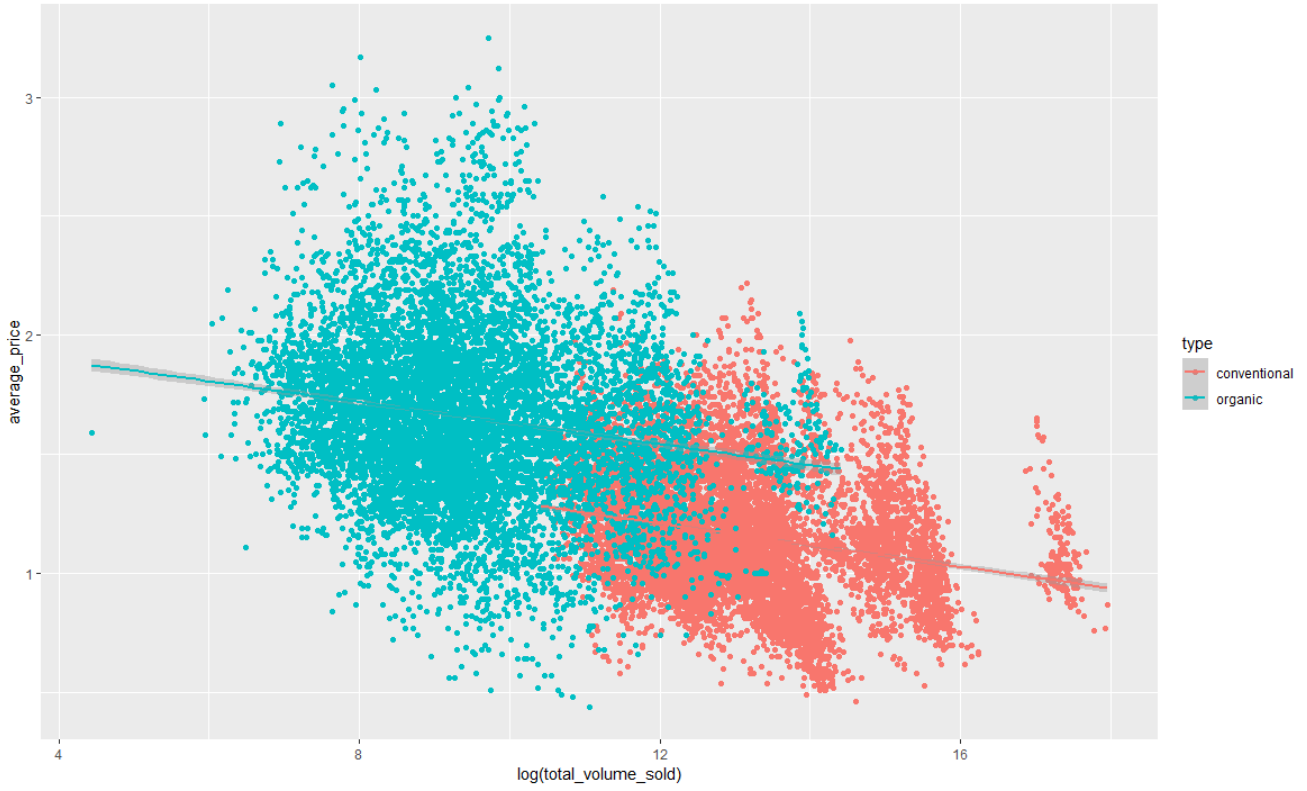
Exercise 1:

```
> #Exercise 1:
> library(ggplot2)
> library(dplyr)
> setwd("C:/Users/PC/OneDrive - VietNam National University - HCM INTERNAT
IONAL UNIVERSITY/Desktop/DA/Lab/LAB6")
> avocado_data <- read.csv("avocado.csv")
> average_price = avocado_data$AveragePrice
> total_volume_sold = avocado_data$Total.Volume
> avocado_data %>%
+   ggplot(
+     aes(
+       x =  average_price,
+       y = log(total_volume_sold),
+       # use a different color for each type of avocado
+       color = type
+     )
+   ) +
+   geom_point() +
+   geom_smooth(method = "lm")
`geom_smooth()` using formula = 'y ~ x'
```

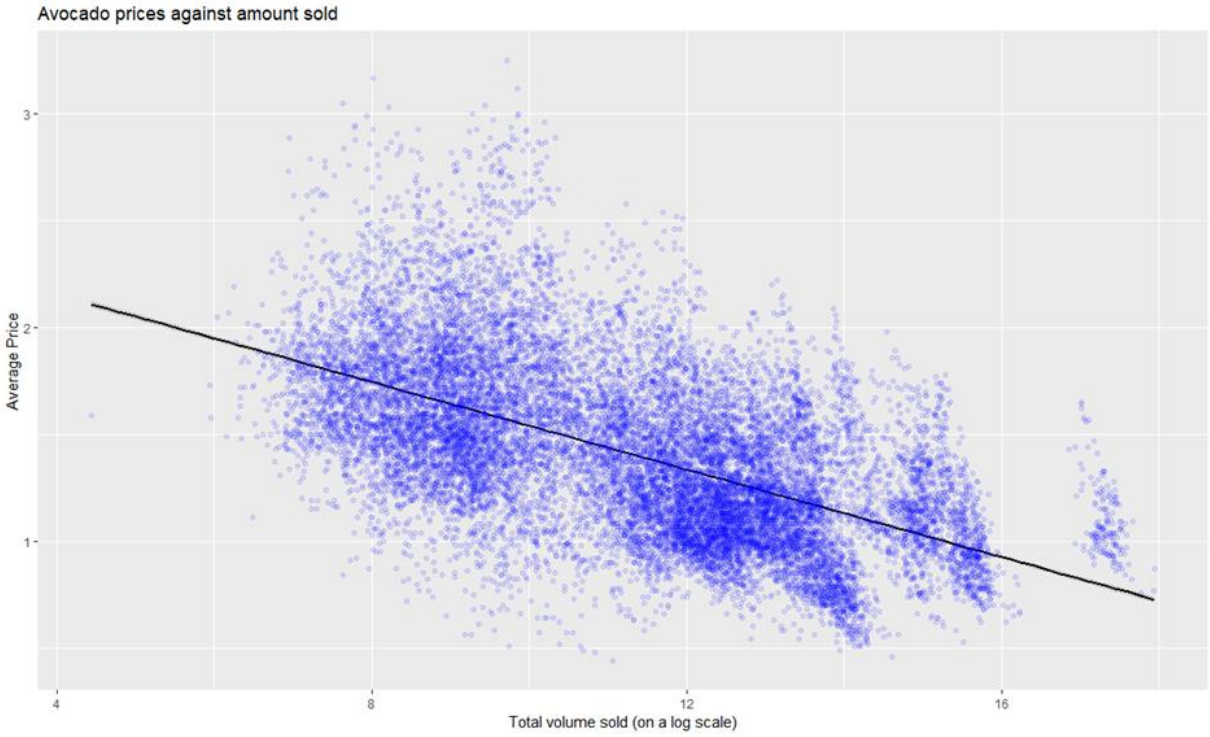
Exercise 2:

```
> #Exercise 2:
> avocado_data %>%
+   ggplot(
+     aes(
+       x =  log(total_volume_sold),
+       y = average_price,
+       # use a different color for each type of avocado
+       color = type
+     )
+   ) +
+   geom_point() +
+   geom_smooth(method = "lm")
`geom_smooth()` using formula = 'y ~ x'
```



Exercise 3:

```
> #Exercise 3:
> avocado_data %>%
+   ggplot(
+     mapping = aes(
+       # notice that we use the log (try without it to understand why)
+       x = log(total_volume_sold),
+       y = average_price
+     )
+   ) +
+   geom_smooth(method = "lm", color = "black") +
+   geom_point(alpha = 0.1, color = "blue") +
+   labs(
+     x = "Total volume sold (on a log scale)",
+     y = "Average Price"
+   )+
+   ggtitle("Avocado prices against amount sold")
`geom_smooth()` using formula = 'y ~ x'
```



Exercise 4:

```
> #Exercise 4:
> avocado_data %>%
+   ggplot(
+     aes(
+       x = log(total_volume_sold),
+       y = average_price,
+     )
+   ) +
+   geom_point(alpha = 0.1, color = "blue") +
+   geom_smooth(
+     # tell the smoother to deal with avocados types separately
+     aes(group = type, linetype = type),
+     method = "lm",
+     color = "black"
+   )+
+   labs(
+     x = "Total volume sold (on a log scale)",
+     y = "Average Price"
+   )+
+   ggtitle("Avocado prices against amount sold")
`geom_smooth()` using formula = 'y ~ x'
```