Lecture 1 - Introduction to R

Code	Description	Examples
Basic Operators + - * / ^	Add, subtract, multiply, divide, exponentiation	1+1 2 3^2 9
Parentheses ()	Specify order of operations	(1+2)*2 6 1+(2*2) 5
Assignment = <-	Assign a value to a variable (object) <- and = are the same	a <- 24/6 a 4
Boolean Operators > <= >= == != !()	Make logical tests Less then Greater than Less than or equal to Greater than or equal to Equal to Not equal to Not (negates)	1 < 2 TRUE 1 > 2 FALSE 3 <= 3 TRUE 1 >= 2 FALSE TRUE == FALSE FALSE 1 != 2 TRUE !(FALSE) TRUE
Math Functions sin() exp() log() abs()	Just some examples: trigonometric sine exponentiation: e^() natural log absolute value	sin(pi/2) 1 exp(1) 2.718… log(1) 0 abs(-10) 10
Importing data read.csv()	Import a .csv file (create using save as… in excel and choosing .csv as the file type)	births <- read.csv("C:/path/to/file/births.csv")
Referencing variables $ []	Reference variable in a dataset Reference an observation in a variable	births$GEST biths$GEST[2] (second obs. value)

Lecture 2 – R Programming I

Code	Description	Examples
Inspect a dataset		
View()	View as a spreadsheet	View(births)
names()	Print variable names	names(births)
dim()	Print the number of observations, variables	dim(births) [1] 122513 117
Inspect a variable		
$	Reference the variable	births$GEST
summary()	Get basic statistics	summary(births$GEST)
mean()	Find other stats	mean(births$GEST)
sd()		
range()		
var()		
quantile()		
sum()		
table()	Make a table	table(births$GEST)
length()	Find # of observations	length(births$GEST)
Inspect objects		
class()	Find object's type	class(births) class(letters)
str()	Print FULL object structure	str(births)
Simple plots		
hist()	histogram distribution	hist(births$GEST)
boxplot()	box-and-whiskers plot	boxplot(births$GEST)
barplot(table())	make a bar graph (of a table)	barplot(table(births$GEST))
Subsetting: [BOOLEAN EXP]	Subset a variable based on the value of a Boolean expression:	G <- births$GEST W <- births$WIC
x[x== …]	X WHERE X is …	G[G==99] length(G[G==99])
y[y== …]	X WHERE Y is …	G[W=="Y"] mean(G[W=="Y"] mean(G[W=="N"]

Lecture 3 – R Programming II

Code	Description	Examples
Using Functions f(a=, b=)	Use function f with arguments a and b set	Hist(births$GEST, col="blue")
Create a sequence : seq(from=, to=, by=) rep(x=, times=) c()	Sequence by 1 Specified sequence Repeated sequence Arbitrary sequence	1:10 seq(from=1,to=100, by=5) seq(x=5, times=10) c(1,1,2,3,5,8,13)
Recoding x[x == …] <- VALUE	Replace specific values of a variable WHERE a Boolean expression is true.	G <- births$GEST G[G==99] = NA G[G >=52] = 42
Classify a variable X[y == …] <- VALUE	Create a categorical variable from a numeric variable. (First create a blank variable, then fill it in using recoding)	GC <- rep(NA,times= length(G)) GC[G<20] <- "low" GC[G>=20 & G<40] <- "medium" GC[G>=40] <- "high"
Build a dataset data.frame()	Create a data frame from multiple variables (of the same length!)	exposure <- c(1,1,0) outcome <- c(1,0,0) dat <- data.frame(exposure,outcome)
Missing data NA is.na()	Missing data value is the value missing?	#Sanitize missing: x[is.na(x)] <- 0
Save and Load Data setwd() save(x, file=) load(x) write.csv(x, file=)	Point R to a folder Save in RData format Load from RData format Save as a .csv	setwd("D:/folder") save(births,file= "births.RData") load("births.RData") write.csv(births, file="births.csv")