

R

droplevels()

- ・ factor 型のデータ（カテゴリー変数）で、
- ・ データ処理中に、使われないレベルが残ってしまった場合、
- ・ 使われないレベルを削除する

例：PID のうち 20 で始まるものだけを選んでサブセットを作る場合

■ 元データ

■ 20 で始まる人だけを選ぶ

```
dplyr::filter(startsWith(as.character(PID), "20"))
```

- ・ これだけやると以下ようになる
- ・ PID: Factor に残っている
- ・ 値はないので、データとしては存在しないのに、、、
- ・ これがトラブルのもとになるので、存在しない level を削除する

```
%>% mutate(PID=droplevels(PID))
```

- ・ これで、構造もきれいになる。