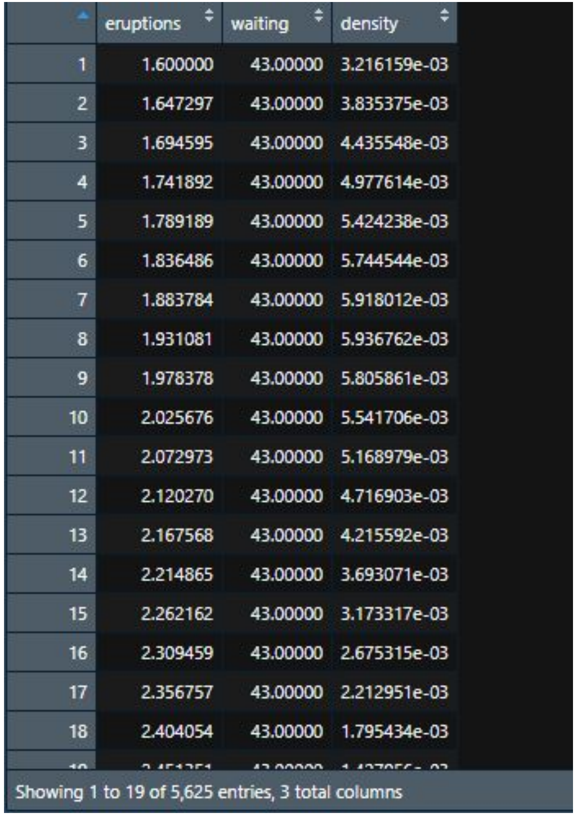
# 2.1 Skill Check

Geom_Raster, Geom_Tile, Geom_Contour

# Our dataset for geom_raster() and geom_contour()

Faithfuld, a dataset on the waiting time between eruptions and the duration of the eruption of the Old Faithful Geyser.

Continuous, numerical data types



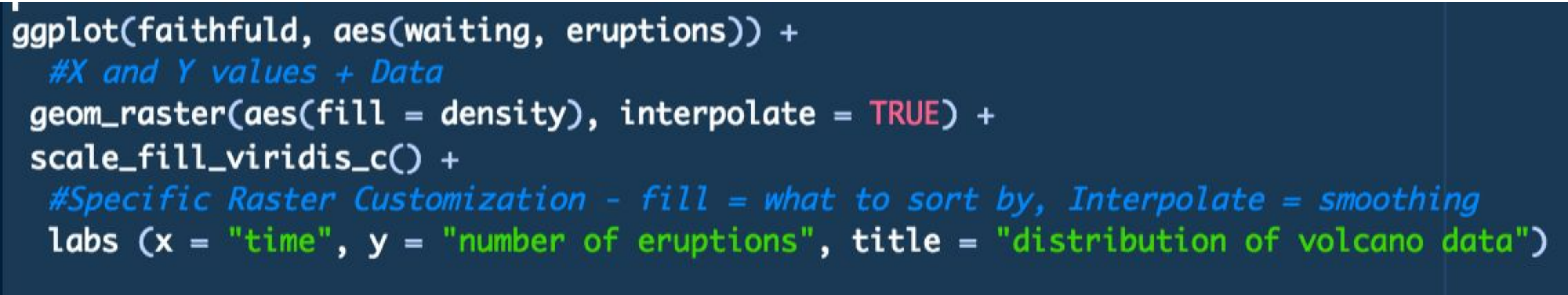
| | eruptions | waiting | density |
|---|---|---|---|
| 1 | 1.600000 | 43.00000 | 3.216159e-03 |
| 2 | 1.647297 | 43.00000 | 3.835375e-03 |
| 3 | 1.694595 | 43.00000 | 4.435548e-03 |
| 4 | 1.741892 | 43.00000 | 4.977614e-03 |
| 5 | 1.789189 | 43.00000 | 5.424238e-03 |
| 6 | 1.836486 | 43.00000 | 5.744544e-03 |
| 7 | 1.883784 | 43.00000 | 5.918012e-03 |
| 8 | 1.931081 | 43.00000 | 5.936762e-03 |
| 9 | 1.978378 | 43.00000 | 5.805861e-03 |
| 10 | 2.025676 | 43.00000 | 5.541706e-03 |
| 11 | 2.072973 | 43.00000 | 5.168979e-03 |
| 12 | 2.120270 | 43.00000 | 4.716903e-03 |
| 13 | 2.167568 | 43.00000 | 4.215592e-03 |
| 14 | 2.214865 | 43.00000 | 3.693071e-03 |
| 15 | 2.262162 | 43.00000 | 3.173317e-03 |
| 16 | 2.309459 | 43.00000 | 2.675315e-03 |
| 17 | 2.356757 | 43.00000 | 2.212951e-03 |
| 18 | 2.404054 | 43.00000 | 1.795434e-03 |
| 19 | 2.451351 | 43.00000 | [illegible] |

Showing 1 to 19 of 5,625 entries, 3 total columns

# Geom_Raster()

- A way to visualize data in a heat map style that factors in statistics like density in order to construct the plot and determine the “fill”
    - The “Fill” is a way to visualize the data by density
- In our context, the general time between eruptions can be seen
- Similar to Geom_Tile, but all rectangles are the same size

Our working example:



```
ggplot(faithfuld, aes(waiting, eruptions)) +
  #X and Y values + Data
 geom_raster(aes(fill = density), interpolate = TRUE) +
 scale_fill_viridis_c() +
  #Specific Raster Customization - fill = what to sort by, Interpolate = smoothing
  labs (x = "time", y = "number of eruptions", title = "distribution of volcano data")
```

# Geom_Raster () cont.

- **Syntax:** ggplot(“DataName”, aes(“X, Y”)) +

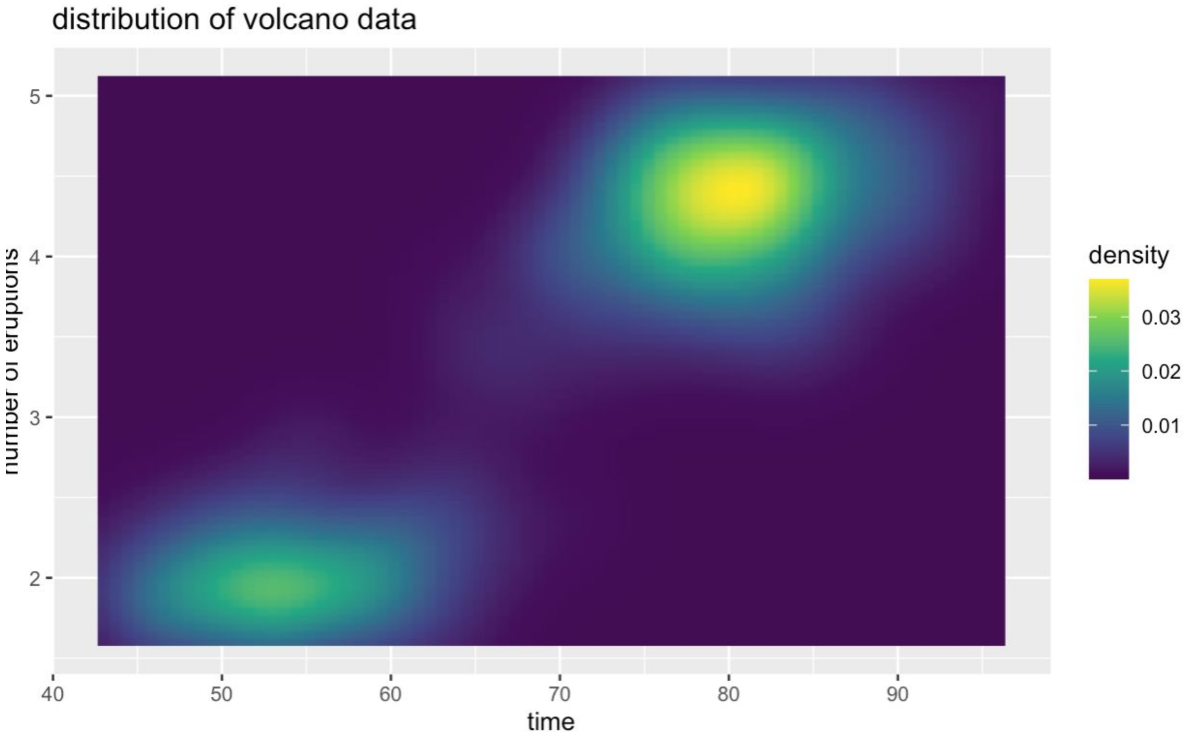
  geom_raster(aes(fill = “ScalarStatistic”), interpolate = TRUE/FALSE) +

  scale_fill_”ColorSelection”() +

  labs (x = “X title”, y = “Y title”, title = “Overall Title”)

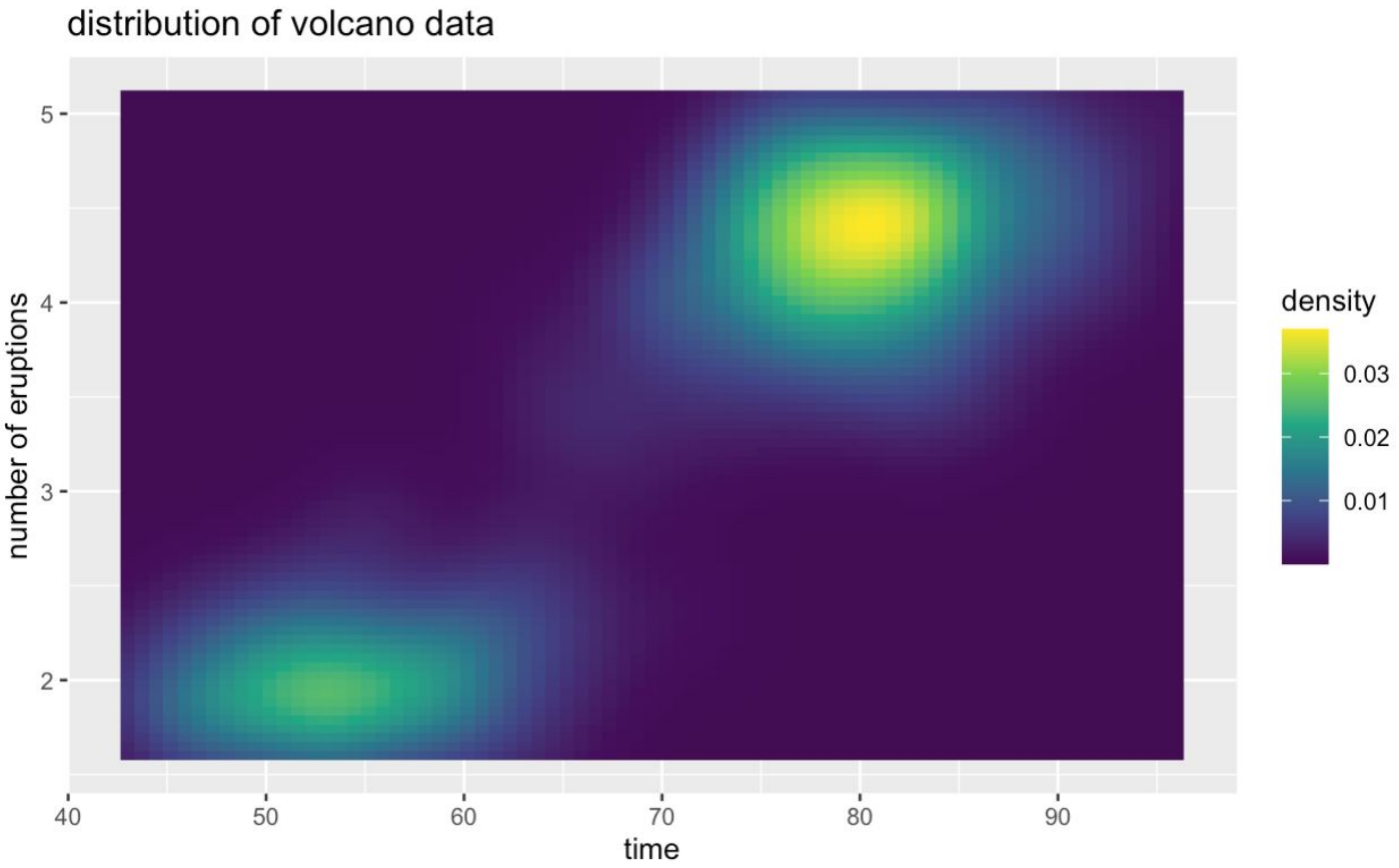
  - “ScalarStatistic” can include scaling statistics like magnitude, density, etc
    - The colors are determined by the fill
  - Interpolate smooths the graph and reduces the “blockiness”
  - Scale_fill_ is a way to color and customize the graph by color
  - Labs labels the graph (similar to other ggplot types)

# Geom_Raster Example



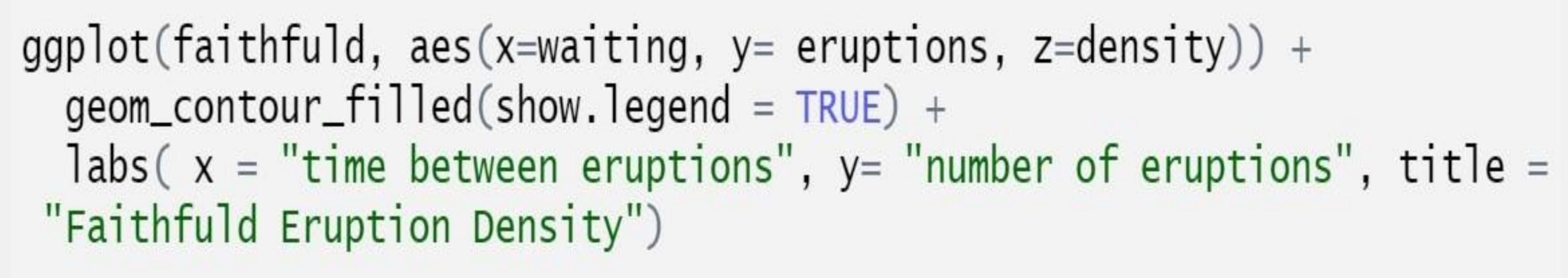
The interpolation = TRUE can be seen in the smoothness of the graph and the color change can be seen indicating the shift in density

# Notice the ‘blockiness’ of the visuals when interpolate = FALSE



# Geom_Contour

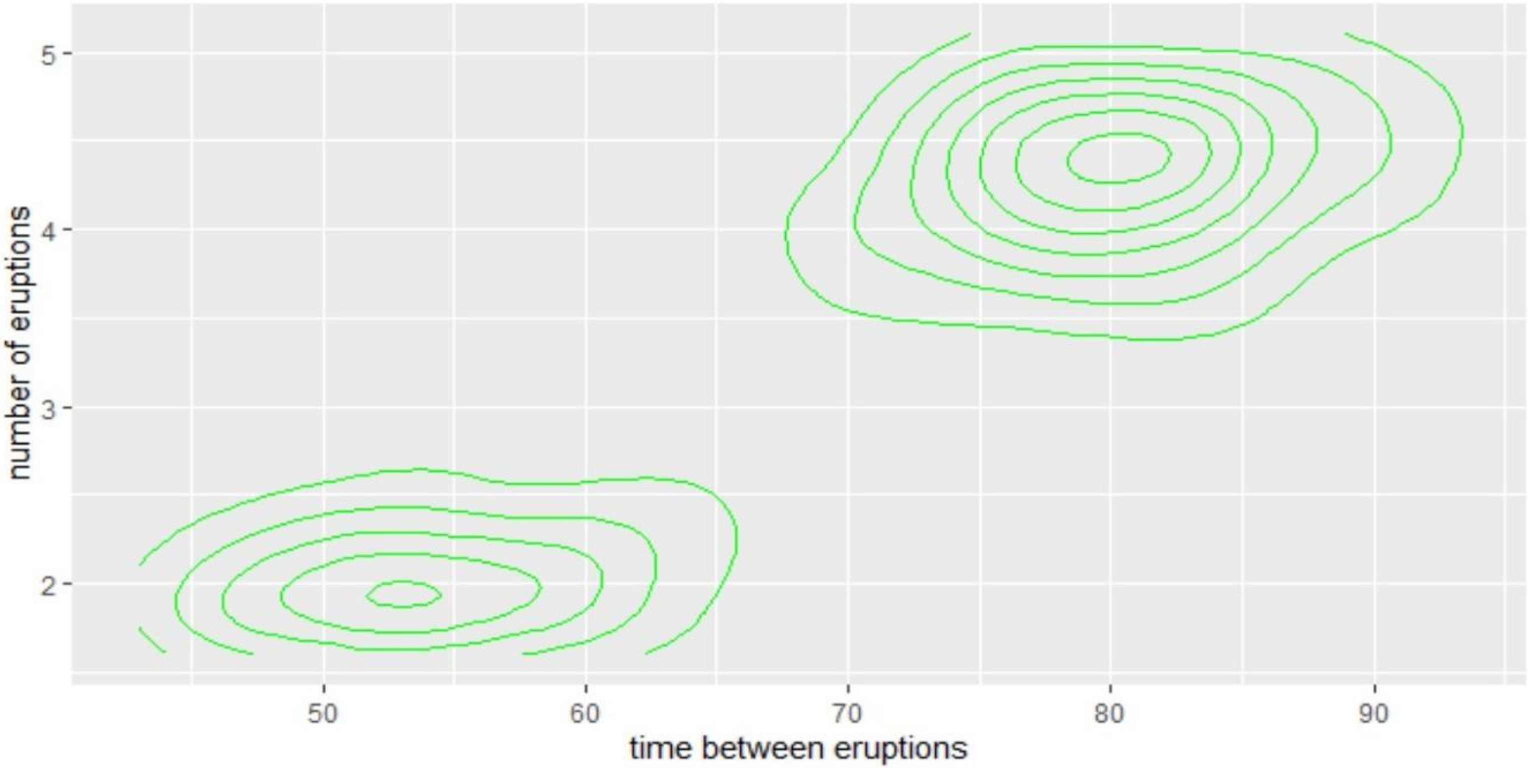
- Geom_Contour is similar to Geom_Raster, but it allows for the addition of levels
  - The levels can block out the data into different bins for better organization
  - It can be used more for more discrete variables
- Geom_Contour is useful for creating a 2D visualization of a 3D data set
- Can be filled or unfilled
  - For our dataset, filled provides better information that is easier to understand
- **Syntax:** ggplot(“DataName”, aes(“X”, “Y”, z = “ScalarStatistic”), ) + geom_contour_filled(Bins = n)
  - Changing the number of bins changes the amount of levels that are present on the visualization



```
ggplot(faithfuld, aes(x=waiting, y= eruptions, z=density)) +
  geom_contour_filled(show.legend = TRUE) +
  labs( x = "time between eruptions", y= "number of eruptions", title =
 "Faithfuld Eruption Density")
```

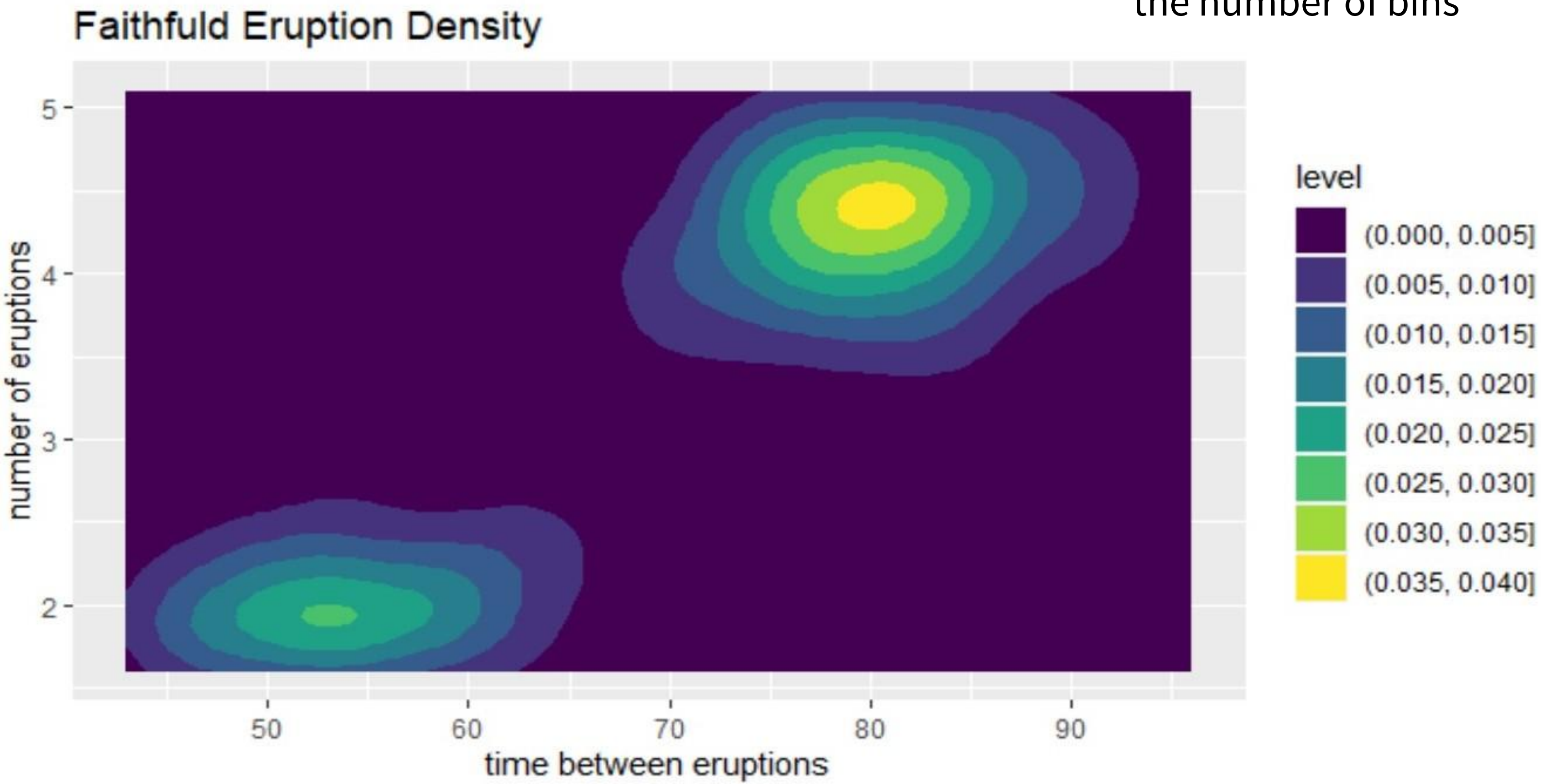
# Geom_Contour Example

This graph is okay, but it doesn’t provide the level of detail a filled version would



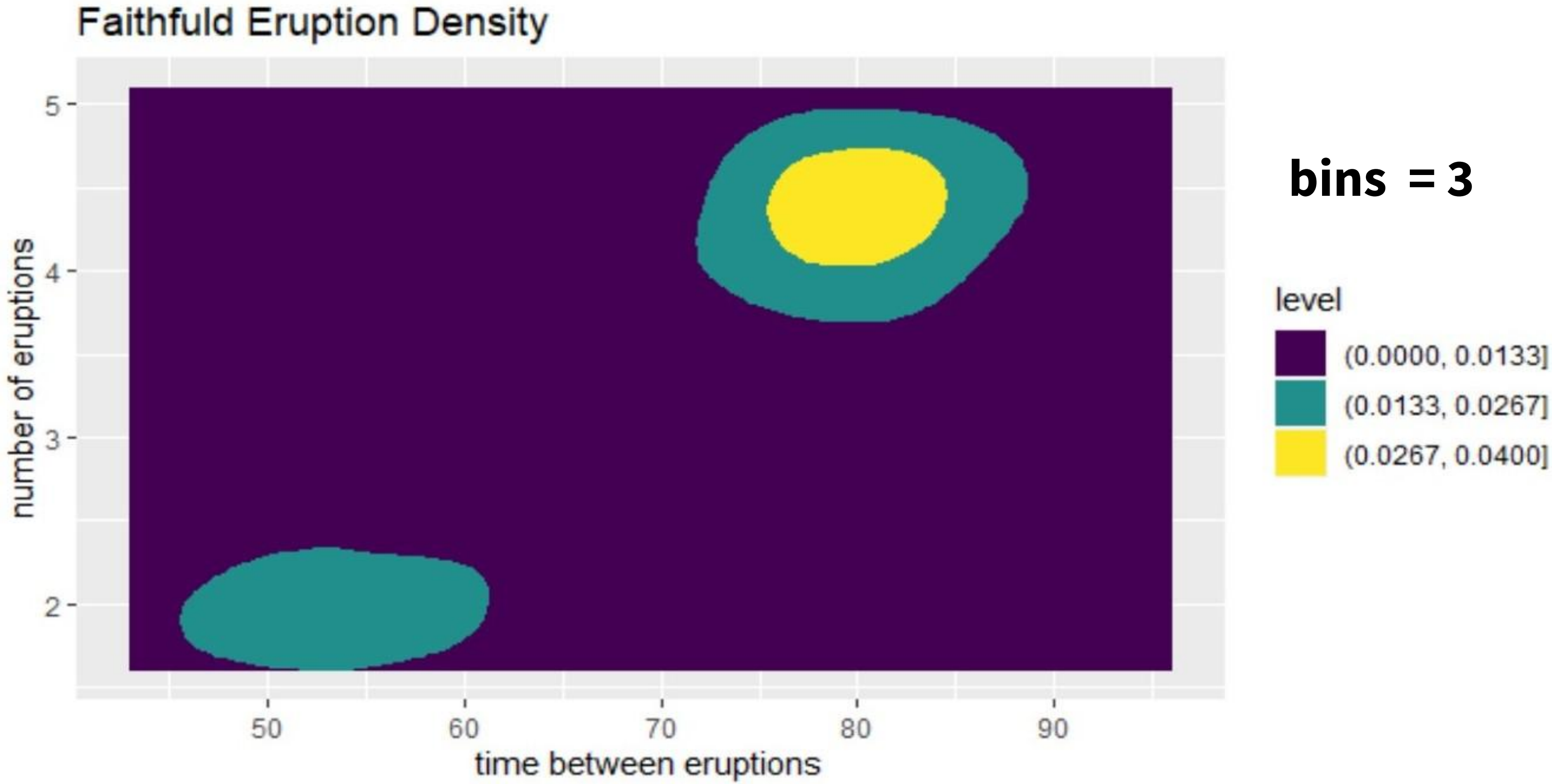
# Geom_Contour_Filled Example

This is the geom_contour_filled graph generated **without** specifying the number of bins



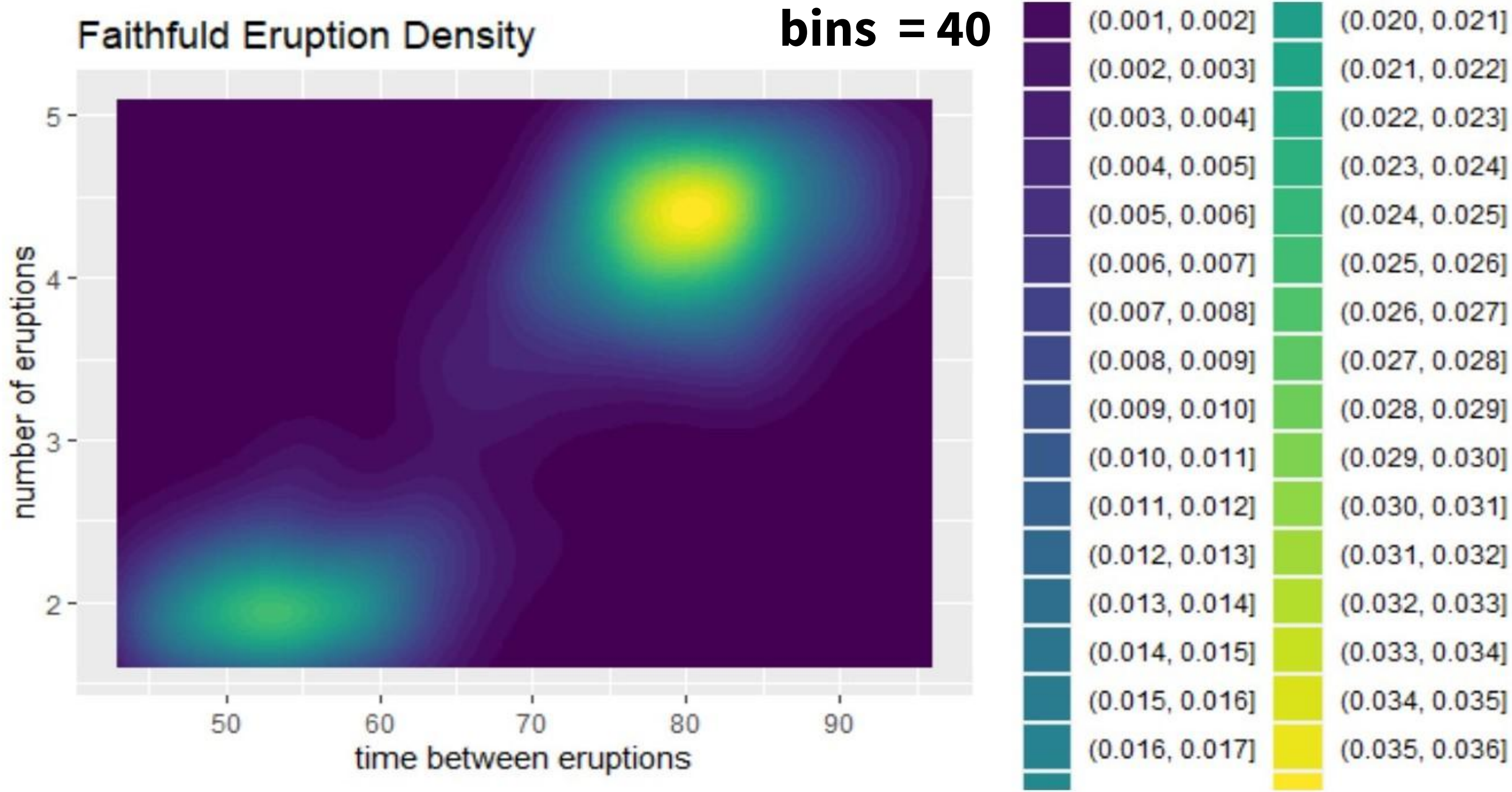
**What happens when we change the levels?**

# Too few bins does not show enough data to be useful



**bins = 3**

# Too many bins blurs data and defeats the purpose of geom_countour

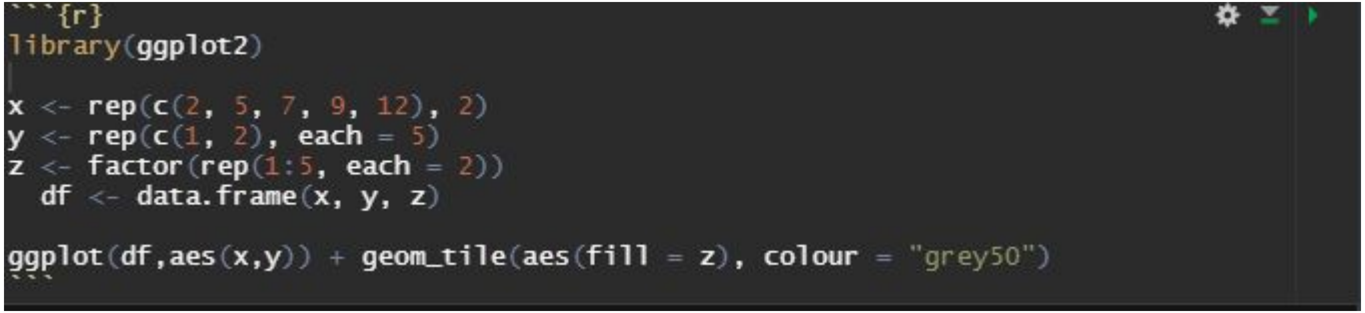


bins = 40

# Geom_tile()

- Can be viewed as a primitive/manual version of the geom_raster and geom_contour plot styles.
- Assigning a (x,y) coordinate as a “z” colored rectangle
- Takes numeric/integer inputs from a data frame
    - All lengths of the vectors must be the same for ‘z’ to assign its color to each respective (x,y) coordinate
    - length(vector.x) = 10
    - length(vector.y) = 10
    - length(vector.z) should be 10!

Our working example :



~~~
```{r}
library(ggplot2)

x <- rep(c(2, 5, 7, 9, 12), 2)
y <- rep(c(1, 2), each = 5)
z <- factor(rep(1:5, each = 2))
  df <- data.frame(x, y, z)

ggplot(df,aes(x,y)) + geom_tile(aes(fill = z), colour = "grey50")
```
~~~

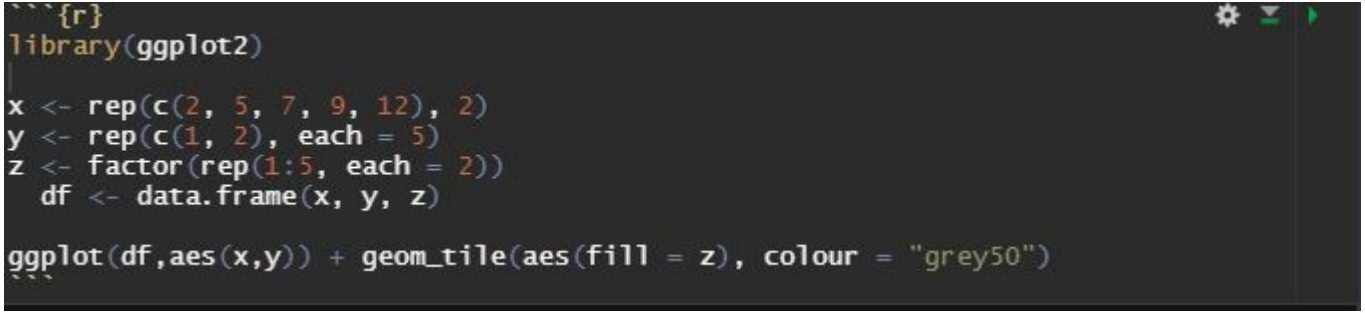
# Geom_tile()

Syntax:

ggplot(“data.name”, aes(“variable”, “variable2”)) + geom_tile(aes(fill = “variable3”), colour = “color.name”)

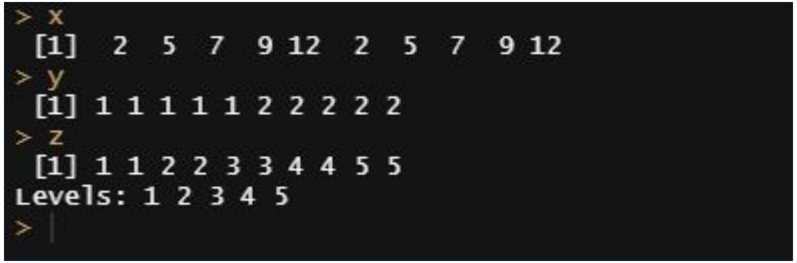
- aes = aesthetic mapping for ggplot... your arguments/variables turn mapped as a visual
- fill = the variable that controls the color of your rectangles from (x,y)
- colour = what color you want the rectangle outline to be

# My simple data for geom_tile()...



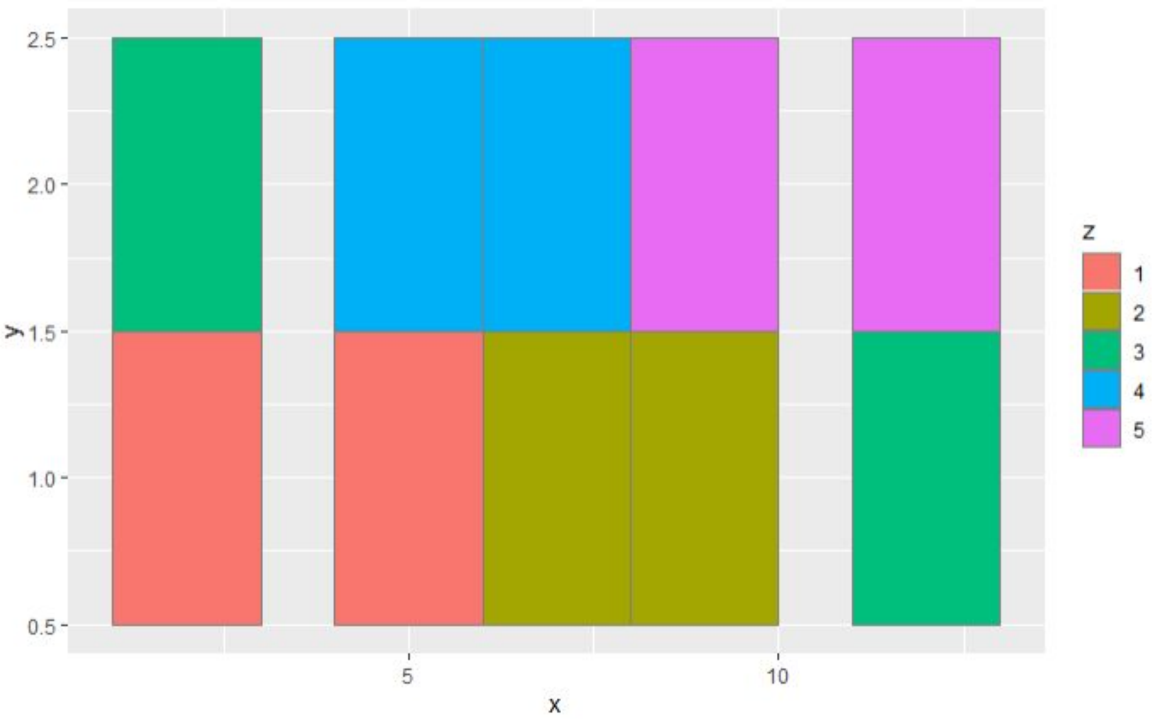
~~~
```{r}
library(ggplot2)

x <- rep(c(2, 5, 7, 9, 12), 2)
y <- rep(c(1, 2), each = 5)
z <- factor(rep(1:5, each = 2))
  df <- data.frame(x, y, z)

ggplot(df,aes(x,y)) + geom_tile(aes(fill = z), colour = "grey50")
```
~~~

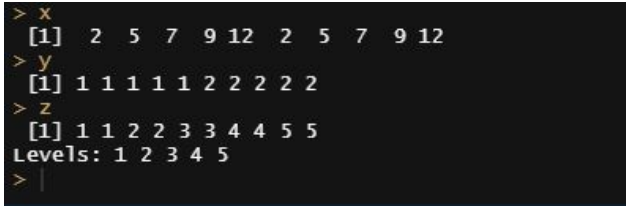
## x, y, and z vectors expanded.. This will come in handy later!



```
> x
 [1]  2  5  7  9 12  2  5  7  9 12
> y
 [1] 1 1 1 1 1 2 2 2 2 2
> z
 [1] 1 1 2 2 3 3 4 4 5 5
Levels: 1 2 3 4 5
> 
```

# My simple data frame visualized!

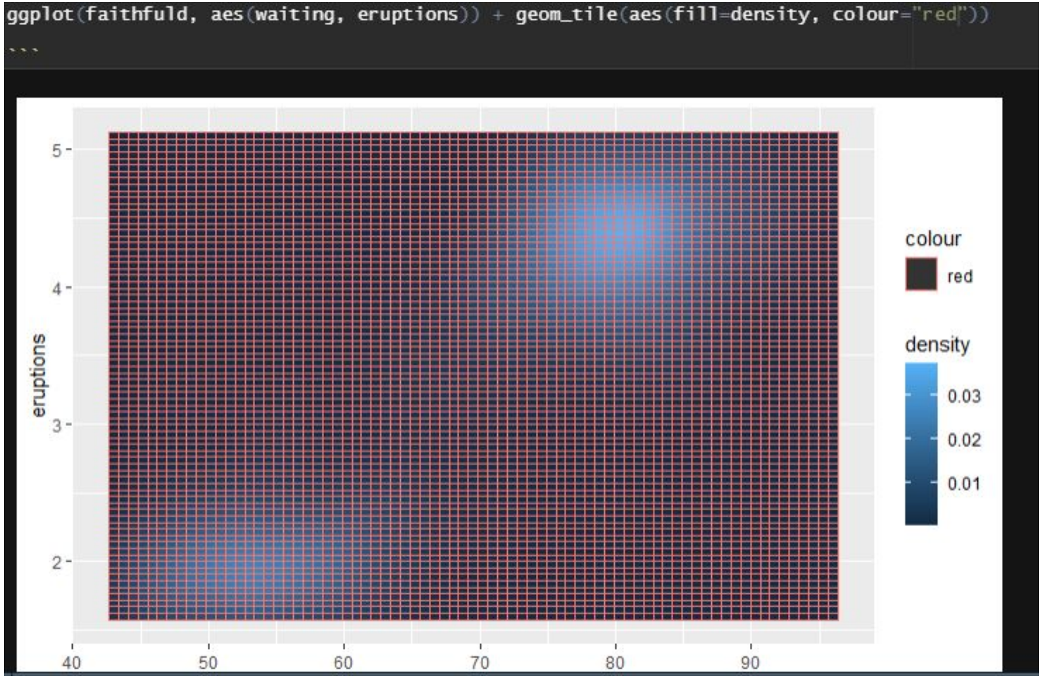




```
> x
 [1]  2  5  7  9 12  2  5  7  9 12
> y
 [1] 1 1 1 1 1 2 2 2 2 2
> z
 [1] 1 1 2 2 3 3 4 4 5 5
Levels: 1 2 3 4 5
>
```

- Read as (x,y, z(color))
- Each coordinate gets a rectangle with a corresponding ‘z’ color
- For example, (2, 1, 1(red)) is the bottom left rectangle

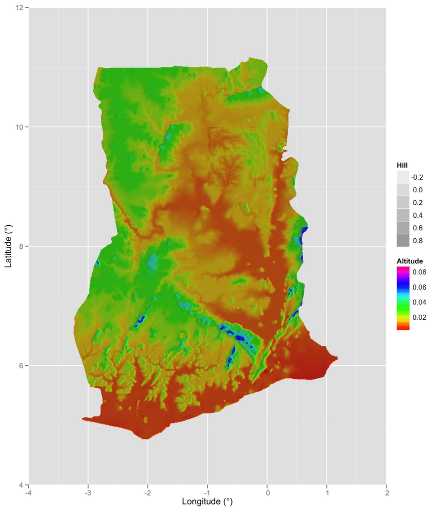
# Visualizing data frame “faithfuld” with geom_tile()

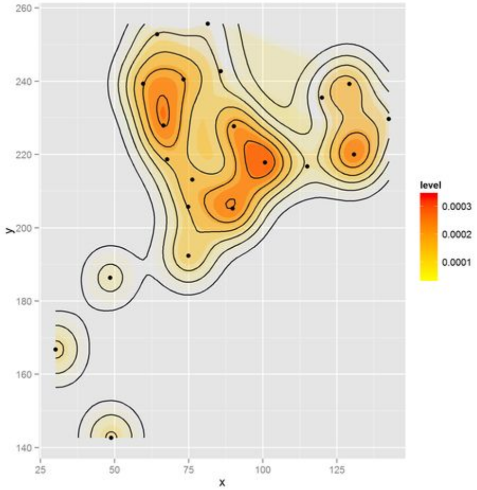


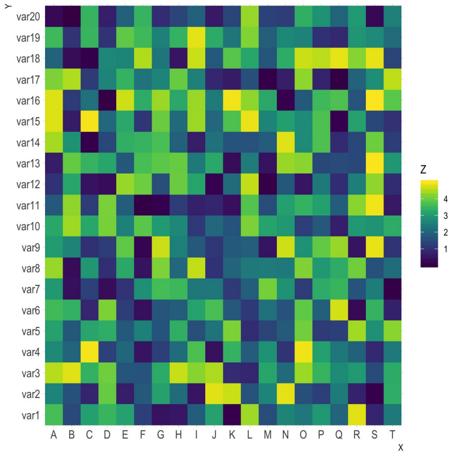
```
ggplot(faithfuld, aes(waiting, eruptions)) + geom_tile(aes(fill=density, colour="red"))
```
```

- x = waiting
- y = eruptions
- color (z) based off of our density

- Thousands of individual blocks can work together into a complex heat map!
```

# Summary

- These three plotting styles are most useful for creating heat-map esque visuals
- Geom_raster, geom_contour, geom_tile are very similar to each other in their own ways and can be combined in some cases to make more clear graphs
- Unique and appropriate applications for each style







# References

https://ggplot2.tidyverse.org/reference/geom_tile.html
https://ggplot2.tidyverse.org/reference/geom_contour.html
https://plotly.com/ggplot2/geom_raster/
https://datacarpentry.org/r-raster-vector-geospatial/02-raster-plot/
https://rdrr.io/cran/ggplot2/man/geom_tile.html
https://ggplot2.tidyverse.org/reference/geom_contour.html
https://www.rdocumentation.org/packages/ggplot2/versions/0.9.0/topics/stat_contour