**Notes for very short intro to R**

**Link để download dữ liệu:**

https://github.com/tuanvnguyen/Regression-Book

library(tidyverse)

library(dplyr)

library(ggplot2)

library(table1)

**# Reading data**

os = read.csv(.... "Osteo data.csv")

# Using mutate

newdata = os %>% mutate(

feet = Height\*0.0328084,

pound = Weight\*2.20462)

**# Recoding data**

t = os %>% mutate

(bmig = cut(bmi,

breaks = c(-Inf, 18.5, 24.9, 29.9, Inf),

labels = c("Underweight", "Normal", "Overweight", "Obese"))

**# Selecting variables**

newdata = os %>% dplyr::select(id, lean.mass, age)

# Select observations

newdata = os %>% filter(age > 50)

newdata = os %>% filter(age > 50 & osteo==2)

**# Summarising data**

os %>% summarise(mean(age), sd(age), mean(pcfat), sd(pcfat))

os %>% summarise(mean.age = mean(age), mean.pcfat = mean(pcfat))

**# Putting togther**

os %>% dplyr::select(age, lean.mass, fat.mass, pcfat, bmi, osta, osteo.group) %>% group\_by(osteo.group) %>% summarise\_all(mean)

**# Merging data**

df1 = os[, c("id", "age", "bmi")]

df2 = os[, c("id", "pcfat", "osteo.group")]

df = full\_join(df1, df2)

# Using table1

table1(~ age + lean.mass + fat.mass + pcfat + osteo.group, data=os)

table1(~ age + lean.mass + fat.mass + pcfat | osteo.group, data=os)

**# Using ggplot2**

ggplot(data=os aes(x = pcfat)) + geom\_histogram(fill="blue", col="white)

ggplot(data=os aes(x = osteo.group, y = pcfat, col=osteo.group)) + geom\_boxplot()

ggplot(data=os aes(x = bmi, y = pcfat, col=osteo.group)) + geom\_point()

ggplot(data=os aes(x = bmi, y = pcfat, col=osteo.group)) + geom\_point() + labs(x="Body mass index", y="Percent body fat")

**# Using GGally**

library(GGally)

os %>% dplyr::select(age, lean.mass, fat.mass, pcfat) %>% ggapirs()