```
# sampling 50 customers from the original data.

# loading data
library(readxl)
input_all <- read_excel("input_node.xlsx", sheet = "Customer_data")

# slicing
distribution_center <- input_all[1, c(1,3,4,5,6)]
distribution_center[1, c(4, 5)] <- 0
customer_data <- input_all[ -1, c(1,3,4,5,6)]

# sampling
set.seed(3)    # using seed 3
sample_ind <- sample(x = seq(length(customer_data$ID)), size = 50,
replace = FALSE)
customer_sample <- customer_data[sample_ind, ]
customer_sample$pack_total_weight <-
as.numeric(customer_sample$pack_total_weight)
customer_sample$pack_total_volume <-
as.numeric(customer_sample$pack_total_volume)

# integration
customer_sample <- rbind(distribution_center, customer_sample)

# writing data to .xlsx file
# library(xlsx)
# write.xlsx2(customer_sample, "customer_sample.xlsx", sheetName =
"customer_sample", append=FALSE)

# write .csv
write.csv(customer_sample, "customer_sample.csv", row.names=FALSE)

# read distance and time
dist_time <- read.delim2("input_distance-time.txt", sep = ",")
dist_time_sample <- dist_time[FALSE, -1]
for (from in customer_sample$ID){
  for (to in customer_sample$ID){
    if(from == to){
      next
    }
    ind <- which(dist_time$from_node == from & dist_time$to_node == to)
    dist <- dist_time[ind, which(names(dist_time)=="distance")]
    travel_time <- dist_time[ind, which(names(dist_time)=="spend_tm")]
    dist_time_sample <- rbind(dist_time_sample, c(from, to, dist,
travel_time))
  }
}
names(dist_time_sample) <- c("from_node", "to_node", "distance",
"travel_time")
rm(from, to, ind, dist, travel_time)
```

```
# changing unit to Kilometer and Hour
dist_time_sample$distance <- dist_time_sample$distance / 1000
dist_time_sample$travel_time <- dist_time_sample$travel_time / 60

# write to file
write.csv(dist_time_sample, "dist_time_sample.csv", row.names=FALSE)
test_data <- read.csv("dist_time_sample.csv")
```