

RWorksheet_Labanero#6

Carl Dave Labanero BSIT-2B

2023-12-06

```
data <- data.frame(  
  Student = 1:10,  
  Pre_test = c(55, 54, 47, 57, 51, 61, 57, 54, 63, 58),  
  Post_test = c(61, 60, 56, 63, 56, 63, 59, 56, 62, 61)  
)  
data
```

##	Student	Pre_test	Post_test
## 1	1	55	61
## 2	2	54	60
## 3	3	47	56
## 4	4	57	63
## 5	5	51	56
## 6	6	61	63
## 7	7	57	59
## 8	8	54	56
## 9	9	63	62
## 10	10	58	61

```
install.packages("Hmisc")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'  
## (as 'lib' is unspecified)
```

```
library(Hmisc)
```

```
##  
## Attaching package: 'Hmisc'  
## The following objects are masked from 'package:base':  
##  
##      format.pval, units
```

```
install.packages("pastecs")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'  
## (as 'lib' is unspecified)
```

```
library(pastecs)
```

```
stat.desc(data)
```

##	Student	Pre_test	Post_test
## nbr.val	10.0000000	10.00000000	10.00000000
## nbr.null	0.0000000	0.00000000	0.00000000
## nbr.na	0.0000000	0.00000000	0.00000000

```
## min      1.0000000  47.0000000  56.0000000
## max     10.0000000  63.0000000  63.0000000
## range    9.0000000  16.0000000   7.0000000
## sum     55.0000000 557.0000000 597.0000000
## median   5.5000000  56.0000000  60.5000000
## mean     5.5000000 55.7000000 59.7000000
## SE.mean   0.9574271  1.46855938  0.89504811
## CI.mean.0.95 2.1658506  3.32211213  2.02473948
## var      9.1666667 21.56666667  8.01111111
## std.dev   3.0276504  4.64399254  2.83039063
## coef.var   0.5504819  0.08337509  0.04741023
```

```
fertt <- c(10, 10, 10, 20, 20, 50, 10, 20, 10, 50, 20, 50, 20, 10)
```

```
fertilizer_levels <- factor(fertt, levels = c(10, 20, 50), ordered = TRUE)
```

```
fertilizer_levels
```

```
## [1] 10 10 10 20 20 50 10 20 10 50 20 50 20 10
## Levels: 10 < 20 < 50
```

```
exercise_levels <- c("l", "n", "n", "i", "l", "l", "n", "n", "i", "l")
```

```
exercise_factor <- factor(exercise_levels, levels = c("n", "l", "i"), labels = c("none", "light", "intense"))
```

```
exercise_factor
```

```
## [1] light none none intense light light none none intense
## [10] light
## Levels: none light intense
```

```
state <- c("tas", "sa", "qld", "nsw", "nsw", "nt", "wa", "wa", "qld",
          "vic", "nsw", "vic", "qld", "qld", "sa", "tas", "sa", "nt",
          "wa", "vic", "qld", "nsw", "nsw", "wa", "sa", "act", "nsw",
          "vic", "vic", "act")
```

```
state_factor <- factor(state, levels = c("act", "nsw", "nt", "qld", "sa", "tas", "vic", "wa"))
```

```
state_factor
```

```
## [1] tas sa qld nsw nsw nt wa wa qld vic nsw vic qld qld sa tas sa nt wa
## [20] vic qld nsw nsw wa sa act nsw vic vic act
## Levels: act nsw nt qld sa tas vic wa
```

#The factor variable state_factor now categorizes each individual's state of origin into one of these s

```
state <- c("tas", "sa", "qld", "nsw", "nsw", "nt", "wa", "wa", "qld",
          "vic", "nsw", "vic", "qld", "qld", "sa", "tas", "sa", "nt",
          "wa", "vic", "qld", "nsw", "nsw", "wa", "sa", "act", "nsw",
          "vic", "vic", "act")
```

```
incomes <- c(60, 49, 40, 61, 64, 60, 59, 54, 62, 69, 70, 42, 56, 61, 61, 61, 58, 51, 48,
            65, 49, 49, 41, 48, 52, 46, 59, 46, 58, 43)
```

```
state_factor <- factor(state, levels = c("act", "nsw", "nt", "qld", "sa", "tas", "vic", "wa"))
```

```
inc_means <- tapply(incomes, state_factor, mean)
```

```
inc_means
```

```
##      act      nsw      nt      qld      sa      tas      vic      wa
## 44.50000 57.33333 55.50000 53.60000 55.00000 60.50000 56.00000 52.25000
```

#The calculated sample mean incomes for tax accountants across various Australian states highlight income differences.

```
state <- c("tas", "sa", "qld", "nsw", "nsw", "nt", "wa", "wa", "qld",
          "vic", "nsw", "vic", "qld", "qld", "sa", "tas", "sa", "nt",
          "wa", "vic", "qld", "nsw", "nsw", "wa", "sa", "act", "nsw",
          "vic", "vic", "act")
```

```
incomes <- c(60, 49, 40, 61, 64, 60, 59, 54, 62, 69, 70, 42, 56, 61, 61, 61, 58, 51, 48,
            65, 49, 49, 41, 48, 52, 46, 59, 46, 58, 43)
```

```
state_factor <- factor(state, levels = c("act", "nsw", "nt", "qld", "sa", "tas", "vic", "wa"))
```

```
stdError <- function(x) sqrt(var(x) / length(x))
```

```
inc_std_errors <- tapply(incomes, state_factor, stdError)
```

```
inc_std_errors
```

```
##      act      nsw      nt      qld      sa      tas      vic      wa
## 1.500000 4.310195 4.500000 4.106093 2.738613 0.500000 5.244044 2.657536
```

#The variability or uncertainty in sample means is measured by standard errors. Here, the standard error for each state is shown.

```
#install.packages("datasets")
```

```
library(datasets)
```

```
data("Titanic")
```