

Training day 2

Ripasso

Data import

```
library(readxl)
mdata <- read_excel("media-data.xlsx")
```

Data manipulation

Select

```
library(dplyr)

tv_radio_data <- mdata %>%
  select(Week, Start, Qty, Eventi, Temperatura,
         Tv_15 = TV_15, Tv_30, Radio_15, Radio_30)
```

Filter

```
tv_radio_data_17 <- tv_radio_data %>%
  filter(Start >= "2017-01-01" & Start <= "2017-12-31")
```

New topics

Mutate

Trasformazioni delle variabili di partenza

```
tr_full <- tv_radio_data %>%
  mutate(Tv_15 = round(Tv_15, 0),
         Tv_30 = round(Tv_30, 0))
```

Operazioni algebriche tra variabili

```
tr_full <- tr_full %>%
  mutate(Tv = Tv_15 + Tv_30)
```

Case when

```
tr_full <- tr_full %>%
  mutate(If_Tv = case_when(
    Tv > 0 ~ 1,
    TRUE ~ 0
  ))
```

Rifare esercizio con Radio

Mutate_if

```
tr_full <- tr_full %>%  
  mutate_if(is.numeric, round, digits = 0)
```

```
tr_full <- tr_full %>%  
  mutate(Radio = Radio_15 + Radio_30,  
         Tv_Radio = Tv + Radio,  
         If_radio = case_when(  
           Tv_Radio > 0 ~ 1,  
           TRUE ~ 0)  
  )
```

Lubridate

```
library(lubridate)  
  
tr_full <- tr_full %>%  
  mutate(Giorno = day(Start),  
         Mese_num = month(Start),  
         Mese = month(Start, label = T),  
         Anno = year(Start))
```

Join

```
season_data <- data.frame(  
  Mese_num = 1:12,  
  Stagione = c("Alta_Stagione", "Alta_Stagione", "Alta_Stagione",  
               "Alta_Stagione", "Alta_Stagione", "Alta_Stagione",  
               "Bassa_Stagione", "Bassa_Stagione", "Bassa_Stagione",  
               "Alta_Stagione", "Alta_Stagione", "Alta_Stagione")  
)  
  
tr_full <- tr_full %>%  
  left_join(season_data)
```

Write data

```
write.csv(tr_full, file = "tv_radio_data.csv")
```

Read data

```
tv_radio_data <- read.csv("tv_radio_data.csv")
```

```
tv_radio_data <- tv_radio_data %>%
  mutate(Start = ymd(Start)) %>%
  filter(!is.na(Week))
```

Analisi delle vendite

In questa sezione analizziamo le vendite realizzate nell'ultimo triennio

Totale venduto

Il totale dei pezzi venduti nel triennio è pari a:

Vendite per anno

```
tv_radio_data %>%
  group_by(Anno) %>%
  summarise(Vendite = sum(Qty))
```

Grafico vendite

```
library(ggplot2)
```

```
qplot(x = Start, y = Qty, data = tv_radio_data, geom = "line") +
  ylim(3000, 6000) +
  theme_minimal() +
  labs(title = "Andamento vendite nel triennio")
```

```
# qplot(x = Start, y = Qty, data = tv_radio_data, geom = "line",
#       colour = Anno) +
#   ylim(3000, 6000) +
#   theme_minimal() +
#   labs(title = "Andamento vendite nel triennio")
```

```
qplot(x = Week, y = Qty, data = tv_radio_data, geom = "line",
      colour = Anno, facets = .~Anno) +
  ylim(3000, 6000) +
  theme_minimal() +
  labs(title = "Andamento vendite nel triennio")
```

Alta vs bassa stagione

```
tv_radio_data %>%
  group_by(Anno, Stagione) %>%
  summarise(Vendite = sum(Qty),
            Settimane = n())
```