## [1] "data.frame"

## [1] 5000 36

## [1] 5000

## [1] 36

## [1] "hotel" "is\_canceled"   
## [3] "lead\_time" "arrival\_date\_year"   
## [5] "arrival\_date\_month" "arrival\_date\_week\_number"   
## [7] "arrival\_date\_day\_of\_month" "stays\_in\_weekend\_nights"   
## [9] "stays\_in\_week\_nights" "adults"   
## [11] "children" "babies"   
## [13] "meal" "country"   
## [15] "market\_segment" "distribution\_channel"   
## [17] "is\_repeated\_guest" "previous\_cancellations"   
## [19] "previous\_bookings\_not\_canceled" "reserved\_room\_type"   
## [21] "assigned\_room\_type" "booking\_changes"   
## [23] "deposit\_type" "agent"   
## [25] "company" "days\_in\_waiting\_list"   
## [27] "customer\_type" "adr"   
## [29] "required\_car\_parking\_spaces" "total\_of\_special\_requests"   
## [31] "reservation\_status" "reservation\_status\_date"   
## [33] "name" "email"   
## [35] "phone.number" "credit\_card"

##   
## Adjuntando el paquete: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

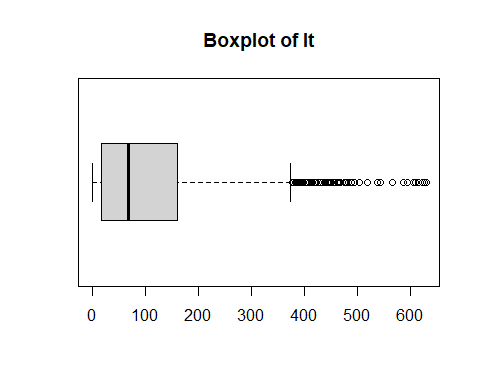
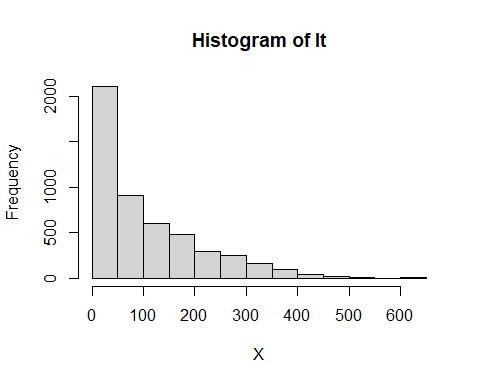
## 'data.frame': 5000 obs. of 18 variables:  
## $ hotel : Factor w/ 2 levels "City Hotel","Resort Hotel": 2 1 1 1 1 2 1 1 1 1 ...  
## $ is\_canceled : Factor w/ 2 levels "0","1": 1 2 2 1 2 1 1 1 2 1 ...  
## $ lead\_time : int 37 121 50 76 73 26 394 14 47 8 ...  
## $ arrival\_date\_year : Factor w/ 3 levels "2015","2016",..: 1 3 1 2 3 3 3 1 1 1 ...  
## $ arrival\_date\_month : Factor w/ 12 levels "1","10","11",..: 11 11 11 8 7 11 8 12 2 12 ...  
## $ arrival\_date\_week\_number : Factor w/ 53 levels "1","2","3","4",..: 35 32 35 21 17 34 20 39 44 37 ...  
## $ arrival\_date\_day\_of\_month: Factor w/ 31 levels "1","10","11",..: 16 2 18 9 16 14 8 19 18 31 ...  
## $ stays\_in\_weekend\_nights : int 2 2 0 0 2 1 0 2 2 0 ...  
## $ stays\_in\_week\_nights : int 3 5 4 2 2 5 4 1 4 4 ...  
## $ adults : int 2 2 2 1 1 2 2 1 2 2 ...  
## $ children : int 1 0 0 0 0 0 0 0 1 0 ...  
## $ babies : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ meal : Factor w/ 5 levels "BB","FB","HB",..: 1 1 1 1 1 1 1 4 3 1 ...  
## $ country : Factor w/ 89 levels "","AGO","ARE",..: 71 7 71 32 71 40 5 6 71 22 ...  
## $ days\_in\_waiting\_list : int 0 0 0 12 0 0 0 0 0 0 ...  
## $ adr : num 143 112 62 130 117 ...  
## $ reservation\_status : Factor w/ 3 levels "Canceled","Check-Out",..: 2 1 1 2 1 2 2 2 1 2 ...  
## $ arrival\_date : Date, format: "2015-08-23" "2017-08-10" ...

numeric\_cols <- dd %>% select(where(is.numeric))

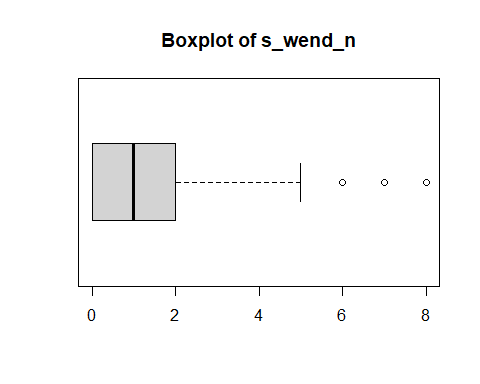
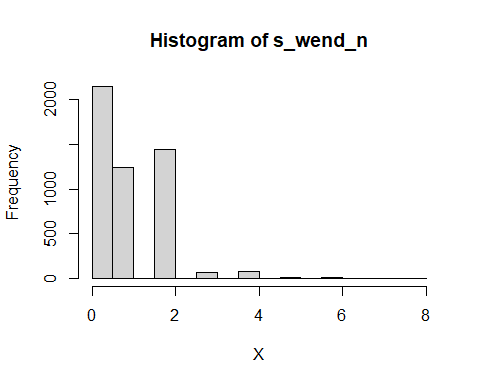
#seleccion some rows #Dselection<-dd[Condition, ]

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

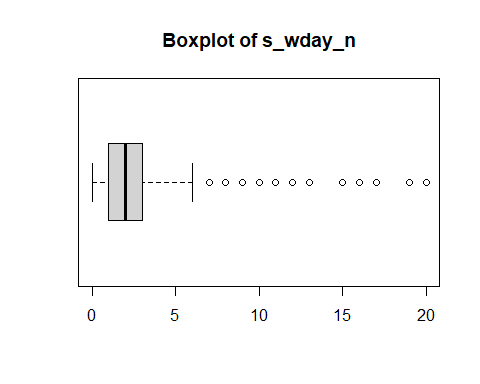
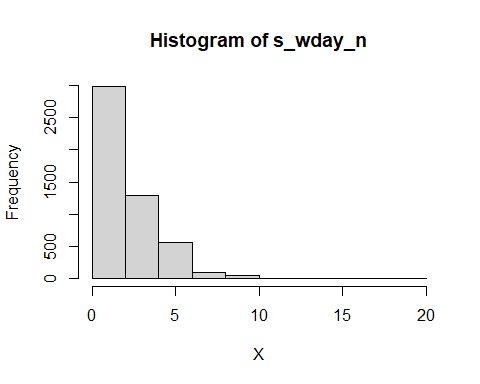
## [1] "variable 1 : hotel"  
## [1] "variable 2 : can"  
## [1] "variable 3 : lt"



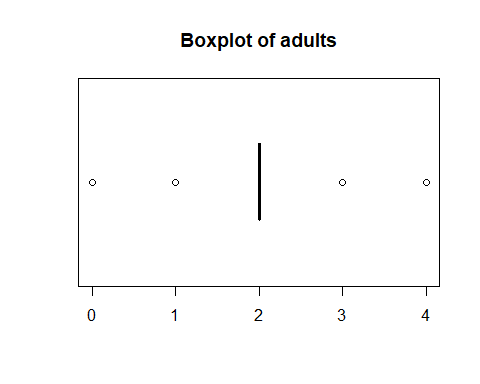
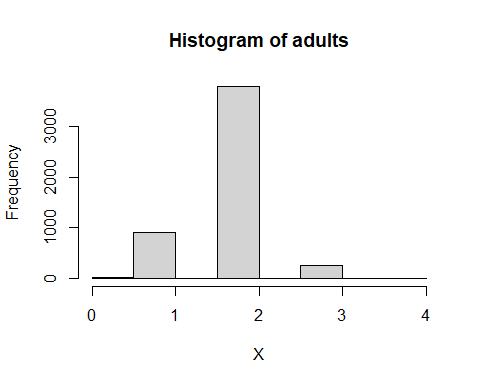
## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0 18.0 69.0 105.3 160.0 629.0   
## Standard Deviation:  
## [1] 109.1437  
## Variation Coefficient:  
## [1] 1.036601  
## [1] "variable 4 : arr\_y"  
## [1] "variable 5 : arr\_m"  
## [1] "variable 6 : arr\_wn"  
## [1] "variable 7 : arr\_dm"  
## [1] "variable 8 : s\_wend\_n"



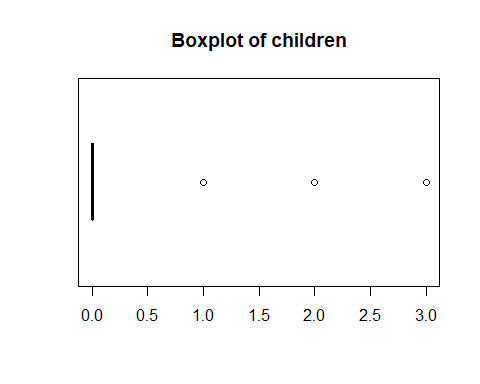
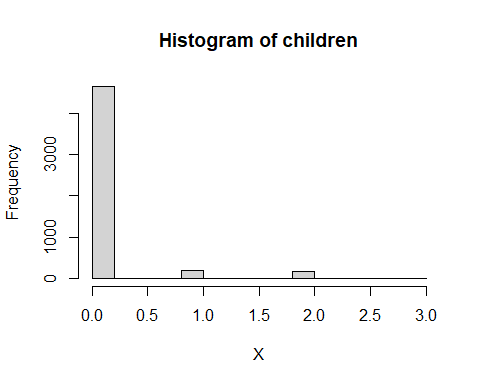
## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0000 0.0000 1.0000 0.9478 2.0000 8.0000   
## Standard Deviation:  
## [1] 0.9966312  
## Variation Coefficient:  
## [1] 1.051521  
## [1] "variable 9 : s\_wday\_n"



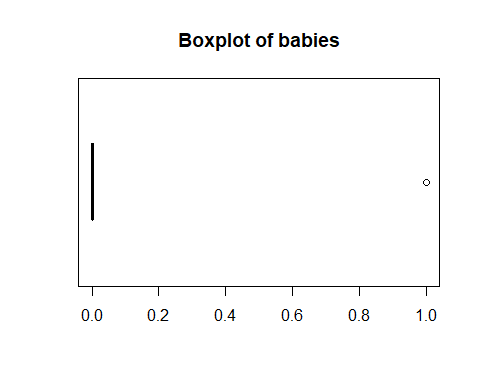
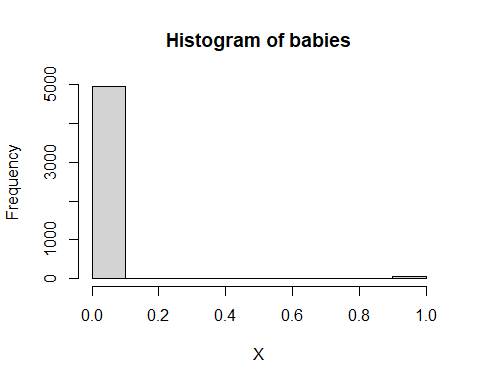
## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.000 1.000 2.000 2.522 3.000 20.000   
## Standard Deviation:  
## [1] 1.876654  
## Variation Coefficient:  
## [1] 0.7441723  
## [1] "variable 10 : adults"



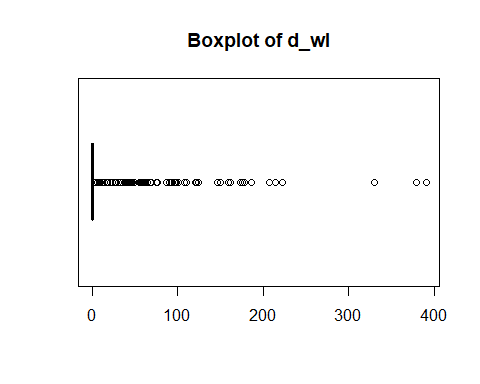
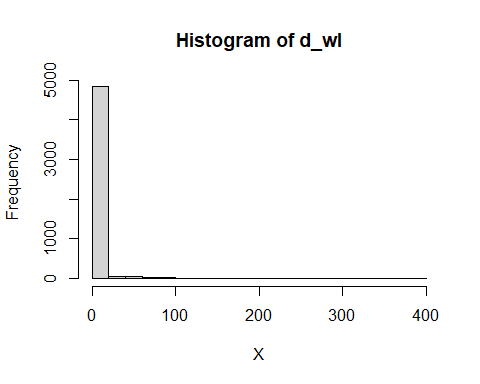
## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.000 2.000 2.000 1.862 2.000 4.000   
## Standard Deviation:  
## [1] 0.4836456  
## Variation Coefficient:  
## [1] 0.2596894  
## [1] "variable 11 : children"



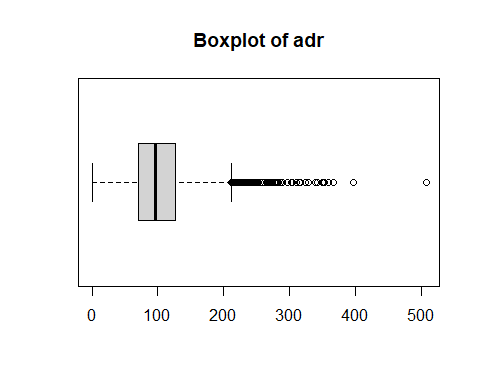
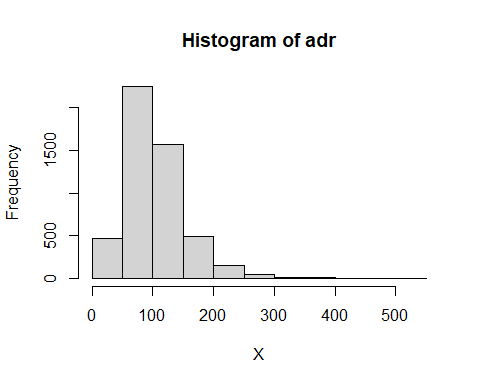
## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0000 0.0000 0.0000 0.1024 0.0000 3.0000   
## Standard Deviation:  
## [1] 0.397927  
## Variation Coefficient:  
## [1] 3.886006  
## [1] "variable 12 : babies"



## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0000 0.0000 0.0000 0.0076 0.0000 1.0000   
## Standard Deviation:  
## [1] 0.08685476  
## Variation Coefficient:  
## [1] 11.42826  
## [1] "variable 13 : meal"  
## [1] "variable 14 : country"  
## [1] "variable 15 : d\_wl"



## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 0.00 0.00 2.71 0.00 391.00   
## Standard Deviation:  
## [1] 19.2902  
## Variation Coefficient:  
## [1] 7.11763  
## [1] "variable 16 : adr"



## [1] "Extended Summary Statistics"  
## Extended Summary Statistics:  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.0 71.0 96.0 103.3 127.3 508.0   
## Standard Deviation:  
## [1] 48.83188  
## Variation Coefficient:  
## [1] 0.4726477  
## [1] "variable 17 : res\_s"  
## [1] "variable 18 : arr"