Visualization of vitamin C content per dollar on various charts in R

```
pkgs <- sort(c('tidyverse', 'ggplot2')) #creates a vector of packages that are not already installed in the system
pkgs_install <- pkgs[!(pkgs %in% installed.packages()[,"Package"])] #filters out packages and creates a new list of packages that need to be installed
if(length(pkgs_install)) install.packages(pkgs_install) #packages are installed before running any code

library(tidyverse)  #Loading tidyverse package to make available for use in R
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.1.2     ✔ readr     2.1.4
## ✔ forcats   1.0.0     ✔ stringr   1.5.0
## ✔ ggplot2   3.4.2     ✔ tibble    3.2.1
## ✔ lubridate 1.9.2     ✔ tidyr     1.3.0
## ✔ purrr     1.0.1
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
fruit <- tibble(    #creates a tibble named "fruit"
  name = c("apple", "banana", "orange"), #sets up a vector variable called "name" with three string values
  price = c(2.5, 2.0, 3.5), #assigns a vector of prices for three items
  vitamin_c = c(20, 45, 250)) #initializes a vector called "vitamin_c" with three numeric values; 20, 45, and 250
fruit #creating a tibble called "fruit" with columns for name, price, and vitamin C content for apples, bananas, and oranges
```

```
## # A tibble: 3 × 3
##   name   price vitamin_c
##   <chr>  <dbl>     <dbl>
## 1 apple    2.5        20
## 2 banana   2          45
## 3 orange   3.5       250
```

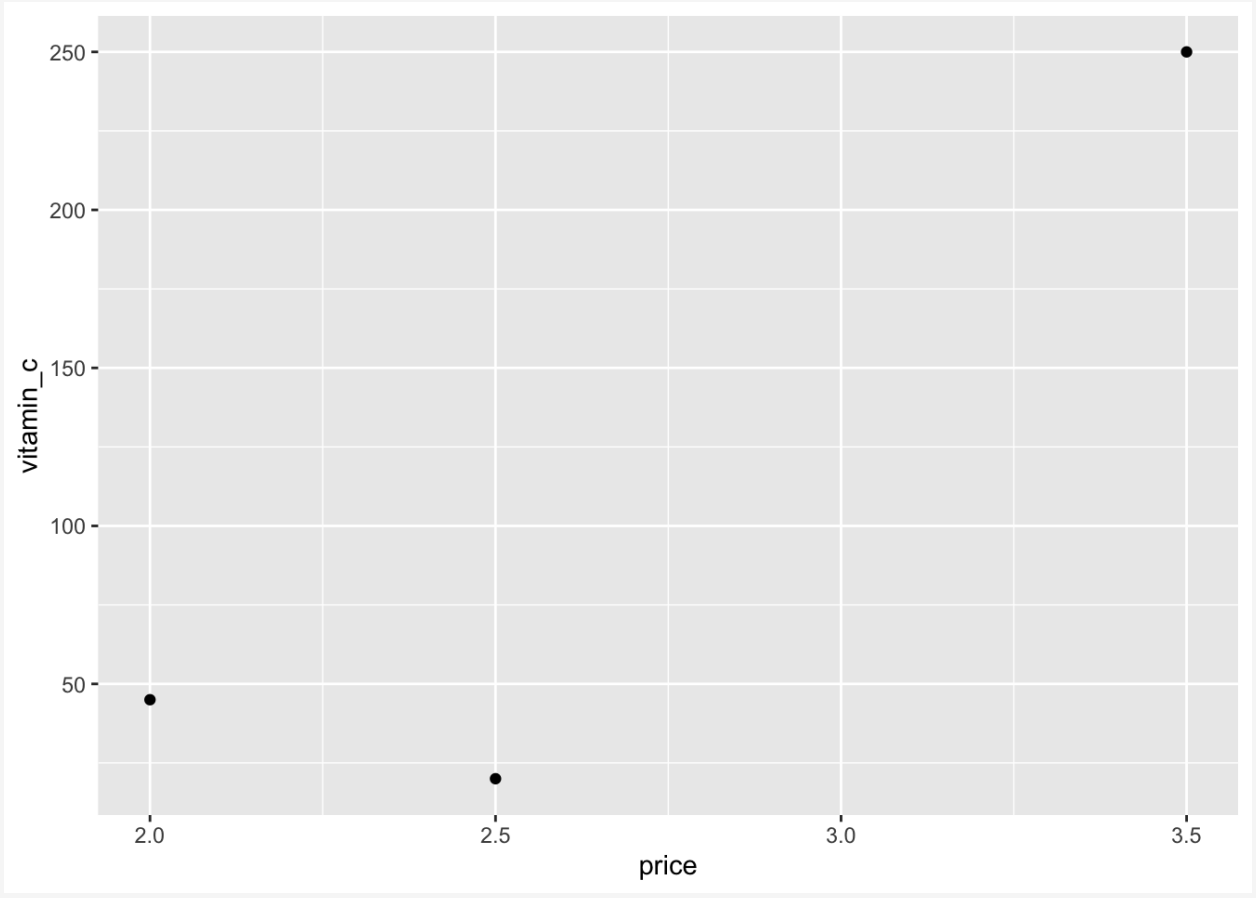
```
affordable_vitamin_c_sources <- fruit %>%  #filters a data frame called "fruit" to find sources of vitamin C that are affordable
  mutate(vitamin_c_per_dollar = vitamin_c / price) %>%  #adds a new column to a data frame, represents the amount of vitamin C per dollar of price
  filter(vitamin_c_per_dollar > 20) %>%   #filters vitamin C
  arrange(desc(vitamin_c_per_dollar))   #arranges the dataset in decending order
affordable_vitamin_c_sources #creating a new table of fruits with a vitamin C to price ratio greater than 20, sorted in descending order of the ratio.
```

```
## # A tibble: 2 × 4
##   name   price vitamin_c vitamin_c_per_dollar
##   <chr>  <dbl>     <dbl>                <dbl>
```

```
## 1 orange   3.5       250              71.4
## 2 banana   2         45               22.5
```

Part1

```
ggplot(fruit, aes(x = price, y = vitamin_c)) +  #sets up a ggplot object using the "fruit" dataset and maps the "price" variable to the x-axis and the "vitamin_c" variable to the y-axis
geom_point()  #creating a scatter plot of the relationship between price and vitamin C content of a fruit dataset using ggplot
```

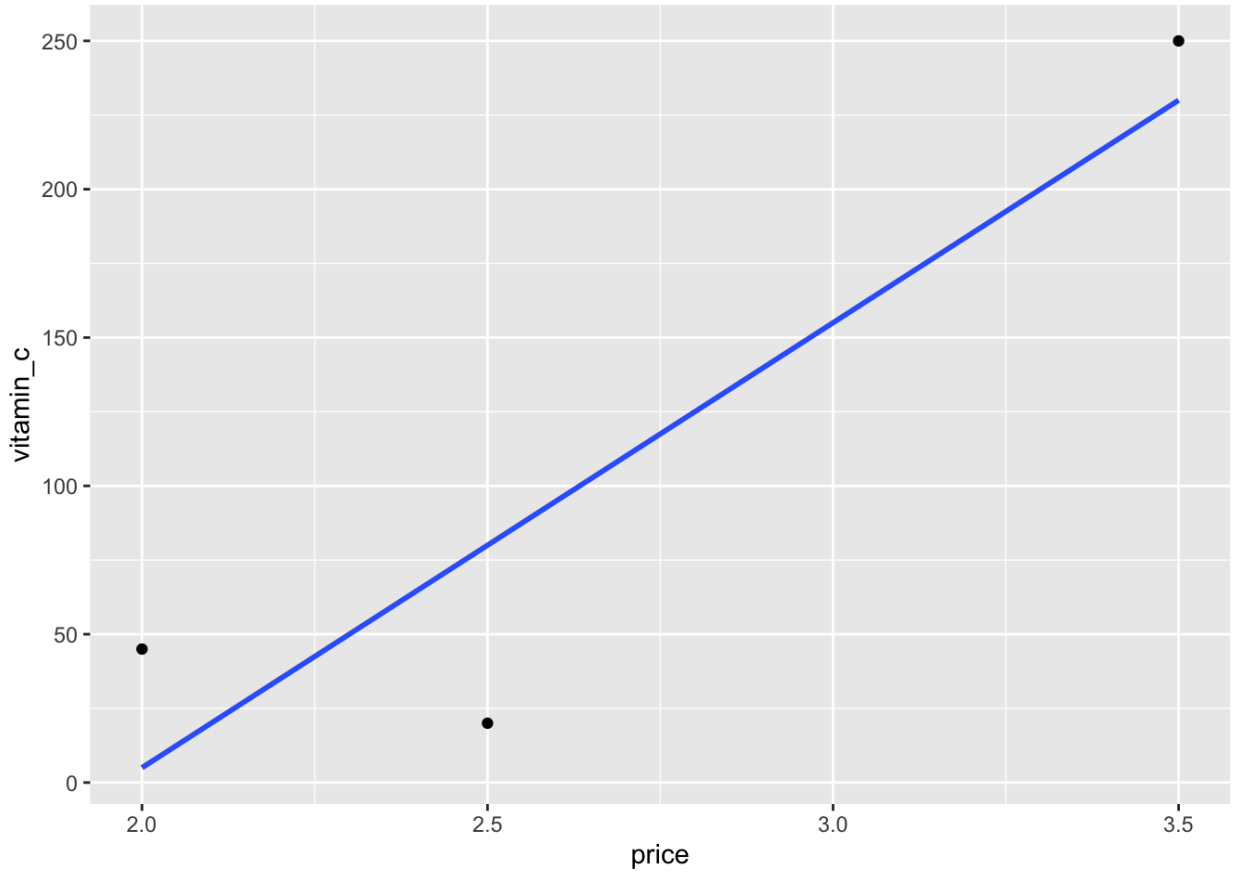


The above scatterplot is visualizing the relationship between the price and vitamin C content for 3 different fruits. Each point on the chart represents one fruit, with the x-axis showing the price of the fruit and the y-axis showing the vitamin C content. Additionally, the chart can be used to compare the vitamin C content of the different fruits, as well as to identify any unusual data points.

Part2

```
ggplot(fruit, aes(x = price, y = vitamin_c)) +  #sets up a ggplot object using the "fruit" dataset and maps the "price" variable to the x-axis and the "vitamin_c" variable to the y-axis
  geom_point() +  #adds a layer of points to the ggplot object
  geom_smooth(method = "lm", se = FALSE) #creating a scatter plot with a linear regression line between price and vitamin_c values for a dataset called "fruit" using ggplot in R
```

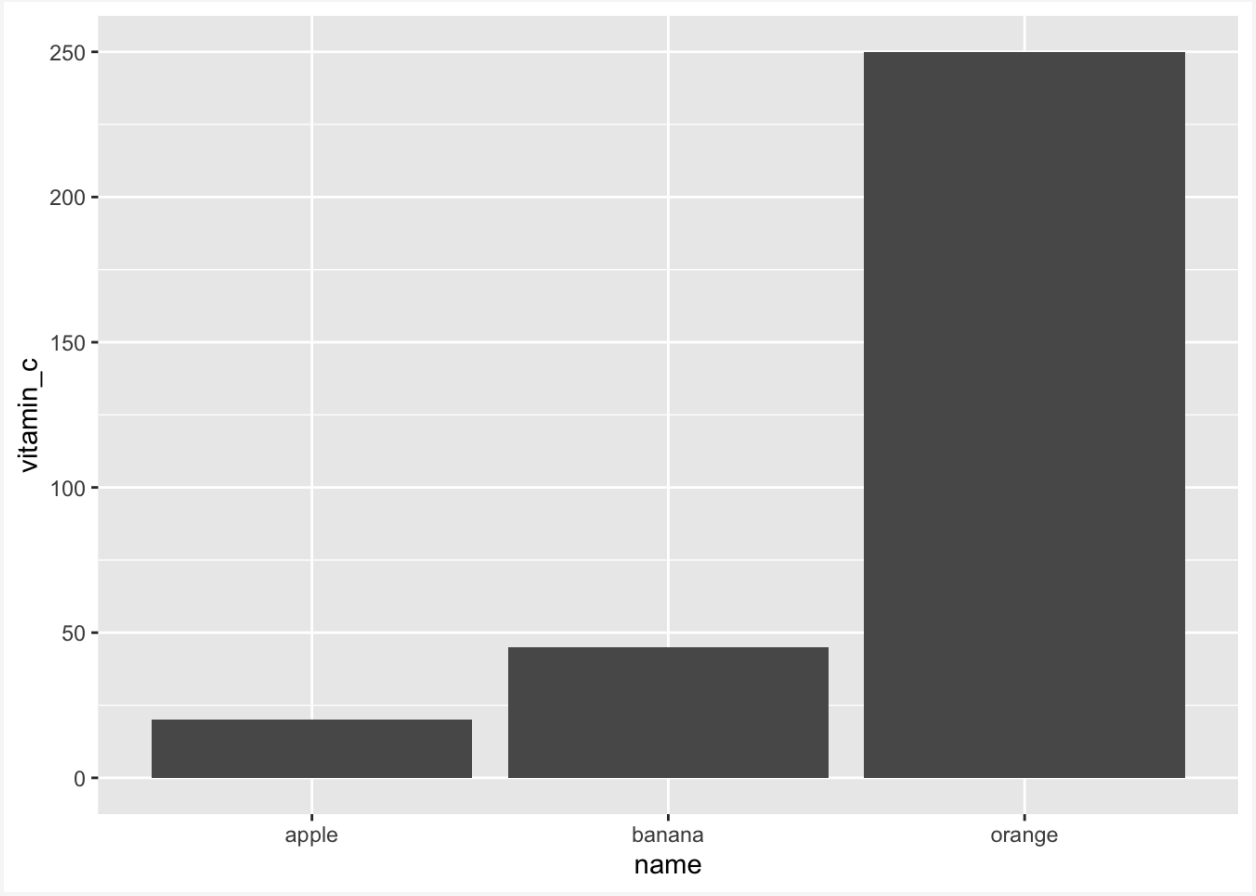
`geom_smooth()` using formula = 'y ~ x'



Adding a regression line using geom_smooth with the method “lm” linear model. The trend line, drawn using linear regression, suggests a positive correlation between the two variables, as price increases, so does vitamin_c content. However, there appears to be a fair amount of variability in the data points, indicating that other factors may also influence vitamin_c levels in fruit.

Part3

```
ggplot(fruit, aes(x = name, y = vitamin_c)) +  #sets up a ggplot object using the "fruit" dataset and maps the "price" variable to the x-axis and the "vitamin_c" variable to the y-axis
  geom_bar(stat = "identity") #generating a bar plot using the "name" variable as x-axis and "vitamin_c" variable as y-axis
```



Visualizing the vitamin C content per dollar of each fruit using a bar chart. The bar chart shows the vitamin C content per dollar of each fruit. The height of each bar represents the amount of vitamin C for each fruit, and the length of the bars gives a visual comparison of the fruit's value for money the longer the bar, the higher the amount of vitamin C per dollar spent. The chart allows the comparison of different fruits' values and their vitamin C content, and the viewer can quickly identify which fruits provide the most vitamin C per dollar spent.

#THE END