Appendix I: Figure Code

```
train <- data.frame(read.csv("train.csv"))
test <- data.frame(read.csv("test.csv"))
library(ggplot2)
library(dplyr)
library(glmnet)
library(gridExtra)
library(car)

#-----------------------------------------
#       Figure 1
#-----------------------------------------
train$month <- factor(train$month
                      , levels = c("jan", "feb", 'mar', 'apr', 'may', 'jun'
                                   , 'jul', 'aug','sep', 'oct', 'nov', 'dec'))
train$day <- factor(train$day
                    , levels = c('sun','mon','tue','wed','thu','fri','sat'))
df1 <- data.frame(train %>% group_by(month) %>% summarize(lmonth = n()))
df2 <- data.frame(train %>% group_by(day) %>% summarize(lday = n()))
train$sqISI <- sqrt(train$ISI)

p1 <- ggplot(data = df1, aes(x = month, y = lmonth)) +
  geom_bar(stat = "identity", fill = "orangered2") +
  theme_minimal() +
  labs(y = "Number of Fires", x = "Month", main = "")

pdf("month_bar.pdf")
p1
dev.off()

p2 <- ggplot(data = df2, aes(x = day, y = lday)) +
  geom_bar(stat = "identity", fill = "orangered2") +
  theme_minimal() +
  labs(y = "Number of Fires", x = "Day", main = "")

pdf("day_bar.pdf")
p2
dev.off()

#-----------------------------------------
#       Figure 2
#-----------------------------------------

p3 <- ggplot(data = train, aes(x = "", y = ISI)) +
  geom_boxplot(fill = "blue", alpha = 0.5) +
  theme_minimal() +
  labs(y = "ISI", x = "", main = "") +
  coord_flip()

pdf("ISI_box.pdf")
p3
```

```r
dev.off()

#-----------------------------------------
#       Figure 3
#-----------------------------------------

p4 <- ggplot(data = train, aes(x = FFMC, y = sqISI)) +
  geom_point(color = "blue", alpha = 0.7) +
  theme_minimal() +
  labs(y = "Square Root ISI", x = "FFMC", main = "")

pdf("FFMC_ISI_scatter.pdf")
p4
dev.off()

#-----------------------------------------
#       Figure 5
#-----------------------------------------

f1 <- formula(sqISI ~ temp + wind + summer + rainvnorain)
f2 <- formula(sqISI ~ temp + wind + summer + rainvnorain + FFMC)

X1 <- model.matrix(f1,train)
X2 <- model.matrix(f2,train)
Y <- as.matrix(train$sqISI)

fit1 <- glmnet(X1, Y, alpha = 1)
fit2 <- glmnet(X2, Y, alpha = 1)

pdf("penalization_plots.pdf")
par(mfrow = c(1, 2), mai = c(1, 0.5, 0.1, 0.1))
plot_glmnet(fit2, ylim = c(0, .6))
plot_glmnet(fit1, ylim = c(0, .6))
dev.off()

#-----------------------------------------
#       Figure 6
#-----------------------------------------

train <- data.frame(train %>% group_by(tFFMC) %>% mutate(weight = n()))
m1 <- lm(sqISI ~temp + wind + summer + rainvnorain
         , data = train)
m2 <- lm(sqISI ~temp + wind + summer + rainvnorain
         , weights = weight
         , data = train)
m3 <- lm(sqISI ~ temp + wind + summer + rainvnorain + tFFMC
         , data = train)
df <- data.frame(r1 = as.vector(qqnorm(resid(m1), plot = F))
                 , r2 = as.vector(qqnorm(resid(m2), plot = F))
                 , r3 = as.vector(qqnorm(resid(m3), plot = F)))

p5 <- ggplot(m1, aes(qqnorm(.stdresid)[[1]], .stdresid)) +
  geom_point(na.rm = TRUE, col = "steelblue", alpha = 0.7) +
```

```
  geom_abline() +
  xlab("Theoretical Quantiles") +
  ylab("Standardized Residuals") +
  ggtitle("No FFMC") +
  theme_bw() +
  coord_cartesian(ylim = c(-3, 3))

p6 <- ggplot(m2, aes(qqnorm(.stdresid)[[1]], .stdresid)) +
  geom_point(na.rm = TRUE, col = "steelblue", alpha = 0.7) +
  geom_abline()+xlab("Theoretical Quantiles") +
  ylab("Standardized Residuals") +
  ggtitle("Weighted FFMC") +
  theme_bw() +
  coord_cartesian(ylim = c(-3, 3))

p7 <- ggplot(m2, aes(qqnorm(.stdresid)[[1]], .stdresid)) +
  geom_point(na.rm = TRUE, col = "steelblue", alpha = 0.7) +
  geom_abline()+xlab("Theoretical Quantiles") +
  ylab("Standardized Residuals") +
  ggtitle("Covariate FFMC") +
  theme_bw() +
  coord_cartesian(ylim = c(-3, 3))

pdf("FFMC-QQ.pdf")
grid.arrange(p5, p6, p7, nrow = 1, ncol = 3)
dev.off()

#-----------------------------------------
#       Figure 7
#-----------------------------------------

train <- data.frame(train %>% group_by(tFFMC) %>% mutate(weight = n()))
m <- lm(lgISI ~ summer + wind + temp + rainvnorain
        , data = train)

pdf("final_model_diag.pdf")
par(mfrow = c(2, 2))
plot(m)
dev.off()

pdf("final_avp.pdf")
avPlots(m)
dev.off()

#-----------------------------------------
#       Figure 8
#-----------------------------------------

m1 <- lm(sqISI ~ summer + wind + temp + rainvnorain
         , data = train)

B <- 1000
ResidualBootstrapM1 <- t(replicate(B, {
```

```
  yb <- fitted(m1) + resid(m1)[sample.int(nrow(train), replace = TRUE)]
  boot <- model.matrix(m1)
  coef(lm(yb ~ boot - 1))
}))

hist(ResidualBootstrapM1[,1], xlab = "summer")
hist(ResidualBootstrapM1[,2], xlab ="wind")
hist(ResidualBootstrapM1[,3], xlab ="temp")
hist(ResidualBootstrapM1[,4], xlab ="rainvnorain")
df <- data.frame(ResidualBootstrapM1)

p8 <- ggplot(df, aes(x = bootsummer1)) +
  geom_histogram(fill = "orangered2", color = "steelblue") +
  theme_bw() +
  labs(x = "Summer", y = "", main = "")

p9 <- ggplot(df, aes(x = bootwind)) +
  geom_histogram(fill = "orangered2", color = "steelblue") +
  theme_bw() +
  labs(x = "Wind", y = "")

p10 = ggplot(df, aes(x = boottemp)) +
  geom_histogram(fill = "orangered2", color = "steelblue") +
  theme_bw() +
  labs(x = "Temperature", y = "")

p11 = ggplot(df, aes(x = bootrainvnorain1)) +
  geom_histogram(fill = "orangered2", color = "steelblue") +
  theme_bw() +
  labs(x = "Rain", y = "")

pdf("boot.pdf")
grid.arrange(p8, p9, p10, p11, nrow = 2, ncol = 2)
dev.off()

#-----------------------------------------
#       Figure 9
#-----------------------------------------

test$sqISI <- sqrt(test$ISI)

test$summer <- as.factor(test$summer)
test$rainvnorain <- as.factor(test$rainvnorain)

test_mod <- lm(sqISI ~ summer + wind + temp + rainvnorain ,data = test)

plotdf <- data.frame(cbind(test$sqISI^2, test_mod$fitted.values^2))
names(plotdf) <- c("ActualValues", "FittedValues")

p12 <- ggplot(plotdf, aes(x= FittedValues, y = ActualValues)) +
  geom_point(col = "steelblue", alpha = .7) +
  geom_segment(aes(x = 0, y = 0, xend = 15, yend = 15), col = "black") +
  theme_bw()
```

```r
pdf("pred_plot.pdf")
p12
dev.off()

pdf("TestSetDiagnostics.pdf")
par(mfrow = c(2, 2))
plot(test_mod)
dev.off()
```