Practical Aspects of Database Design

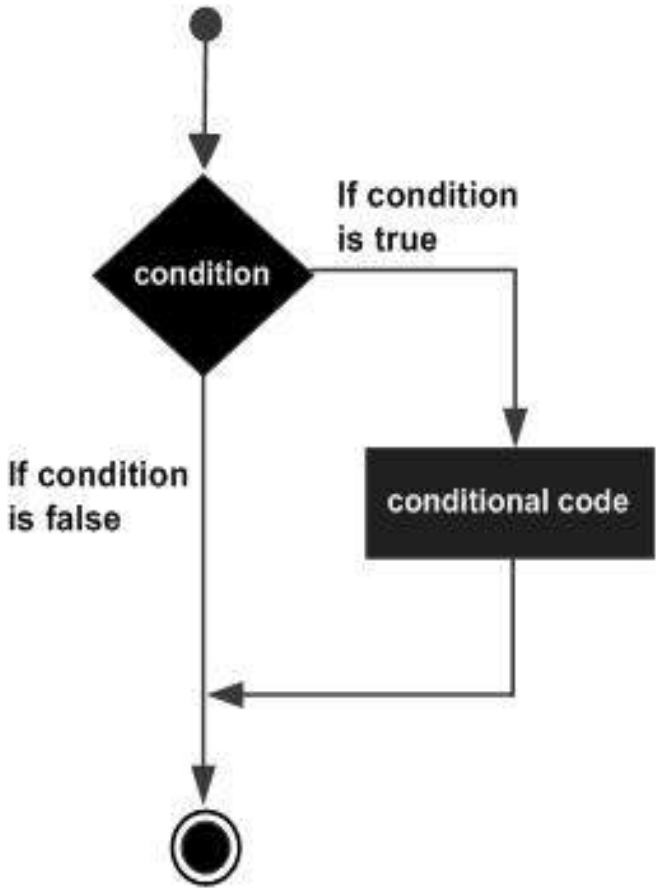
L3 - R Session II

Stevens Institute of Technology
Fall, 2017

Decision Making

If statement

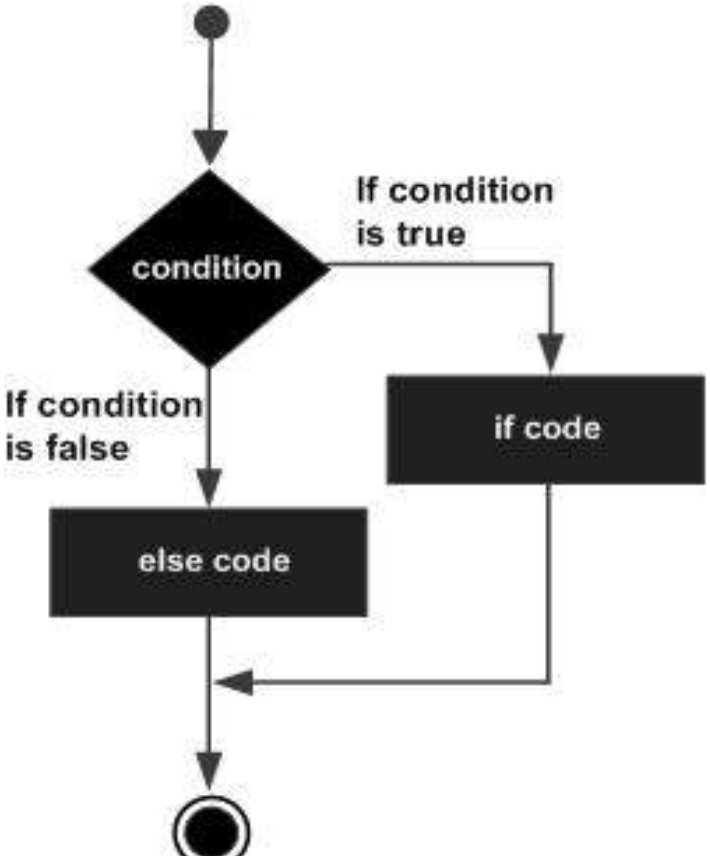
```
if(boolean expression){
// statement(s) will execute if the boolean expression is true.
}
```



Decision Making

If else statement

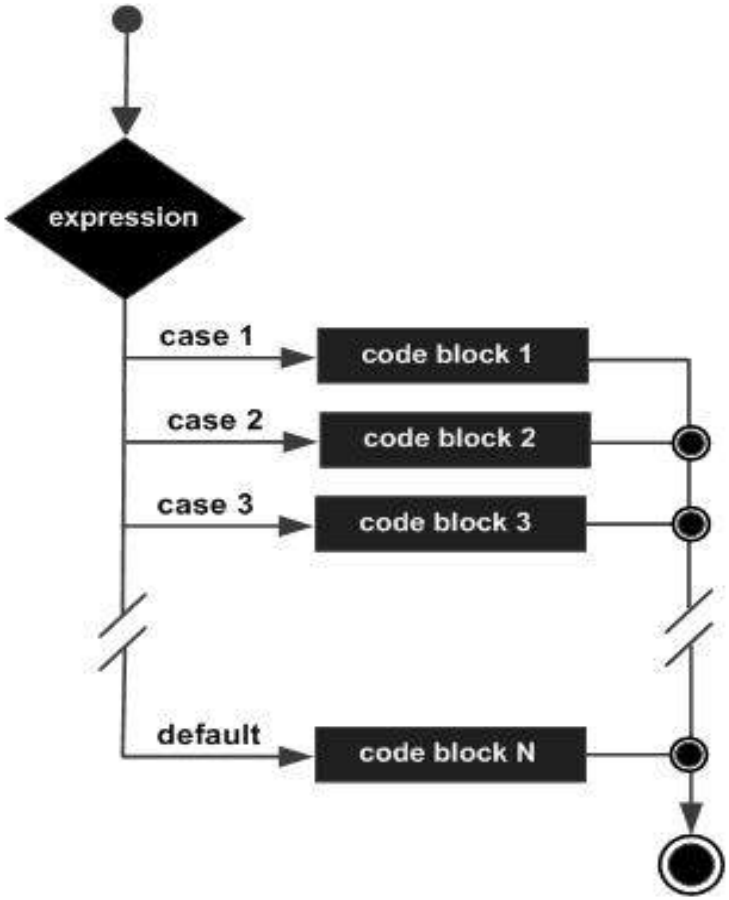
```
if(boolean expression) {
// statement(s) will execute if the boolean expression is true. }
else {
// statement(s) will execute if the boolean expression is false.
```



Decision Making

Switch statement

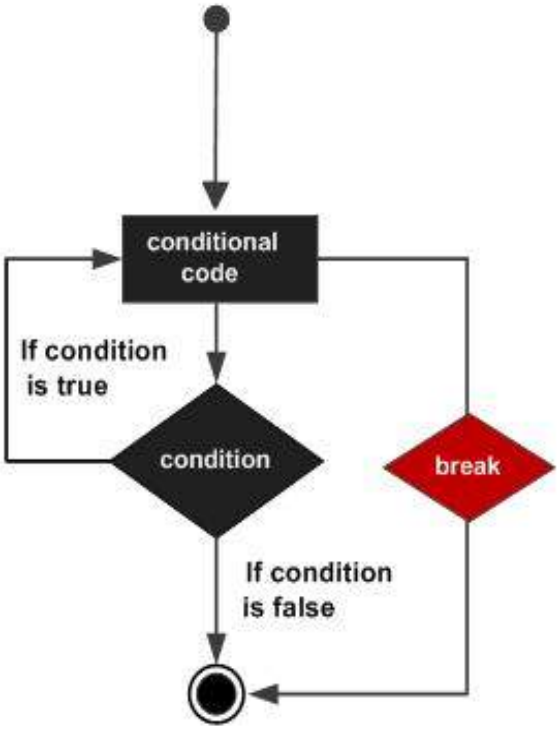
switch(expression, case1, case2, case3....)



Loop

Repeat loop

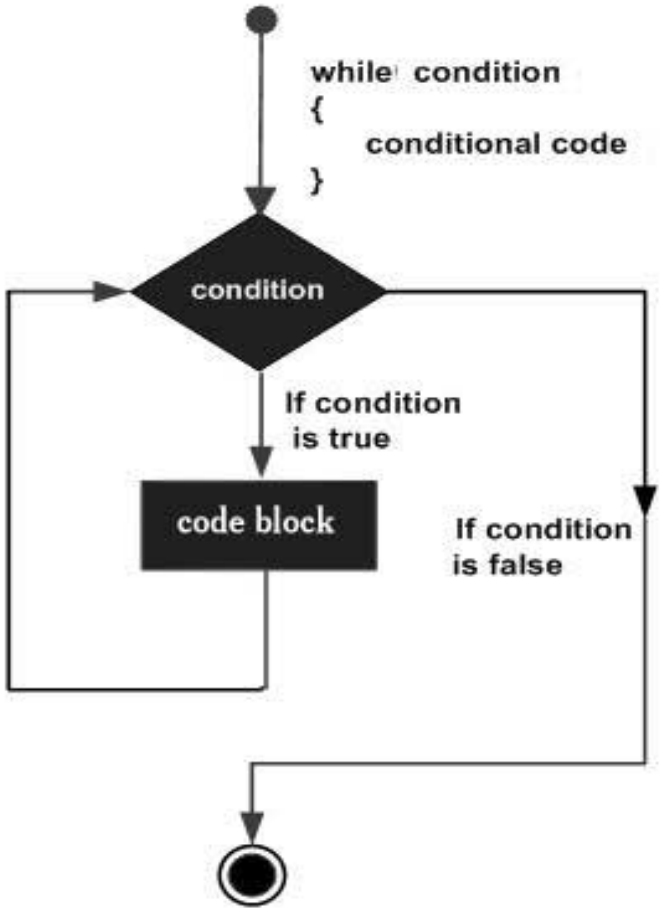
```
repeat { commands
if(condition) {
break }
}
```



Loop

While loop

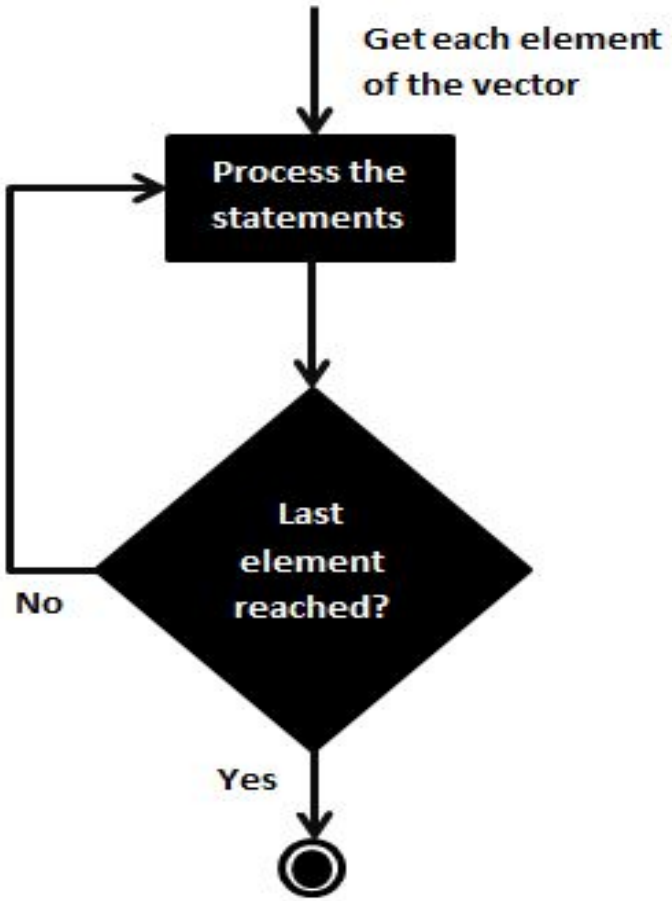
```
while (test expression){ statement}
```



Loop

For loop

```
for(value in vector) {
statement
}
```

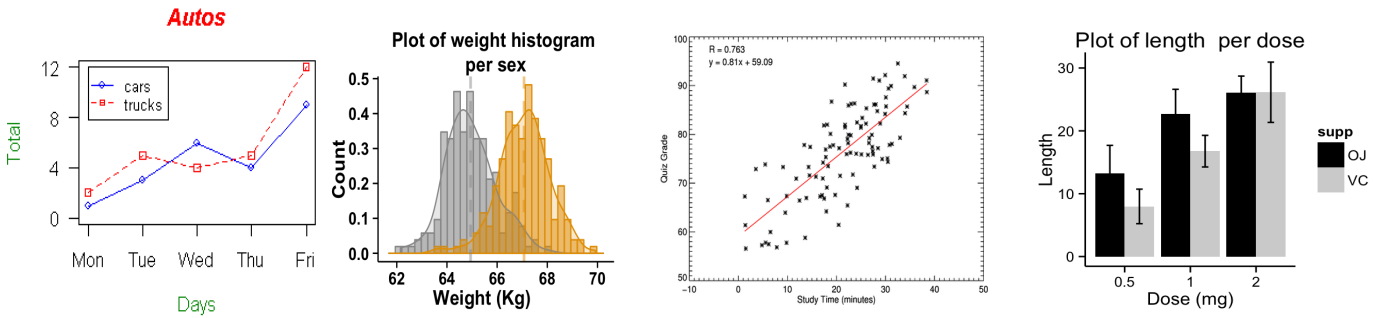


Read Files

- Read from RData
  - read.table()
- Read from txt
  - read.table()
- Read from excel
  - package XLConnect
  - loadWorkbook()
  - readWorksheet()
- Read from csv
  - read.csv()
- Read from web
  - readLines()
  - grep()
  - substring()

Basic Graph

- plot: study trend of data.
- histogram: study distribution of frequency.
- boxplot: study distribution of a variable.
- scatter plot: study outliers and relationships of multiple data sets.
- bar chart: compare multiple data sets.
- multiple datasets on one plot
- multiple graphs on one image



Basic Graph

ggplot2.qplot

- Mastering the ggplot2 language can be challenging. There is a helper function called qplot() that can hide much of this complexity when creating standard graphs.
- format of qplot:

```
qplot(x, y, data=, color=, shape=, size=, alpha=, geom=,
method=, formula=, facets=, xlim=, ylim= xlab=, ylab=,
main=, sub=)
```

  - x,y: specifies variables placed on horizontal/vertical axis
  - alpha: transparency for overlapping elements expressed as a fraction between 0 (complete transparency) and 1.
  - geom: geom values include "point", "smooth", "boxplot", "line", "histogram", "density", "bar" et.al.
  - farcets: creates a trellis graph by specifying conditioning variables. Its value is expressed as rowvar ~ colvar