# Do you know the {janitor} package?

# Do you know the {janitor} package?

# and `janitor::tabyl()`?

# Do you know the {janitor} package?

# and `janitor::tabyl()`?

## 'A fully-featured alternative to table()'?

# Let's have a look at how to produce some report-ready tables!

```r
library(tidyverse)
```

```r
library(tidyverse)
datasets::Titanic
```

```
, , Age = Child, Survived = No

      Sex
Class  Male Female
  1st     0      0
  2nd     0      0
  3rd    35     17
  Crew    0      0

, , Age = Adult, Survived = No

      Sex
Class  Male Female
  1st   118      4
  2nd   154     13
  3rd   387     89
  Crew  670      3

, , Age = Child, Survived = Yes

      Sex
Class  Male Female
  1st     5      1
  2nd    11     13
  3rd    13     14
  Crew    0      0

, , Age = Adult, Survived = Yes

      Sex
Class  Male Female
  1st    57    140
  2nd    14     80
  3rd    75     76
  Crew  192     20
```

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame()
```

| | Class | Sex | Age | Survived | Freq |
|---|---|---|---|---|---|
| 1 | 1st | Male | Child | No | 0 |
| 2 | 2nd | Male | Child | No | 0 |
| 3 | 3rd | Male | Child | No | 35 |
| 4 | Crew | Male | Child | No | 0 |
| 5 | 1st | Female | Child | No | 0 |
| 6 | 2nd | Female | Child | No | 0 |
| 7 | 3rd | Female | Child | No | 17 |
| 8 | Crew | Female | Child | No | 0 |
| 9 | 1st | Male | Adult | No | 118 |
| 10 | 2nd | Male | Adult | No | 154 |
| 11 | 3rd | Male | Adult | No | 387 |
| 12 | Crew | Male | Adult | No | 670 |
| 13 | 1st | Female | Adult | No | 4 |
| 14 | 2nd | Female | Adult | No | 13 |
| 15 | 3rd | Female | Adult | No | 89 |
| 16 | Crew | Female | Adult | No | 3 |
| 17 | 1st | Male | Child | Yes | 5 |
| 18 | 2nd | Male | Child | Yes | 11 |
| 19 | 3rd | Male | Child | Yes | 13 |
| 20 | Crew | Male | Child | Yes | 0 |
| 21 | 1st | Female | Child | Yes | 1 |
| 22 | 2nd | Female | Child | Yes | 13 |
| 23 | 3rd | Female | Child | Yes | 14 |
| 24 | Crew | Female | Child | Yes | 0 |
| 25 | 1st | Male | Adult | Yes | 57 |
| 26 | 2nd | Male | Adult | Yes | 14 |
| 27 | 3rd | Male | Adult | Yes | 75 |
| 28 | Crew | Male | Adult | Yes | 192 |
| 29 | 1st | Female | Adult | Yes | 140 |
| 30 | 2nd | Female | Adult | Yes | 80 |
| 31 | 3rd | Female | Adult | Yes | 76 |
| 32 | Crew | Female | Adult | Yes | 20 |

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names()
```

| | class | sex | age | survived | freq |
|---|---|---|---|---|---|
| 1 | 1st | Male | Child | No | 0 |
| 2 | 2nd | Male | Child | No | 0 |
| 3 | 3rd | Male | Child | No | 35 |
| 4 | Crew | Male | Child | No | 0 |
| 5 | 1st | Female | Child | No | 0 |
| 6 | 2nd | Female | Child | No | 0 |
| 7 | 3rd | Female | Child | No | 17 |
| 8 | Crew | Female | Child | No | 0 |
| 9 | 1st | Male | Adult | No | 118 |
| 10 | 2nd | Male | Adult | No | 154 |
| 11 | 3rd | Male | Adult | No | 387 |
| 12 | Crew | Male | Adult | No | 670 |
| 13 | 1st | Female | Adult | No | 4 |
| 14 | 2nd | Female | Adult | No | 13 |
| 15 | 3rd | Female | Adult | No | 89 |
| 16 | Crew | Female | Adult | No | 3 |
| 17 | 1st | Male | Child | Yes | 5 |
| 18 | 2nd | Male | Child | Yes | 11 |
| 19 | 3rd | Male | Child | Yes | 13 |
| 20 | Crew | Male | Child | Yes | 0 |
| 21 | 1st | Female | Child | Yes | 1 |
| 22 | 2nd | Female | Child | Yes | 13 |
| 23 | 3rd | Female | Child | Yes | 14 |
| 24 | Crew | Female | Child | Yes | 0 |
| 25 | 1st | Male | Adult | Yes | 57 |
| 26 | 2nd | Male | Adult | Yes | 14 |
| 27 | 3rd | Male | Adult | Yes | 75 |
| 28 | Crew | Male | Adult | Yes | 192 |
| 29 | 1st | Female | Adult | Yes | 140 |
| 30 | 2nd | Female | Adult | Yes | 80 |
| 31 | 3rd | Female | Adult | Yes | 76 |
| 32 | Crew | Female | Adult | Yes | 20 |

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq)
```

```
      class    sex   age survived
3       3rd   Male Child       No
3.1     3rd   Male Child       No
3.2     3rd   Male Child       No
3.3     3rd   Male Child       No
3.4     3rd   Male Child       No
3.5     3rd   Male Child       No
3.6     3rd   Male Child       No
3.7     3rd   Male Child       No
3.8     3rd   Male Child       No
3.9     3rd   Male Child       No
3.10    3rd   Male Child       No
3.11    3rd   Male Child       No
3.12    3rd   Male Child       No
3.13    3rd   Male Child       No
3.14    3rd   Male Child       No
3.15    3rd   Male Child       No
3.16    3rd   Male Child       No
3.17    3rd   Male Child       No
3.18    3rd   Male Child       No
3.19    3rd   Male Child       No
3.20    3rd   Male Child       No
3.21    3rd   Male Child       No
3.22    3rd   Male Child       No
3.23    3rd   Male Child       No
3.24    3rd   Male Child       No
3.25    3rd   Male Child       No
3.26    3rd   Male Child       No
3.27    3rd   Male Child       No
3.28    3rd   Male Child       No
3.29    3rd   Male Child       No
3.30    3rd   Male Child       No
3.31    3rd   Male Child       No
```

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class)
```

```
survived 1st 2nd 3rd Crew
      No 122 167 528  673
     Yes 203 118 178  212
```

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col"))
```

| survived | 1st | 2nd | 3rd | Crew | Total |
|---|---|---|---|---|---|
| No | 122 | 167 | 528 | 673 | 1490 |
| Yes | 203 | 118 | 178 | 212 | 711 |
| Total | 325 | 285 | 706 | 885 | 2201 |

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col")
```

```
survived       1st       2nd       3rd      Crew     Total
      No 0.3753846 0.5859649 0.7478754 0.760452 0.676965
     Yes 0.6246154 0.4140351 0.2521246 0.239548 0.323035
   Total 1.0000000 1.0000000 1.0000000 1.000000 1.000000
```

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col") %>
  janitor::adorn_pct_formatting()
```

| survived | 1st | 2nd | 3rd | Crew | Total |
|---|---|---|---|---|---|
| No | 37.5% | 58.6% | 74.8% | 76.0% | 67.7% |
| Yes | 62.5% | 41.4% | 25.2% | 24.0% | 32.3% |
| Total | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% |

```r
library(tidyverse)
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col") %>
  janitor::adorn_pct_formatting() %>%
# add Ns also known as counts
  janitor::adorn_ns(position = "front")
```

| survived | 1st | 2nd | 3rd | Crew | Total |
|---|---|---|---|---|---|
| No | 122 (37.5%) | 167 (58.6%) | 528 (74.8%) | 673 (76.0%) | 1490 (67.7%) |
| Yes | 203 (62.5%) | 118 (41.4%) | 178 (25.2%) | 212 (24.0%) | 711 (32.3%) |
| Total | 325 (100.0%) | 285 (100.0%) | 706 (100.0%) | 885 (100.0%) | 2201 (100.0%) |

```r
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class, sex) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col") %>
  janitor::adorn_pct_formatting() %>%
  janitor::adorn_ns(position = "front")
```

```
$Male
 survived          1st          2nd          3rd          Crew          Total
       No 118  (65.6%) 154  (86.0%) 422  (82.7%) 670  (77.7%) 1364  (78.8%)
      Yes  62  (34.4%)  25  (14.0%)  88  (17.3%) 192  (22.3%)  367  (21.2%)
    Total 180 (100.0%) 179 (100.0%) 510 (100.0%) 862 (100.0%) 1731 (100.0%)

$Female
 survived          1st          2nd          3rd          Crew          Total
       No   4   (2.8%)  13  (12.3%) 106  (54.1%)  3  (13.0%) 126  (26.8%)
      Yes 141  (97.2%)  93  (87.7%)  90  (45.9%) 20  (87.0%) 344  (73.2%)
    Total 145 (100.0%) 106 (100.0%) 196 (100.0%) 23 (100.0%) 470 (100.0%)
```

```r
datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class, sex) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col") %>
  janitor::adorn_pct_formatting() %>%
  janitor::adorn_ns(position = "front")
```

```
$Male
 survived          1st          2nd          3rd          Crew          Total
       No 118  (65.6%) 154  (86.0%) 422  (82.7%) 670  (77.7%) 1364  (78.8%)
      Yes  62  (34.4%)  25  (14.0%)  88  (17.3%) 192  (22.3%)  367  (21.2%)
    Total 180 (100.0%) 179 (100.0%) 510 (100.0%) 862 (100.0%) 1731 (100.0%)

$Female
 survived          1st          2nd          3rd          Crew          Total
       No   4   (2.8%)  13  (12.3%) 106  (54.1%)  3  (13.0%) 126  (26.8%)
      Yes 141  (97.2%)  93  (87.7%)  90  (45.9%) 20  (87.0%) 344  (73.2%)
    Total 145 (100.0%) 106 (100.0%) 196 (100.0%) 23 (100.0%) 470 (100.0%)
```

# A bit more about how it works.

# A bit more about how it works.

## Let's see tabyl carrying along the raw counts as meta data using `str()` on each step in the same pipeline.

# A bit more about how it works.

**Let's see tabyl carrying along the raw counts as meta data using `str()` on each step in the same pipeline.**

**This allows tabyl() to calculate other quantities of interest, like percentages and percentages on totals wtih its adorn_*() functions.**

```
output <- datasets::Titanic

output     %>%
str()
```

```
'table' num [1:4, 1:2, 1:2, 1:2] 0 0 35 0 0 0 17 0 118 154 ...
- attr(*, "dimnames")=List of 4
 ..$ Class   : chr [1:4] "1st" "2nd" "3rd" "Crew"
 ..$ Sex     : chr [1:2] "Male" "Female"
 ..$ Age     : chr [1:2] "Child" "Adult"
 ..$ Survived: chr [1:2] "No" "Yes"
```

```
output <- datasets::Titanic %>%
  data.frame()

output     %>%
str()
```

```
'data.frame':    32 obs. of  5 variables:
 $ Class   : Factor w/ 4 levels "1st","2nd","3rd",..: 1 2 3 4 1 2 3 4 1 2 ...
 $ Sex     : Factor w/ 2 levels "Male","Female": 1 1 1 1 2 2 2 2 1 1 ...
 $ Age     : Factor w/ 2 levels "Child","Adult": 1 1 1 1 1 1 1 1 2 2 ...
 $ Survived: Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
 $ Freq    : num  0 0 35 0 0 0 17 0 118 154 ...
```

```
output <- datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names()

output     %>%
str()
```

```
'data.frame':    32 obs. of  5 variables:
 $ class   : Factor w/ 4 levels "1st","2nd","3rd",..: 1 2 3 4 1 2 3 4 1 2 ...
 $ sex     : Factor w/ 2 levels "Male","Female": 1 1 1 1 2 2 2 2 1 1 ...
 $ age     : Factor w/ 2 levels "Child","Adult": 1 1 1 1 1 1 1 1 2 2 ...
 $ survived: Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
 $ freq    : num  0 0 35 0 0 0 17 0 118 154 ...
```

```r
output <- datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq)

output     %>%
str()
```

```
'data.frame':    2201 obs. of  4 variables:
 $ class   : Factor w/ 4 levels "1st","2nd","3rd",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ sex     : Factor w/ 2 levels "Male","Female": 1 1 1 1 1 1 1 1 1 1 ...
 $ age     : Factor w/ 2 levels "Child","Adult": 1 1 1 1 1 1 1 1 1 1 ...
 $ survived: Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...
```

```r
output <- datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class)

output     %>%
str()
```

```
Classes 'tabyl' and 'data.frame':    2 obs. of  5 variables:
 $ survived: Factor w/ 2 levels "No","Yes": 1 2
 $ 1st     : num  122 203
 $ 2nd     : num  167 118
 $ 3rd     : num  528 178
 $ Crew    : num  673 212
 - attr(*, "core")='data.frame':    2 obs. of  5 variables:
  ..$ survived: Factor w/ 2 levels "No","Yes": 1 2
  ..$ 1st     : num  122 203
  ..$ 2nd     : num  167 118
  ..$ 3rd     : num  528 178
  ..$ Crew    : num  673 212
 - attr(*, "tabyl_type")= chr "two_way"
 - attr(*, "var_names")=List of 2
  ..$ row: chr "survived"
  ..$ col: chr "class"
```

```
output <- datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col"))

output     %>%
str()
```

```
Classes 'tabyl' and 'data.frame':    3 obs. of  6 variables:
 $ survived: chr  "No" "Yes" "Total"
 $ 1st     : num  122 203 325
 $ 2nd     : num  167 118 285
 $ 3rd     : num  528 178 706
 $ Crew    : num  673 212 885
 $ Total   : num  1490 711 2201
 - attr(*, "core")='data.frame':    2 obs. of  5 variables:
  ..$ survived: Factor w/ 2 levels "No","Yes": 1 2
  ..$ 1st     : num  122 203
  ..$ 2nd     : num  167 118
  ..$ 3rd     : num  528 178
  ..$ Crew    : num  673 212
 - attr(*, "tabyl_type")= chr "two_way"
 - attr(*, "var_names")=List of 2
  ..$ row: chr "survived"
  ..$ col: chr "class"
 - attr(*, "totals")= chr  "row" "col"
```

```r
output <- datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col")

output     %>%
str()
```

```
Classes 'tabyl' and 'data.frame':    3 obs. of  6 variables:
 $ survived: chr  "No" "Yes" "Total"
 $ 1st     : num  0.375 0.625 1
 $ 2nd     : num  0.586 0.414 1
 $ 3rd     : num  0.748 0.252 1
 $ Crew    : num  0.76 0.24 1
 $ Total   : num  0.677 0.323 1
 - attr(*, "core")='data.frame':    2 obs. of  5 variables:
  ..$ survived: Factor w/ 2 levels "No","Yes": 1 2
  ..$ 1st     : num  122 203
  ..$ 2nd     : num  167 118
  ..$ 3rd     : num  528 178
  ..$ Crew    : num  673 212
 - attr(*, "tabyl_type")= chr "two_way"
 - attr(*, "var_names")=List of 2
  ..$ row: chr "survived"
  ..$ col: chr "class"
 - attr(*, "totals")= chr  "row" "col"
```

```r
output <- datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col") %>
  janitor::adorn_pct_formatting()

output     %>%
str()
```

```
Classes 'tabyl' and 'data.frame':    3 obs. of  6 variables:
 $ survived: chr  "No" "Yes" "Total"
 $ 1st     : chr  "37.5%" "62.5%" "100.0%"
 $ 2nd     : chr  "58.6%" "41.4%" "100.0%"
 $ 3rd     : chr  "74.8%" "25.2%" "100.0%"
 $ Crew    : chr  "76.0%" "24.0%" "100.0%"
 $ Total   : chr  "67.7%" "32.3%" "100.0%"
 - attr(*, "core")='data.frame':    2 obs. of  5 variables:
  ..$ survived: Factor w/ 2 levels "No","Yes": 1 2
  ..$ 1st     : num  122 203
  ..$ 2nd     : num  167 118
  ..$ 3rd     : num  528 178
  ..$ Crew    : num  673 212
 - attr(*, "tabyl_type")= chr "two_way"
 - attr(*, "var_names")=List of 2
  ..$ row: chr "survived"
  ..$ col: chr "class"
 - attr(*, "totals")= chr  "row" "col"
```

```r
output <- datasets::Titanic %>%
  data.frame() %>%
  janitor::clean_names() %>%
  tidyr::uncount(weights = freq) %>%
  janitor::tabyl(survived, class) %>%
  janitor::adorn_totals(c("row", "col")) %>%
  janitor::adorn_percentages(denominator = "col") %>
  janitor::adorn_pct_formatting() %>%
  janitor::adorn_ns(position = "front")

output     %>%
str()
```

```
Classes 'tabyl' and 'data.frame':    3 obs. of  6 variables:
 $ survived: chr  "No" "Yes" "Total"
 $ 1st     : chr  "122  (37.5%)" "203  (62.5%)" "325 (100.0%)"
 $ 2nd     : chr  "167  (58.6%)" "118  (41.4%)" "285 (100.0%)"
 $ 3rd     : chr  "528  (74.8%)" "178  (25.2%)" "706 (100.0%)"
 $ Crew    : chr  "673  (76.0%)" "212  (24.0%)" "885 (100.0%)"
 $ Total   : chr  "1490  (67.7%)" " 711  (32.3%)" "2201 (100.0%)"
 - attr(*, "core")='data.frame':    2 obs. of  5 variables:
  ..$ survived: Factor w/ 2 levels "No","Yes": 1 2
  ..$ 1st     : num  122 203
  ..$ 2nd     : num  167 118
  ..$ 3rd     : num  528 178
  ..$ Crew    : num  673 212
 - attr(*, "tabyl_type")= chr "two_way"
 - attr(*, "var_names")=List of 2
  ..$ row: chr "survived"
  ..$ col: chr "class"
 - attr(*, "totals")= chr  "row" "col"
```

**https://github.com/sfirke/janitor**