```
#BUSPH GH 811

#Date: November 13, 2018


#****************************************#

#****Data Viz & Linear Regression Demo****#

#****************************************#


# This demo covers

# - Data Viz with ggplot2

# - Linear modeling

# - Linear model diagnostics

# - Intercept & slope transformations


library("ggplot2")

setwd("C:/Users/acstokes/Desktop/GH 811/Spring 2018/Data")

w <- read.csv(file="WDS2012.csv", head=TRUE, sep=",")


# Simple scatter plot

p <- ggplot(data=w, aes(x=le, y=tfr))

p + geom_point()

p + geom_point(size=3, color = "steelblue", alpha=1/2)

p + geom_point(size=3, color = "steelblue", alpha=1/2) +
stat_smooth(method="loess", se=TRUE)

p + geom_point(size=3, color = "steelblue", alpha=1/2) +
stat_smooth(method="lm", se=TRUE)


# Faceting

p <- ggplot(data=w, aes(x=le, y=tfr))
```

```
p + geom_point(size=2, shape=1)

p + geom_point(size=2, shape=1) + facet_grid(.~area)

p + geom_point(size=2, shape=1) + facet_grid(area~.)


# Add a 3rd dimension to the original plot!

p <- ggplot(data=w, aes(x=le, y=tfr, color=area))

p + geom_point(size=3)

p + geom_point(size=3) + stat_smooth(method="lm", se=FALSE)


# Add labels to above plot

q <- p + geom_point(size=3) + stat_smooth(method="lm", se=FALSE)

q <- q + labs(title="Fertility vs. Life Expectancy by World
Region")

q <- q + labs(x="Life Expectancy (years)", y="Total Fertility
Rate")

q


#Basic linear model

fit <- lm(tfr ~ le, data = w)


#Intercept transformation

fit2 <- lm(tfr ~ I(le - mean(le)), data = w)


#Slope transformation

fit3 <- lm(tfr ~ I(le * (1/10)), data = w)


#Model predictions

newx <- c(50, 60, 70, 80, 83)
```

```
coef(fit)[1] + coef(fit)[2] * newx

predict(fit, newdata = data.frame(le = newx))
```