

# Diamond sizes Fall '24

Gopinath Achuthan

2024

*#Note, Each gray box below is a code chunk. You need to insert a code chunk and put your R code in it. By setting echo = FALSE. this comment and any code will not show in my output document. If it were TRUE, the comment and code would appear.*

*#The include = FALSE function hides both the code and output in my output document. See what happens when you replace include=FALSE with echo=FALSE. In your final product, make sure this code chunk isn't displayed.*

*#You need to install these packages first to be able to use the functions within them (if you have not installed them before). You can install them from the Tools tab or write a new code chunk: install.packages("package\_name").*

```
library(ggplot2)
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.4.1
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##   filter, lag
```

```
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
diamonds<- as.data.frame(diamonds)
head(diamonds, 5)
```

```
##   carat    cut color clarity depth table price     x     y     z
## 1  0.23   Ideal     E    SI2  61.5    55   326  3.95  3.98  2.43
## 2  0.21 Premium     E    SI1  59.8    61   326  3.89  3.84  2.31
## 3  0.23    Good     E    VS1  56.9    65   327  4.05  4.07  2.31
## 4  0.29 Premium     I    VS2  62.4    58   334  4.20  4.23  2.63
## 5  0.31    Good     J    SI2  63.3    58   335  4.34  4.35  2.75
```

*#Remember, the head command tell R to show us the top rows. Note where the first 5 rows show up when you knit. They are not in your output word document because you have include=FALSE in this code chunk heading.*

We have data about 53940 diamonds. Only 749 are larger than 2.2 carats. The distribution of the remainder is shown below:

