Python ↔ R Core Data Structures Cheat Sheet

Numeric Vectors / Arrays

R	Python
x <- c(1,2,3,4)	x = np.array([1,2,3,4])
x * 2	x * 2
x[x > 2]	x[x > 2]

Lists / Dictionaries

R	Python
list(name='Alice', age=30)	{'name':'Alice','age':30}

Matrices

R	Python
matrix(1:6, nrow=2)	np.array([[1,2,3],[4,5,6]])
m %*% t(m)	m @ m.T

Data Frames

R	Python
data.frame(name=c('A','B'))	pd.DataFrame({'name':['A','B']})
df$age	df['age']
subset(df, age>24)	df[df['age']>24]
df$new <- df$x+df$y	df['new']=df['x']+df['y']

Categoricals / Factors

R	Python
factor(c('M','F'))	pd.Categorical(['M','F'])

Sequences

R	Python
seq(0,10,by=2)	np.arange(0,11,2)

Missing Values

R	Python
NA	np.nan
is.na(x)	np.isnan(x)

Inspecting Data

R	Python
dim(df)	df.shape
str(df)	df.info()
summary(df)	df.describe()