

data preparation 6 - 10 / 3 / 2023

## Script per la preparazione dei dati

```
cols_to_keep <- c("Giorno",
                  "Comune di partenza",
                  "Comune di arrivo",
                  "Tipologia viaggiatore",
                  "Viaggi",
                  "00-03",
                  "03-06",
                  "06-09",
                  "09-12",
                  "12-15",
                  "15-18",
                  "18-21",
                  "21-24")

#read data and filter relevant columns
df <- read_excel('./data/EXPORT FVG _ 6-10 MARZO 2023.xlsx') %>%
  select(all_of(cols_to_keep))

# remove whitespace in column names
colnames(df) <- gsub(" ", "_", colnames(df))

print(df)

## # A tibble: 257,985 x 13
##   Giorno      Comune_di_partenza Comune_di_arrivo Tipologia_viaggiatore
##   <dtm>      <chr>                <chr>            <chr>
## 1 2023-03-06 00:00:00 Porcia      Pordenone      Residente Comunale d~
## 2 2023-03-06 00:00:00 Pordenone   Porcia         Residente Comunale d~
## 3 2023-03-06 00:00:00 Pordenone   Cordenons      Residente Comunale d~
## 4 2023-03-06 00:00:00 Cordenons   Pordenone      Residente Comunale d~
## 5 2023-03-06 00:00:00 Tavagnacco   Udine          Residente Comunale d~
## 6 2023-03-06 00:00:00 Udine      Tavagnacco     Residente Comunale d~
## 7 2023-03-06 00:00:00 Trieste   Muggia         Residente Comunale d~
## 8 2023-03-06 00:00:00 Muggia     Trieste        Residente Comunale d~
## 9 2023-03-06 00:00:00 Sacile      Fontanafredda  Residente Comunale d~
## 10 2023-03-06 00:00:00 Fontanafredda Sacile         Residente Comunale d~
## # i 257,975 more rows
## # i 9 more variables: Viaggi <dbl>, `00-03` <dbl>, `03-06` <dbl>,
## #   `06-09` <dbl>, `09-12` <dbl>, `12-15` <dbl>, `15-18` <dbl>, `18-21` <dbl>,
## #   `21-24` <dbl>

# filter relevant rows
print("BEROFRE filtering:")

## [1] "BEROFRE filtering:"
```

```

table(df$Tipologia_viaggiatore)

##
##      Residente Comunale di ritorno da altro Comune di Provincia di residenza
##                                     34628
## Residente Comunale di ritorno da altro Comune di Regione (esterno alla Provincia)
##                                     23002
##      Residente Comunale diretto in altro Comune di Provincia di residenza
##                                     35418
## Residente Comunale diretto in altro Comune di Regione (esterno alla Provincia)
##                                     24462
##      Viaggiatore Estero diretto in qualsiasi altro Comune (interno alla Regione)
##                                     18993
## Viaggiatore Nazionale diretto in qualsiasi altro Comune (interno alla Regione)
##                                     39014
##      Viaggiatore Provinciale diretto in altro Comune di Provincia di residenza
##                                     37063
## Viaggiatore Provinciale diretto in altro Comune di Regione (esterno alla Provincia)
##                                     13853
##      Viaggiatore Regionale diretto in altro Comune di Provincia di residenza
##                                     14733
## Viaggiatore Regionale diretto in altro Comune di Regione (esterno alla Provincia)
##                                     16819

rows_to_keep <- c(
  'Residente Comunale diretto in altro Comune di Provincia di residenza',
  'Residente Comunale di ritorno da altro Comune di Provincia di residenza')

df <- df %>%
  filter(Tipologia_viaggiatore %in% rows_to_keep)
print("AFTER filtering:")

## [1] "AFTER filtering:"

table(df$Tipologia_viaggiatore)

##
## Residente Comunale di ritorno da altro Comune di Provincia di residenza
##                                     34628
## Residente Comunale diretto in altro Comune di Provincia di residenza
##                                     35418

df %>% write.csv("dati_6_10_marzo_2023.csv")

```