Diamantes (Práctica AFI)

Gerardo Alcalá

## Análisis descriptivo: Dataset de diamantes.

El Dataset de diamantes se encuentra disponible en (<http://ggplot2.tidyverse.org/reference/diamonds.html>)[<http://ggplot2.tidyverse.org/reference/diamonds.html>] y está compuesto por 53,940 diamantes de los cuales se dispone de los siguientes atributos:

#### Lista de atributos del dataset.

* **Variables continuas:**
  + price
  + carat
  + depth
  + table
  + x
  + y
  + z
* **Variables categóricas:**
  + cut (ordinal)
  + color (ordinal)
  + clarity (ordinal)

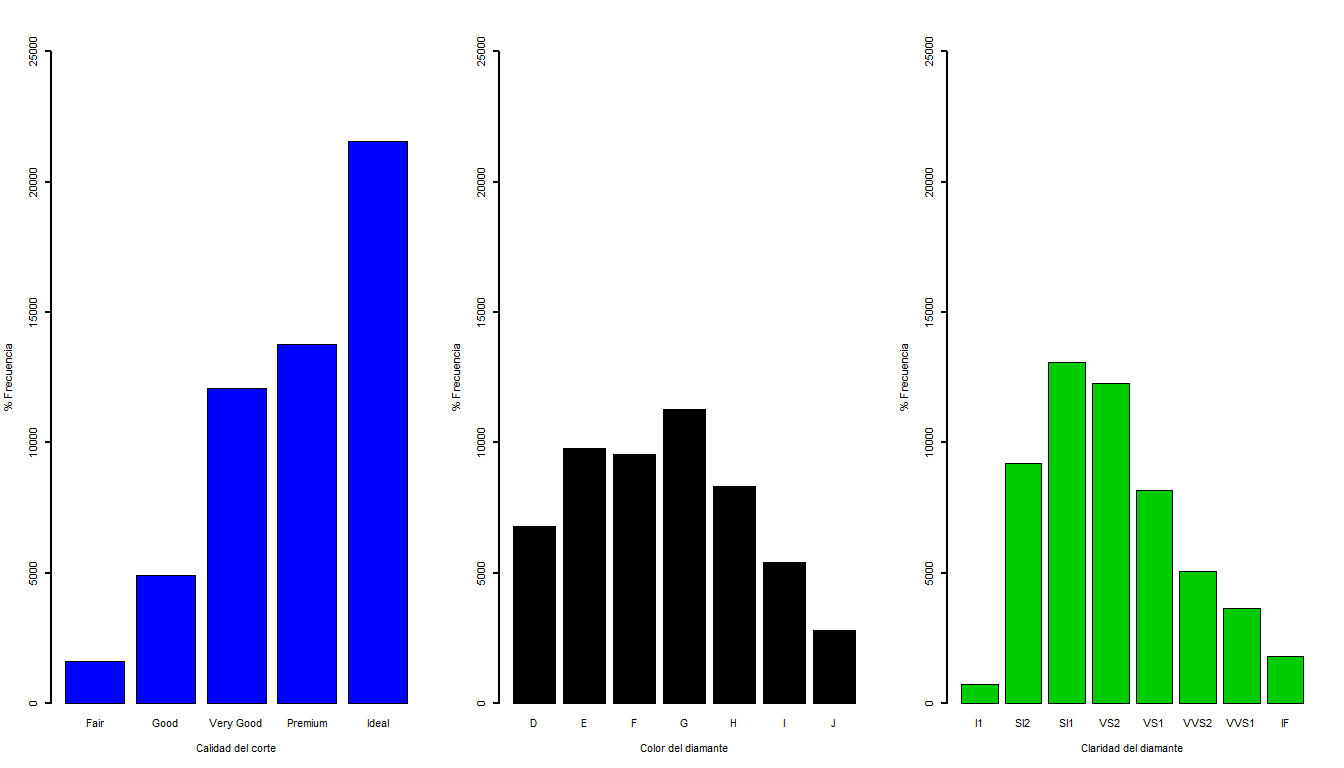
### Análisis exploratorio de los datos.

Las variables categóricas de la base de datos son ordinales y sólo son 3. A continuación presentamos las tablas de frecuencias:

|  |  |  |
| --- | --- | --- |
| cut | Frecuencia | Relativa |
| Fair | 1610 | 2.98 |
| Good | 4906 | 9.10 |
| Very Good | 12082 | 22.40 |
| Premium | 13791 | 25.57 |
| Ideal | 21551 | 39.95 |

|  |  |  |
| --- | --- | --- |
| color | Frecuencia | Relativa |
| D | 6775 | 12.56 |
| E | 9797 | 18.16 |
| F | 9542 | 17.69 |
| G | 11292 | 20.93 |
| H | 8304 | 15.39 |
| I | 5422 | 10.05 |
| J | 2808 | 5.21 |

|  |  |  |
| --- | --- | --- |
| clarity | Frecuencia | Relativa |
| I1 | 741 | 1.37 |
| SI2 | 9194 | 17.04 |
| SI1 | 13065 | 24.22 |
| VS2 | 12258 | 22.73 |
| VS1 | 8171 | 15.15 |
| VVS2 | 5066 | 9.39 |
| VVS1 | 3655 | 6.78 |
| IF | 1790 | 3.32 |



## 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.