DAY ONE

YOU ARE THE AUTHOR

2023-10-11

Contents

MATRIX

Matrix Creation

```
A = matrix(c(5,6,2,8,9,2,4,5,1),ncol = 3, nrow = 3, byrow = T)
A
```

```
     [,1] [,2] [,3]
[1,]    5    6    2
[2,]    8    9    2
[3,]    4    5    1
```

Getting the determinant of a matrix

```
det(A)
```

```
[1] 3
```

Inverse

```
solve(A)
```

```
           [,1]       [,2] [,3]
[1,] -0.3333333  1.3333333   -2
[2,]  0.0000000 -1.0000000    2
[3,]  1.3333333 -0.3333333   -1
```

Matrix Operation

```
B <- matrix(c(3,4,6,2,3,4,5,2,5), ncol = 3, nrow = 3, byrow = T)
C <- matrix(c(8,5,3,2,3,5,9,3,3), ncol = 3, nrow = 3, byrow = T)
```

View the matrix

```
B
```

```
     [,1] [,2] [,3]
[1,]    3    4    6
[2,]    2    3    4
[3,]    5    2    5
```

```
C
```

```
     [,1] [,2] [,3]
[1,]    8    5    3
[2,]    2    3    5
[3,]    9    3    3
```

Matrix Addition

```
B+C
```

```
     [,1] [,2] [,3]
[1,]   11    9    9
[2,]    4    6    9
[3,]   14    5    8
```

Matrix Subtraction

```
B-C
```

```
     [,1] [,2] [,3]
[1,]   -5   -1    3
[2,]    0    0   -1
[3,]   -4   -1    2
```

Matrix Division

```
B/C
```

```
          [,1]      [,2]     [,3]
[1,] 0.3750000 0.8000000 2.000000
[2,] 1.0000000 1.0000000 0.800000
[3,] 0.5555556 0.6666667 1.666667
```

Matrix Multiplication

```
B*C
```

```
     [,1] [,2] [,3]
[1,]   24   20   18
[2,]    4    9   20
[3,]   45    6   15
```

```
B%*%C
```

```
     [,1] [,2] [,3]
[1,]   86   45   47
[2,]   58   31   33
[3,]   89   46   40
```

Getting the Identity Matrix

```
zapsmall(solve(A)%*%A)
```

```
     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1
```

Mathematical Operations

Addition

```
y = 45+65
y
```

```
[1] 110
```

Subtraction

```
x = 563-546
x
```

```
[1] 17
```

Division

```
m = 563/87
m
```

```
[1] 6.471264
```

Multiplication

```r
t = 56*56
t
```

```
[1] 3136
```

Squares and Square roots

```r
sqrt(81)
```

```
[1] 9
```

```r
sqrt(225)
```

```
[1] 15
```

```r
225^0.5
```

```
[1] 15
```

```r
5^2
```

```
[1] 25
```

Exponentials and Logarithmic

```r
log10(100)
```

```
[1] 2
```

To be checked!!!!!!

```r
exp(2)
```

```
[1] 7.389056
```

Data Importation (Comma Seperated Values, csv)

```r
data <- read.csv("Gapminder.csv")
head(data,5)
```

```
      country year population continent life_exp  gdp_cap ln_population
1 Afghanistan 1952    8425333      Asia   28.801 779.4453      6.925587
2 Afghanistan 1957    9240934      Asia   30.332 820.8530      6.965716
3 Afghanistan 1962   10267083      Asia   31.997 853.1007      7.011447
4 Afghanistan 1967   11537966      Asia   34.020 836.1971      7.062129
5 Afghanistan 1972   13079460      Asia   36.088 739.9811      7.116590
  ln_life_exp ln_gdpPercap
1    1.459408     6.658583
2    1.481901     6.710344
3    1.505109     6.748878
4    1.531734     6.728864
5    1.557363     6.606625
```

```
tail(data,5)
```

```
      country year population continent life_exp  gdp_cap ln_population
1700 Zimbabwe 1987    9216418    Africa   62.351 706.1573      6.964562
1701 Zimbabwe 1992   10704340    Africa   60.377 693.4208      7.029560
1702 Zimbabwe 1997   11404948    Africa   46.809 792.4500      7.057093
1703 Zimbabwe 2002   11926563    Africa   39.989 672.0386      7.076515
1704 Zimbabwe 2007   12311143    Africa   43.487 469.7093      7.090298
     ln_life_exp ln_gdpPercap
1700    1.794843     6.559838
1701    1.780872     6.541637
1702    1.670329     6.675129
1703    1.601941     6.510316
1704    1.638359     6.152114
```

Check the structure of the data

```
str(data)
```

```
'data.frame':   1704 obs. of  9 variables:
 $ country      : chr  "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan" ...
 $ year         : int  1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 ...
 $ population   : int  8425333 9240934 10267083 11537966 13079460 14880372 12881816 13867957 16317921 2
 $ continent    : chr  "Asia" "Asia" "Asia" "Asia" ...
 $ life_exp     : num  28.8 30.3 32 34 36.1 ...
 $ gdp_cap      : num  779 821 853 836 740 ...
 $ ln_population: num  6.93 6.97 7.01 7.06 7.12 ...
 $ ln_life_exp  : num  1.46 1.48 1.51 1.53 1.56 ...
 $ ln_gdpPercap : num  6.66 6.71 6.75 6.73 6.61 ...
```

Manual Data Entry

```
age <- c(45,65,34,32,23,25,56,76,45,22,21,45,34,56,54)
age
```

```
 [1] 45 65 34 32 23 25 56 76 45 22 21 45 34 56 54
```

```r
height <- c(122,134,144,165,155,133,123,132,145,154,166,134,121,154,165)
height
```

```
 [1] 122 134 144 165 155 133 123 132 145 154 166 134 121 154 165
```

Column Binding

```r
height_age <- cbind(age, height)
height_age
```

```
      age height
 [1,]  45    122
 [2,]  65    134
 [3,]  34    144
 [4,]  32    165
 [5,]  23    155
 [6,]  25    133
 [7,]  56    123
 [8,]  76    132
 [9,]  45    145
[10,]  22    154
[11,]  21    166
[12,]  45    134
[13,]  34    121
[14,]  56    154
[15,]  54    165
```

Data Framing

```r
mydata <- data.frame(age, height)
head(mydata,5)
```

```
  age height
1  45    122
2  65    134
3  34    144
4  32    165
5  23    155
```

Descriptive Statistics

```r
library(stargazer)
library(gtsummary)
stargazer(data[,-2], type = "text")
```

```
===================================================================
Statistic       N       Mean         St. Dev.       Min        Max
-------------------------------------------------------------------
population    1,704 29,601,212.000 106,157,897.000 60,011  1,318,683,096
life_exp      1,704     59.474         12.917      23.599     82.603
gdp_cap       1,704   7,215.327       9,857.455    241.166  113,523.100
ln_population 1,704     6.847          0.697       4.778      9.120
ln_life_exp   1,704     1.763          0.101       1.373      1.917
ln_gdpPercap  1,704     8.159          1.241       5.485     11.640
-------------------------------------------------------------------
```

Additional Way of Displaying Summary Statistics.

```r
### Load the libraries
library(ggplot2)
library(devtools)
library(predict3d)
library(psych)
library(dplyr)
library(gtsummary)
library(DescTools)
library(nortest)
library(lmtest)
library(sandwich)
```

Display the Summary Statistics

```r
knitr::kable(
  describeBy(data[,-1]) %>% round(2)
)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
year	1	1704	1979.50	17.27	1979.50	1979.50	22.24	1952.00	2.007000e+03	5.500000e+01	0.00	-1.22	0.42
population	2	1704	29601212.32	106157896.74	7023595.50	11399459.45	7841473.62	60011.00	1.318683e+09	1.318623e+09	8.33	77.62	2571683.45
continent*	3	1704	2.33	1.21	2.00	2.27	1.48	1.00	5.000000e+00	4.000000e+00	0.25	-1.34	0.03
life_exp	4	1704	59.47	12.92	60.71	59.92	16.10	23.60	8.260000e+01	5.901000e+01	-0.25	-1.13	0.31
gdp_cap	5	1704	7215.33	9857.45	3531.85	5221.44	4007.61	241.17	1.135231e+05	1.132820e+05	3.85	27.40	238.80
ln_population	6	1704	6.85	0.70	6.85	6.85	0.62	4.78	9.120000e+00	4.340000e+00	0.08	0.47	0.02
ln_life_exp	7	1704	1.76	0.10	1.78	1.77	0.11	1.37	1.920000e+00	5.400000e-01	-0.57	-0.66	0.00
ln_gdpPercap	8	1704	8.16	1.24	8.17	8.14	1.51	5.49	1.164000e+01	6.150000e+00	0.11	-0.95	0.03