Hotel Data Week 4 Data Dive

Reshu Gupta

2024-02-02

1. Initial setup and Configure the data set.
2. Load the data set file in variable hotel\_data files.
3. Data set - Hotels : This data comes from an open hotel booking demand dataset from Antonio, Almeida and Nunes.

knitr::opts\_chunk$set(echo = FALSE)  
# Load the 'dplyr' library  
library(dplyr)

##   
## Attaching package: 'dplyr'

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

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

# Load the data into hotel\_data for further use  
hotel\_data <- read.csv(file.choose())

In this section, 1. Data set ‘hotel\_data’ is summarized. 2. then find the length of dataset - hotel\_data by using nrow() and assign to variable - hotel\_data\_length. 3. Calculate and print the size of subsample (50% of hotel\_data\_lenght).

## hotel is\_canceled lead\_time arrival\_date\_year  
## Length:119390 Min. :0.0000 Min. : 0 Min. :2015   
## Class :character 1st Qu.:0.0000 1st Qu.: 18 1st Qu.:2016   
## Mode :character Median :0.0000 Median : 69 Median :2016   
## Mean :0.3704 Mean :104 Mean :2016   
## 3rd Qu.:1.0000 3rd Qu.:160 3rd Qu.:2017   
## Max. :1.0000 Max. :737 Max. :2017   
##   
## arrival\_date\_month arrival\_date\_week\_number arrival\_date\_day\_of\_month  
## Length:119390 Min. : 1.00 Min. : 1.0   
## Class :character 1st Qu.:16.00 1st Qu.: 8.0   
## Mode :character Median :28.00 Median :16.0   
## Mean :27.17 Mean :15.8   
## 3rd Qu.:38.00 3rd Qu.:23.0   
## Max. :53.00 Max. :31.0   
##   
## stays\_in\_weekend\_nights stays\_in\_week\_nights adults   
## Min. : 0.0000 Min. : 0.0 Min. : 0.000   
## 1st Qu.: 0.0000 1st Qu.: 1.0 1st Qu.: 2.000   
## Median : 1.0000 Median : 2.0 Median : 2.000   
## Mean : 0.9276 Mean : 2.5 Mean : 1.856   
## 3rd Qu.: 2.0000 3rd Qu.: 3.0 3rd Qu.: 2.000   
## Max. :19.0000 Max. :50.0 Max. :55.000   
##   
## children babies meal country   
## Min. : 0.0000 Min. : 0.000000 Length:119390 Length:119390   
## 1st Qu.: 0.0000 1st Qu.: 0.000000 Class :character Class :character   
## Median : 0.0000 Median : 0.000000 Mode :character Mode :character   
## Mean : 0.1039 Mean : 0.007949   
## 3rd Qu.: 0.0000 3rd Qu.: 0.000000   
## Max. :10.0000 Max. :10.000000   
## NA's :4   
## market\_segment distribution\_channel is\_repeated\_guest  
## Length:119390 Length:119390 Min. :0.00000   
## Class :character Class :character 1st Qu.:0.00000   
## Mode :character Mode :character Median :0.00000   
## Mean :0.03191   
## 3rd Qu.:0.00000   
## Max. :1.00000   
##   
## previous\_cancellations previous\_bookings\_not\_canceled reserved\_room\_type  
## Min. : 0.00000 Min. : 0.0000 Length:119390   
## 1st Qu.: 0.00000 1st Qu.: 0.0000 Class :character   
## Median : 0.00000 Median : 0.0000 Mode :character   
## Mean : 0.08712 Mean : 0.1371   
## 3rd Qu.: 0.00000 3rd Qu.: 0.0000   
## Max. :26.00000 Max. :72.0000   
##   
## assigned\_room\_type booking\_changes deposit\_type agent   
## Length:119390 Min. : 0.0000 Length:119390 Length:119390   
## Class :character 1st Qu.: 0.0000 Class :character Class :character   
## Mode :character Median : 0.0000 Mode :character Mode :character   
## Mean : 0.2211   
## 3rd Qu.: 0.0000   
## Max. :21.0000   
##   
## company days\_in\_waiting\_list customer\_type adr   
## Length:119390 Min. : 0.000 Length:119390 Min. : -6.38   
## Class :character 1st Qu.: 0.000 Class :character 1st Qu.: 69.29   
## Mode :character Median : 0.000 Mode :character Median : 94.58   
## Mean : 2.321 Mean : 101.83   
## 3rd Qu.: 0.000 3rd Qu.: 126.00   
## Max. :391.000 Max. :5400.00   
##   
## required\_car\_parking\_spaces total\_of\_special\_requests reservation\_status  
## Min. :0.00000 Min. :0.0000 Length:119390   
## 1st Qu.:0.00000 1st Qu.:0.0000 Class :character   
## Median :0.00000 Median :0.0000 Mode :character   
## Mean :0.06252 Mean :0.5714   
## 3rd Qu.:0.00000 3rd Qu.:1.0000   
## Max. :8.00000 Max. :5.0000   
##   
## reservation\_status\_date  
## Length:119390   
## Class :character   
## Mode :character   
##   
##   
##   
##

## [1] 119390

## [1] 59695

In this section, 5 sub samples have been created of size - subsample\_size (subsample\_size - 50% of hotel\_data\_length).

Scrutinize #1

In this section, Ist Sub sample - hotel\_data\_subsample\_df\_1 has been scrutinize. 1. Length of Sub sample - hotel\_data\_subsample\_df\_1. 2. Print Ist five rows (for few column) of sub sample. 3. Print Last five rows (for few column) of sub sample. Note : By seeing 2 & 3, data consistency can be verified. 4. Print the Internal structure of Sub sample. 5. then summarized - sub sample.

## [1] 59695

## hotel lead\_time meal market\_segment distribution\_channel country  
## 1 Resort Hotel 70 HB Online TA TA/TO CHE  
## 2 Resort Hotel 4 HB Corporate Corporate PRT  
## 3 City Hotel 151 BB Groups TA/TO PRT  
## 4 Resort Hotel 207 BB Offline TA/TO TA/TO GBR  
## 5 City Hotel 311 BB Online TA TA/TO KOR

## hotel lead\_time meal market\_segment distribution\_channel country  
## 59691 City Hotel 278 BB Groups TA/TO PRT  
## 59692 Resort Hotel 223 HB Online TA TA/TO FRA  
## 59693 City Hotel 9 BB Online TA TA/TO ESP  
## 59694 City Hotel 273 BB Offline TA/TO TA/TO PRT  
## 59695 City Hotel 44 BB Online TA TA/TO ROU

## 'data.frame': 59695 obs. of 32 variables:  
## $ hotel : chr "Resort Hotel" "Resort Hotel" "City Hotel" "Resort Hotel" ...  
## $ is\_canceled : int 1 0 0 0 0 0 0 0 0 0 ...  
## $ lead\_time : int 70 4 151 207 311 99 55 32 59 196 ...  
## $ arrival\_date\_year : int 2017 2016 2017 2016 2017 2016 2015 2015 2016 2016 ...  
## $ arrival\_date\_month : chr "May" "May" "March" "June" ...  
## $ arrival\_date\_week\_number : int 20 22 12 27 18 8 35 43 14 38 ...  
## $ arrival\_date\_day\_of\_month : int 17 28 20 26 4 19 26 24 31 17 ...  
## $ stays\_in\_weekend\_nights : int 2 0 1 4 0 0 0 1 0 2 ...  
## $ stays\_in\_week\_nights : int 5 1 1 5 1 1 2 1 3 3 ...  
## $ adults : int 2 1 2 2 2 1 1 1 2 2 ...  
## $ children : int 2 0 0 0 0 0 0 0 0 0 ...  
## $ babies : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ meal : chr "HB" "HB" "BB" "BB" ...  
## $ country : chr "CHE" "PRT" "PRT" "GBR" ...  
## $ market\_segment : chr "Online TA" "Corporate" "Groups" "Offline TA/TO" ...  
## $ distribution\_channel : chr "TA/TO" "Corporate" "TA/TO" "TA/TO" ...  
## $ is\_repeated\_guest : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_cancellations : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_bookings\_not\_canceled: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ reserved\_room\_type : chr "C" "A" "A" "A" ...  
## $ assigned\_room\_type : chr "C" "A" "A" "A" ...  
## $ booking\_changes : int 0 0 0 0 0 0 1 0 0 0 ...  
## $ deposit\_type : chr "No Deposit" "No Deposit" "No Deposit" "No Deposit" ...  
## $ agent : chr "240" "NULL" "37" "40" ...  
## $ company : chr "NULL" "408" "NULL" "NULL" ...  
## $ days\_in\_waiting\_list : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ customer\_type : chr "Transient" "Transient" "Transient-Party" "Contract" ...  
## $ adr : num 173 65 95 59.8 109.8 ...  
## $ required\_car\_parking\_spaces : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ total\_of\_special\_requests : int 2 0 0 1 2 0 0 1 1 1 ...  
## $ reservation\_status : chr "Canceled" "Check-Out" "Check-Out" "Check-Out" ...  
## $ reservation\_status\_date : chr "2017-03-23" "2016-05-29" "2017-03-22" "2016-07-05" ...

## hotel is\_canceled lead\_time arrival\_date\_year  
## Length:59695 Min. :0.0000 Min. : 0.0 Min. :2015   
## Class :character 1st Qu.:0.0000 1st Qu.: 18.0 1st Qu.:2016   
## Mode :character Median :0.0000 Median : 70.0 Median :2016   
## Mean :0.3719 Mean :104.1 Mean :2016   
## 3rd Qu.:1.0000 3rd Qu.:162.0 3rd Qu.:2017   
## Max. :1.0000 Max. :737.0 Max. :2017   
##   
## arrival\_date\_month arrival\_date\_week\_number arrival\_date\_day\_of\_month  
## Length:59695 Min. : 1.00 Min. : 1.00   
## Class :character 1st Qu.:16.00 1st Qu.: 8.00   
## Mode :character Median :28.00 Median :16.00   
## Mean :27.23 Mean :15.78   
## 3rd Qu.:38.00 3rd Qu.:23.00   
## Max. :53.00 Max. :31.00   
##   
## stays\_in\_weekend\_nights stays\_in\_week\_nights adults children   
## Min. : 0.0000 Min. : 0.000 Min. : 0.000 Min. :0.000   
## 1st Qu.: 0.0000 1st Qu.: 1.000 1st Qu.: 2.000 1st Qu.:0.000   
## Median : 1.0000 Median : 2.000 Median : 2.000 Median :0.000   
## Mean : 0.9267 Mean : 2.501 Mean : 1.858 Mean :0.104   
## 3rd Qu.: 2.0000 3rd Qu.: 3.000 3rd Qu.: 2.000 3rd Qu.:0.000   
## Max. :19.0000 Max. :50.000 Max. :55.000 Max. :3.000   
## NA's :3   
## babies meal country market\_segment   
## Min. :0.000000 Length:59695 Length:59695 Length:59695   
## 1st Qu.:0.000000 Class :character Class :character Class :character   
## Median :0.000000 Mode :character Mode :character Mode :character   
## Mean :0.007706   
## 3rd Qu.:0.000000   
## Max. :2.000000   
##   
## distribution\_channel is\_repeated\_guest previous\_cancellations  
## Length:59695 Min. :0.00000 Min. : 0.00000   
## Class :character 1st Qu.:0.00000 1st Qu.: 0.00000   
## Mode :character Median :0.00000 Median : 0.00000   
## Mean :0.03225 Mean : 0.08697   
## 3rd Qu.:0.00000 3rd Qu.: 0.00000   
## Max. :1.00000 Max. :26.00000   
##   
## previous\_bookings\_not\_canceled reserved\_room\_type assigned\_room\_type  
## Min. : 0.0000 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character   
## Mean : 0.1374   
## 3rd Qu.: 0.0000   
## Max. :72.0000   
##   
## booking\_changes deposit\_type agent company   
## Min. : 0.0000 Length:59695 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character Mode :character   
## Mean : 0.2205   
## 3rd Qu.: 0.0000   
## Max. :16.0000   
##   
## days\_in\_waiting\_list customer\_type adr   
## Min. : 0.000 Length:59695 Min. : 0.00   
## 1st Qu.: 0.000 Class :character 1st Qu.: 69.55   
## Median : 0.000 Mode :character Median : 95.00   
## Mean : 2.338 Mean : 101.97   
## 3rd Qu.: 0.000 3rd Qu.: 126.00   
## Max. :391.000 Max. :5400.00   
##   
## required\_car\_parking\_spaces total\_of\_special\_requests reservation\_status  
## Min. :0.00000 Min. :0.0000 Length:59695   
## 1st Qu.:0.00000 1st Qu.:0.0000 Class :character   
## Median :0.00000 Median :0.0000 Mode :character   
## Mean :0.06289 Mean :0.5732   
## 3rd Qu.:0.00000 3rd Qu.:1.0000   
## Max. :2.00000 Max. :5.0000   
##   
## reservation\_status\_date  
## Length:59695   
## Class :character   
## Mode :character   
##   
##   
##   
##

Scrutinize #2 In this section, IInd Sub sample - sub\_sample\_hotel\_data\_2 has been scrutinize. 1. Length of Sub sample - hotel\_data\_subsample\_df\_1. 2. Print Ist five rows (for few column) of sub sample. 3. Print Last five rows (for few column) of sub sample. Note : By seeing 2 & 3, data consistency can be verified. 4. Print the Internal structure of Sub sample. 5. then summarized - sub sample

## [1] 59695

## hotel lead\_time meal market\_segment distribution\_channel country  
## 1 Resort Hotel 70 HB Online TA TA/TO CHE  
## 2 Resort Hotel 4 HB Corporate Corporate PRT  
## 3 City Hotel 151 BB Groups TA/TO PRT  
## 4 Resort Hotel 207 BB Offline TA/TO TA/TO GBR  
## 5 City Hotel 311 BB Online TA TA/TO KOR

## hotel lead\_time meal market\_segment distribution\_channel country  
## 59691 City Hotel 278 BB Groups TA/TO PRT  
## 59692 Resort Hotel 223 HB Online TA TA/TO FRA  
## 59693 City Hotel 9 BB Online TA TA/TO ESP  
## 59694 City Hotel 273 BB Offline TA/TO TA/TO PRT  
## 59695 City Hotel 44 BB Online TA TA/TO ROU

## 'data.frame': 59695 obs. of 32 variables:  
## $ hotel : chr "City Hotel" "City Hotel" "City Hotel" "Resort Hotel" ...  
## $ is\_canceled : int 0 1 0 0 0 0 0 0 0 0 ...  
## $ lead\_time : int 20 315 199 273 224 71 92 222 423 73 ...  
## $ arrival\_date\_year : int 2016 2016 2017 2017 2016 2016 2016 2016 2017 2017 ...  
## $ arrival\_date\_month : chr "November" "September" "August" "June" ...  
## $ arrival\_date\_week\_number : int 47 40 35 22 22 48 46 18 29 31 ...  
## $ arrival\_date\_day\_of\_month : int 15 30 29 2 25 25 11 28 22 2 ...  
## $ stays\_in\_weekend\_nights : int 0 1 0 1 1 0 0 2 1 2 ...  
## $ stays\_in\_week\_nights : int 2 2 5 2 4 2 1 3 1 6 ...  
## $ adults : int 1 2 1 3 2 2 2 2 2 2 ...  
## $ children : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ babies : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ meal : chr "BB" "HB" "SC" "BB" ...  
## $ country : chr "ESP" "PRT" "DEU" "IRL" ...  
## $ market\_segment : chr "Online TA" "Offline TA/TO" "Online TA" "Online TA" ...  
## $ distribution\_channel : chr "TA/TO" "TA/TO" "TA/TO" "TA/TO" ...  
## $ is\_repeated\_guest : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_cancellations : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_bookings\_not\_canceled: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ reserved\_room\_type : chr "D" "A" "A" "G" ...  
## $ assigned\_room\_type : chr "D" "A" "A" "G" ...  
## $ booking\_changes : int 0 0 0 0 4 0 0 0 0 0 ...  
## $ deposit\_type : chr "No Deposit" "Non Refund" "No Deposit" "No Deposit" ...  
## $ agent : chr "9" "31" "9" "242" ...  
## $ company : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ days\_in\_waiting\_list : int 0 69 0 0 0 0 0 0 0 0 ...  
## $ customer\_type : chr "Transient" "Transient" "Transient" "Transient" ...  
## $ adr : num 114 160 99 143 46 ...  
## $ required\_car\_parking\_spaces : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ total\_of\_special\_requests : int 2 0 2 2 0 2 0 0 1 0 ...  
## $ reservation\_status : chr "Check-Out" "Canceled" "Check-Out" "Check-Out" ...  
## $ reservation\_status\_date : chr "2016-11-17" "2016-01-28" "2017-09-03" "2017-06-05" ...

## hotel is\_canceled lead\_time arrival\_date\_year  
## Length:59695 Min. :0.0000 Min. : 0.0 Min. :2015   
## Class :character 1st Qu.:0.0000 1st Qu.: 18.0 1st Qu.:2016   
## Mode :character Median :0.0000 Median : 69.0 Median :2016   
## Mean :0.3721 Mean :103.9 Mean :2016   
## 3rd Qu.:1.0000 3rd Qu.:160.0 3rd Qu.:2017   
## Max. :1.0000 Max. :629.0 Max. :2017   
## arrival\_date\_month arrival\_date\_week\_number arrival\_date\_day\_of\_month  
## Length:59695 Min. : 1.00 Min. : 1.00   
## Class :character 1st Qu.:16.00 1st Qu.: 8.00   
## Mode :character Median :27.00 Median :16.00   
## Mean :27.08 Mean :15.74   
## 3rd Qu.:38.00 3rd Qu.:23.00   
## Max. :53.00 Max. :31.00   
## stays\_in\_weekend\_nights stays\_in\_week\_nights adults children   
## Min. : 0.0000 Min. : 0.000 Min. : 0.000 Min. :0.0000   
## 1st Qu.: 0.0000 1st Qu.: 1.000 1st Qu.: 2.000 1st Qu.:0.0000   
## Median : 1.0000 Median : 2.000 Median : 2.000 Median :0.0000   
## Mean : 0.9296 Mean : 2.503 Mean : 1.856 Mean :0.1058   
## 3rd Qu.: 2.0000 3rd Qu.: 3.000 3rd Qu.: 2.000 3rd Qu.:0.0000   
## Max. :18.0000 Max. :42.000 Max. :40.000 Max. :3.0000   
## babies meal country market\_segment   
## Min. : 0.000000 Length:59695 Length:59695 Length:59695   
## 1st Qu.: 0.000000 Class :character Class :character Class :character   
## Median : 0.000000 Mode :character Mode :character Mode :character   
## Mean : 0.008108   
## 3rd Qu.: 0.000000   
## Max. :10.000000   
## distribution\_channel is\_repeated\_guest previous\_cancellations  
## Length:59695 Min. :0.00000 Min. : 0.00000   
## Class :character 1st Qu.:0.00000 1st Qu.: 0.00000   
## Mode :character Median :0.00000 Median : 0.00000   
## Mean :0.03268 Mean : 0.08406   
## 3rd Qu.:0.00000 3rd Qu.: 0.00000   
## Max. :1.00000 Max. :26.00000   
## previous\_bookings\_not\_canceled reserved\_room\_type assigned\_room\_type  
## Min. : 0.0000 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character   
## Mean : 0.1402   
## 3rd Qu.: 0.0000   
## Max. :69.0000   
## booking\_changes deposit\_type agent company   
## Min. : 0.0000 Length:59695 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character Mode :character   
## Mean : 0.2247   
## 3rd Qu.: 0.0000   
## Max. :21.0000   
## days\_in\_waiting\_list customer\_type adr   
## Min. : 0.000 Length:59695 Min. : 0.0   
## 1st Qu.: 0.000 Class :character 1st Qu.: 70.0   
## Median : 0.000 Mode :character Median : 94.5   
## Mean : 2.334 Mean :101.6   
## 3rd Qu.: 0.000 3rd Qu.:126.0   
## Max. :391.000 Max. :510.0   
## required\_car\_parking\_spaces total\_of\_special\_requests reservation\_status  
## Min. :0.00000 Min. :0.0000 Length:59695   
## 1st Qu.:0.00000 1st Qu.:0.0000 Class :character   
## Median :0.00000 Median :0.0000 Mode :character   
## Mean :0.06195 Mean :0.5784   
## 3rd Qu.:0.00000 3rd Qu.:1.0000   
## Max. :2.00000 Max. :5.0000   
## reservation\_status\_date  
## Length:59695   
## Class :character   
## Mode :character   
##   
##   
##

Scrutinize #3 In this section, IIIrd Sub sample - sub\_sample\_hotel\_data\_3 has been scrutinize. 1. Length of Sub sample - hotel\_data\_subsample\_df\_3. 2. Print Ist five rows (for few column) of sub sample. 3. Print Last five rows (for few column) of sub sample. Note : By seeing 2 & 3, data consistency can be verified. 4. Print the Internal structure of Sub sample. 5. then summarized - sub sample

## [1] 59695

## hotel lead\_time meal market\_segment distribution\_channel country  
## 1 Resort Hotel 70 HB Online TA TA/TO CHE  
## 2 Resort Hotel 4 HB Corporate Corporate PRT  
## 3 City Hotel 151 BB Groups TA/TO PRT  
## 4 Resort Hotel 207 BB Offline TA/TO TA/TO GBR  
## 5 City Hotel 311 BB Online TA TA/TO KOR

## hotel lead\_time meal market\_segment distribution\_channel country  
## 59691 City Hotel 278 BB Groups TA/TO PRT  
## 59692 Resort Hotel 223 HB Online TA TA/TO FRA  
## 59693 City Hotel 9 BB Online TA TA/TO ESP  
## 59694 City Hotel 273 BB Offline TA/TO TA/TO PRT  
## 59695 City Hotel 44 BB Online TA TA/TO ROU

## 'data.frame': 59695 obs. of 32 variables:  
## $ hotel : chr "City Hotel" "Resort Hotel" "Resort Hotel" "Resort Hotel" ...  
## $ is\_canceled : int 1 1 1 0 1 1 0 0 0 0 ...  
## $ lead\_time : int 77 80 143 223 322 16 35 61 7 0 ...  
## $ arrival\_date\_year : int 2017 2015 2015 2016 2016 2016 2016 2016 2016 2015 ...  
## $ arrival\_date\_month : chr "July" "August" "October" "May" ...  
## $ arrival\_date\_week\_number : int 27 33 41 22 21 44 47 12 9 48 ...  
## $ arrival\_date\_day\_of\_month : int 2 11 10 24 19 29 16 19 21 25 ...  
## $ stays\_in\_weekend\_nights : int 2 2 2 0 1 0 0 2 1 0 ...  
## $ stays\_in\_week\_nights : int 0 5 5 5 3 1 1 2 0 3 ...  
## $ adults : int 2 2 2 1 2 2 1 2 1 1 ...  
## $ children : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ babies : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ meal : chr "SC" "BB" "BB" "BB" ...  
## $ country : chr "IRL" "PRT" "PRT" "GBR" ...  
## $ market\_segment : chr "Online TA" "Online TA" "Groups" "Groups" ...  
## $ distribution\_channel : chr "TA/TO" "TA/TO" "TA/TO" "Direct" ...  
## $ is\_repeated\_guest : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_cancellations : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_bookings\_not\_canceled: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ reserved\_room\_type : chr "A" "A" "A" "A" ...  
## $ assigned\_room\_type : chr "A" "A" "A" "A" ...  
## $ booking\_changes : int 0 0 0 1 0 0 0 0 0 1 ...  
## $ deposit\_type : chr "No Deposit" "No Deposit" "Non Refund" "No Deposit" ...  
## $ agent : chr "9" "241" "96" "NULL" ...  
## $ company : chr "NULL" "NULL" "NULL" "223" ...  
## $ days\_in\_waiting\_list : int 0 0 0 0 120 0 0 0 0 0 ...  
## $ customer\_type : chr "Transient" "Transient" "Transient" "Transient-Party" ...  
## $ adr : num 120 98 36 68 80 ...  
## $ required\_car\_parking\_spaces : int 0 0 0 1 0 0 0 0 0 0 ...  
## $ total\_of\_special\_requests : int 2 0 0 0 0 1 2 2 1 0 ...  
## $ reservation\_status : chr "Canceled" "Canceled" "Canceled" "Check-Out" ...  
## $ reservation\_status\_date : chr "2017-05-29" "2015-05-23" "2015-07-08" "2016-05-29" ...

## hotel is\_canceled lead\_time arrival\_date\_year  
## Length:59695 Min. :0.0000 Min. : 0.0 Min. :2015   
## Class :character 1st Qu.:0.0000 1st Qu.: 18.0 1st Qu.:2016   
## Mode :character Median :0.0000 Median : 70.0 Median :2016   
## Mean :0.3669 Mean :104.2 Mean :2016   
## 3rd Qu.:1.0000 3rd Qu.:161.0 3rd Qu.:2017   
## Max. :1.0000 Max. :737.0 Max. :2017   
##   
## arrival\_date\_month arrival\_date\_week\_number arrival\_date\_day\_of\_month  
## Length:59695 Min. : 1.00 Min. : 1.00   
## Class :character 1st Qu.:16.00 1st Qu.: 8.00   
## Mode :character Median :27.00 Median :16.00   
## Mean :27.08 Mean :15.82   
## 3rd Qu.:38.00 3rd Qu.:24.00   
## Max. :53.00 Max. :31.00   
##   
## stays\_in\_weekend\_nights stays\_in\_week\_nights adults children   
## Min. : 0.0000 Min. : 0.000 Min. : 0.000 Min. :0.0000   
## 1st Qu.: 0.0000 1st Qu.: 1.000 1st Qu.: 2.000 1st Qu.:0.0000   
## Median : 1.0000 Median : 2.000 Median : 2.000 Median :0.0000   
## Mean : 0.9347 Mean : 2.508 Mean : 1.854 Mean :0.1026   
## 3rd Qu.: 2.0000 3rd Qu.: 3.000 3rd Qu.: 2.000 3rd Qu.:0.0000   
## Max. :18.0000 Max. :42.000 Max. :20.000 Max. :3.0000   
## NA's :3   
## babies meal country market\_segment   
## Min. : 0.000000 Length:59695 Length:59695 Length:59695   
## 1st Qu.: 0.000000 Class :character Class :character Class :character   
## Median : 0.000000 Mode :character Mode :character Mode :character   
## Mean : 0.008326   
## 3rd Qu.: 0.000000   
## Max. :10.000000   
##   
## distribution\_channel is\_repeated\_guest previous\_cancellations  
## Length:59695 Min. :0.00000 Min. : 0.00000   
## Class :character 1st Qu.:0.00000 1st Qu.: 0.00000   
## Mode :character Median :0.00000 Median : 0.00000   
## Mean :0.03205 Mean : 0.08564   
## 3rd Qu.:0.00000 3rd Qu.: 0.00000   
## Max. :1.00000 Max. :26.00000   
##   
## previous\_bookings\_not\_canceled reserved\_room\_type assigned\_room\_type  
## Min. : 0.0000 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character   
## Mean : 0.1373   
## 3rd Qu.: 0.0000   
## Max. :70.0000   
##   
## booking\_changes deposit\_type agent company   
## Min. : 0.0000 Length:59695 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character Mode :character   
## Mean : 0.2213   
## 3rd Qu.: 0.0000   
## Max. :17.0000   
##   
## days\_in\_waiting\_list customer\_type adr   
## Min. : 0.000 Length:59695 Min. : 0.0   
## 1st Qu.: 0.000 Class :character 1st Qu.: 69.0   
## Median : 0.000 Mode :character Median : 94.5   
## Mean : 2.266 Mean :101.7   
## 3rd Qu.: 0.000 3rd Qu.:126.0   
## Max. :391.000 Max. :510.0   
##   
## required\_car\_parking\_spaces total\_of\_special\_requests reservation\_status  
## Min. :0.00000 Min. :0.00 Length:59695   
## 1st Qu.:0.00000 1st Qu.:0.00 Class :character   
## Median :0.00000 Median :0.00 Mode :character   
## Mean :0.06332 Mean :0.57   
## 3rd Qu.:0.00000 3rd Qu.:1.00   
## Max. :3.00000 Max. :5.00   
##   
## reservation\_status\_date  
## Length:59695   
## Class :character   
## Mode :character   
##   
##   
##   
##

Scrutinize #4 In this section, IVth Sub sample - sub\_sample\_hotel\_data\_4 has been scrutinize. 1. Length of Sub sample - hotel\_data\_subsample\_df\_4. 2. Print Ist five rows (for few column) of sub sample. 3. Print Last five rows (for few column) of sub sample. Note : By seeing 2 & 3, data consistency can be verified. 4. Print the Internal structure of Sub sample. 5. then summarized - sub sample

## [1] 59695

## hotel lead\_time meal market\_segment distribution\_channel country  
## 1 Resort Hotel 70 HB Online TA TA/TO CHE  
## 2 Resort Hotel 4 HB Corporate Corporate PRT  
## 3 City Hotel 151 BB Groups TA/TO PRT  
## 4 Resort Hotel 207 BB Offline TA/TO TA/TO GBR  
## 5 City Hotel 311 BB Online TA TA/TO KOR

## hotel lead\_time meal market\_segment distribution\_channel country  
## 59691 City Hotel 278 BB Groups TA/TO PRT  
## 59692 Resort Hotel 223 HB Online TA TA/TO FRA  
## 59693 City Hotel 9 BB Online TA TA/TO ESP  
## 59694 City Hotel 273 BB Offline TA/TO TA/TO PRT  
## 59695 City Hotel 44 BB Online TA TA/TO ROU

## 'data.frame': 59695 obs. of 32 variables:  
## $ hotel : chr "City Hotel" "Resort Hotel" "City Hotel" "Resort Hotel" ...  
## $ is\_canceled : int 0 0 0 0 0 0 0 1 0 0 ...  
## $ lead\_time : int 0 10 7 87 73 316 74 245 87 84 ...  
## $ arrival\_date\_year : int 2016 2016 2016 2015 2016 2016 2017 2016 2017 2016 ...  
## $ arrival\_date\_month : chr "July" "July" "May" "July" ...  
## $ arrival\_date\_week\_number : int 30 29 22 28 46 22 13 28 28 20 ...  
## $ arrival\_date\_day\_of\_month : int 20 13 24 10 12 26 30 6 13 13 ...  
## $ stays\_in\_weekend\_nights : int 0 0 0 2 0 0 2 2 2 1 ...  
## $ stays\_in\_week\_nights : int 2 4 1 5 1 3 3 4 6 2 ...  
## $ adults : int 2 2 1 2 2 1 2 1 2 3 ...  
## $ children : int 2 0 0 1 0 0 0 0 1 0 ...  
## $ babies : int 0 0 0 1 0 0 0 0 0 0 ...  
## $ meal : chr "BB" "BB" "BB" "FB" ...  
## $ country : chr "BEL" "NLD" "AUT" "PRT" ...  
## $ market\_segment : chr "Online TA" "Online TA" "Online TA" "Offline TA/TO" ...  
## $ distribution\_channel : chr "TA/TO" "TA/TO" "TA/TO" "TA/TO" ...  
## $ is\_repeated\_guest : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_cancellations : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_bookings\_not\_canceled: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ reserved\_room\_type : chr "F" "A" "A" "C" ...  
## $ assigned\_room\_type : chr "F" "A" "D" "C" ...  
## $ booking\_changes : int 0 0 0 0 0 1 0 0 0 0 ...  
## $ deposit\_type : chr "No Deposit" "No Deposit" "No Deposit" "No Deposit" ...  
## $ agent : chr "9" "240" "9" "5" ...  
## $ company : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ days\_in\_waiting\_list : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ customer\_type : chr "Transient" "Transient" "Transient" "Transient" ...  
## $ adr : num 206 171.5 114.8 133.8 79.2 ...  
## $ required\_car\_parking\_spaces : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ total\_of\_special\_requests : int 0 1 1 2 2 0 2 0 1 2 ...  
## $ reservation\_status : chr "Check-Out" "Check-Out" "Check-Out" "Check-Out" ...  
## $ reservation\_status\_date : chr "2016-07-22" "2016-07-17" "2016-05-25" "2015-07-17" ...

## hotel is\_canceled lead\_time arrival\_date\_year  
## Length:59695 Min. :0.0000 Min. : 0.0 Min. :2015   
## Class :character 1st Qu.:0.0000 1st Qu.: 18.0 1st Qu.:2016   
## Mode :character Median :0.0000 Median : 69.0 Median :2016   
## Mean :0.3721 Mean :103.8 Mean :2016   
## 3rd Qu.:1.0000 3rd Qu.:160.0 3rd Qu.:2017   
## Max. :1.0000 Max. :737.0 Max. :2017   
## arrival\_date\_month arrival\_date\_week\_number arrival\_date\_day\_of\_month  
## Length:59695 Min. : 1.00 Min. : 1.00   
## Class :character 1st Qu.:16.00 1st Qu.: 8.00   
## Mode :character Median :28.00 Median :16.00   
## Mean :27.21 Mean :15.89   
## 3rd Qu.:38.00 3rd Qu.:24.00   
## Max. :53.00 Max. :31.00   
## stays\_in\_weekend\_nights stays\_in\_week\_nights adults children   
## Min. : 0.0000 Min. : 0.000 Min. : 0.000 Min. :0.0000   
## 1st Qu.: 0.0000 1st Qu.: 1.000 1st Qu.: 2.000 1st Qu.:0.0000   
## Median : 1.0000 Median : 2.000 Median : 2.000 Median :0.0000   
## Mean : 0.9294 Mean : 2.498 Mean : 1.856 Mean :0.1028   
## 3rd Qu.: 2.0000 3rd Qu.: 3.000 3rd Qu.: 2.000 3rd Qu.:0.0000   
## Max. :16.0000 Max. :40.000 Max. :50.000 Max. :3.0000   
## babies meal country market\_segment   
## Min. :0.00000 Length:59695 Length:59695 Length:59695   
## 1st Qu.:0.00000 Class :character Class :character Class :character   
## Median :0.00000 Mode :character Mode :character Mode :character   
## Mean :0.00794   
## 3rd Qu.:0.00000   
## Max. :9.00000   
## distribution\_channel is\_repeated\_guest previous\_cancellations  
## Length:59695 Min. :0.00000 Min. : 0.00000   
## Class :character 1st Qu.:0.00000 1st Qu.: 0.00000   
## Mode :character Median :0.00000 Median : 0.00000   
## Mean :0.03344 Mean : 0.08765   
## 3rd Qu.:0.00000 3rd Qu.: 0.00000   
## Max. :1.00000 Max. :26.00000   
## previous\_bookings\_not\_canceled reserved\_room\_type assigned\_room\_type  
## Min. : 0.0000 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character   
## Mean : 0.1409   
## 3rd Qu.: 0.0000   
## Max. :71.0000   
## booking\_changes deposit\_type agent company   
## Min. : 0.0000 Length:59695 Length:59695 Length:59695   
## 1st Qu.: 0.0000 Class :character Class :character Class :character   
## Median : 0.0000 Mode :character Mode :character Mode :character   
## Mean : 0.2167   
## 3rd Qu.: 0.0000   
## Max. :16.0000   
## days\_in\_waiting\_list customer\_type adr   
## Min. : 0.00 Length:59695 Min. : 0.00   
## 1st Qu.: 0.00 Class :character 1st Qu.: 68.95   
## Median : 0.00 Mode :character Median : 94.50   
## Mean : 2.29 Mean : 101.78   
## 3rd Qu.: 0.00 3rd Qu.: 125.81   
## Max. :391.00 Max. :5400.00   
## required\_car\_parking\_spaces total\_of\_special\_requests reservation\_status  
## Min. :0.00000 Min. :0.0000 Length:59695   
## 1st Qu.:0.00000 1st Qu.:0.0000 Class :character   
## Median :0.00000 Median :0.0000 Mode :character   
## Mean :0.06335 Mean :0.5742   
## 3rd Qu.:0.00000 3rd Qu.:1.0000   
## Max. :8.00000 Max. :5.0000   
## reservation\_status\_date  
## Length:59695   
## Class :character   
## Mode :character   
##   
##   
##

Scrutinize #5 In this section, Vth Sub sample - sub\_sample\_hotel\_data\_5 has been scrutinize. 1. Length of Sub sample - hotel\_data\_subsample\_df\_5. 2. Print Ist five rows(for few column) of sub sample. 3. Print Last five rows(for few column) of sub sample. Note : By seeing 2 & 3, data consistency can be verified. 4. Print the Internal structure of Sub sample. 5. then summarized - sub sample

## [1] 59695

## hotel meal market\_segment distribution\_channel country  
## 1 Resort Hotel HB Online TA TA/TO CHE  
## 2 Resort Hotel HB Corporate Corporate PRT  
## 3 City Hotel BB Groups TA/TO PRT  
## 4 Resort Hotel BB Offline TA/TO TA/TO GBR  
## 5 City Hotel BB Online TA TA/TO KOR

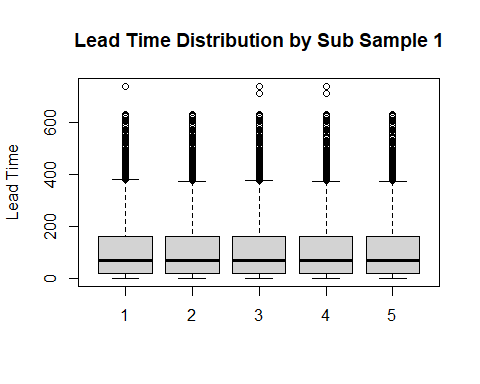
## hotel meal market\_segment distribution\_channel country  
## 59691 City Hotel BB Groups TA/TO PRT  
## 59692 Resort Hotel HB Online TA TA/TO FRA  
## 59693 City Hotel BB Online TA TA/TO ESP  
## 59694 City Hotel BB Offline TA/TO TA/TO PRT  
## 59695 City Hotel BB Online TA TA/TO ROU

## 'data.frame': 59695 obs. of 32 variables:  
## $ hotel : chr "City Hotel" "Resort Hotel" "City Hotel" "City Hotel" ...  
## $ is\_canceled : int 1 0 1 1 1 1 0 0 0 0 ...  
## $ lead\_time : int 158 1 44 56 80 45 1 34 76 306 ...  
## $ arrival\_date\_year : int 2016 2016 2016 2017 2016 2017 2017 2016 2017 2016 ...  
## $ arrival\_date\_month : chr "May" "September" "June" "March" ...  
## $ arrival\_date\_week\_number : int 22 39 25 12 46 2 11 40 17 34 ...  
## $ arrival\_date\_day\_of\_month : int 24 20 18 21 12 11 14 25 28 14 ...  
## $ stays\_in\_weekend\_nights : int 0 0 1 0 1 1 0 2 2 2 ...  
## $ stays\_in\_week\_nights : int 2 1 1 3 1 4 1 1 5 1 ...  
## $ adults : int 1 1 2 2 2 2 1 2 2 2 ...  
## $ children : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ babies : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ meal : chr "BB" "BB" "BB" "BB" ...  
## $ country : chr "PRT" "PRT" "PRT" "PRT" ...  
## $ market\_segment : chr "Groups" "Corporate" "Offline TA/TO" "Offline TA/TO" ...  
## $ distribution\_channel : chr "TA/TO" "Corporate" "TA/TO" "TA/TO" ...  
## $ is\_repeated\_guest : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_cancellations : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ previous\_bookings\_not\_canceled: int 0 0 0 0 0 0 0 0 0 0 ...  
## $ reserved\_room\_type : chr "A" "A" "D" "A" ...  
## $ assigned\_room\_type : chr "A" "A" "D" "A" ...  
## $ booking\_changes : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ deposit\_type : chr "Non Refund" "No Deposit" "No Deposit" "Non Refund" ...  
## $ agent : chr "37" "282" "6" "86" ...  
## $ company : chr "NULL" "NULL" "NULL" "NULL" ...  
## $ days\_in\_waiting\_list : int 31 0 0 0 0 0 0 0 0 0 ...  
## $ customer\_type : chr "Transient" "Transient-Party" "Transient" "Transient" ...  
## $ adr : num 130 65 96.3 85 93.6 ...  
## $ required\_car\_parking\_spaces : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ total\_of\_special\_requests : int 0 0 0 0 0 1 1 1 0 3 ...  
## $ reservation\_status : chr "Canceled" "Check-Out" "Canceled" "Canceled" ...  
## $ reservation\_status\_date : chr "2016-01-18" "2016-09-21" "2016-05-05" "2017-01-24" ...

## hotel is\_canceled lead\_time arrival\_date\_year  
## Length:59695 Min. :0.0000 Min. : 0.0 Min. :2015   
## Class :character 1st Qu.:0.0000 1st Qu.: 18.0 1st Qu.:2016   
## Mode :character Median :0.0000 Median : 70.0 Median :2016   
## Mean :0.3695 Mean :104.2 Mean :2016   
## 3rd Qu.:1.0000 3rd Qu.:160.0 3rd Qu.:2017   
## Max. :1.0000 Max. :629.0 Max. :2017   
##   
## arrival\_date\_month arrival\_date\_week\_number arrival\_date\_day\_of\_month  
## Length:59695 Min. : 1.0 Min. : 1.00   
## Class :character 1st Qu.:16.0 1st Qu.: 8.00   
## Mode :character Median :28.0 Median :16.00   
## Mean :27.2 Mean :15.82   
## 3rd Qu.:38.0 3rd Qu.:23.00   
## Max. :53.0 Max. :31.00   
##   
## stays\_in\_weekend\_nights stays\_in\_week\_nights adults   
## Min. : 0.0000 Min. : 0.000 Min. : 0.000   
## 1st Qu.: 0.0000 1st Qu.: 1.000 1st Qu.: 2.000   
## Median : 1.0000 Median : 2.000 Median : 2.000   
## Mean : 0.9331 Mean : 2.509 Mean : 1.854   
## 3rd Qu.: 2.0000 3rd Qu.: 3.000 3rd Qu.: 2.000   
## Max. :18.0000 Max. :42.000 Max. :40.000   
##   
## children babies meal country   
## Min. : 0.0000 Min. : 0.000000 Length:59695 Length:59695   
## 1st Qu.: 0.0000 1st Qu.: 0.000000 Class :character Class :character   
## Median : 0.0000 Median : 0.000000 Mode :character Mode :character   
## Mean : 0.1001 Mean : 0.007857   
## 3rd Qu.: 0.0000 3rd Qu.: 0.000000   
## Max. :10.0000 Max. :10.000000   
## NA's :5   
## market\_segment distribution\_channel is\_repeated\_guest  
## Length:59695 Length:59695 Min. :0.00000   
## Class :character Class :character 1st Qu.:0.00000   
## Mode :character Mode :character Median :0.00000   
## Mean :0.03181   
## 3rd Qu.:0.00000   
## Max. :1.00000   
##   
## previous\_cancellations previous\_bookings\_not\_canceled reserved\_room\_type  
## Min. : 0.00000 Min. : 0.0000 Length:59695   
## 1st Qu.: 0.00000 1st Qu.: 0.0000 Class :character   
## Median : 0.00000 Median : 0.0000 Mode :character   
## Mean : 0.08883 Mean : 0.1297   
## 3rd Qu.: 0.00000 3rd Qu.: 0.0000   
## Max. :26.00000 Max. :71.0000   
##   
## assigned\_room\_type booking\_changes deposit\_type agent   
## Length:59695 Min. : 0.0000 Length:59695 Length:59695   
## Class :character 1st Qu.: 0.0000 Class :character Class :character   
## Mode :character Median : 0.0000 Mode :character Mode :character   
## Mean : 0.2237   
## 3rd Qu.: 0.0000   
## Max. :21.0000   
##   
## company days\_in\_waiting\_list customer\_type adr   
## Length:59695 Min. : 0.000 Length:59695 Min. : 0.0   
## Class :character 1st Qu.: 0.000 Class :character 1st Qu.: 69.0   
## Mode :character Median : 0.000 Mode :character Median : 94.5   
## Mean : 2.321 Mean : 101.5   
## 3rd Qu.: 0.000 3rd Qu.: 125.3   
## Max. :391.000 Max. :5400.0   
##   
## required\_car\_parking\_spaces total\_of\_special\_requests reservation\_status  
## Min. :0.00000 Min. :0.0000 Length:59695   
## 1st Qu.:0.00000 1st Qu.:0.0000 Class :character   
## Median :0.00000 Median :0.0000 Mode :character   
## Mean :0.06332 Mean :0.5675   
## 3rd Qu.:0.00000 3rd Qu.:1.0000   
## Max. :8.00000 Max. :5.0000   
##   
## reservation\_status\_date  
## Length:59695   
## Class :character   
## Mode :character   
##   
##   
##   
##

## Box plot for a numeric variable -‘lead\_time’ in subsample.

boxplot(hotel\_data\_subsample\_df\_1$lead\_time, hotel\_data\_subsample\_df\_2$lead\_time, hotel\_data\_subsample\_df\_3$lead\_time, hotel\_data\_subsample\_df\_4$lead\_time, hotel\_data\_subsample\_df\_5$lead\_time, names = c("1", "2", "3", "4", "5"), main = "Lead Time Distribution by Sub Sample 1", ylab = "Lead Time")

 #Anomaly Detection for sub sample data frame#1 : - hotel\_data\_subsample\_df\_1

## # A tibble: 2 × 2  
## hotel count  
## <chr> <int>  
## 1 City Hotel 39739  
## 2 Resort Hotel 19956

#Anomaly Detection for data frame#2 : - hotel\_data\_subsample\_df\_2

## # A tibble: 2 × 2  
## hotel count  
## <chr> <int>  
## 1 City Hotel 39685  
## 2 Resort Hotel 20010

#Anomaly Detection for data frame#3 : - hotel\_data\_subsample\_df\_3

## # A tibble: 2 × 2  
## hotel count  
## <chr> <int>  
## 1 City Hotel 39702  
## 2 Resort Hotel 19993

#Anomaly Detection for data frame#4 : - hotel\_data\_subsample\_df\_4

## # A tibble: 2 × 2  
## hotel count  
## <chr> <int>  
## 1 City Hotel 39484  
## 2 Resort Hotel 20211

#Anomaly Detection for data frame#5 : - hotel\_data\_subsample\_df\_5

## # A tibble: 2 × 2  
## hotel count  
## <chr> <int>  
## 1 City Hotel 39446  
## 2 Resort Hotel 20249

# Consistency Analysis on data frame #1

##   
## AGO ALB AND ARE ARG ARM ASM AUS AUT AZE BDI BEL BEN   
## 157 8 4 34 124 4 1 238 652 4 3 1182 3   
## BGD BGR BHR BHS BIH BLR BOL BRA BRB CAF CHE CHL CHN   
## 8 36 5 1 9 18 4 1064 1 3 823 38 517   
## CMR CN COL COM CPV CRI CUB CYP CZE DEU DMA DNK DOM   
## 3 626 34 2 12 10 2 30 75 3647 1 200 9   
## DZA ECU EGY ESP EST ETH FIN FRA FRO GAB GBR GEO GGY   
## 62 10 14 4259 34 1 243 5214 2 3 6193 12 1   
## GHA GIB GLP GNB GRC GTM GUY HKG HRV HUN IDN IMN IND   
## 1 4 1 4 71 3 1 7 52 127 15 1 80   
## IRL IRN IRQ ISL ISR ITA JAM JEY JOR JPN KAZ KEN KIR   
## 1612 47 5 29 335 1935 6 4 12 88 7 6 1   
## KNA KOR KWT LAO LBN LBY LCA LIE LKA LTU LUX LVA MAC   
## 1 50 10 1 8 4 1 1 5 46 154 29 8   
## MAR MCO MDG MDV MEX MLI MLT MMR MNE MOZ MRT MUS MWI   
## 134 3 1 4 45 2 4 1 2 34 1 2 2   
## MYS NCL NGA NLD NOR NULL NZL OMN PAK PAN PER PHL POL   
## 12 1 15 1017 312 206 36 6 5 4 16 21 481   
## PRI PRT PRY PYF QAT ROU RUS SAU SEN SGP SRB STP SUR   
## 5 24374 2 1 5 229 305 12 6 15 59 1 2   
## SVK SVN SWE SYC TGO THA TJK TUN TUR TWN TZA UKR URY   
## 27 27 500 1 1 21 6 18 148 24 4 36 16   
## USA UZB VEN VNM ZAF ZWE   
## 1018 2 10 2 38 1

# Consistency Analysis on data frame #2

##   
## ABW AGO ALB AND ARE ARG ARM AUS AUT AZE BEL BGD BGR   
## 3 179 8 2 37 88 1 201 654 11 1193 9 49   
## BHR BHS BIH BLR BOL BRA BRB BWA CAF CHE CHL CHN CIV   
## 6 1 7 14 9 1114 2 1 2 853 31 530 3   
## CMR CN COL COM CPV CRI CUB CYP CZE DEU DNK DOM DZA   
## 3 672 41 2 6 11 4 24 83 3566 218 9 53   
## ECU EGY ESP EST ETH FIN FRA FRO GAB GBR GEO GGY GHA   
## 20 11 4370 37 2 217 5220 4 1 6094 9 4 1   
## GIB GLP GNB GRC HKG HRV HUN IDN IND IRL IRN IRQ ISL   
## 10 2 4 78 18 60 128 23 83 1665 48 7 33   
## ISR ITA JAM JEY JOR JPN KAZ KIR KOR KWT LBN LBY LIE   
## 325 1922 3 3 15 107 10 1 77 5 10 2 3   
## LKA LTU LUX LVA MAC MAR MCO MDG MDV MEX MKD MLT MMR   
## 2 37 140 32 10 125 2 1 4 42 5 10 1   
## MNE MOZ MRT MUS MYS NAM NGA NIC NLD NOR NPL NULL NZL   
## 2 34 1 3 14 1 14 1 989 277 1 239 41   
## OMN PAK PAN PER PHL PLW POL PRI PRT PRY PYF QAT ROU   
## 8 13 2 11 14 1 443 9 24249 1 2 4 250   
## RUS SAU SEN SGP SLV SRB SUR SVK SVN SWE SYR TGO THA   
## 294 19 6 18 3 50 2 30 31 537 2 3 27   
## TJK TMP TUN TUR TWN TZA UKR UMI URY USA VEN VGB VNM   
## 4 1 26 115 29 1 38 2 20 1013 11 1 6   
## ZAF ZMB ZWE   
## 32 2 5

# Consistency Analysis on data frame #3

##   
## ABW AGO ALB AND ARE ARG ARM ASM ATA AUS AUT AZE BEL   
## 2 188 6 1 34 104 4 2 2 217 626 7 1178   
## BEN BGD BGR BHR BHS BIH BLR BOL BRA BRB BWA CAF CHE   
## 1 6 41 4 3 4 16 4 1066 2 1 1 872   
## CHL CHN CIV CMR CN COL CPV CRI CUB CYM CYP CZE DEU   
## 31 507 3 6 676 37 11 9 3 1 29 95 3753   
## DNK DOM DZA ECU EGY ESP EST ETH FIN FJI FRA FRO GAB   
## 229 7 55 21 16 4179 40 2 229 2 5303 2 3   
## GBR GEO GHA GIB GLP GNB GRC GTM GUY HKG HND HRV HUN   
## 6020 9 4 9 3 6 53 3 1 16 1 45 121   
## IDN IMN IND IRL IRN ISL ISR ITA JAM JEY JOR JPN KAZ   
## 18 3 84 1692 26 40 322 1916 6 4 11 94 4   
## KEN KIR KNA KOR KWT LAO LBN LBY LCA LKA LTU LUX LVA   
## 5 1 2 78 9 1 13 6 1 6 31 141 32   
## MAC MAR MCO MDV MEX MKD MLT MNE MOZ MUS MWI MYS MYT   
## 4 140 2 6 39 6 11 4 38 6 1 19 1   
## NGA NLD NOR NULL NZL OMN PAK PAN PER PHL PLW POL PRI   
## 20 1073 328 213 33 4 11 3 12 11 2 450 6   
## PRT PRY QAT ROU RUS SAU SDN SEN SGP SLV SMR SRB STP   
## 24152 2 7 235 305 24 1 8 14 2 3 45 2   
## SVK SVN SWE SYR THA TJK TMP TUN TUR TWN TZA UGA UKR   
## 25 24 523 2 31 6 3 22 131 11 5 1 39   
## UMI URY USA UZB VEN VNM ZAF ZMB ZWE   
## 1 16 1086 2 11 3 29 1 2

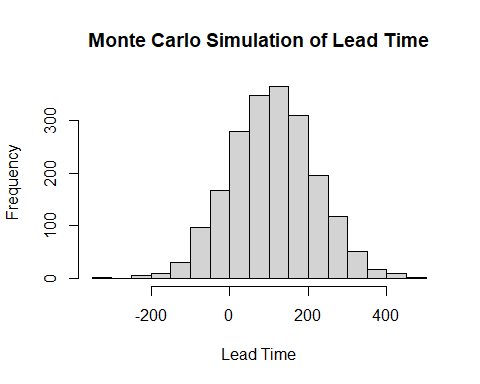
# Consistency Analysis on data frame #4

##   
## ABW AGO ALB AND ARE ARG ARM ASM ATA ATF AUS AUT AZE   
## 2 165 9 2 35 102 3 1 1 1 201 638 8   
## BDI BEL BEN BGD BGR BHR BHS BIH BLR BOL BRA BRB CAF   
## 2 1221 2 10 41 4 1 13 17 4 1136 2 3   
## CHE CHL CHN CIV CMR CN COL CPV CRI CUB CYP CZE DEU   
## 872 42 502 1 8 664 35 8 8 1 23 95 3695   
## DMA DNK DOM DZA ECU EGY ESP EST FIN FRA FRO GAB GBR   
## 1 218 6 56 14 16 4267 47 230 5178 1 2 5993   
## GEO GGY GIB GLP GNB GRC GTM HKG HND HRV HUN IDN IND   
## 22 2 11 1 5 80 3 17 1 42 121 26 85   
## IRL IRN IRQ ISL ISR ITA JAM JEY JOR JPN KAZ KEN KHM   
## 1685 45 6 40 316 1851 2 4 8 89 9 3 1   
## KOR KWT LBN LBY LCA LIE LKA LTU LUX LVA MAC MAR MCO   
## 81 11 13 6 1 1 4 41 149 24 8 115 1   
## MDV MEX MKD MLT MNE MOZ MRT MUS MYS NCL NGA NLD NOR   
## 5 54 1 8 2 36 1 4 13 2 19 1087 289   
## NPL NULL NZL OMN PAK PAN PER PHL POL PRI PRT PRY PYF   
## 1 228 36 8 8 6 19 24 410 9 24307 3 1   
## QAT ROU RUS SAU SDN SEN SGP SLV SRB SUR SVK SVN SWE   
## 7 245 283 21 1 7 18 2 49 3 30 23 521   
## SYC THA TJK TMP TUN TUR TWN TZA UGA UKR URY USA VEN   
## 1 25 5 1 16 119 26 6 2 39 16 1043 6   
## VNM ZAF ZMB ZWE   
## 9 47 1 5

# Consistency Analysis on data frame #5

##   
## ABW AGO AIA ALB AND ARE ARG ARM ASM ATA AUS AUT AZE   
## 3 201 1 4 6 24 113 4 1 3 213 663 6   
## BDI BEL BEN BGD BGR BHR BHS BIH BLR BOL BRA BRB CAF   
## 1 1205 1 5 42 2 1 4 13 4 1063 3 1   
## CHE CHL CHN CIV CMR CN COL COM CPV CRI CUB CYM CYP   
## 858 27 490 4 6 628 34 1 8 12 6 1 18   
## CZE DEU DJI DMA DNK DOM DZA ECU EGY ESP EST ETH FIN   
## 77 3569 2 2 229 4 48 17 15 4389 51 2 209   
## FJI FRA FRO GAB GBR GEO GGY GHA GIB GLP GNB GRC GTM   
## 2 5188 1 2 6006 9 1 1 11 1 7 48 1   
## HKG HND HRV HUN IDN IMN IND IRL IRN IRQ ISL ISR ITA   
## 10 1 52 110 18 1 103 1718 34 5 27 359 1896   
## JAM JEY JOR JPN KAZ KEN KIR KOR KWT LBN LIE LKA LTU   
## 5 4 11 102 7 2 1 64 6 14 4 6 43   
## LUX LVA MAC MAR MCO MDV MEX MKD MLI MLT MNE MOZ MUS   
## 155 27 7 158 2 8 48 4 1 16 2 36 3   
## MWI MYS MYT NAM NGA NLD NOR NULL NZL OMN PAK PAN PER   
## 2 13 1 3 17 1057 279 256 37 4 5 6 14   
## PHL POL PRI PRT PRY PYF QAT ROU RUS SAU SDN SEN SGP   
## 24 469 2 24150 2 2 4 251 309 21 1 6 20   
## SLV SRB STP SUR SVK SVN SWE TGO THA TJK TUN TUR TWN   
## 1 53 1 3 31 31 537 2 25 3 18 130 40   
## TZA UKR URY USA UZB VEN VGB VNM ZAF ZMB   
## 2 33 20 1100 2 14 2 4 45 1

#Monte Carlo Simulations of data set hotel\_data\_subsample\_df\_1



##Monte Carlo Simulations ##

Monte Carlo Simulations histograms diagram is graphical representationof frequency distribution of data set which were based on randam numbers genetated using Monte Carlo simulations.

##

In above histograms diagram, The Title of Graph is - “Monte Carlo Simulation of Lead Time” which is indicating that the plot represt the distribution of Lead Times(on X Axis) derived from Monte Carlo simulation. The X-axis label - Lead Time specifies the variable being measured.

##

The hight of each bar represnts the frequency of lead times falling within a specific range. Higher Bar means a higher occurence of lead time with in the range.

##

The histogram can be compared to expected or historical lead time data. Discrepancies may highlight areas where the simulation differs from real-world observations, prompting further investigation or refinement of the simulation model.

##

The insights gained from the histogram can be used for decision support, especially in scenarios where lead time variability is a critical factor. Understanding the distribution helps in making informed decisions and developing strategies to manage lead time uncertainties.

Thank You.!!!