> # Q1

> #(1)

> setwd("P:/R")

> county <- read.csv("county\_data.csv", stringsAsFactors = FALSE, na.strings = "")

> county\_pop <- county[c("State", "County", "TotalPop")]

> county\_pop$State <- as.factor(county\_pop$State)

> max\_county\_pop <- aggregate(county\_pop[, c("TotalPop")], list(county\_pop$State), max)

> index <- match(max\_county\_pop$x, county$TotalPop)

> max\_county\_pop$x <- county\_pop$County[index]

> population <- aggregate(county\_pop[, c("TotalPop")], by=list(county\_pop$State), max)

> max\_county\_pop <- cbind(max\_county\_pop, population$x)

> names(max\_county\_pop) <- c("State", "County", "TotalPop")

> total\_pop\_178942 <- county\_pop[county\_pop$TotalPop == 178942,]

> max\_county\_pop$County[43] <- total\_pop\_178942[2,2]

> max\_county\_pop

State County TotalPop

1 Alabama Jefferson 659026

2 Alaska Anchorage Municipality 299107

3 Arizona Maricopa 4018143

4 Arkansas Pulaski 390463

5 California Los Angeles 10038388

6 Colorado El Paso 655024

7 Connecticut Fairfield 939983

8 Delaware New Castle 549643

9 District of Columbia District of Columbia 647484

10 Florida Miami-Dade 2639042

11 Georgia Fulton 983903

12 Hawaii Honolulu 984178

13 Idaho Ada 417501

14 Illinois Cook 5236393

15 Indiana Marion 926335

16 Iowa Polk 452369

17 Kansas Johnson 566814

18 Kentucky Jefferson 755809

19 Louisiana East Baton Rouge 444690

20 Maine Cumberland 286119

21 Maryland Montgomery 1017859

22 Massachusetts Middlesex 1556116

23 Michigan Wayne 1778969

24 Minnesota Hennepin 1197776

25 Mississippi Hinds 245874

26 Missouri St. Louis 1001327

27 Montana Yellowstone 153692

28 Nebraska Douglas 537655

29 Nevada Clark 2035572

30 New Hampshire Hillsborough 403972

31 New Jersey Bergen 926330

32 New Mexico Bernalillo 673943

33 New York Kings 2595259

34 North Carolina Mecklenburg 990288

35 North Dakota Cass 162500

36 Ohio Cuyahoga 1263189

37 Oklahoma Oklahoma 754480

38 Oregon Multnomah 768418

39 Pennsylvania Philadelphia 1555072

40 Puerto Rico San Juan 371400

41 Rhode Island Providence 630459

42 South Carolina Greenville 474903

43 South Dakota Minnehaha 178942

44 Tennessee Shelby 937750

45 Texas Harris 4356362

46 Utah Salt Lake 1078958

47 Vermont Chittenden 159711

48 Virginia Fairfax 1128722

49 Washington King 2045756

50 West Virginia Kanawha 190781

51 Wisconsin Milwaukee 955939

52 Wyoming Laramie 95431

>

>

> #(2)

> rank <- order(county\_pop[, "TotalPop"], decreasing=TRUE)

> county\_pop\_ordered <- county\_pop[rank,]

> quant <- quantile(county\_pop\_ordered$TotalPop, probs=c(.9))

> top\_10\_pct <- county\_pop\_ordered[county\_pop\_ordered$TotalPop >= quant,]

> median <- median(top\_10\_pct$TotalPop)

> mean <- mean(top\_10\_pct$TotalPop)

> names(median) <- c("Median")

> names(mean) <- c("Mean")

> mean

Mean

656545.9

> median

Median

427097

>

>

> # Q2

> x <- matrix(rnorm(100), nrow=10, ncol=10)

> min <- apply(x, MARGIN=1, FUN=min)

> names(min) <- c("Row 1 Min", "Row 2 Min", "Row 3 Min", "Row 4 Min", "Row 5 Min", "Row 6 Min", "Row 7 Min", "Row 8 Min", "Row 9 Min", "Row 10 Min")

> min

Row 1 Min Row 2 Min Row 3 Min Row 4 Min Row 5 Min Row 6 Min Row 7 Min

-1.9533663 -1.3197614 -1.8250610 -1.1691599 -0.4024638 -0.2931149 -1.5775512

Row 8 Min Row 9 Min Row 10 Min

-1.5607806 -1.9411920 -2.3207887

>

>

>

> # Q3

> fun <- function(x, y) {

+ x <- county$Men/county$Women

+ x2 <- x \*\* 2

+ y <- county$Unemployment

+ reg <- lm(y ~ x + x2)

+ return(reg$coefficients)

+ }

> fun(x, y)

(Intercept) x x2

22.106582 -21.149591 7.081153