NGUYỄN THANH HÒA _SE183091

TEST_1

#1:

```
---
Nguyễn Thanh Hòa -SE183091
Test_1
---

### 1. Install packages and library
```{r}
install.packages("RODBC")
install.packages("odbc")
install.packages("readr")
install.packages("DBI")
```

```{r}
library(RODBC)
library(odbc)
library(readr)
library(DBI)
```

Fold code onnect data Server
```

#2. Connect data Server

```
### 2. Connect data Server

```{r}
server <- "localhost"
dbname <- "ChicagoPublicSchools_test1"
user <- "sa"
password <- "123456789"

# Correct the connection string format
connecttion_string <- paste0("Driver={SQL Server};",
                             "Server=", server, ";",
                             "Database=", dbname, ";",
                             "Uid=", user, ";",
                             "Pwd=", password, ";",
                             "charset=UTF-8")

# Establish the connection
con <- DBI::dbConnect(odbc::odbc(), .connection_string =
connecttion_string)

# Check if the connection is successful
if (is.null(con)) {
  cat("Connection failed.\n")
} else {
  cat("Connection successful.\n")
}
```
```

Connection successful.

3. Import data into the database

3. Import data into the database

```{r}
chicago_schools <-
read.csv("D:/FPT/Kihoc/fall2024/DRS301m/lab/test/ChicagoPublicS
chools.csv")
dbWriteTable(con,
             "chicago_public_schools",
             chicago_schools,
             overwrite = TRUE)

print("Quá trình nhập dữ liệu đã thành công")
```

```
[1] "Quá trình nhập dữ liệu đã thành công"
```

```{r}
row_count <- dbGetQuery(con,
                        "SELECT COUNT(*) AS count FROM
chicago_public_schools")
cat("So luong ban ghi da dc nhap:", row_count$count, "\n")
```

```
So luong ban ghi da dc nhap: 566
```

```{r}
cat("Qua trinh nhap du lieu da hoan thanh\n")
```

```
Qua trinh nhap du lieu da hoan thanh
```

4. Summarize the data

4. Summarize the data

```{r}
chicago_data <- dbReadTable(con,"chicago_public_schools")
summary(chicago_data)
```

```
    School_ID      NAME_OF_SCHOOL
Elementary..Middle..or.High.School Street_Address      City
State
 Min.   :400018    Length:566          Length:566
Length:566          Length:566           Length:566
 1st Qu.:609873    Class :character    Class :character
Class :character    Class :character    Class :character
 Median :610053    Mode  :character    Mode  :character
Mode  :character    Mode  :character    Mode  :character
 Mean   :609682
 3rd Qu.:610213
 Max.   :610544


     ZIP_Code      Phone_Number              Link
Network_Manager    Collaborative_Name
Adequate_Yearly_Progress_Made_
 Min.   :60605    Length:566          Length:566
Length:566          Length:566           Length:566
 1st Qu.:60618    Class :character    Class :character    Class
:character    Class :character    Class :character
 Median :60625    Mode  :character    Mode  :character    Mode
:character    Mode  :character    Mode  :character
 Mean   :60630
 3rd Qu.:60639
 Max.   :60827
```

```
 Track_Schedule        CPS_Performance_Policy_Status
CPS_Performance_Policy_Level HEALTHY_SCHOOL_CERTIFIED
Safety_Icon
 Length:566            Length:566                    Length:566
Length:566               Length:566
 Class :character    Class :character              Class
:character            Class :character       Class
:character
 Mode  :character    Mode  :character              Mode
:character              Mode  :character       Mode
:character
```

```
   SAFETY_SCORE   Family_Involvement_Icon
Family_Involvement_Score Environment_Icon    Environment_Score
Instruction_Icon
 Min.    : 1.0    Length:566                  Length:566
Length:566           Min.    : 1.00    Length:566
 1st Qu.:35.0    Class :character           Class :character
Class :character    1st Qu.:37.00    Class :character
 Median :48.0    Mode  :character           Mode  :character
Mode  :character    Median :47.00    Mode  :character
 Mean    :49.5
Mean    :47.77
 3rd Qu.:61.0
3rd Qu.:58.00
 Max.    :99.0
Max.    :99.00
 NA's    :53
NA's    :53
 Instruction_Score Leaders_Icon        Leaders_Score
Teachers_Icon        Teachers_Score      Parent_Engagement_Icon
 Min.    : 1.00    Length:566          Length:566
Length:566          Length:566          Length:566
 1st Qu.:37.00     Class :character    Class :character     Class
:character    Class :character    Class :character
 Median :47.00     Mode  :character    Mode  :character     Mode
:character    Mode  :character    Mode  :character
 Mean    :48.29
 3rd Qu.:59.00
 Max.    :99.00
 NA's    :53
 Parent_Engagement_Score Parent_Environment_Icon
Parent_Environment_Score AVERAGE_STUDENT_ATTENDANCE
 Length:566                  Length:566                  Length:566
Length:566
 Class :character           Class :character            Class
:character           Class :character
 Mode  :character           Mode  :character            Mode
:character           Mode  :character
```

```
 Rate_of_Misconducts__per_100_students_
Average_Teacher_Attendance
Individualized_Education_Program_Compliance_Rate
 Min.    :   0.000                        Length:566
Length:566
 1st Qu.:   4.525                         Class :character
Class :character
 Median :  12.250                         Mode  :character
Mode   :character
 Mean   :  21.086
 3rd Qu.:  26.975
 Max.   : 251.600


 Pk_2_Literacy__      Pk_2_Math__
Gr3_5_Grade_Level_Math__ Gr3_5_Grade_Level_Read__
Gr3_5_Keep_Pace_Read__
 Length:566           Length:566           Length:566
Length:566               Length:566
 Class :character     Class :character     Class :character
Class :character         Class :character
 Mode  :character     Mode  :character     Mode  :character
Mode  :character         Mode  :character


 Gr3_5_Keep_Pace_Math__ Gr6_8_Grade_Level_Math__
Gr6_8_Grade_Level_Read__ Gr6_8_Keep_Pace_Math_
Gr6_8_Keep_Pace_Read__
 Length:566             Length:566             Length:566
Length:566             Length:566
 Class :character       Class :character       Class
:character             Class :character       Class :character
 Mode  :character       Mode  :character       Mode
:character             Mode  :character       Mode  :character
```

```
 Gr_8_Explore_Math__ Gr_8_Explore_Read__ ISAT_Exceeding_Math__
ISAT_Exceeding_Reading__ ISAT_Value_Add_Math
ISAT_Value_Add_Read
 Length:566          Length:566          Min.   :  0.000
Min.   :  0.00       Min.   :-3.50000    Min.   :-5.00000
 Class :character    Class :character    1st Qu.:  9.175
1st Qu.:  6.10       1st Qu.:-0.60000    1st Qu.:-0.60000
 Mode  :character    Mode  :character    Median : 16.000
Median : 10.95       Median : 0.00000    Median : 0.00000
                                         Mean   : 20.356
Mean   : 15.67       Mean   : 0.06859    Mean   : 0.02265
                                         3rd Qu.: 24.900
3rd Qu.: 19.52       3rd Qu.: 0.70000    3rd Qu.: 0.62500
                                         Max.   :100.000
Max.   :100.00       Max.   : 3.60000    Max.   : 4.90000
                                         NA's   :90
NA's   :90           NA's   :98          NA's   :98
 ISAT_Value_Add_Color_Math ISAT_Value_Add_Color_Read
Students_Taking__Algebra__ Students_Passing__Algebra__
X9th.Grade.EXPLORE..2009.
 Length:566                Length:566
Length:566                 Length:566
Length:566
 Class :character          Class :character          Class
:character                 Class :character          Class
:character
 Mode  :character          Mode  :character          Mode
:character                 Mode  :character          Mode
:character
```

```
 X9th.Grade.EXPLORE..2010. X10th.Grade.PLAN..2009.
X10th.Grade.PLAN..2010. Net_Change_EXPLORE_and_PLAN
X11th.Grade.Average.ACT..2011.
 Length:566                Length:566                Length:566
Length:566                 Length:566
 Class :character          Class :character          Class
:character                 Class :character          Class :character
 Mode  :character          Mode  :character          Mode
:character                 Mode  :character          Mode  :character
```

```
 Net_Change_PLAN_and_ACT College_Eligibility__
Graduation_Rate__  College_Enrollment_Rate__
COLLEGE_ENROLLMENT General_Services_Route
 Length:566                Length:566                Length:566
Length:566                 Min.   : 21.0    Min.     :29.00
 Class :character          Class :character       Class
:character   Class :character            1st Qu.: 342.2     1st
Qu.:34.00
 Mode  :character          Mode  :character       Mode
:character   Mode  :character            Median : 504.0
Median :39.00

Mean    : 626.1      Mean    :39.08

3rd Qu.: 790.5      3rd Qu.:45.00

Max.    :4368.0     Max.    :49.00
```

```
 Freshman_on_Track_Rate__   X_COORDINATE          Y_COORDINATE
Latitude          Longitude         COMMUNITY_AREA_NUMBER
 Length:566                 Min.    :1118114    Min.    :1817242
Min.    :41.65    Min.    :-87.84    Min.    : 1.00
 Class :character           1st Qu.:1152166    1st Qu.:1860248
1st Qu.:41.77    1st Qu.:-87.72    1st Qu.:23.00
 Mode  :character           Median :1163553    Median :1886889
Median :41.85    Median :-87.68    Median :31.50
                           Mean    :1163016    Mean    :1885663
Mean    :41.84    Mean    :-87.68    Mean    :37.71
                           3rd Qu.:1174256    3rd Qu.:1911652
3rd Qu.:41.91    3rd Qu.:-87.64    3rd Qu.:58.00
                           Max.    :1202811    Max.    :1950960
Max.    :42.02    Max.    :-87.53    Max.    :77.00

 COMMUNITY_AREA_NAME         Ward            Police_District
Location
 Length:566              Min.    : 1.00    Min.    : 1.00
Length:566
 Class :character        1st Qu.:10.00    1st Qu.: 7.00    Class
:character
 Mode  :character        Median :21.00    Median :10.00    Mode
:character
                         Mean    :22.06    Mean    :11.71
                         3rd Qu.:32.75    3rd Qu.:17.00
                         Max.    :50.00    Max.    :25.00
```

5. Retrieve table structure information

```{r}
columns_query <- dbGetQuery(con,
                            "SELECT School_ID
                            FROM chicago_public_schools
                            WHERE School_ID = 610281")

columns_query
```

Description: df [1 × 1]

School_ID <int>
610281

1 row

6. Rename columns in the chicago_public_schools table

```{r}
dbExecute(con, "EXEC sp_rename
'chicago_public_schools.[Elementary..Middle..or.High.School]',
'School_level', 'COLUMN';")
```

[1] 0

7. Count the number of elementary schools

```{r}
query_elementary_count <- dbGetQuery(con,
                                     "SELECT
COUNT(School_level) AS count
                                     FROM
chicago_public_schools
                                     WHERE School_level =
'ES';")
query_elementary_count
```

Description: df [1 × 1]

count <int>
462

1 row

8. Find the highest safety score

```{r}
query_highest_safety_score <- dbGetQuery(con,"
                                         SELECT
MAX(SAFETY_SCORE) AS highest_safety_score
                                         FROM
chicago_public_schools")
query_highest_safety_score
```

Description: df [1 × 1]

highest_safety_score <int>
99

1 row

9. Write R code to clean the columns names of a data frame called chicago_schools.

```{r}
clean_column_names <- function(names) {
  names <- gsub("+", "", names)
  names <- gsub("[._]+$", "", names)
  names <- gsub("\\.", "", names)
  return(names)
}
```

10. Check for missing data

```{r}
missing_data <- colSums(is.na(chicago_data))
missing_data
```

```
                      School_ID                          NAME_OF_SCHOOL
                              0                                       0
 Elementary..Middle..or.High.School                      Street_Address
                              0                                       0
                           City                                   State
                              0                                       0
                       ZIP_Code                            Phone_Number
                              0                                       0
                           Link                         Network_Manager
                              0                                       0
              Collaborative_Name           Adequate_Yearly_Progress_Made_
                              0                                       0
                 Track_Schedule             CPS_Performance_Policy_Status
                              0                                       0
    CPS_Performance_Policy_Level             HEALTHY_SCHOOL_CERTIFIED
                              0                                       0
                    Safety_Icon                            SAFETY_SCORE
                              0                                      53
        Family_Involvement_Icon                Family_Involvement_Score
                              0                                       0
               Environment_Icon                       Environment_Score
                              0                                      53
               Instruction_Icon                       Instruction_Score
                              0                                      53
                   Leaders_Icon                           Leaders_Score
                              0                                       0
                  Teachers_Icon                          Teachers_Score
                              0                                       0
        Parent_Engagement_Icon                Parent_Engagement_Score
                              0                                       0
       Parent_Environment_Icon               Parent_Environment_Score
                              0                                       0
      AVERAGE_STUDENT_ATTENDANCE  Rate_of_Misconducts__per_100_students_
                              0                                       0
     Average_Teacher_Attendance Individualized_Education_Program_Compliance_Rate
                              0                                       0
                  Pk_2_Literacy__                              Pk_2_Math__
```

```
                              0
Gr3_5_Grade_Level_Math__        Gr3_5_Grade_Level_Read__
                     0                               0
Gr3_5_Keep_Pace_Read__          Gr3_5_Keep_Pace_Math__
                     0                               0
Gr6_8_Grade_Level_Math__        Gr6_8_Grade_Level_Read__
                     0                               0
Gr6_8_Keep_Pace_Math_           Gr6_8_Keep_Pace_Read__
                     0                               0
Gr_8_Explore_Math__             Gr_8_Explore_Read__
                     0                               0
ISAT_Exceeding_Math__           ISAT_Exceeding_Reading__
                    90                              90
ISAT_Value_Add_Math             ISAT_Value_Add_Read
                    98                              98
ISAT_Value_Add_Color_Math       ISAT_Value_Add_Color_Read
                     0                               0
Students_Taking__Algebra__      Students_Passing__Algebra__
                     0                               0
X9th.Grade.EXPLORE..2009.       X9th.Grade.EXPLORE..2010.
                     0                               0
X10th.Grade.PLAN..2009.         X10th.Grade.PLAN..2010.
                     0                               0
Net_Change_EXPLORE_and_PLAN     X11th.Grade.Average.ACT..2011.
                     0                               0
Net_Change_PLAN_and_ACT         College_Eligibility__
                     0                               0
Graduation_Rate__               College_Enrollment_Rate__
                     0                               0
COLLEGE_ENROLLMENT              General_Services_Route
                     0                               0
Freshman_on_Track_Rate__        X_COORDINATE
                     0                               0
Y_COORDINATE                    Latitude
                     0                               0
Longitude                       COMMUNITY_AREA_NUMBER
                     0                               0

                     0                               0
COMMUNITY_AREA_NAME             Ward
                     0                               0
Police_District                 Location
                     0                               0


                     0                               0
COMMUNITY_AREA_NAME             Ward
                     0                               0
Police_District                 Location
                     0                               0
```