Day2

R 조건문 : 프로그래밍의 아주 기본적인 문법!

“순서도” Sequence. Algorithm!

Thinking  
R 주석처리방법  
#, Ctrl+Shift+C  
또는 Comment = “ 주석~ “ 형태 이용

자료구조(Data structure)  
C언어에서는 매우 중요한 개념.  
Python, R에서는 어떤가? 저장장치 가격 하락, CPU성능/저장용량 상승…  
중요도가 떨어졌었지만, 여전히 중요하다! 고속처리, 대용량 처리… 빅데이터에서 중요한듯?  
자료를 어떻게 넣을 것인가에 관한 개념. 자리를 어떻게 구성할 것인가.  
Java, python은 자료구조의 핵심 개념을 전부 다루지 않는다고…  
C, C++에서의 자료구조가 핵심이라고.

R 데이터 유형과 데이터 구조 구분하기  
Numaric/ Chracter/ Null, …

R의 7가지 자료구조

1. Scalar
2. Vector
3. Matrix
4. Array
5. Data Frame
6. List
7. Factor(요인?)

Tip : Python pandas는 Vector와 Data Frame를 다룬다.(즉 이 둘이 매우 중요)

1. Scalar스칼라  
   자료가 1개인 자료구조  
   자료 구조보다 유형이 중요하다. (실수, 정수, 문자, 불대수, …)
2. Vector벡터  
   # Vector  
   # 데이터가 한 개의 열로만 나열될 수 있다.  
   # 원소는 같은 data type을 가져야 한다.

####### Confusing point #######  
# 스칼라, 벡터, 리스트, 데이터프레임, ... => Data Structure  
# integer, numeric, character, logical, ... => Data type  
# R에서 Vector는 Indexing이 1부터 시작한다. 주의해야함!  
Python의 List Indexing과 혼동하면 안됨!  
# Vector의 이름(?)도 정해줄 수 있다…!

#Naming은 일괄적으로 가능하다. 데이터 대입기능이 다른 언어에서 보던 것과 매우 다름  
#List가 Vector 보다 흔하지만, Vector가 처리속도가 List 보다 매우 빠르다고함.  
#Vector는 Database에서 더 유용하다고 함.

c( )의 c는 Combine을 의미

#문자 sequence도 만들 수 있다?

#한글 작업용은 따로 받아야함.

#R에서 한글 깨질 때 : Tool > Project > UTF-8...?

#뭐 구글링하면 방법이 나오겠지

1. Matrix  
   # 계란판을 연상 - 엑셀과 유사  
   # 같은type의 데이터가 들어가야함.
2. Array  
   # 여러층의 Matrix 형태.  
   # 모든 원소의 data type은 동일하다.  
   # Data 분석때 많이 쓰지는 않는다고.   
   # pandas에 배열과 비슷한게 있지만, '기본' Python에는 배열이 존재하지 않는다.  
   # 데이터 분석할 때 가치가 있다.  
   # 3차원은 사용이 어렵다. -> 범용성은 떨어진다.
3. Data Frame  
   # 이제부터 다른 종류의 Data type을 다룰 수 있다.  
   # 가장 많이 사용하는 데이터 형태!  
   # 엑셀과 유사한 형태. 데이터의 유형과 상관없으며  
   # 2차원 형태를 가짐.  
   # Qeustion : 심도는 낮을까? DF는 array처럼 3차원을 구현할 수 없을까?
4. List  
   Array보다 처리속도가 ‘느리다’ -> **어떻게, 왜?**List에는 1~4번까지 모두 들어갈 수 있다. (Data Frame까지!)
5. Factor  
   # 담겨져 있는 데이터가 어떻게 구성되어 있는가?  
   # 요인들이 어떻게 구성되어 있는가?  
   # 순위를 가질 수 있고, 순위가 없을 수 있다.  
   # 지금까지 Scalar를 제외하고 모두 순서가 있었다.  
   # Factor는 문자처럼 보이지만 실제로는 숫자이다.(??)  
   # 수준(level)이라고 알려진 사전에 정의된 값만 다룰 수 있다.  
   데이터가 어떻게 들어가있는지에 대한 물음에 답을 하기 위한 자료구조!

Problem : 왜 linter, 추천?, 자동완성 기능이 작동하지 않지??? 다른 옵션이 뭐 있는지 자동으로 보고싶은데?

예약어 색은 어떻게 바꾸는가?

##Confusing Check  
data\_frame은 rbind, cbind로 합치고  
vector는 c(combine), union으로 합치고  
그렇다면, Matrix, Array는?

#굳이 c 남용할 필요가 없는거 같은데? 왜 1개 인자에 대해 c를 쓰는거지? 나중에 Vector~Matrix쪽 복습하면서 찾아봐야함!

#Thinking : Clear environment : rm(list = ls()) -> 콘솔창에  
<https://community.rstudio.com/t/how-to-clear-the-r-environment/14303>

Integer(0) ??  
<https://stackoverflow.com/questions/6451152/how-to-catch-integer0>

3dim of Data Frame  
<https://stackoverflow.com/questions/32766990/creating-a-three-dimensional-array-from-a-data-frame-in-r>

class(), typeof(), mode()  
<https://buillee.tistory.com/83>

R data type(자료형, 데이터 유형)  
<https://nittaku.tistory.com/329>

R 반복문  
<https://ordo.tistory.com/46>

R Tutorial  
<https://www.tutorialspoint.com/r/r_operators.htm>

R 제어문  
<https://kyun2.tistory.com/65>