



Pratica 4

Roteiro 4 - Correlação

<https://luisfelipecbr.github.io/mti2020/roteiros/roteiro4.html>

Codigo e dados disponivel em →

https://github.com/wesleyclzns/analise_dados_planejamento_territorial_UFABC

Importando e elendo os dados

```
#Baseado no roteiro 4
#Caminho para a pasta: D:\CM\ADPT

# INSTALAR EM NOVOS PCs ANTES DE RODAR O SCRIPT
#install.packages("tidyverse")
#install.packages("corrplot")
#install.packages("Hmisc")

#RODAR SEMPRE PARA CARREGAR A LIB
library(tidyverse)
library(corrplot)
library(Hmisc)

getwd()
#"D:/CM/ADPT/scripts"

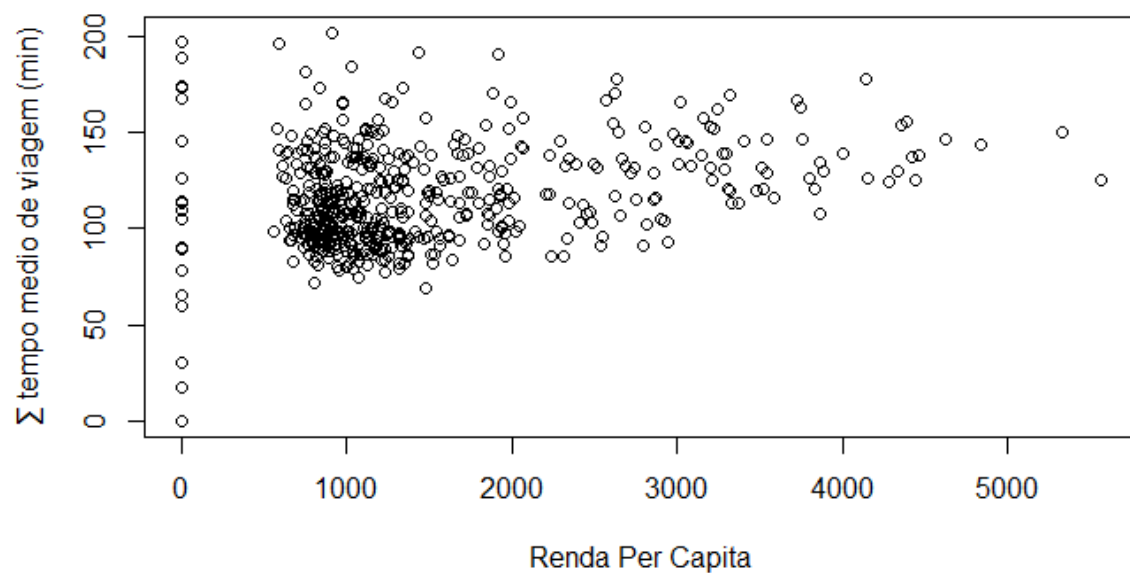
od <- read.csv("D:/CM/ADPT/dados/od.csv")
head(od)
names(od) #117 colunas
```

Criando uma nova coluna com a soma dos tipos de modais

```
od["tmv_somaTotalModal"] <- od$tmv_coletivo + od$tmv_individual + od$tmv_pe + od$tmv_bike
od$tmv_somaTotalModal
```

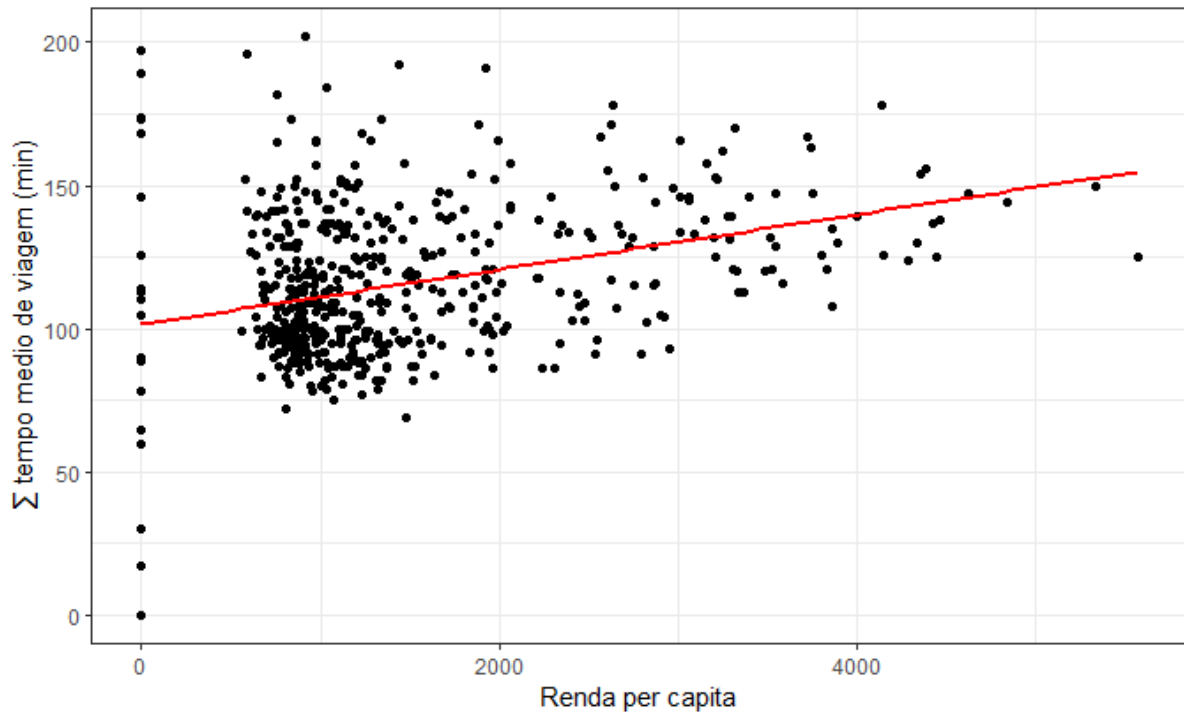
Diagrama de dispersão

```
plot(x = od$renPerCap,  
     y = od$tmv_somaTotalModal,  
     xlab = "Renda Per Capita",  
     ylab = "Σ tempo medio de viagem (min)")
```



Dispersão com o GG

```
ggplot(data = od, aes(x = renPerCap, y = tmv_somaTotalModal)) +  
  geom_point() + geom_smooth(data =  
    lm(formula = tmv_somaTotalModal ~ renPerCap, data = od),  
    method = "lm", col = "red", se = FALSE) + theme_bw() +  
  xlab("Renda per capita") +  
  ylab("Σ tempo medio de viagem (min)")
```



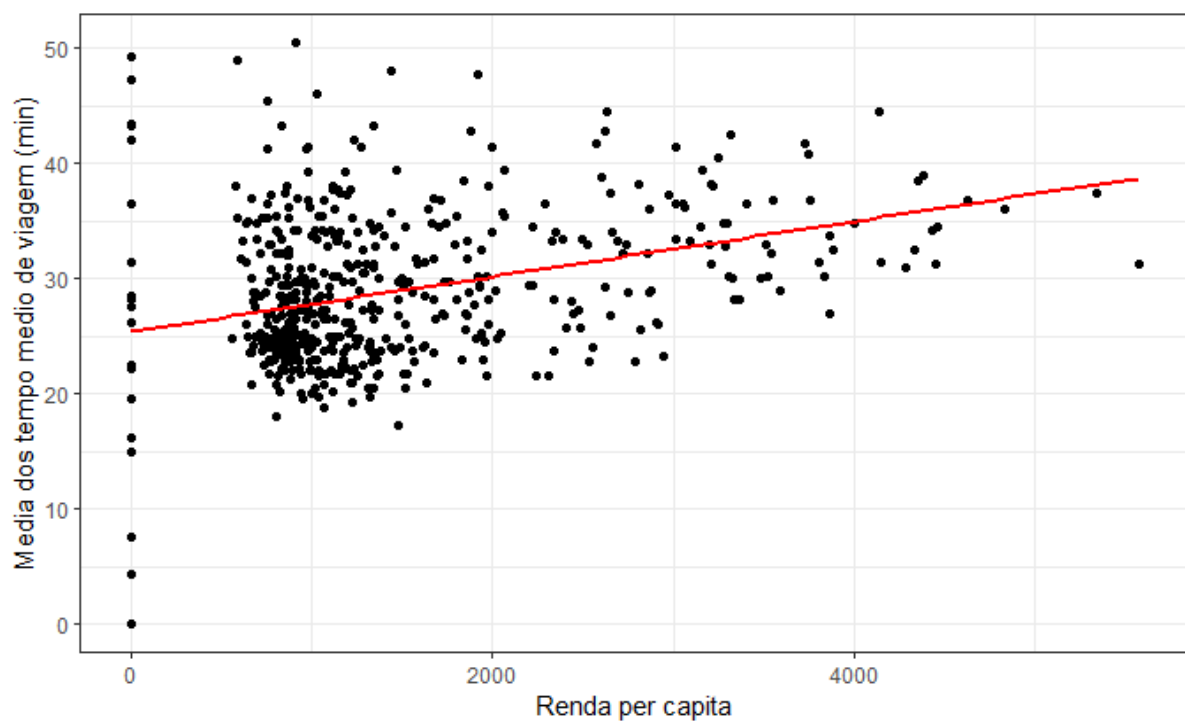
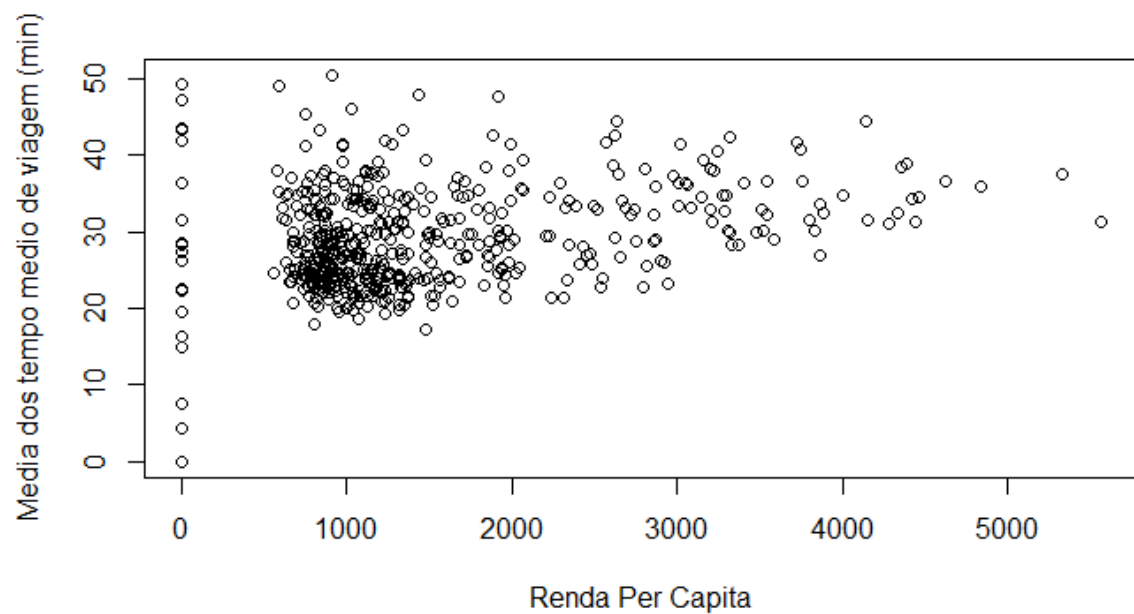
Acredito que não deveria estar fazendo o somatorio do TMV_total e sim a Media das Medias do TMV_total

Criando uma nova coluna com a media do TMV e Gerando um Plot e um GGplot

```
#Renda media / Media do Tempo Medio de Viagem
od["tmv_mediaTotalModal"] <- od$tmv_somaTotalModal/4
od$tmv_mediaTotalModal

plot(x = od$renPerCap,
      y = od$tmv_mediaTotalModal,
      xlab = "Renda Per Capita",
      ylab = "Media dos tempo medio de viagem (min)")

ggplot(data = od, aes(x = renPerCap, y = tmv_mediaTotalModal)) + geom_point() + geom_smooth(data = lm(formula = tmv_mediaTotalModal ~ renPerCap, data = od), method = "lm", col = "red", se = FALSE) + theme_bw() + xlab("Renda per capita") + ylab("Media dos tempo medio de viagem (min)")
```



aparentemente nada mudou

Correlação e Teste de Significancia

```

#Coeficiente de correlação
cor(x = od$renPerCap,
    y = od$tmv_mediaTotalModal,
    method = "pearson",
    use = "complete.obs")
## Resultado - > 0.3256155

cor.test(x = od$renPerCap,
    y = od$tmv_mediaTotalModal,
    method = "pearson",
    alternative = "two.sided",
    conf.level = 0.95)

# t = 7.8153, df = 515, p-value = 3.108e-14
# alternative hypothesis: true correlation is not equal to 0
# 95 percent confidence interval: 0.2462958 0.4006023
# sample estimates: cor 0.3256155

```

Matriz de Correlação

```

od["mo_trabTotal"] <- od$mo_traIndus + od$mo_trabCome + od$mo_trabServ
od$tmv_mo_trabTotal

od["crianças"] <- od$popIdd_0.3 + od$popIdd_4.6 + od$popIdd_7.10
od["adolescentes"] <- od$popIdd_11.14 + od$popIdd_15.17
od["jovens"] <- od$popIdd_18.22 + od$popIdd_23.29
od["adultos"] <- od$popIdd_30.39 + od$popIdd_40.49 + od$popIdd_50.59
od["idosos"] <- od$popIdd_60.mais
od[1:6,c("nomeZona", "crianças", "adolescentes", "jovens", "adultos", "idosos")]

#Matriz de Correlação
names(od)
od %>%
  select(populacao, matriEscolar, empregos, renMedFam, res_assSem, res_autonomo, res_t
rabFamiliar, vp_dirigindoAuto, vp_passAuto, vp_taxiCom, vp_taxiNCom, vp_dirigindoMoto,
vp_passMoto, vp_bike, vp_pe, tmv_coletivo, tmv_individual, tmv_pe, tmv_bike, mo_educac
ao, mo_procEmp, mo_lazer, mo_saude, mo_trabTotal, crianças, adolescentes, jovens, adult
os, idosos, popFem, popMasc, renPerCap, renMedianFam, renMedFam) %>%
  cor(method = "pearson",
    use = "complete.obs") %>%
  round(digits = 2) %>%
  corrplot(method = c("pie"),
    tl.cex = 1.5, cl.cex = 1.5)

```

