

R

ggplot2

ggstatsplot

<https://indrajeetpatil.github.io/ggstatsplot/articles/ggstatsplot.html>

<https://indrajeetpatil.github.io/ggstatsplot/>

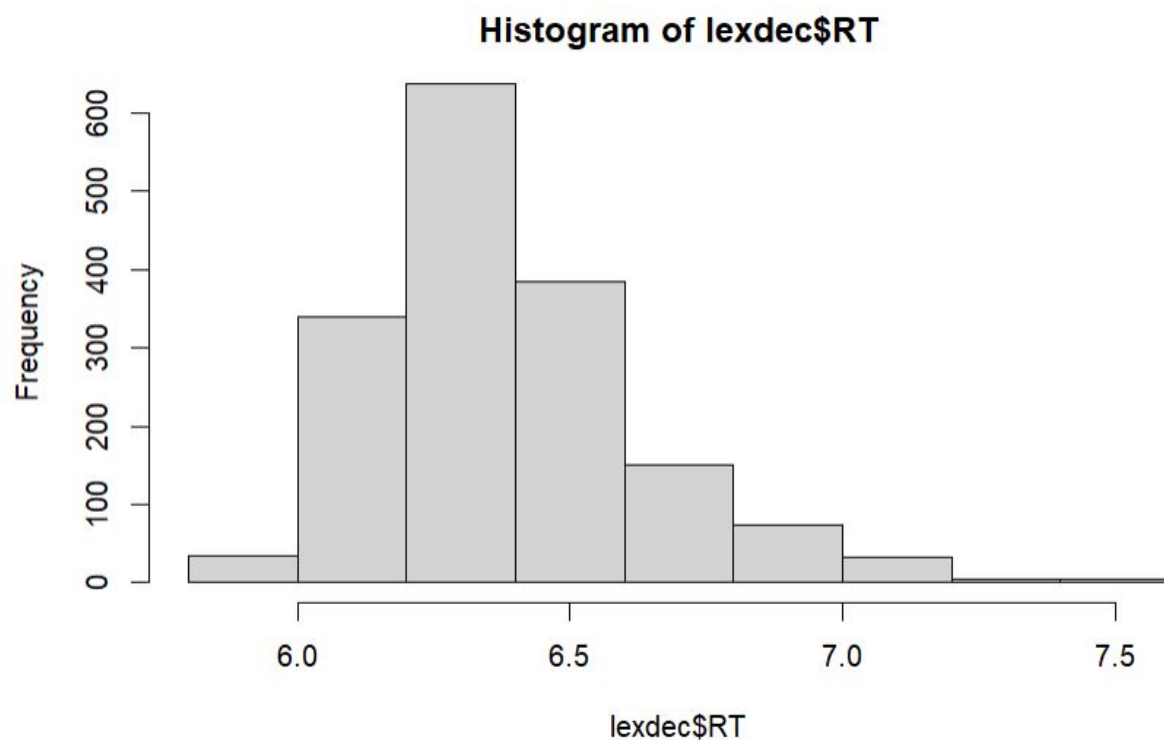
- ・統計分析値とともにグラフも描いてくれる。
- ・統計部分の見方
 - ・ <https://indrajeetpatil.github.io/ggstatsplot/articles/ggstatsplot.html#benefits>

例えばヒストグラム

- ・サンプルのデータは、[languageR](#) 内の lexdec

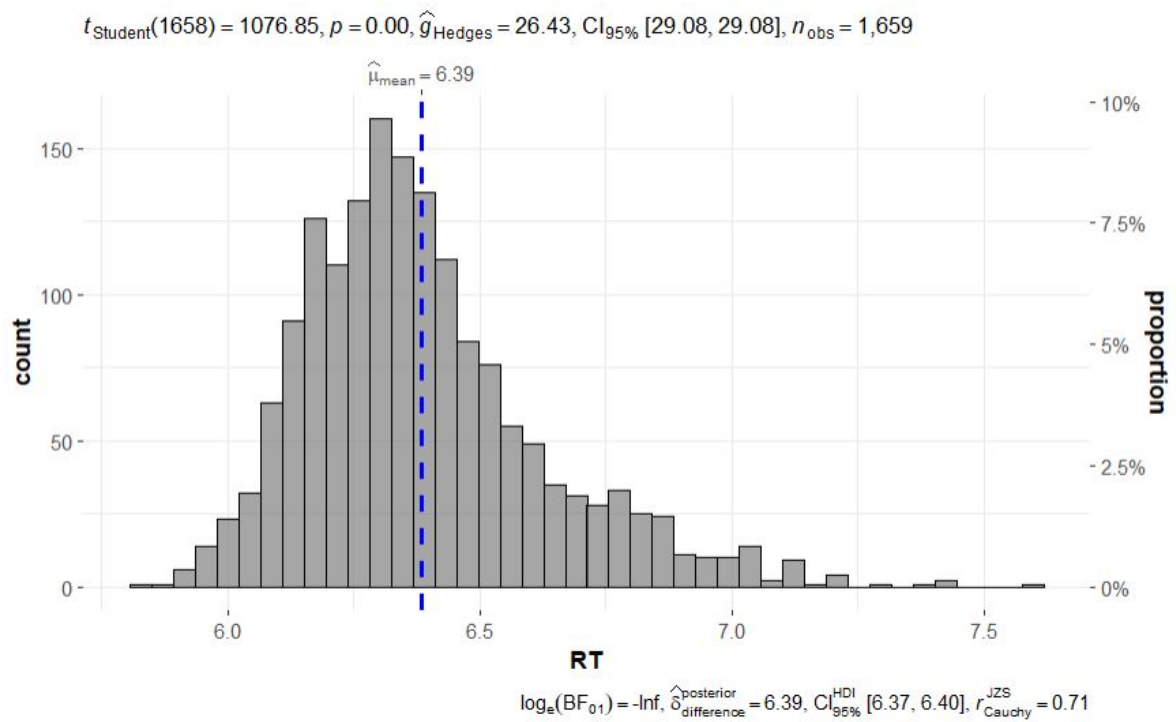
■ 普通はこのように ↓

```
hist(lexdec$RT)
```



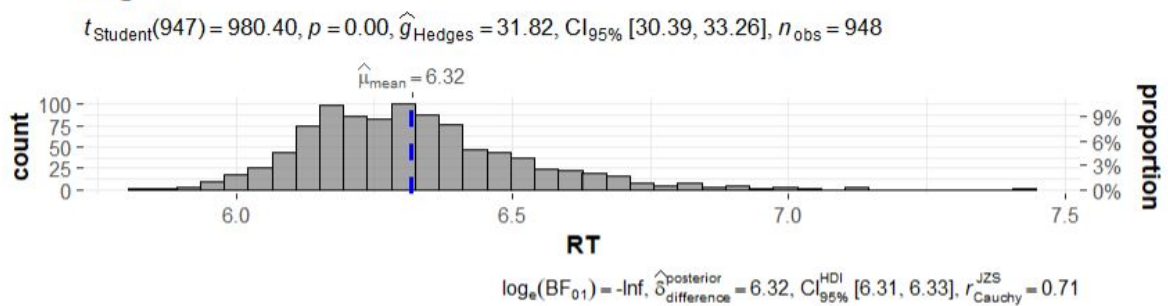
■ gghistostats を使うとこのように ↓

```
gghistostats(lexdec, x=RT)
```

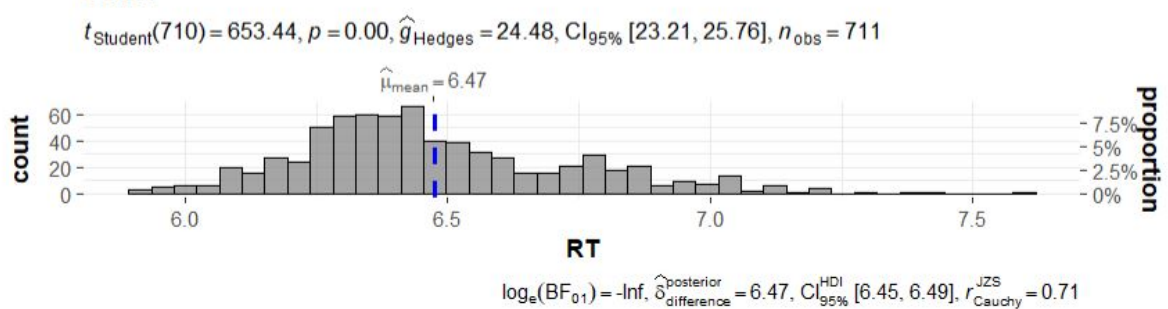


■ グループ分けして見ることも

English

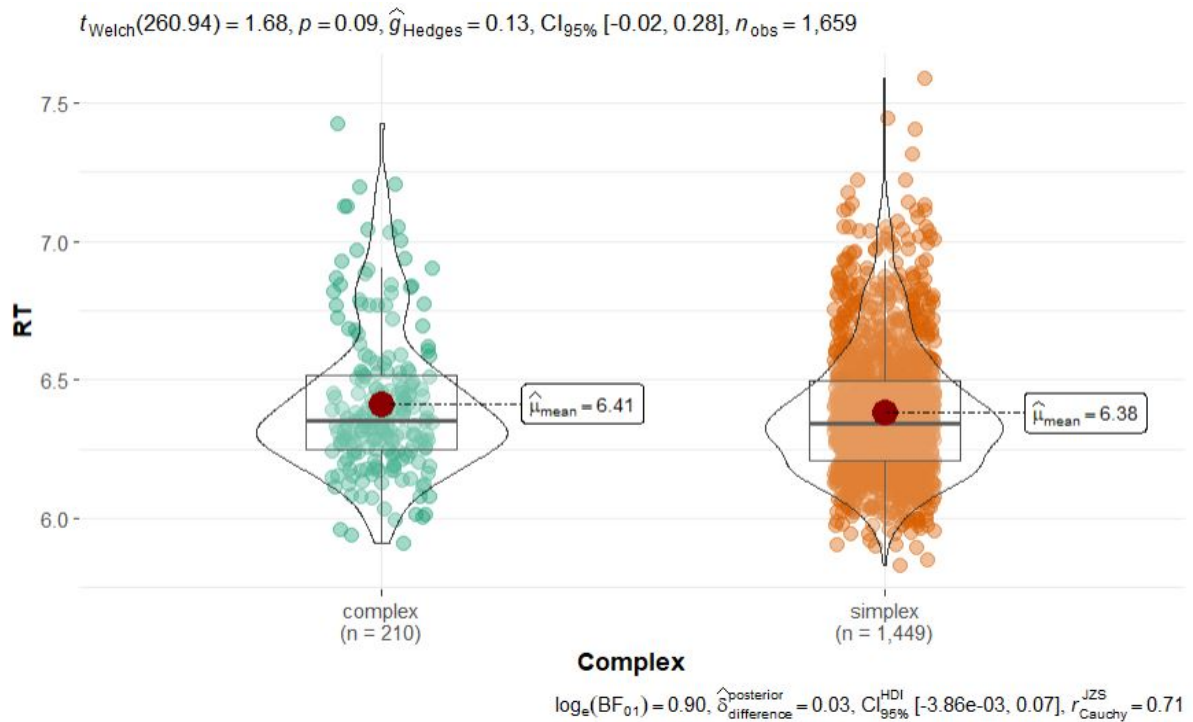


Other



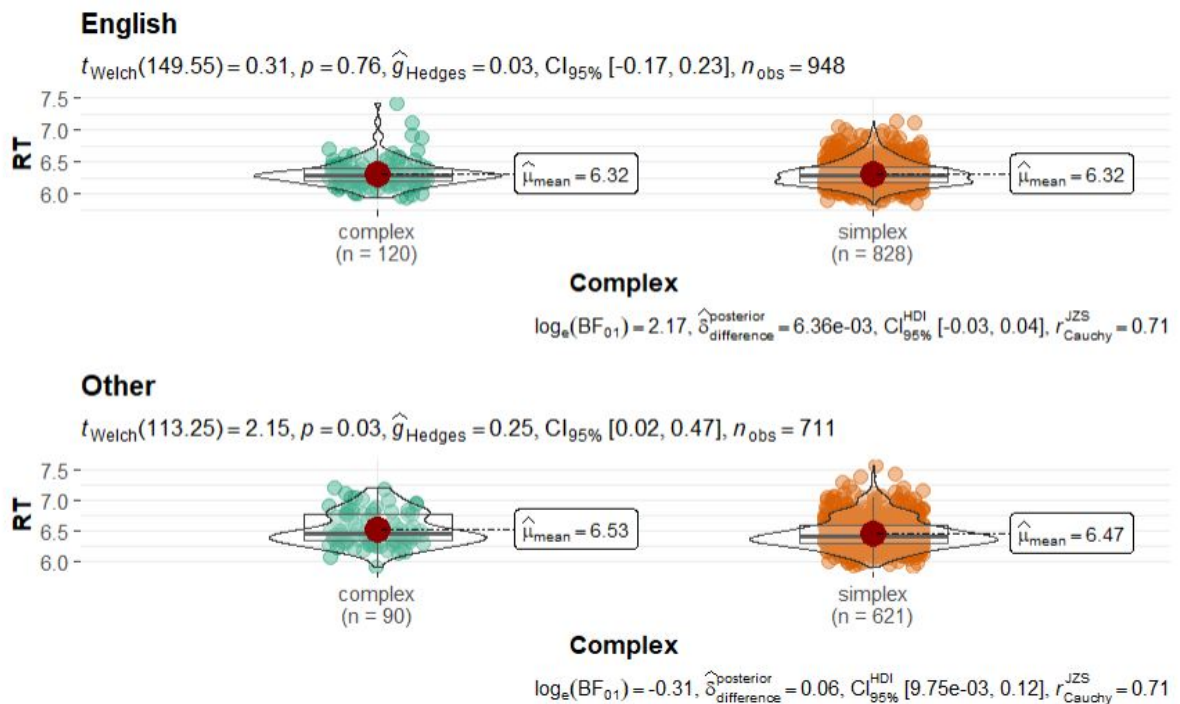
二群の差

■ ggbetweenstats



- ・語の複雑さにより反応時間に、違いがあるとは言えない。
- ・しかし、母語話者と非母語話者のデータが混ざっている...

■ grouped_ggbetweenstats



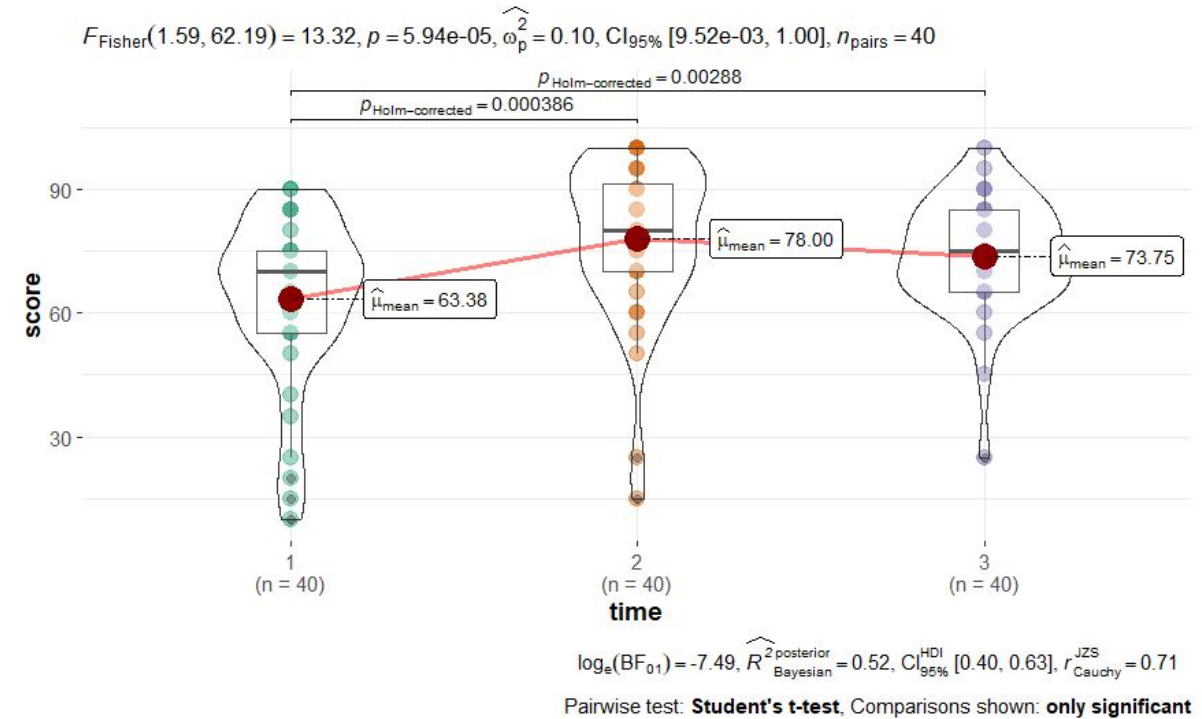
- ・英語母語話者は、 $p=0.76$ で、複雑な語と単純な語の差は、有意ではない。
- ・非母語話者は、 $p=0.03$ で、有意。

二群の差（対応あり）

■ サンプルデータは、杉浦・岩崎 (2003)

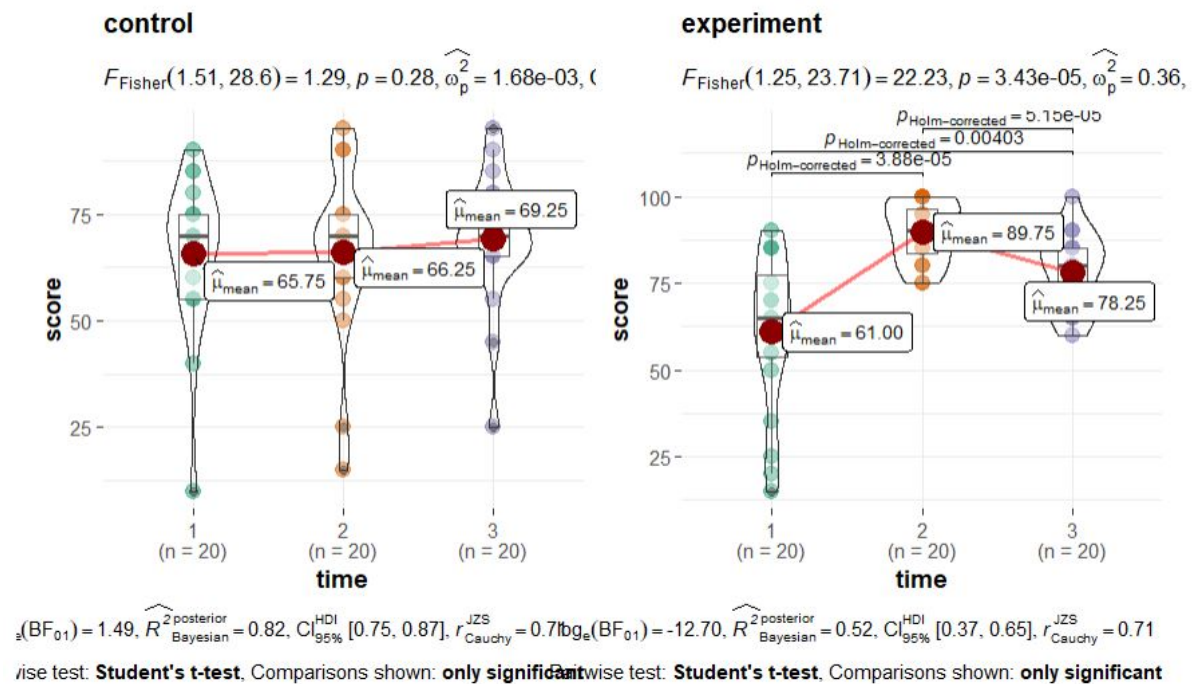
■ ggwithinstats

- ・ 事前、事後、遅延での正答率の差



■ grouped_ggwithinstats

- ・ 統制項目（学習しなかった項目）と実験項目（学習した項目）を分けてみる
- ・ オプションの outlier をつけると、外れ値を表示
- ・ 外れ値のラベルを指定すれば、値ではなく、外れていた人を表示



- ・ 学習しなかった項目は、スコアは変化が見られない。
- ・ 学習した項目は、学習直後にスコアが上がり、遅延テストではスコアが下がる。
 - ・ 下がるといっても学習前よりはスコアは高い。

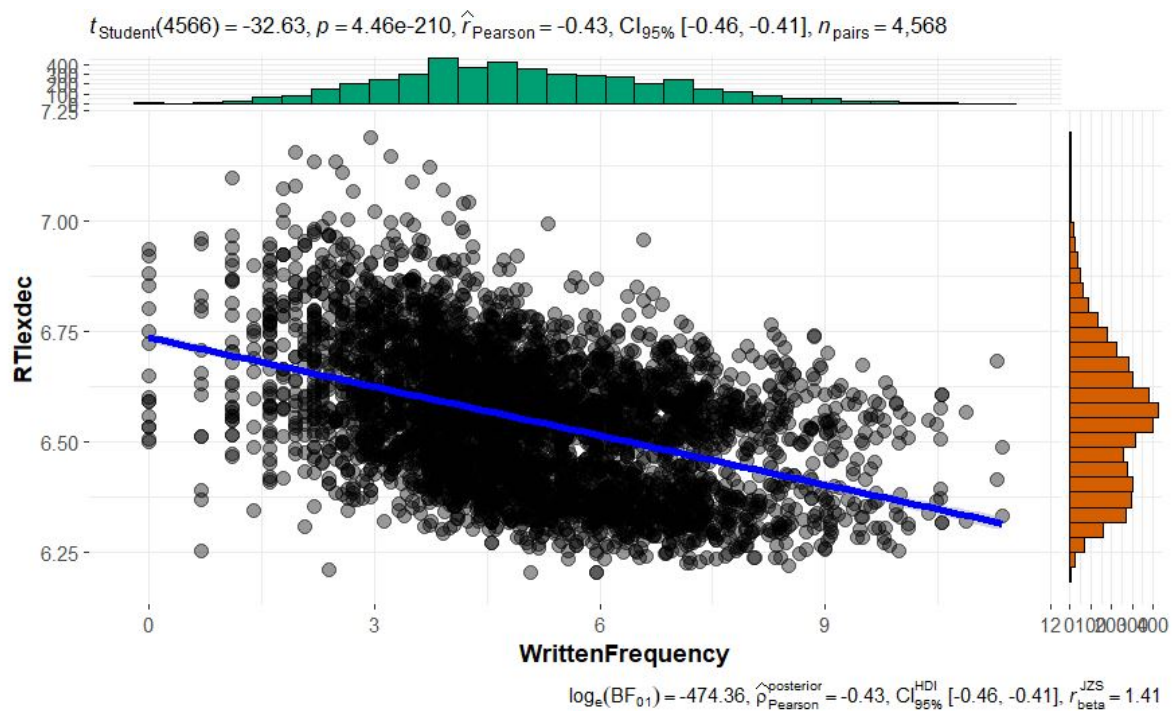
ノンパラメトリックな場合

type = "nonparametric",

を入れるだけ！

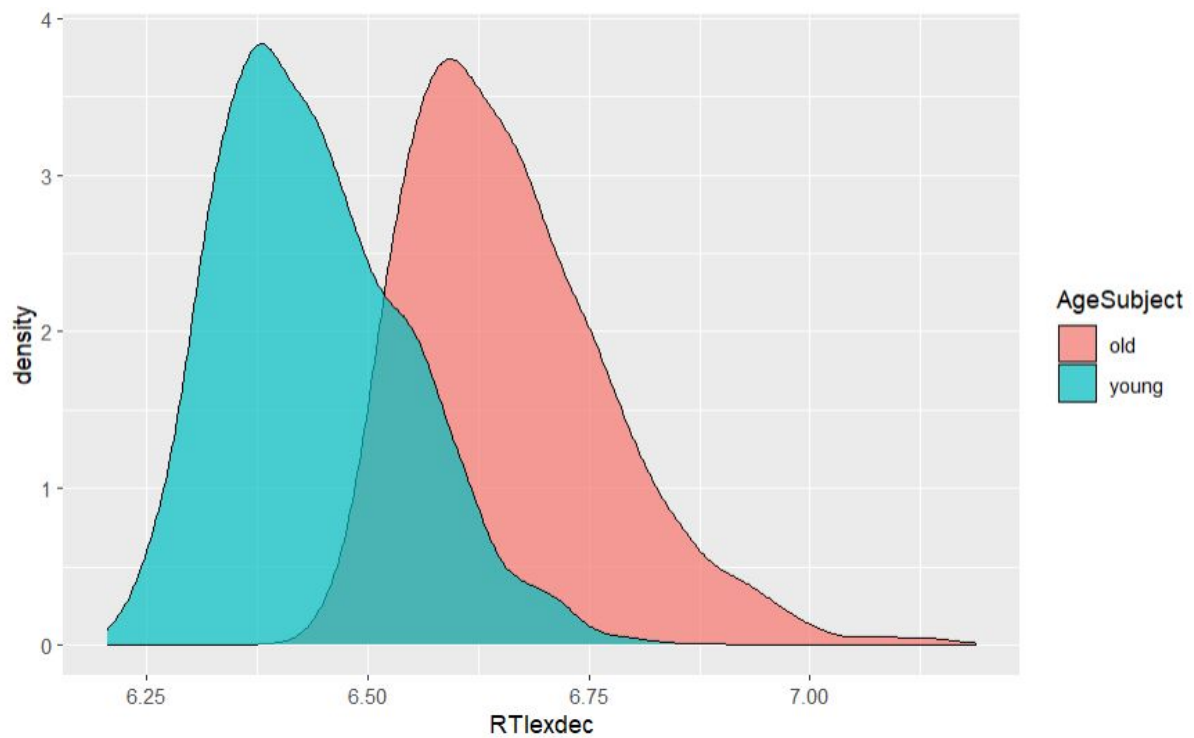
散布図

■ ggscatterstats



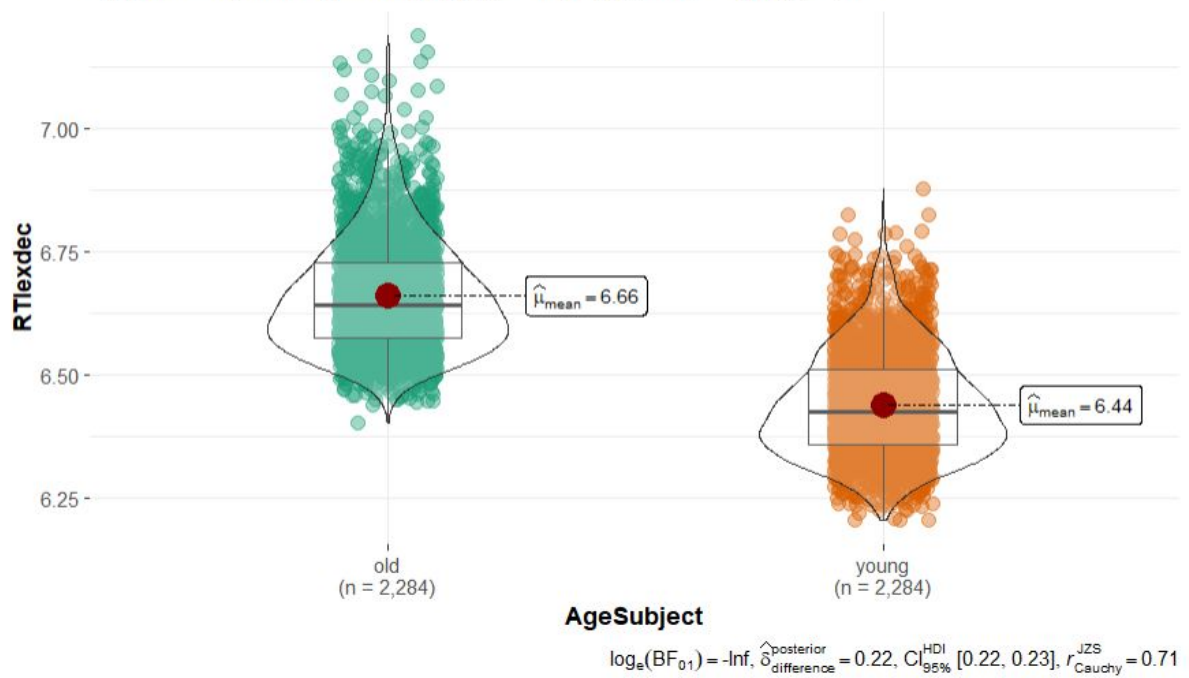
- ・ 相関係数は -0.43
- ・ しかし、どうも反応時間に二つ山があるみたい

■ 年齢 (AgeSubject) が、old と young とある。ひょっとして、年齢で反応時間が違うかも。



■ 二群の差は、ggbetweenstats

$t_{\text{Welch}}(4534.55) = 67.47, p = 0.00, \hat{g}_{\text{Hedges}} = 2.00, \text{CI}_{95\%} [1.76, 2.31], n_{\text{obs}} = 4,568$



■ grouped_ggscatterstats

- ・ 二つのグループに分けてみる

