Assignment\_2

Mario Colon

February 24, 2018

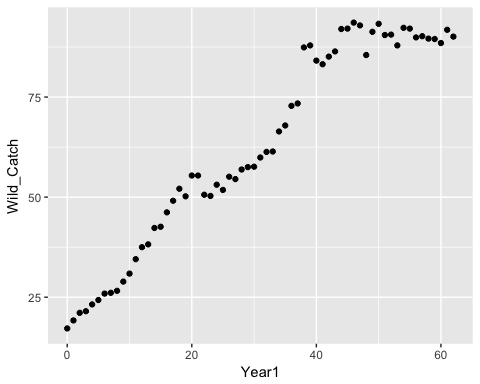
# Load Packages

library(tidyverse)  
library(Ecdat)  
library(plm)  
library(lmtest)  
library(quantreg)  
library(car)  
library(boot)  
library(nlstools)

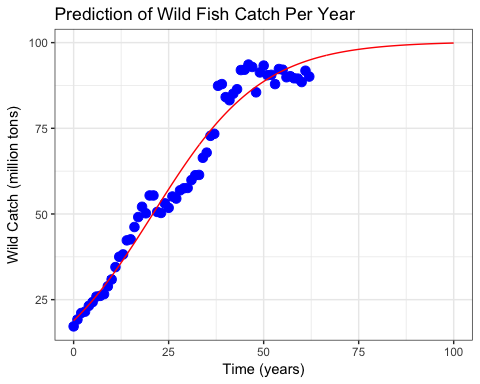
# Task 3

Fish stuff

fish <- read\_csv("fish\_catch.csv")  
  
fish\_fun <- fish %>%   
 select(Wild\_Catch, Farmed\_Fish, Total\_Fish\_Produced, Year1)  
  
fish\_fun  
  
ggplot(fish\_fun, aes(x = Year1, y = Wild\_Catch)) +  
 geom\_point()



r\_est <- lm(log(fish\_fun$Wild\_Catch[1:7]) ~ fish\_fun$Year1[1:7])  
r\_est  
  
#Ct = A / (1 + Be^-rt)  
  
#Estimate for K ~ 88  
#Estimate for C0 ~ 17.2  
#Estimate for B (K-C0)/C0 = 4.2  
  
#Estimate for A ~ 88  
#Estimate for B ~ 4.2  
#Estimate for r ~ .06407  
  
fish\_fit <- nls(Wild\_Catch ~ A / (1 + B\*exp(-r\*Year1)),  
 data = fish\_fun,  
 start = list(A = 88, B = 4.2, r = .06407),  
 trace = TRUE  
 )  
  
  
A <- coef(fish\_fit)[1]  
B <- coef(fish\_fit)[2]  
r <- coef(fish\_fit)[3]  
  
time\_seq <- seq(0, 100, length = 300)  
time\_seq  
  
  
fish\_pred <- A/(1+B\*exp(-r\*time\_seq))  
fish\_pred  
  
pred\_df <- data.frame(time\_seq, fish\_pred)  
pred\_df  
  
ggplot(fish\_fun, aes(x=Year1, y=Wild\_Catch)) +  
 geom\_point(color = "blue", size = 3) +  
 theme\_bw() +  
 geom\_line(data = pred\_df, aes(x = time\_seq, y = fish\_pred), color = "red") +  
 xlab("Time (years)") +  
 ylab("Wild Catch (million tons)") +  
 ggtitle("Prediction of Wild Fish Catch Per Year")



**Figure 1. Prediction of Wild Fish Catch** Wild fish caught per year has increased logistically over time (Time adjusted; 1950 = 0). Despite more fish caught, the rate of capture has decreased over time, with a carrying capacity of 100 million tons of wild fish caught per year possible in the near future.

*Final Equation: Wild Fish Catch = 88 / (1 + 4.2e^-.06407t)*

# Task 4

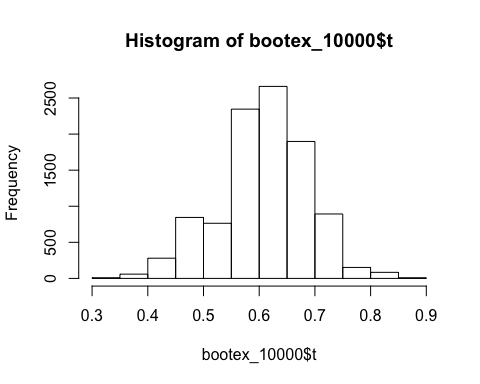
In the study, 22 of 36 surveyed UCSB community members (61%) identifying as genderqueer responded that they had personally experienced “exclusionary, offensive, hostile or intimidating conduct” (compared to 26% and 19% for those identifying as women and men, respectively).

Create a vector reflecting the collected survey data, then find the bootstrapped 95% confidence interval for the proportion of genderqueer individuals experiencing exclusionary conduct based on the 2014 UCSB survey data.

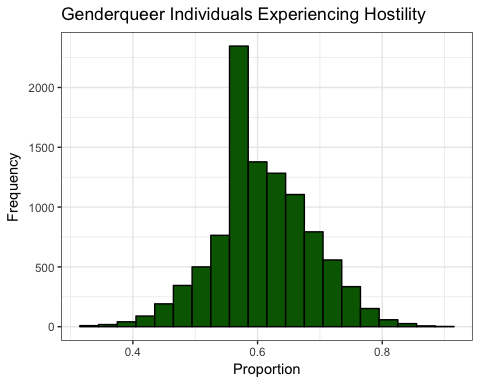
#n=35  
#nq = 22 = 61%  
  
n=36  
exclusion <- sample(1, n, TRUE)  
exclusion[sample(1:36, .39 \* n, FALSE)] <- 0  
table(exclusion)

## exclusion  
## 0 1   
## 14 22

mean\_exclusion <- function(x,i) {mean (x[i])}  
  
bootex\_10000 <- boot(exclusion, mean\_exclusion, R = 10000)  
hist(bootex\_10000$t)



bootex\_graph <- data.frame(bootex\_10000$t)  
  
ggplot(bootex\_graph, aes(bootex\_10000$t)) +  
 geom\_histogram(binwidth = .03, fill = "darkgreen", colour = "black") +  
 xlab("Proportion") +  
 ylab ("Frequency") +  
 ggtitle ("Genderqueer Individuals Experiencing Hostility") +  
 theme\_bw()



ex\_ci <- boot.ci(bootex\_10000, conf = .95, type = "perc")  
ex\_ci

## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS  
## Based on 10000 bootstrap replicates  
##   
## CALL :   
## boot.ci(boot.out = bootex\_10000, conf = 0.95, type = "perc")  
##   
## Intervals :   
## Level Percentile   
## 95% ( 0.4444, 0.7500 )   
## Calculations and Intervals on Original Scale

**Figure 2. Genderqueer individuals experiencing hostility** Results of bootstraped data (CI = 95%) from survey exploring the proportions of individuals who identify as genderqueer (61%) who experienced “exclusionary, offensive, hostile or intimidating conduct” (n = 36) at UCSB. Source: 2014 UCSB Campus Climate Project Final Report (prepared by Rankin & Associates Consulting

# Task 5

### [d]

library(tidyverse)

### [e]

ca\_pest <- read\_csv("ca\_pest.csv")  
ca\_pop <- read\_csv("ca\_pop.csv")

### [f - i]

ca\_merge <- full\_join(ca\_pest, ca\_pop) %>%   
 mutate(PesticidePounds\_SqMile = PestPounds/CountyArea) %>%   
 mutate(Population1 = Population/1000) %>%   
 mutate(MedFamilyIncome1 = MedFamilyIncome/1000) %>%   
 arrange(desc(PesticidePounds\_SqMile)) %>%   
 select(County, PesticidePounds\_SqMile, Population1, MedFamilyIncome1)

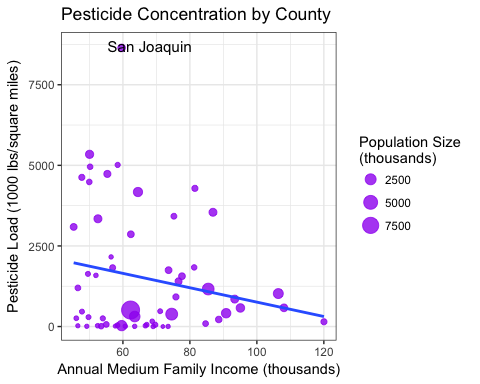
## Joining, by = "County"

ca\_merge

## # A tibble: 58 x 4  
## County PesticidePounds\_SqMile Population1 MedFamilyIncome1  
## <chr> <dbl> <dbl> <dbl>  
## 1 San Joaquin 8643 701 59.6  
## 2 Fresno 5342 949 50.0  
## 3 Sutter 5013 95.1 58.4  
## 4 Kings 4956 151 50.2  
## 5 Stanislaus 4734 523 55.4  
## 6 Merced 4627 262 47.7  
## 7 Madera 4485 152 50.0  
## 8 Santa Cruz 4286 267 81.5  
## 9 Sacramento 4173 1450 64.5  
## 10 Ventura 3544 836 86.9  
## # ... with 48 more rows

## [j]

ggplot(ca\_merge, aes(x = MedFamilyIncome1, y = PesticidePounds\_SqMile, size = Population1)) +  
 geom\_point(colour = "purple", alpha = .8) +  
 labs(size = "Population Size \n(thousands)") +  
 ylab("Pesticide Load (1000 lbs/square miles)") +  
 xlab("Annual Medium Family Income (thousands)") +  
 ggtitle("Pesticide Concentration by County") +  
 annotate("text", x = 68, y = 8700, label = "San Joaquin", size = 4) +  
 geom\_smooth(method = lm, se = FALSE, show.legend = FALSE) +  
 theme\_bw()



**Figure 3. Pesticide Concentration by County** Pesticide concentrations per county decreases as medium famiy income increase. Outlier of San Joaquin County exists, witha pesticide use of 8,645,000 lbs/square mile a medium family income of $59,600 per year Source: Pesticide data: 2014 CA pesticide application (by county) from the California Department of Pesticide Regulation & California population & income data: U.S. Census Bureau. American Community Survey, 2010- 2014 American Community Survey 5-Year Estimates