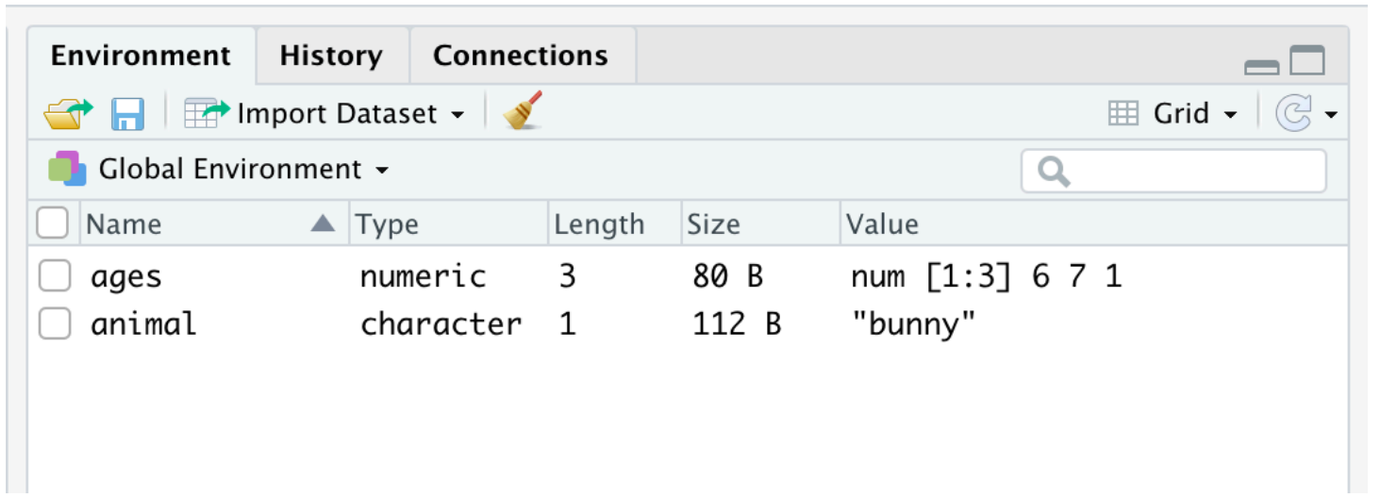
Packages

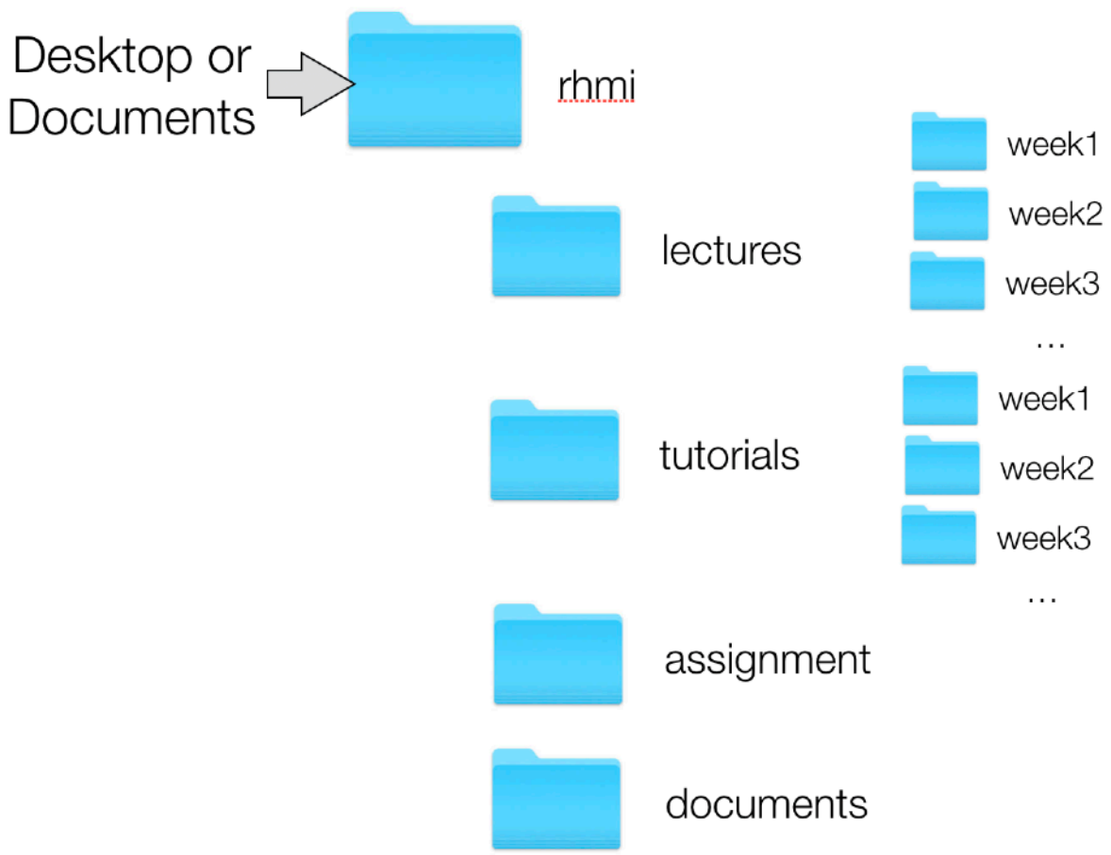
Packages: 5000+ available online

install	**load**
put on computer	make available to R
`install.packages("lsr")`	`library(lsr)`

Environment panel



Loading and saving data



"here" package lets you specify a file path easily

`read_csv()` loads up a csv file

`write_csv()` saves a csv file

```
loc <- here("folders/filename.csv")
data <- read_csv(file = loc)
```

Data manipulation

```
> data
         name colour height bunnyrank bearrank doggyrank
1       bunny   grey     20          1         3          2
2      gladly purple     18          3         1          2
3      flopsy  black     20          1         2          3
4      shadow    red     20          1         3          2
5         lfb purple     24          1         2          3
6 cuddly paws   <NA>     24         NA        NA         NA
7      doggie   blue     17          2         3          1
```

`data$height` selects the variable `height` from the tibble called `data`

`data[1,]` selects the first entire row from the tibble

`data$height[1]` selects the first case (row) of `height` from the tibble

`data[ c(1,4,7), c("age","gender")]` selects rows 1,4,7 and age/gender columns

`data[ 1:3, 2:4]` selects rows 1,2, and 3 and columns 2, 3, and 4

`data$tall <- data$height > 19` creates a new variable called `tall` which is true if `height` is over 19

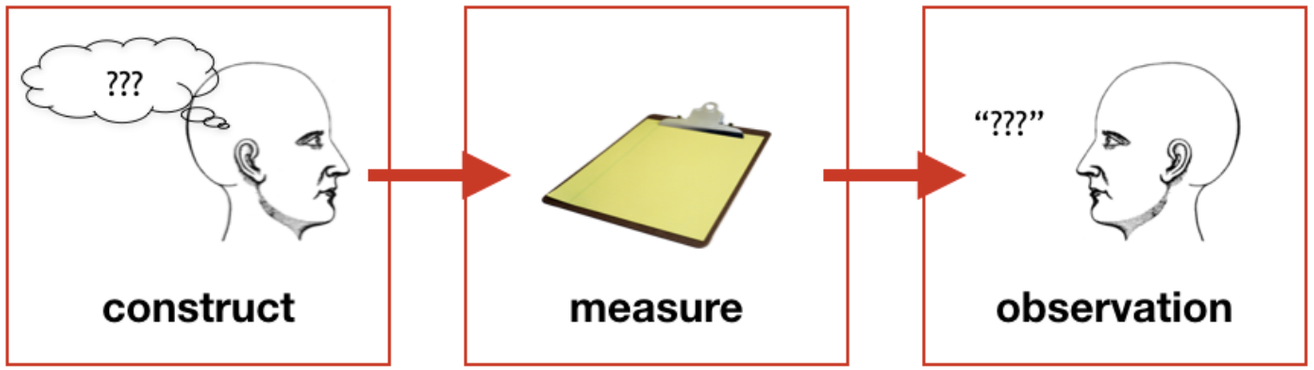
`data$tall <- NULL` removes the variable `tall`

`data[data$height>19,]` selects the rows from data for which the height is over 19

Measurement

1

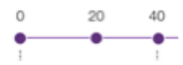
operationalisation measurement



Ex) intelligence WAIS IQ score

2

kinds of variables: matter for what tests you do!

	categorical	non-categorical		
	nominal	ordinal	interval	ratio
				
continuous?			✔	✔
discrete?	✔	✔	✔	✔
natural ordering?		✔	✔	✔
differences meaningful?			✔	✔
zero means zero?				✔
ratios meaningful?				✔

3

reliability: measuring the same thing twice

	test-retest	inter-rater	internal consistency
	time 1 → time 2 		"eat this?" "like this?" "breakfast?" "dinner?"
what's different?	time	people	version

4

variable types

explains	needs explaining
predictor	outcome
independent variable	dependent variable
treatment	response

Projects and Markdown

1 R Markdown lets you embed code chunks

````
You can also embed plots, for example:

```{r pressure, echo=FALSE}
plot(pressure)
```
````

Tells it there is R code here

R code

Name of code chunk

Optional arguments

2

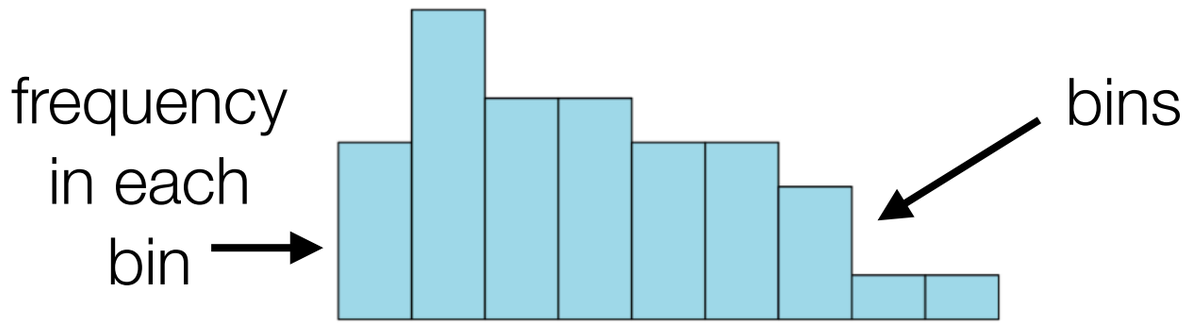
Format	Description	Example
Italic	Italicised text in between the *	*Italic*
Bold	Bolds text in between the **	**Bold**
Pi rounded is `round(3.14)`	Lets you run the R code inline in between the `	Pi rounded is 3
# Header 1	Makes a really big header	Header 1
## Header 2 (down to 4)	Makes progressively smaller headers	etc.
[Link](https://google.com)	Lets you embed links	Link

Descriptive statistics

1

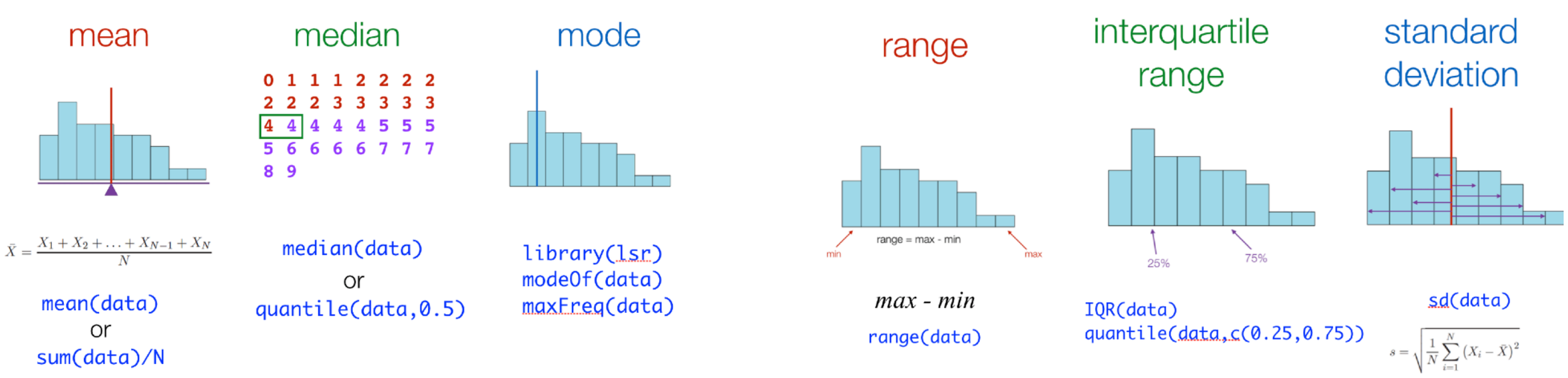
Histograms

```
hist(x=data,breaks=12,
    main="title",col="blue")
```



2 Central tendency: where is it?

3 Spread: how wide is it?



$\bar{X} = \frac{X_1 + X_2 + \ldots + X_{N-1} + X_N}{N}$

```
mean(data)
```
or
```
sum(data)/N
```

```
median(data)
```
or
```
quantile(data,0.5)
```

```
library(lsr)
modeOf(data)
maxFreq(data)
```

max - min

```
range(data)
```

```
IQR(data)
quantile(data,c(0.25,0.75))
```

```
sd(data)
```

$s = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})^2}$

4 Summary functions

Shows part of tibble

```
head(data)
glimpse(data)
```

Lots of descriptive statistics

```
summary(data)
```