Kendall-tau & Pearson Correlation Comparisons

Robert M Flight

16 September, 2020

Purpose

Quick report comparing our the missing data aware Kendall-tau and previous implementation of weighted Pearson correlations.

Data

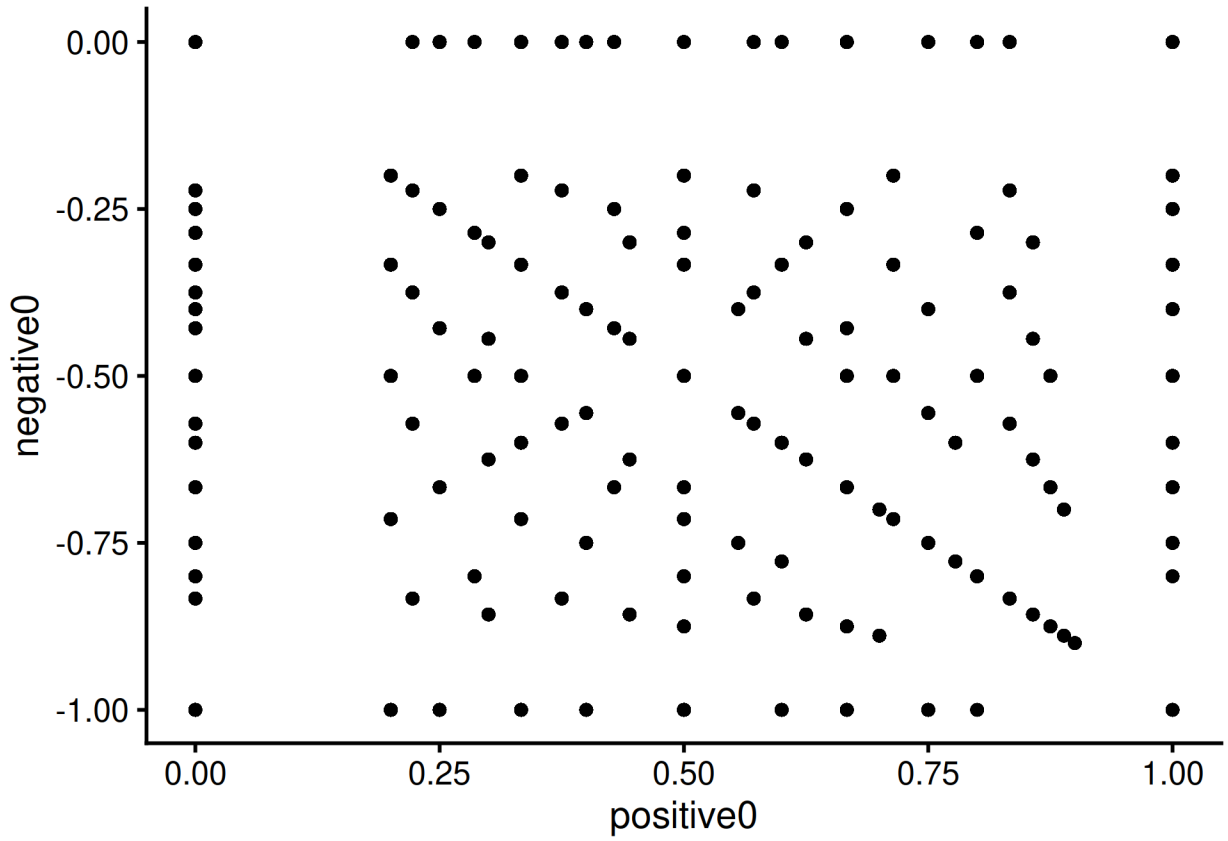
Positive and negative correlating vectors where:

- x = 1 - 10
- y = 1 - 10
- y2 = 10 - 1
- Values in x and y replaced by NA values, from 1 to 10 in either one

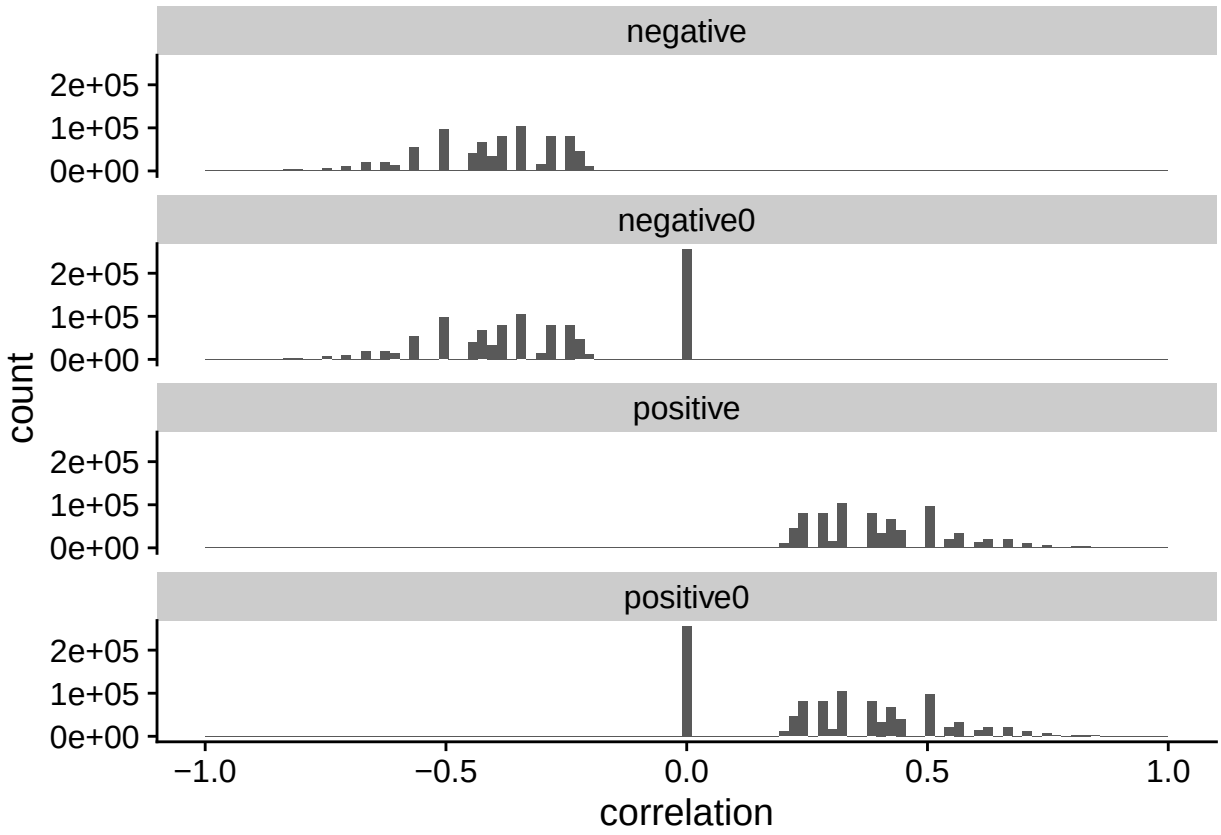
Pearson correlations used the "locally-it-weighted" function, Kendall-tau were calculated using "global", without scaling for "max-cor".

Results

Pearson



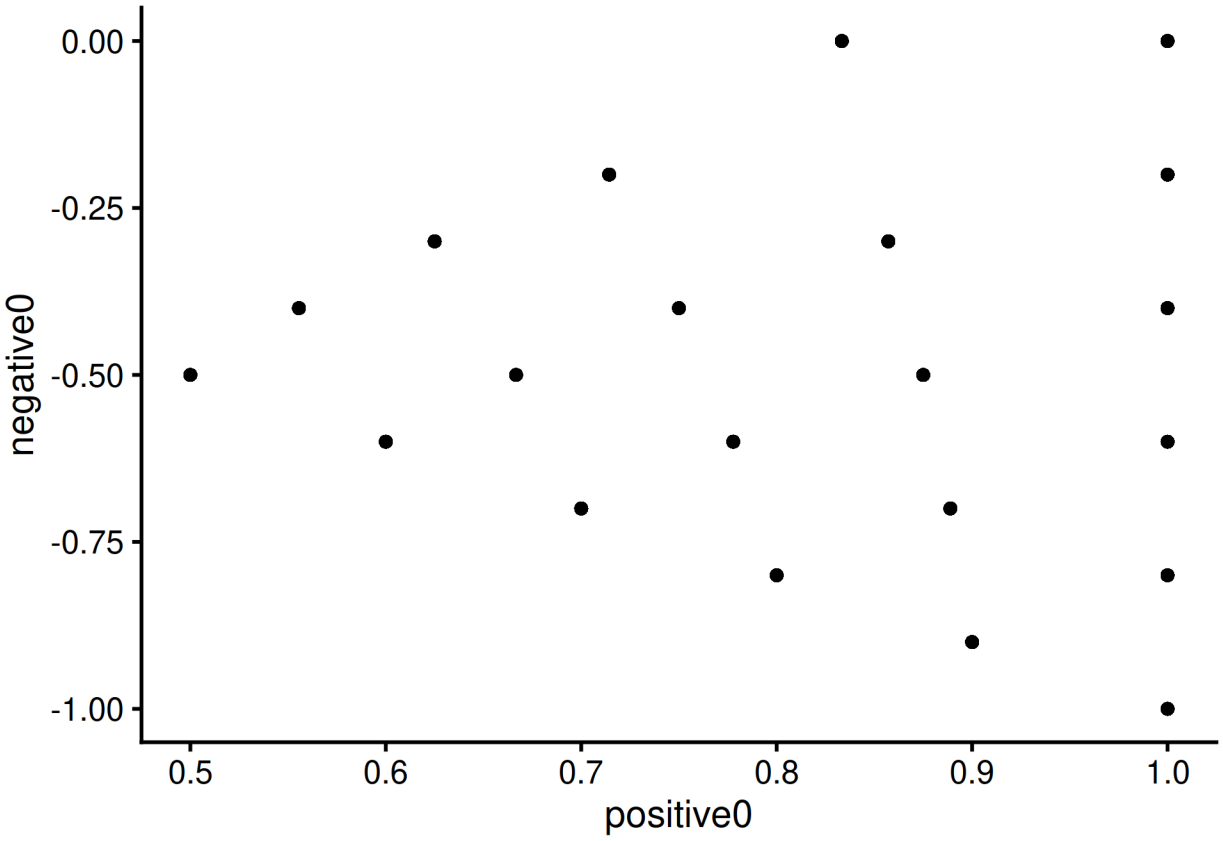

```
## Warning: Removed 511758 rows containing non-finite values (stat_bin).
```



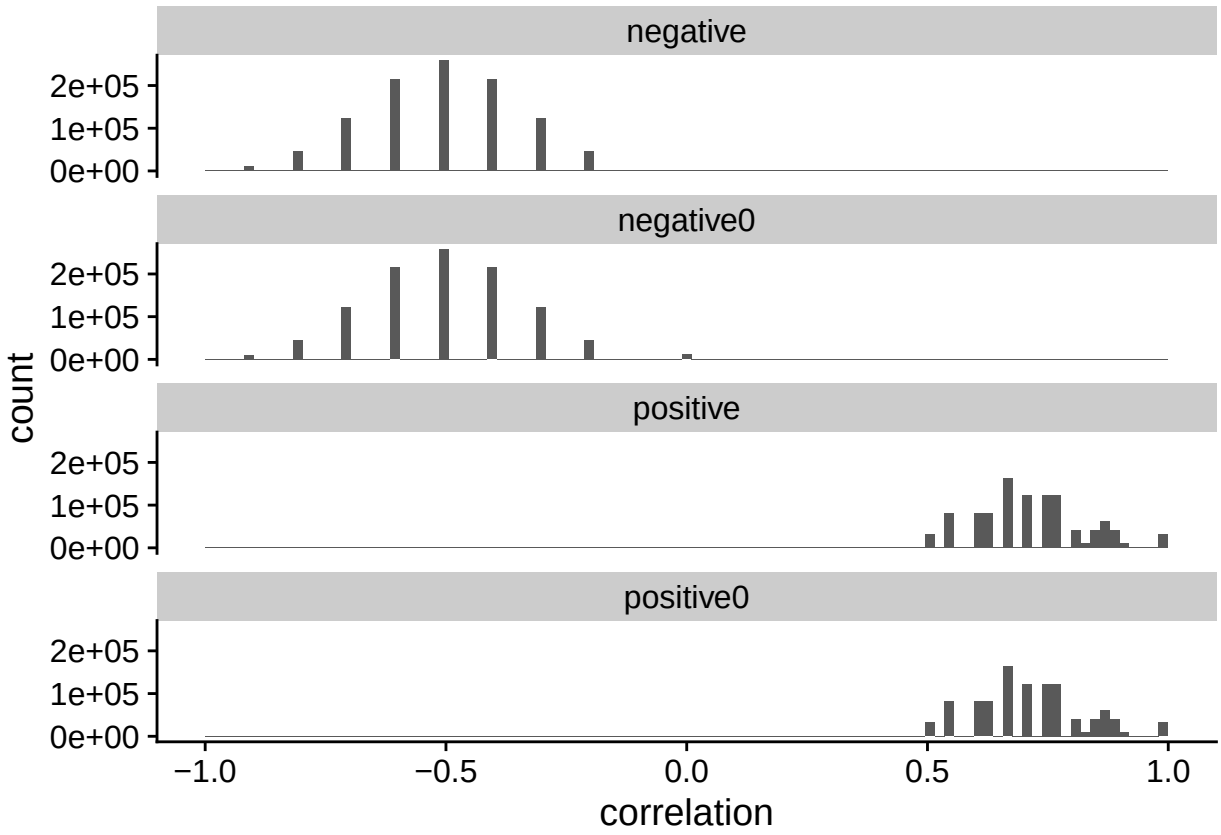

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

type	mean
negative	-0.4036856
negative0	-0.3051761
positive	0.4036856
positive0	0.3051761

Replace Low Samples Only



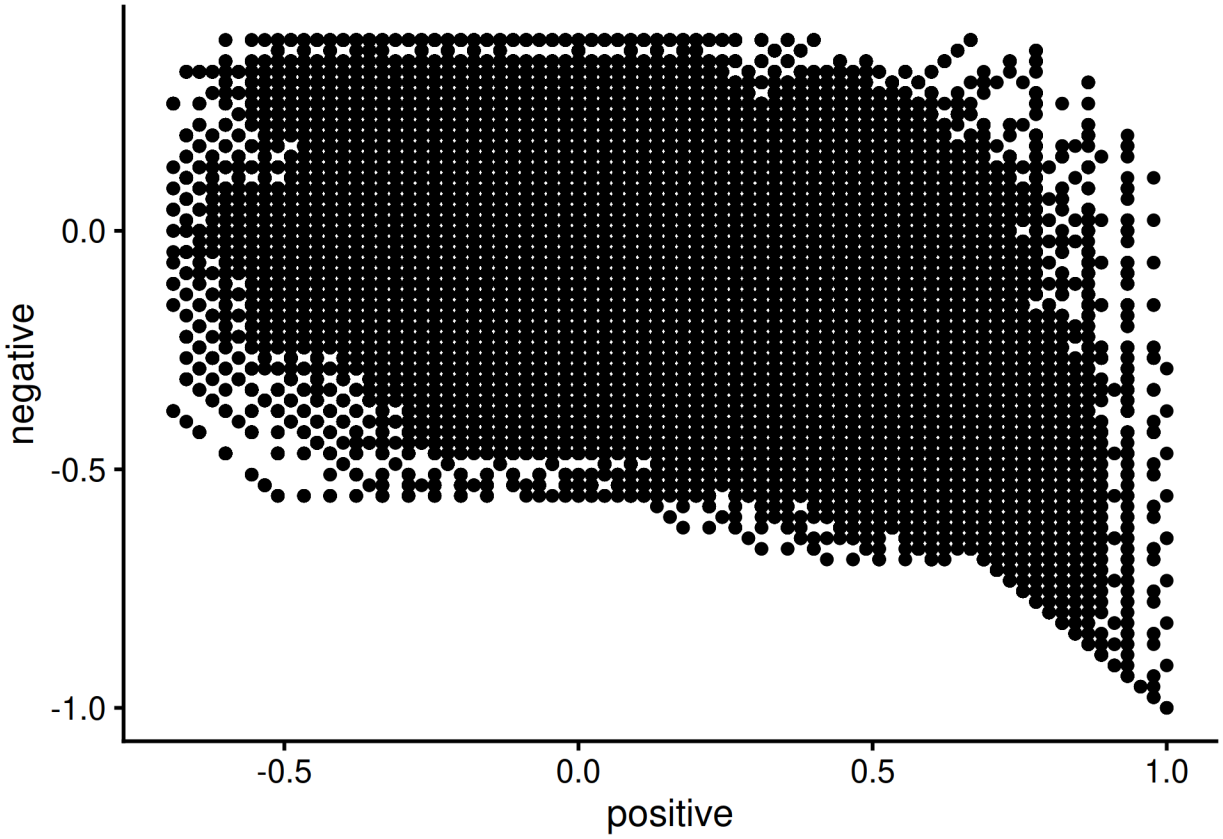

```
## Warning: Removed 11264 rows containing non-finite values (stat_bin).
```

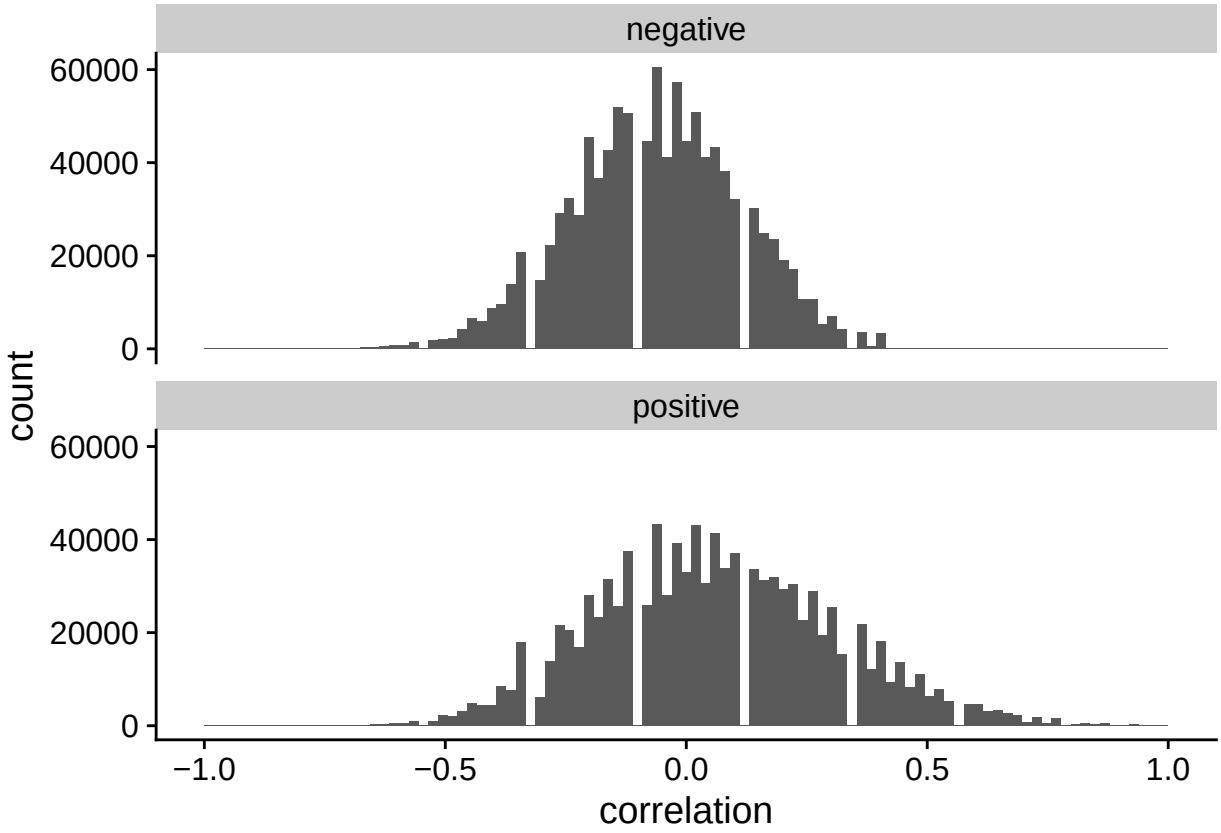



```
## `summarise()` ungrouping output (override with `.groups` argument)
```

type	mean
negative	-0.5044418
negative0	-0.4990230
positive	0.7194240
positive0	0.7194240

Kendall-tau

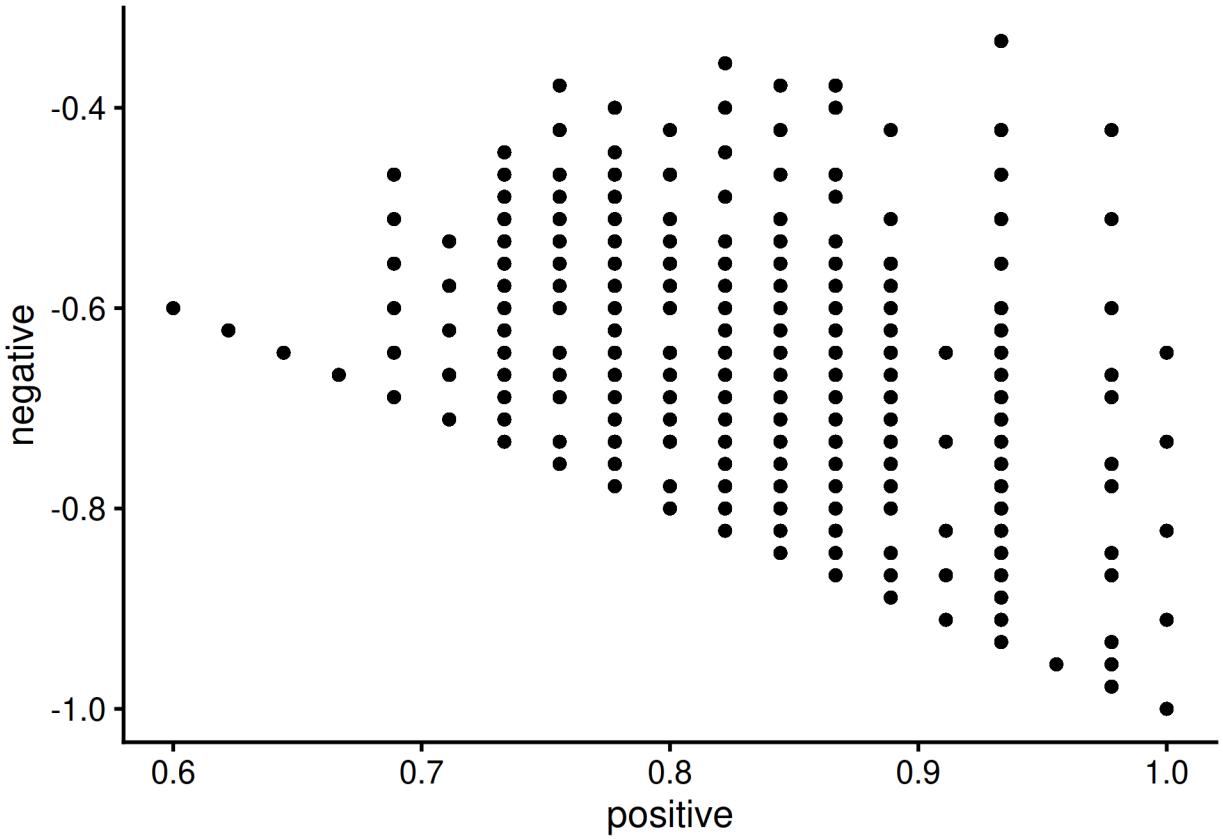


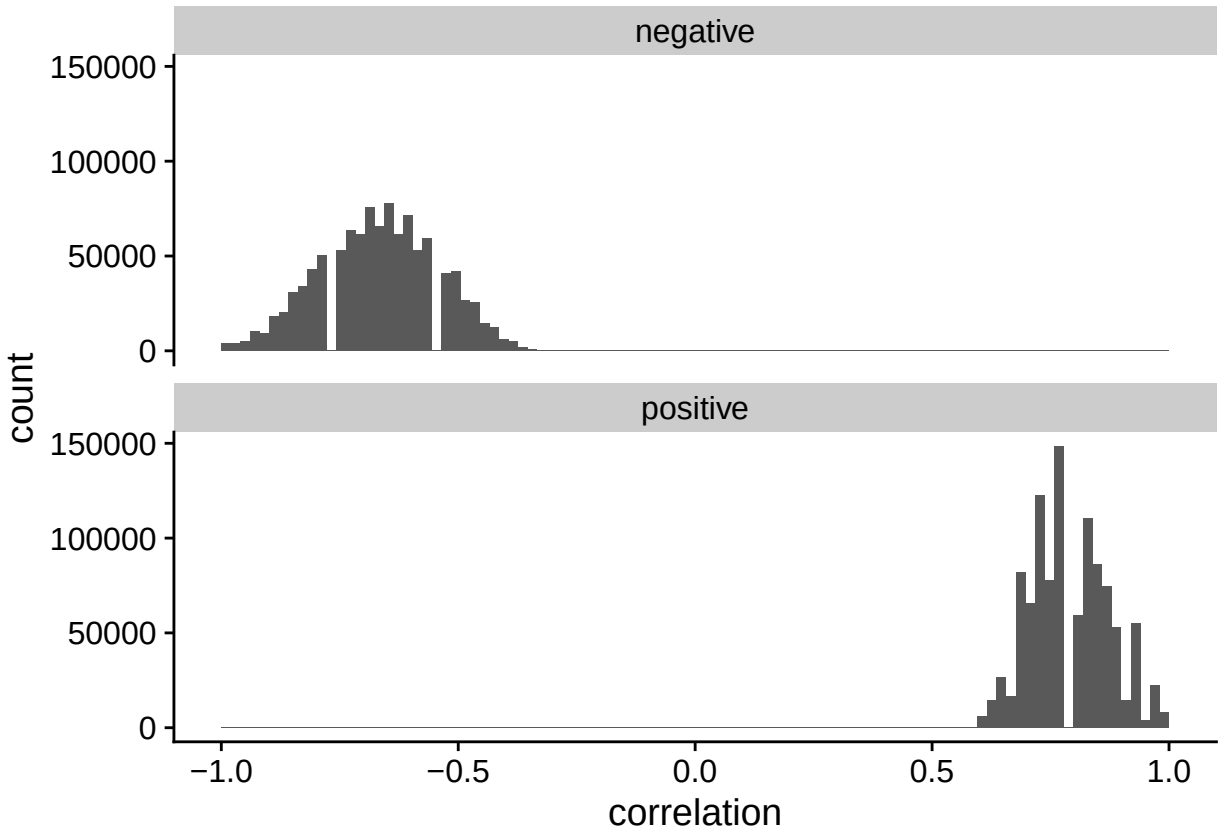



```
## `summarise()` ungrouping output (override with `.groups` argument)
```

type	mean
negative	-0.0624991
positive	0.0624991

Replace Low Entries Only


-0.4
-0.6
-0.8
-1.0
negative
0.6
0.7
0.8
0.9
1.0
positive



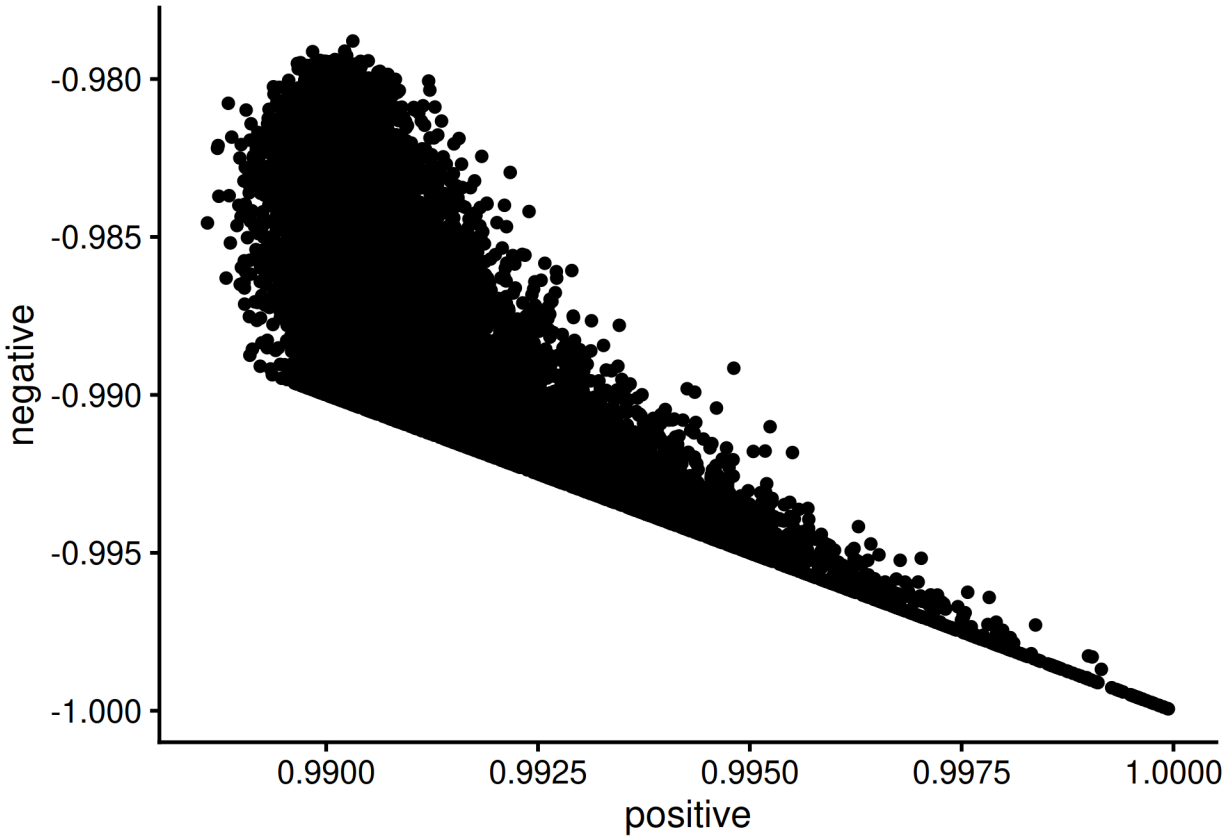

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

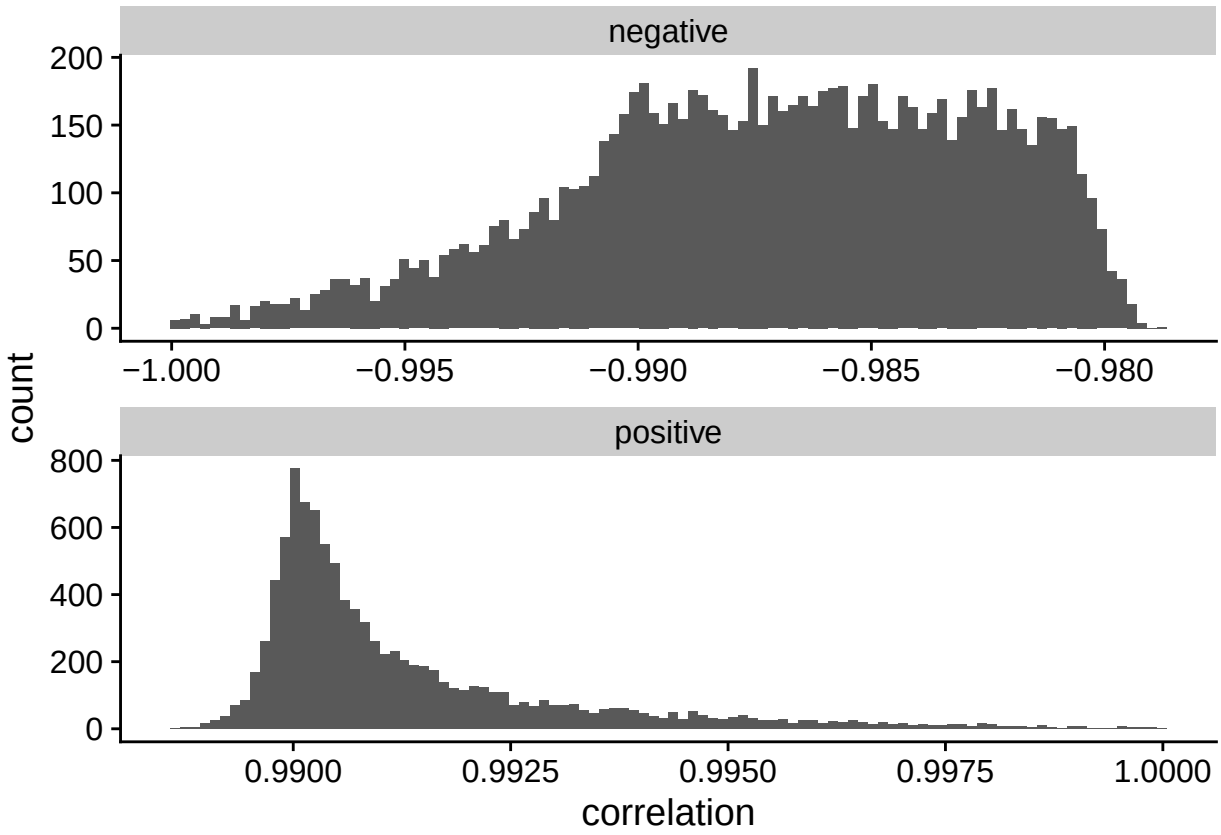
type	mean
negative	-0.6666663
positive	0.7916665

Weird, that didn't quite work like we expected.

What about the *more realistic samples?*

Realistic Samples


-0.980
-0.985
-0.990
-0.995
-1.000
negative
0.9900
0.9925
0.9950
0.9975
1.0000
positive




```
## `summarise()` ungrouping output (override with `.groups` argument)
```

type	mean
negative	-0.9870464
positive	0.9912785