Assignment 4

Corinna Hong

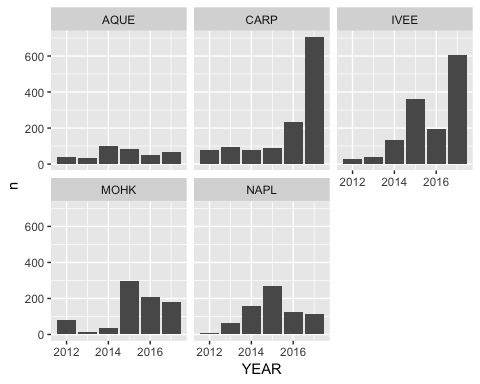
November 14, 2018

1. Lobster abundance and fishing pressure (2012 - 2017)

# Create five graphs of each location with years on the x-axis and count of lobsters on y-axis  
  
abundance\_summary <- lobster\_abundance %>%   
 group\_by(SITE) %>%   
 count(YEAR) # Count how many observations there are in each year (data is grouped by site)  
abundance\_summary

## # A tibble: 30 x 3  
## # Groups: SITE [5]  
## SITE YEAR n  
## <fct> <int> <int>  
## 1 AQUE 2012 38  
## 2 AQUE 2013 32  
## 3 AQUE 2014 100  
## 4 AQUE 2015 83  
## 5 AQUE 2016 48  
## 6 AQUE 2017 67  
## 7 CARP 2012 78  
## 8 CARP 2013 93  
## 9 CARP 2014 78  
## 10 CARP 2015 90  
## # ... with 20 more rows

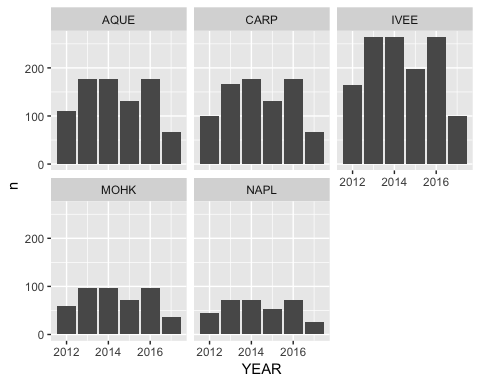
abundance\_col <- ggplot(abundance\_summary, aes(x = YEAR, y = n)) +  
 geom\_col() +  
 facet\_wrap(~ SITE)  
abundance\_col



# Create graphs of trap buoys versus year at each site  
  
trap\_summary <- lobster\_traps %>%   
 group\_by(SITE) %>%   
 count(YEAR) # Count how many observations there are in each year (data is grouped by site)  
trap\_summary

## # A tibble: 30 x 3  
## # Groups: SITE [5]  
## SITE YEAR n  
## <chr> <int> <int>  
## 1 AQUE 2012 110  
## 2 AQUE 2013 176  
## 3 AQUE 2014 176  
## 4 AQUE 2015 132  
## 5 AQUE 2016 176  
## 6 AQUE 2017 66  
## 7 CARP 2012 100  
## 8 CARP 2013 166  
## 9 CARP 2014 176  
## 10 CARP 2015 132  
## # ... with 20 more rows

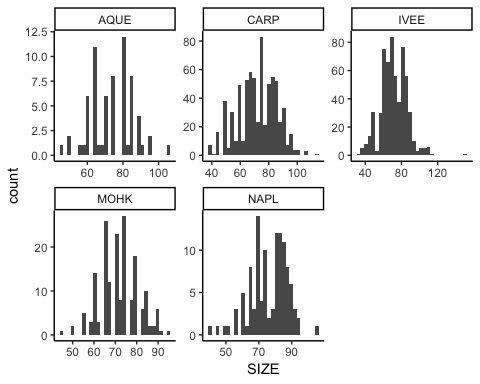
trap\_col <- ggplot(trap\_summary, aes(x = YEAR, y = n)) +  
 geom\_col() +  
 facet\_wrap(~ SITE)  
trap\_col



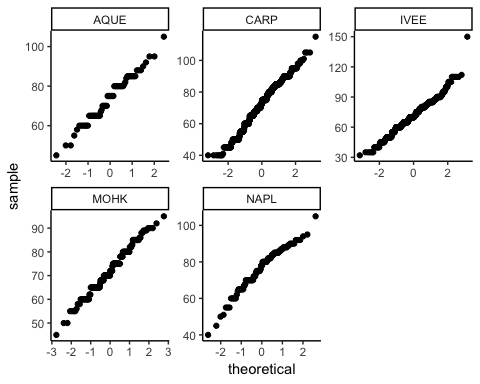
1. Compare mean lobster size by site in 2017

# New dataframe for only year 2017  
  
size\_2017 <- lobster\_abundance %>%   
 filter(YEAR == 2017) %>%   
 select(SITE, SIZE)   
  
  
  
# Data exploration  
  
#Histograms  
  
size\_2017\_hist <- ggplot(size\_2017, aes(x = SIZE)) +  
 geom\_histogram() +  
 facet\_wrap(~ SITE, scale = "free") + # Create a histogram, split graphic visualization by site. Give each histogram its own y-axis scale  
 theme\_classic()  
size\_2017\_hist

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



# QQ-Plots  
  
size\_2017\_qq <- ggplot(size\_2017, aes(sample = SIZE)) +  
 geom\_qq() +  
 facet\_wrap(~ SITE, scale = "free") + # Create a Q-Q plot, split graphic visualization by site. Give each Q-Q Plot its own y-axis scale  
 theme\_classic()  
size\_2017\_qq



# Question: Is there a significant difference in mean lobster size between the five sites?  
  
# H0: There is no significant difference in mean lobster size between the five sites  
# HA: There is a significant difference in mean lobster size between the five sites  
  
  
  
# Levene's Test   
  
# H0: There are no differences in varance across groups (variances are equal)  
# HA: There are differences in variances across groups (variances are not equal)  
  
lobster\_levene <- leveneTest(SIZE ~ SITE, data = size\_2017)  
lobster\_levene

## Levene's Test for Homogeneity of Variance (center = median)  
## Df F value Pr(>F)   
## group 4 8.3893 1.065e-06 \*\*\*  
## 1663   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

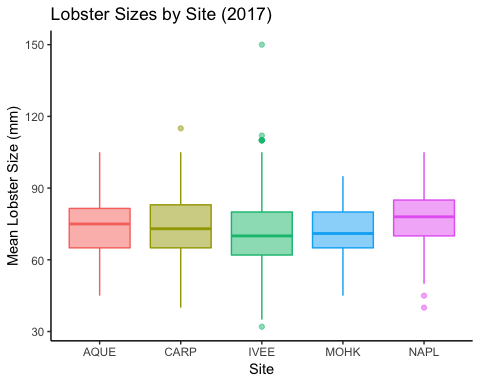
# p < 0.05. Reject the null hypothesis. Our variances are differenct (Variances are not equal)  
  
  
  
# What are the actual variances?  
  
variances <- size\_2017 %>%   
 group\_by(SITE) %>%   
 summarize(  
 mean = mean(SIZE),  
 sd = sd(SIZE),  
 variance = var(SIZE)  
 )  
  
# Our largest variance is less than 4x larger than our smallest variance  
  
  
  
# ONE-WAY ANOVA:  
  
lobster\_aov <- aov(SIZE ~ SITE, data = size\_2017)  
lobster\_sum\_aov <- summary(lobster\_aov)  
lobster\_sum\_aov

## Df Sum Sq Mean Sq F value Pr(>F)   
## SITE 4 2355 588.6 3.424 0.0085 \*\*  
## Residuals 1663 285871 171.9   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

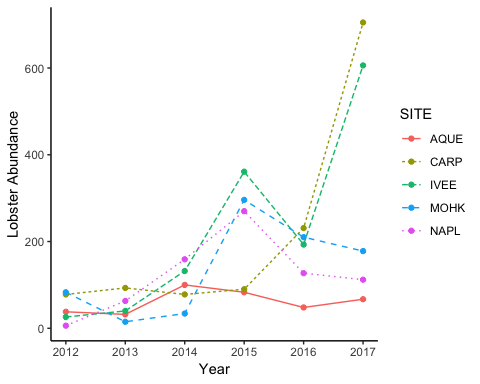
lobster\_ph <- TukeyHSD(lobster\_aov)  
lobster\_ph

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = SIZE ~ SITE, data = size\_2017)  
##   
## $SITE  
## diff lwr upr p adj  
## CARP-AQUE -1.6657352 -6.24294710 2.911477 0.8582355  
## IVEE-AQUE -2.4433772 -7.05292315 2.166169 0.5968998  
## MOHK-AQUE -1.8955224 -7.02720717 3.236162 0.8514711  
## NAPL-AQUE 2.3366205 -3.19311600 7.866357 0.7775633  
## IVEE-CARP -0.7776420 -2.76097123 1.205687 0.8216104  
## MOHK-CARP -0.2297872 -3.23309697 2.773523 0.9995765  
## NAPL-CARP 4.0023556 0.36042398 7.644287 0.0228728  
## MOHK-IVEE 0.5478548 -2.50450730 3.600217 0.9882889  
## NAPL-IVEE 4.7799976 1.09751057 8.462485 0.0037001  
## NAPL-MOHK 4.2321429 -0.08607271 8.550358 0.0579286

lobster\_box <- ggplot(size\_2017, aes(x = SITE, y = SIZE)) +  
 geom\_boxplot(aes(fill = SITE, colour = SITE), alpha = 0.5, show.legend = FALSE) +  
 ylab("Mean Lobster Size (mm)") +  
 ggtitle("Lobster Sizes by Site (2017)") +  
 xlab("Site") +  
 theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank(),  
 panel.background = element\_blank(), axis.line = element\_line(colour = "black"))  
lobster\_box



# Created separate datasets to attempt to merge them.   
  
lobster\_counts <- lobster\_abundance %>%  
 group\_by(SITE, YEAR) %>%   
 summarize(  
 count = length(SIZE)  
 )  
  
lobster\_line <- ggplot(lobster\_counts, aes(x = YEAR, y = count, group = SITE)) +   
 geom\_line(aes(linetype = SITE, color = SITE))+  
 theme\_classic()+  
 geom\_point(aes(color = SITE))+  
 labs(x = "Year", y = "Lobster Abundance")  
   
lobster\_line



traps\_counts <- lobster\_traps %>%   
 group\_by(SITE, YEAR) %>%   
 summarize(  
 sum = sum(TRAPS))  
  
traps\_counts

## # A tibble: 30 x 3  
## # Groups: SITE [?]  
## SITE YEAR sum  
## <chr> <int> <int>  
## 1 AQUE 2012 509  
## 2 AQUE 2013 813  
## 3 AQUE 2014 685  
## 4 AQUE 2015 676  
## 5 AQUE 2016 816  
## 6 AQUE 2017 179  
## 7 CARP 2012 788  
## 8 CARP 2013 1039  
## 9 CARP 2014 1164  
## 10 CARP 2015 568  
## # ... with 20 more rows

traps\_line <- ggplot(traps\_counts, aes(x = YEAR, y = sum, group = SITE)) +   
 geom\_line(aes(linetype = SITE, color = SITE))+  
 theme\_classic()+  
 geom\_point(aes(color = SITE))+  
 labs(x = "Year", y = "Traps Adundance")  
  
traps\_line

