01 Basic R for Finance

Boni

9/1/2020

R Basic

Working with Matrix

We can create matrix from one atomic vector.

```
x_vec <- c(9, 4, 6, 20, 19, 29)
x_mat <- matrix(data = x_vec, nrow = 2, ncol = 3, byrow = TRUE)
x_mat
```

```
##      [,1] [,2] [,3]
## [1,]    9    4    6
## [2,]   20   19   29
```

```
cor(x_mat)
```

```
##      [,1] [,2] [,3]
## [1,]    1    1    1
## [2,]    1    1    1
## [3,]    1    1    1
```

We can also create matrix from 2 vectors.

```
y_vec = rnorm(6, 23, 98)

matrix_yx <- cbind(x_vec, y_vec)
matrix_yx
```

```
##      x_vec     y_vec
## [1,]     9  18.36635
## [2,]     4 132.20773
## [3,]     6  15.55785
## [4,]    20 -55.87040
## [5,]    19  31.67186
## [6,]    29 137.86723
```

```r
cor(matrix_yx)
```

```
##            x_vec      y_vec
## x_vec 1.00000000 0.05831526
## y_vec 0.05831526 1.00000000
```

Working with DataFrame

```r
cash_flow <- rnorm(10, 4, 9)
year <- runif(10, 5, 45)
company <- c("Google", "Google", "Microsoft", "Microsoft", "Apple", "Apple", "Google", "Google", "Micros
company_data <- data.frame(company, cash_flow, year)
company_data
```

```
##      company  cash_flow      year
## 1     Google  1.9677696  8.332475
## 2     Google  1.3755069 39.351794
## 3  Microsoft 15.5039733 30.910408
## 4  Microsoft -6.0063092 39.845423
## 5      Apple 10.6872099 40.381282
## 6      Apple -7.1780426 25.087020
## 7     Google  0.2780684 20.642924
## 8     Google 23.3630900 15.429620
## 9  Microsoft -0.1574589 28.228571
## 10 Microsoft  7.5155508 21.333907
```

```r
# sub-setting the DF
company_data[1:3,1, drop=FALSE]
```

```
##     company
## 1    Google
## 2    Google
## 3 Microsoft
```

```r
company_data$cash_flow
```

```
##  [1]  1.9677696  1.3755069 15.5039733 -6.0063092 10.6872099 -7.1780426
##  [7]  0.2780684 23.3630900 -0.1574589  7.5155508
```

```r
subset(company_data, cash_flow < 6.00)
```

```
##     company  cash_flow      year
## 1    Google  1.9677696  8.332475
## 2    Google  1.3755069 39.351794
## 4 Microsoft -6.0063092 39.845423
## 6     Apple -7.1780426 25.087020
## 7    Google  0.2780684 20.642924
## 9 Microsoft -0.1574589 28.228571
```

```r
# Delete a column
company_data$year = NULL
company_data
```

```
##       company  cash_flow
## 1      Google  1.9677696
## 2      Google  1.3755069
## 3   Microsoft 15.5039733
## 4   Microsoft -6.0063092
## 5       Apple 10.6872099
## 6       Apple -7.1780426
## 7      Google  0.2780684
## 8      Google 23.3630900
## 9   Microsoft -0.1574589
## 10  Microsoft  7.5155508
```

Working with Factor

```r
# Factor will make R treat string as integer.
investment <- c("stock", "bonds", "stock", "bonds", "stock")
investment_factor <- factor(investment)
# Summary of factor is more informative.
summary(investment)
```

```
##    Length     Class      Mode
##         5 character character
```

```r
summary(investment_factor)
```

```
## bonds stock
##     2     3
```

```r
# Change the levels.
investment
```

```
## [1] "stock" "bonds" "stock" "bonds" "stock"
```

```r
levels(investment) <- c("B", "S")
investment
```

```
## [1] "stock" "bonds" "stock" "bonds" "stock"
## attr(,"levels")
## [1] "B" "S"
```

```r
# cut() will factoring the numbers based on interval.
price <- runif(100, 20, 100)
breaks <- c(0, 20, 40, 60, 80, 100)
ranking_grouped <- cut(price, breaks = breaks)
head(ranking_grouped)
```

```
## [1] (80,100] (80,100] (20,40]  (20,40]  (80,100] (60,80]
## Levels: (0,20] (20,40] (40,60] (60,80] (80,100]
```

```r
levels(ranking_grouped) <- c("very_low", "low", "medium", "high", "very_high")
# Plot the factor
plot(ranking_grouped)
```