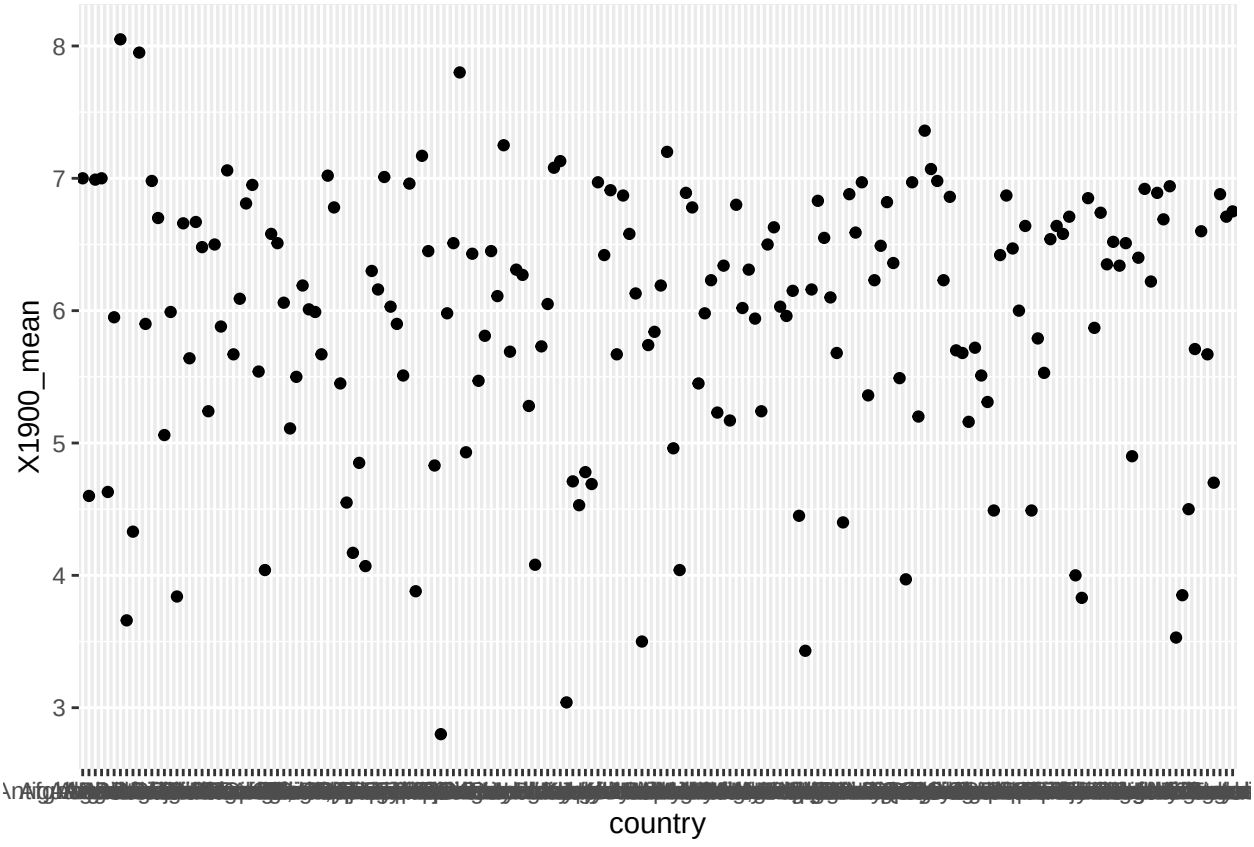
```
summary(babies$X1800)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   4.040   5.670   6.190   6.111   6.718   8.100
```

```
range(babies$X1900)
```

```
## [1] 2.80 8.05
```

```
library(dplyr)
babies.X1900_by_country = babies %>%
  group_by(country) %>%
  summarise(X1900_mean = mean(X1900),
            X1900_median = median(X1900),
            n = n()) %>%
  arrange(country)

ggplot(aes(x = country, y = X1900_mean), data = babies.X1900_by_country) +
  geom_point()
```

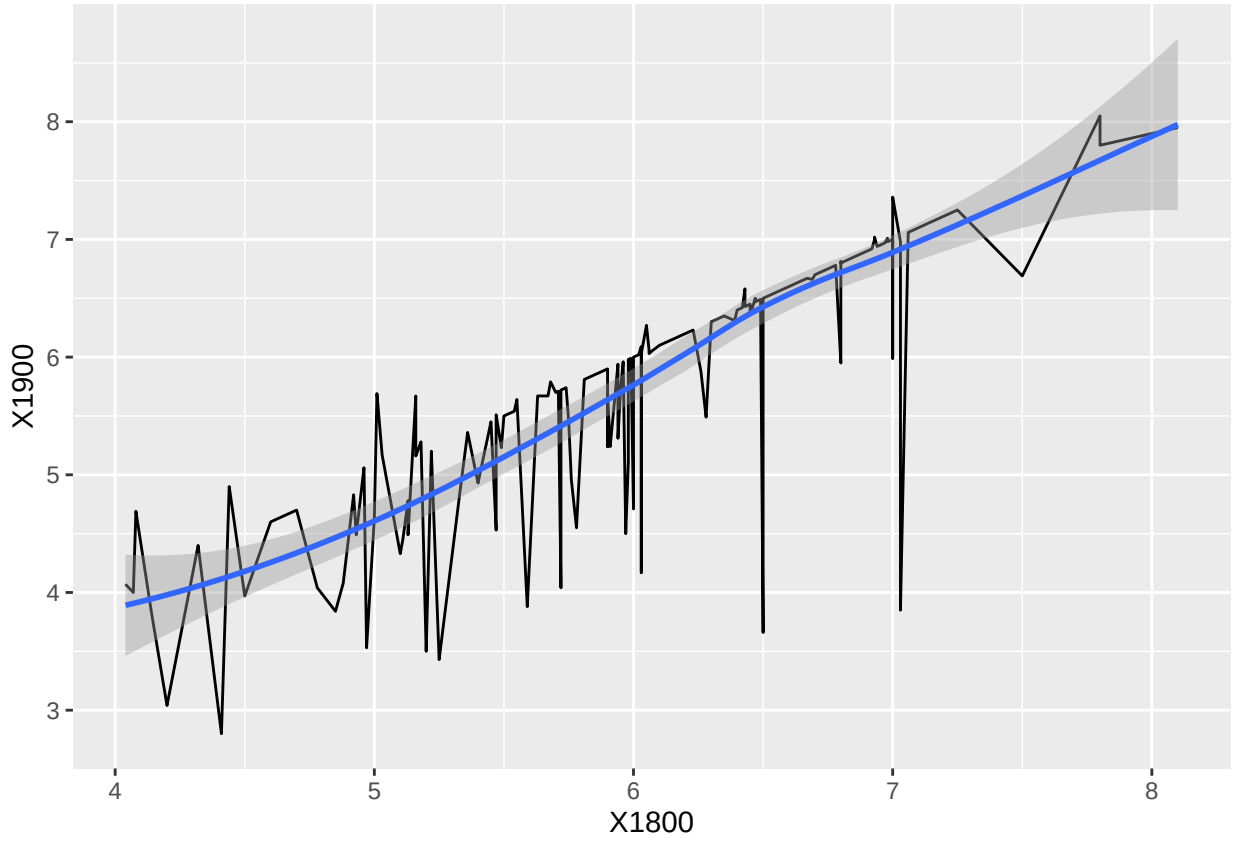


```
ggsave('point2.png')
```

```
## Saving 6.5 x 4.5 in image
```

```
ggplot(aes(x = X1800, y = X1900), data = babies) +
  geom_line() +
  geom_smooth()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
ggsave('line.png')
```

```
## Saving 6.5 x 4.5 in image
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
with(babies, cor.test(X1800, X1900))
```

```
##
##  Pearson's product-moment correlation
##
## data:  X1800 and X1900
```

```
## t = 22.64, df = 182, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.8157684 0.8927681
## sample estimates:
##       cor
## 0.8590527
```

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.