# MILESTONE 4

library(readr)

sales <- read\_csv("vgsales.csv")

View(sales)

hist(sales$NA\_Sales, xlab = "Revenue (in millions)", ylab = "Number of Games", main = "Games Sold in North America")

hist(sales$EU\_Sales, xlab = "Revenue (in millions)", ylab = "Number of Games", main = "Games Sold in Europe")

hist(sales$JP\_Sales, xlab = "Revenue (in millions)", ylab = "Number of Games", main = "Games Sold in Japan")

barplot(table(sales$Genre),xlab = "Genre", ylab = "Number of Games", main = "Sales by Genre",las = 2)

# MILESTONE 6

Mean\_NA\_Sales <-mean(vgsales$NA\_Sales) 0.2646674

Mean\_EU\_Sales <-mean(vgsales$EU\_Sales) 0.146652

Mean\_JP\_Sales <-mean(vgsales$JP\_Sales) 0.07778166

Mean\_Other\_Sales <-mean(vgsales$Other\_Sales) 0.04806302

Mean\_Global\_Sales <-mean(vgsales$Global\_Sales) 0.5374407

Median\_NA\_Sales <- median(vgsales$NA\_Sales) 0.08

Median\_EU\_Sales <- median(vgsales$EU\_Sales) 0.02

Median\_JP\_Sales <- median(vgsales$JP\_Sales) 0

Median\_Other\_Sales <- median(vgsales$Other\_Sales) 0.01

Median\_Global\_Sales <- median(vgsales$Global\_Sales) 0.17

Stndrd\_Dev\_NA\_Sales <- sd(vgsales$NA\_Sales) 0.816683

Stndrd\_Dev\_EU\_Sales <- sd(vgsales$EU\_Sales) 0.5053512

Stndrd\_Dev\_JP\_Sales <- sd(vgsales$JP\_Sales) 0.3092906

Stndrd\_Dev\_Other\_Sales <- sd(vgsales$Other\_Sales) 0.1885884

Stndrd\_Dev\_Global\_Sales <- sd(vgsales$Global\_Sales) 1.555028

Variance\_NA\_Sales <- var(vgsales$NA\_Sales) 0.6669712

Variance\_EU\_Sales <- var(vgsales$EU\_Sales) 0.2553799

Variance\_JP\_Sales <- var(vgsales$JP\_Sales) 0.0956607

Variance\_Other\_Sales <- var(vgsales$Other\_Sales) 0.03556559

Variance\_Global\_Sales <- var(vgsales$Global\_Sales) 2.418112

# MILESTONE 7

plot(NA\_Sales~JP\_Sales, data = vgsales, main = "Possible Correlation Between JP and NA Sales”, xlab = "JP (millions)", ylab = "NA (millions)")

lm\_vgsales = lm(NA\_Sales~JP\_Sales, data = vgsales)

coef(lm\_vgsales) (Intercept) JP\_Sales

0.1722889 1.1876652

abline(lm\_vgsales, col = "green", lwd = 3)

# MILESTONE 8

-qt(0.025,9) 2.262157

# MILESTONE 9

NAvsJP.lm <- lm(JP\_Sales~NA\_Sales, data = vgsales)

coef(NAvsJP.lm) (Intercept) NA\_Sales

0.0326978 0.1703415

summary(NAvsJP.lm) Call:

lm(formula = JP\_Sales ~ NA\_Sales, data = vgsales)

Residuals:

Min 1Q Median 3Q Max

-4.3400 -0.0634 -0.0412 -0.0127 8.2676

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 0.032698 0.002254 14.51 <2e-16 \*\*\*

NA\_Sales 0.170342 0.002626 64.88 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2762 on 16596 degrees of freedom

Multiple R-squared: 0.2023, Adjusted R-squared: 0.2023

F-statistic: 4209 on 1 and 16596 DF, p-value: < 2.2e-16

plot(fitted.values(NAvsJP.lm), resid(NAvsJP.lm))

abline(h=0)

plot(fitted.values(NAvsJP.lm), resid(NAvsJP.lm))

abline(h=0)

hist(resid(NAvsJP.lm))

# MILESTONE 10

qnorm(0.99) 2.326348

# MILESTONE 11

qnorm(0.95) 1.644854