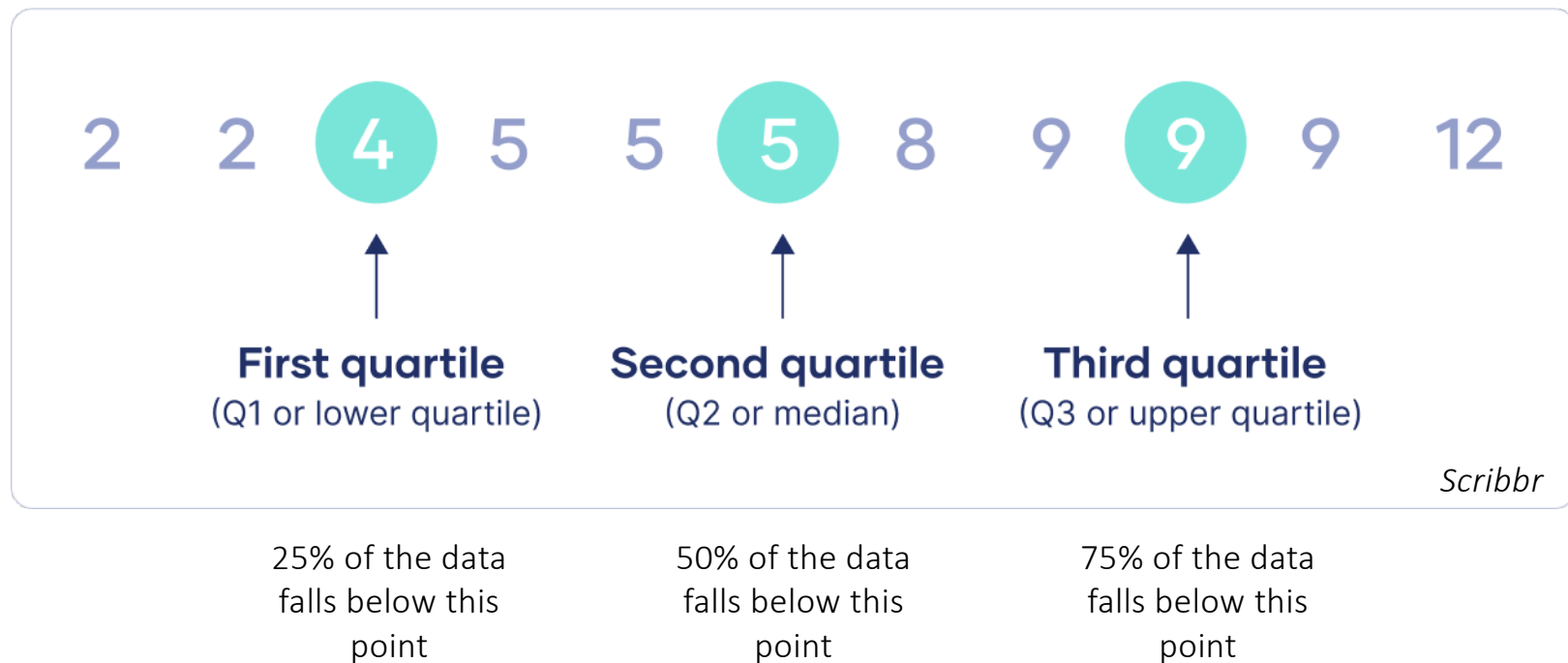


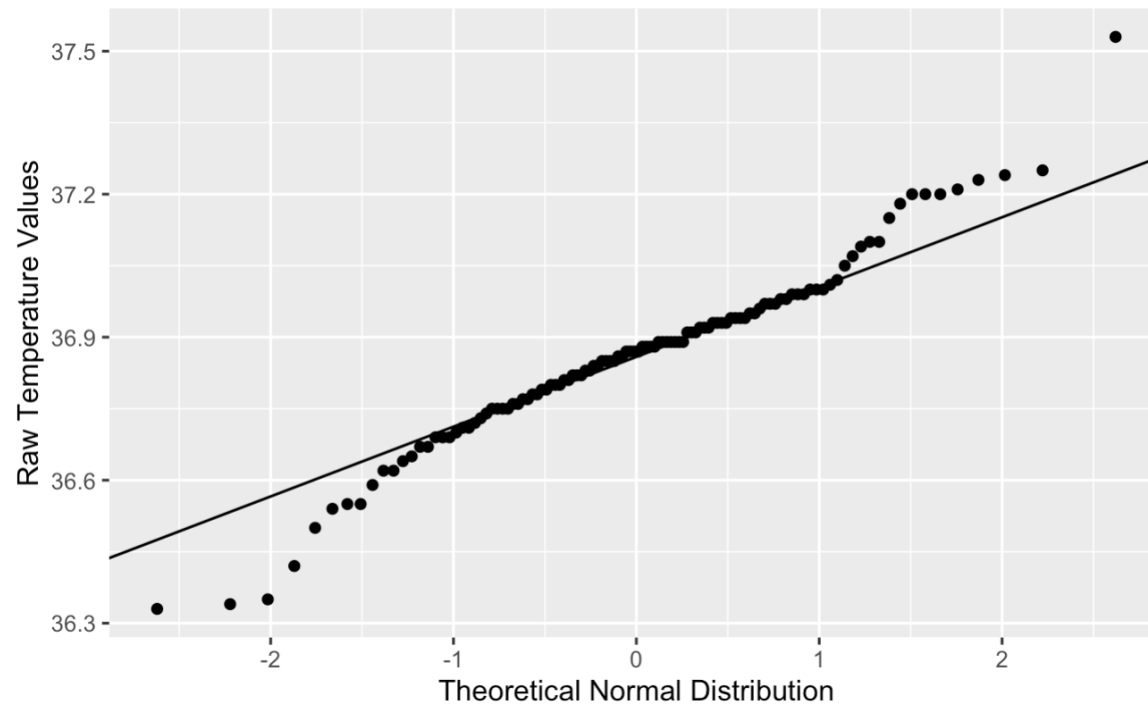
quantile quantile (QQ) plots

What is a quantile?

tells you *where* a data point ranks in a distribution



A QQ plot is **theoretical quantiles** plotted against **sample quantiles**.



How do you calculate a *sample* quantile?

1. sort your data from smallest to largest
2. assign each point a quantile position

$$\frac{i - 0.5}{n}$$

$i = \text{rank}$
 $n = \text{total \# of points}$

If you have 10 data points...

What quantile is the 1st point?

What quantile is the 5th point?

What quantile is the 10th point?

How do you calculate a *theoretical* quantile?

For each quantile position in your data, ask: if the data were perfectly normal, what VALUE would we expect at this quantile?

1. Use the quantile positions from your data
0.05, 0.45, 0.95
2. Use inverse CDF (quantile function) of the normal distribution to get the *z-score* for that quantile → `qnorm ( )`

If you have 10 data points...

Your 1st point is at 0.05 quantile

Your 5th point is at 0.45 quantile

Your 10th point is at 0.95 quantile

```
> qnorm(0.05)
[1] -1.644854
> qnorm(0.45)
[1] -0.1256613
> qnorm(0.95)
[1] 1.644854
```

Try it!

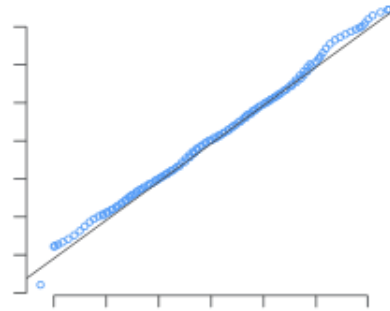
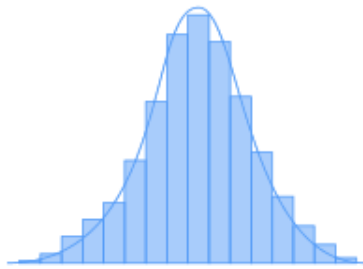
You have 6 test scores: 45, 67, 71, 73, 89, 95

1. calculate sample quantile for each score

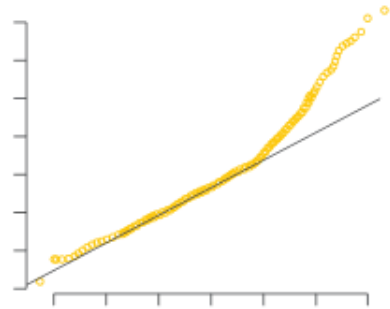
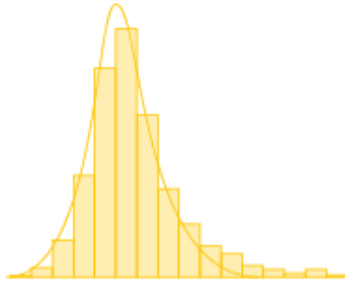
$$\frac{i - 0.5}{n}$$

2. look up the theoretical quantiles in R
3. label X, Y pairs as you would plot on a QQ plot

Normally distributed data



Right-skewed data



Left-skewed data

