

Analysis

2024-11-30

```
data <- read.csv("pandemic_data.csv")
head(data)
```

```
##      peak.inf soc.iso rate.vac quar.dur num.daily
## 1 0.8071379      0.0      0.00         0         15
## 2 0.6471836      0.5      0.00         0         15
## 3 0.2764590      1.0      0.00         0         15
## 4 0.8095392      0.0      0.01         0         15
## 5 0.7619943      0.5      0.01         0         15
## 6 0.1626645      1.0      0.01         0         15
```

```
summary(data) # peak.inf is i+q
```

```
##      peak.inf      soc.iso      rate.vac      quar.dur      num.daily
## Min.   :0.01002  Min.   :0.0  Min.   :0.00  Min.   : 0  Min.   :15
## 1st Qu.:0.27646  1st Qu.:0.0  1st Qu.:0.00  1st Qu.: 0  1st Qu.:15
## Median :0.72210  Median :0.5  Median :0.01  Median : 7  Median :30
## Mean   :0.58065  Mean   :0.5  Mean   :0.01  Mean   : 7  Mean   :30
## 3rd Qu.:0.84296  3rd Qu.:1.0  3rd Qu.:0.02  3rd Qu.:14  3rd Qu.:45
## Max.   :0.93448  Max.   :1.0  Max.   :0.02  Max.   :14  Max.   :45
```

```
# Anova
```

```
anova_model <- aov(peak.inf ~ soc.iso * rate.vac * quar.dur * num.daily, data = data)
summary(anova_model)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## soc.iso         1  4.897   4.897 189.338 < 2e-16 ***
## rate.vac        1  0.010   0.010   0.374  0.5430
## quar.dur        1  0.447   0.447 17.297 9.56e-05 ***
## num.daily       1  0.888   0.888 34.330 1.69e-07 ***
## soc.iso:rate.vac 1  0.000   0.000   0.017  0.8975
## soc.iso:quar.dur 1  0.090   0.090   3.491  0.0662 .
## rate.vac:quar.dur 1  0.000   0.000   0.001  0.9818
## soc.iso:num.daily 1  0.029   0.029   1.127  0.2923
## rate.vac:num.daily 1  0.019   0.019   0.735  0.3944
## quar.dur:num.daily 1  0.001   0.001   0.027  0.8696
## soc.iso:rate.vac:quar.dur 1  0.009   0.009   0.333  0.5658
## soc.iso:rate.vac:num.daily 1  0.001   0.001   0.042  0.8387
## soc.iso:quar.dur:num.daily 1  0.080   0.080   3.107  0.0827 .
## rate.vac:quar.dur:num.daily 1  0.001   0.001   0.021  0.8860
## soc.iso:rate.vac:quar.dur:num.daily 1  0.014   0.014   0.540  0.4652
## Residuals      65  1.681   0.026
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Regression and Box-Cox
```

```
lm_model <- lm(peak.inf ~ soc.iso * rate.vac * quar.dur * num.daily, data = data)
```

```
summary(lm_model)
```

```
##
## Call:
## lm(formula = peak.inf ~ soc.iso * rate.vac * quar.dur * num.daily,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.39881 -0.08998 -0.03392  0.09016  0.32617
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      7.914e-01  1.869e-01   4.235 7.35e-05 ***
## soc.iso         -6.715e-01  2.895e-01  -2.319  0.0235 *
## rate.vac         3.868e+00  1.448e+01   0.267  0.7901
## quar.dur        -8.537e-03  2.068e-02  -0.413  0.6811
## num.daily         4.287e-03  5.767e-03   0.743  0.4599
## soc.iso:rate.vac -1.759e+01  2.243e+01  -0.784  0.4356
## soc.iso:quar.dur -3.533e-03  3.204e-02  -0.110  0.9125
## rate.vac:quar.dur -1.258e+00  1.602e+00  -0.786  0.4350
## soc.iso:num.daily  6.677e-03  8.934e-03   0.747  0.4575
## rate.vac:num.daily -1.199e-01  4.467e-01  -0.268  0.7892
## quar.dur:num.daily  1.828e-04  6.382e-04   0.286  0.7755
## soc.iso:rate.vac:quar.dur  2.229e+00  2.481e+00   0.898  0.3724
## soc.iso:rate.vac:num.daily  4.832e-01  6.921e-01   0.698  0.4875
## soc.iso:quar.dur:num.daily -5.396e-04  9.887e-04  -0.546  0.5870
## rate.vac:quar.dur:num.daily  3.263e-02  4.943e-02   0.660  0.5116
## soc.iso:rate.vac:quar.dur:num.daily -5.625e-02  7.658e-02  -0.735  0.4652
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1608 on 65 degrees of freedom
## Multiple R-squared:  0.7942, Adjusted R-squared:  0.7467
## F-statistic: 16.72 on 15 and 65 DF,  p-value: < 2.2e-16
```

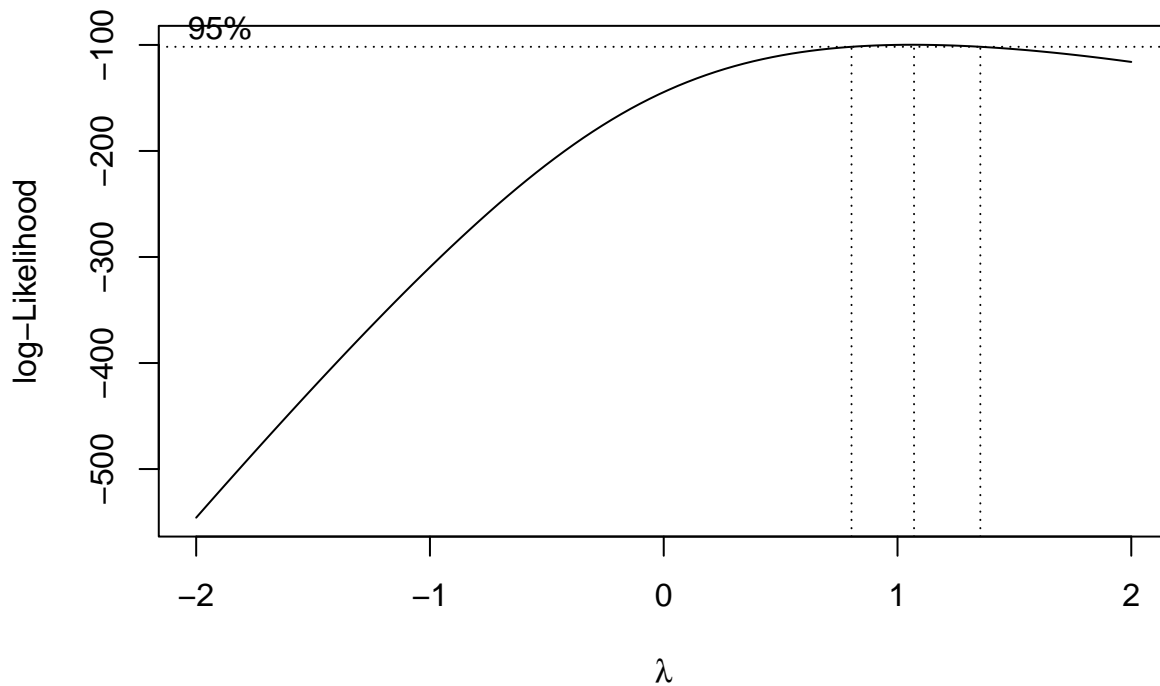
```
confint(lm_model, level = 0.95)
```

```
##              2.5 %      97.5 %
## (Intercept)    0.418174816  1.164610401
## soc.iso        -1.249655121 -0.093282085
## rate.vac       -25.041273167  32.777378637
## quar.dur       -0.049835522  0.032762552
## num.daily      -0.007230679  0.015804824
## soc.iso:rate.vac -62.378235189  27.194035027
## soc.iso:quar.dur -0.067513048  0.060447338
## rate.vac:quar.dur -4.457253629  1.940765672
## soc.iso:num.daily -0.011165802  0.024520645
## rate.vac:num.daily -1.012059176  0.772263157
## quar.dur:num.daily -0.001091752  0.001457280
## soc.iso:rate.vac:quar.dur -2.726966160  7.184802721
## soc.iso:rate.vac:num.daily -0.898927600  1.865332672
## soc.iso:quar.dur:num.daily -0.002514115  0.001434828
## rate.vac:quar.dur:num.daily -0.066097189  0.131349973
```

```
## soc.iso:rate.vac:quar.dur:num.daily -0.209194848 0.096688980
```

```
# Box-Cox Transformation
```

```
boxcox_result <- boxcox(lm_model, lambda = seq(-2, 2, by = 0.1))
```



```
optimal_lambda <- boxcox_result$x[which.max(boxcox_result$y)]
```

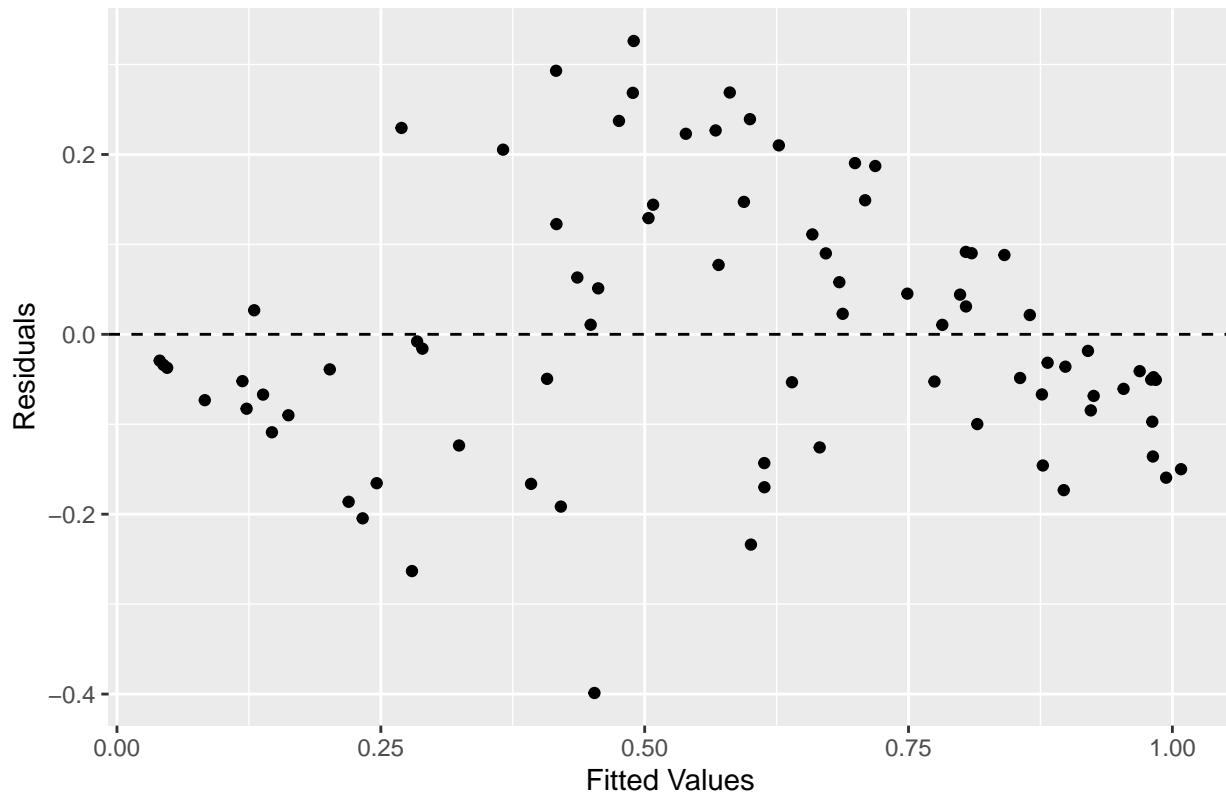
```
optimal_lambda # no transformation
```

```
## [1] 1.070707
```

```
# Residual plots
```

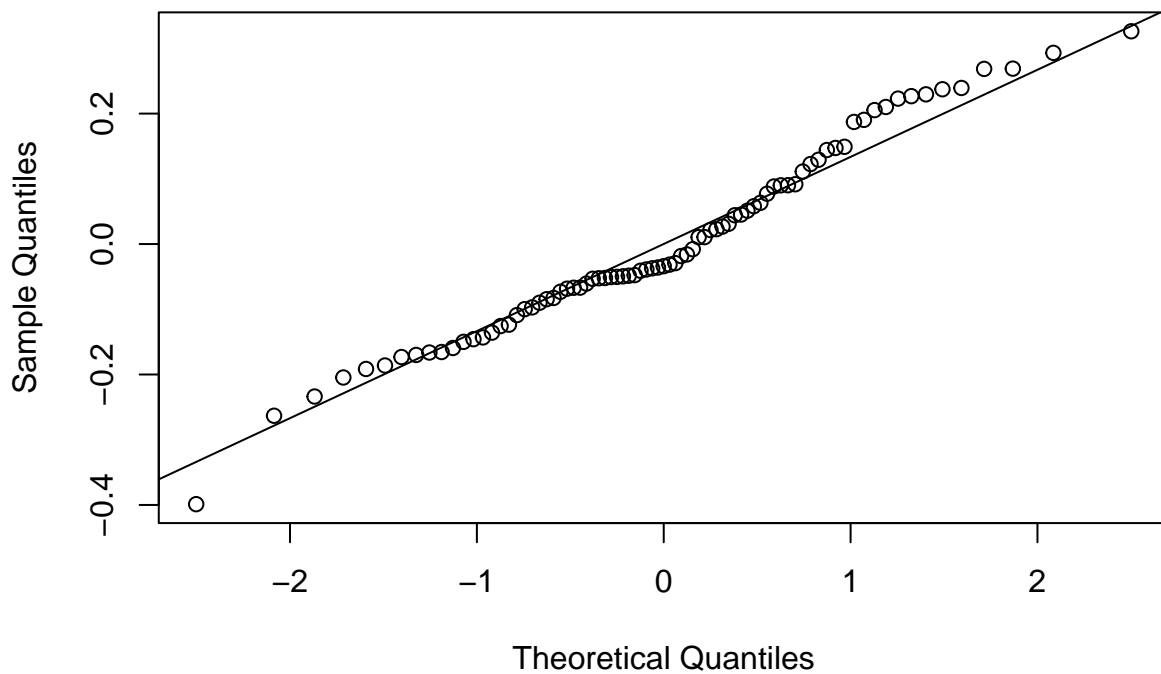
```
residuals_plot <- ggplot(data.frame(fitted = lm_model$fitted.values,  
                                   residuals = lm_model$residuals),  
                          aes(x = fitted, y = residuals)) +  
  geom_point() +  
  geom_hline(yintercept = 0, linetype = "dashed") +  
  labs(title = "Residuals vs Fitted Values", x = "Fitted Values", y = "Residuals")  
print(residuals_plot)
```

Residuals vs Fitted Values



```
# Q-Q Plot  
qqnorm(lm_model$residuals, main = "Normal Q-Q Plot")  
qqline(lm_model$residuals)
```

Normal Q-Q Plot

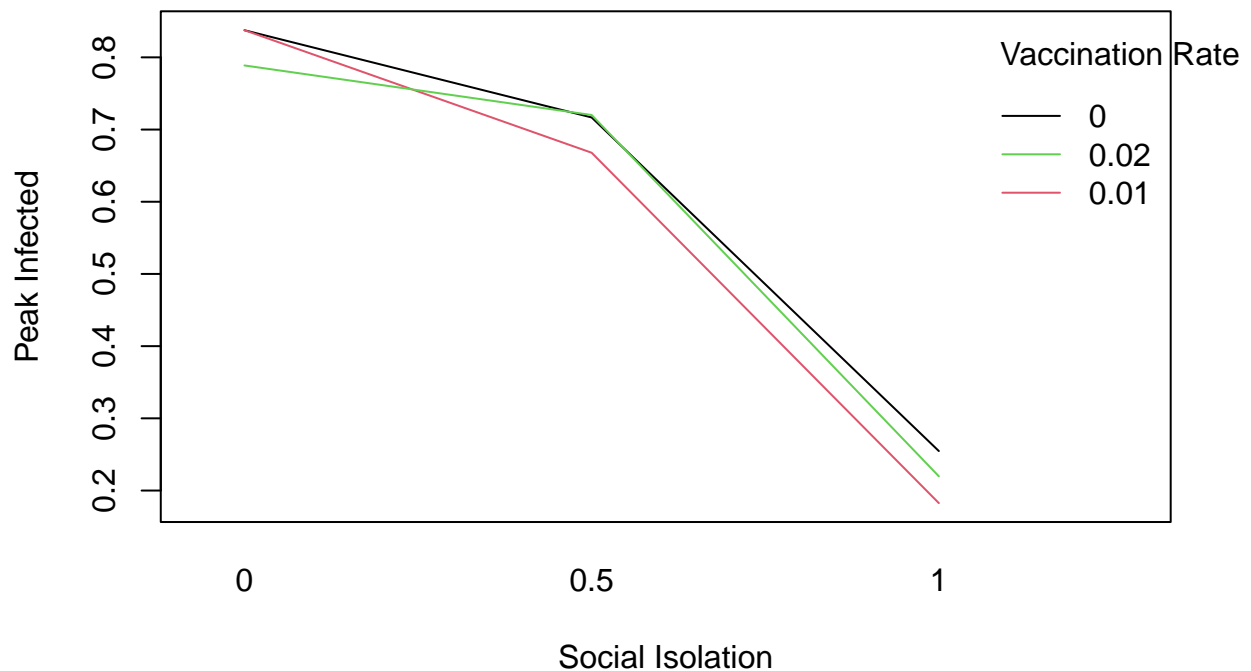


```

# Graphics
# Interaction plots
interaction.plot(
  x.factor = data$soc.iso,
  trace.factor = data$rate.vac,
  response = data$peak.inf,
  col = 1:4,
  lty = 1,
  main = "Interaction Plot: Social Isolation x Vaccination Rate",
  xlab = "Social Isolation",
  ylab = "Peak Infected",
  trace.label = "Vaccination Rate"
)

```

Interaction Plot: Social Isolation x Vaccination Rate

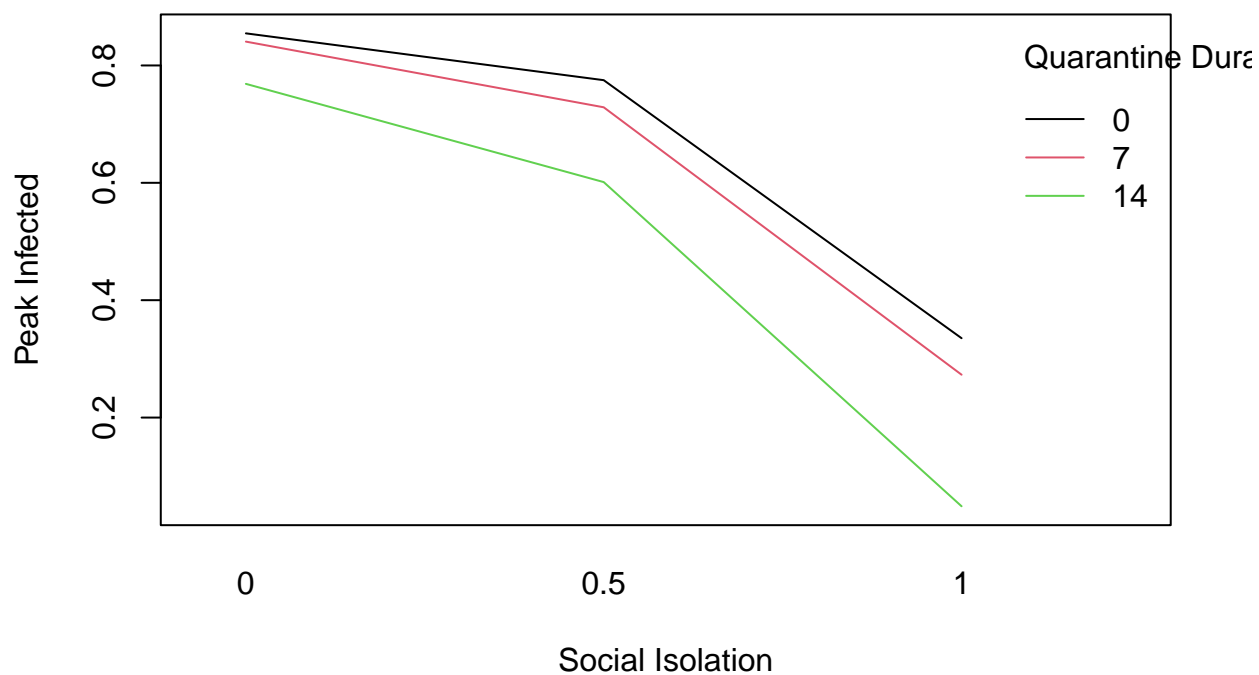


```

interaction.plot(
  x.factor = data$soc.iso,
  trace.factor = data$quar.dur,
  response = data$peak.inf,
  col = 1:4,
  lty = 1,
  main = "Interaction Plot: Social Isolation x Quarantine Duration",
  xlab = "Social Isolation",
  ylab = "Peak Infected",
  trace.label = "Quarantine Duration"
)

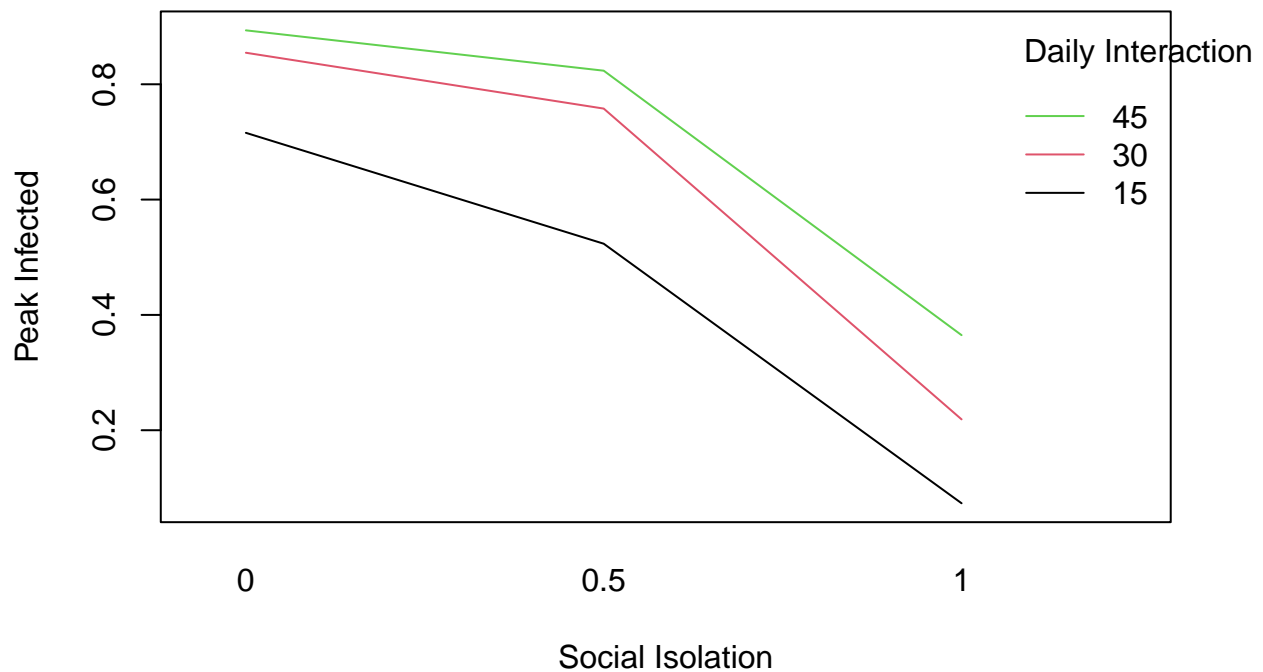
```

Interaction Plot: Social Isolation x Quarantine Duration



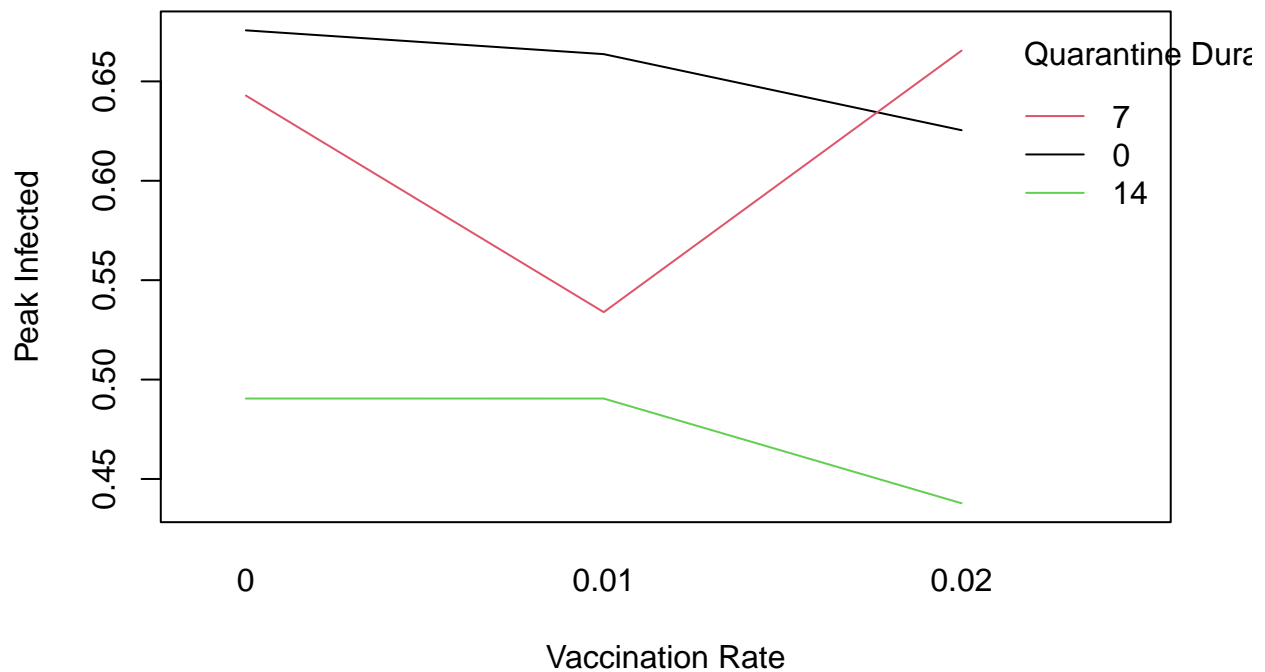
```
interaction.plot(  
  x.factor = data$soc.iso,  
  trace.factor = data$num.daily,  
  response = data$peak.inf,  
  col = 1:4,  
  lty = 1,  
  main = "Interaction Plot: Social Isolation x Daily Interactions",  
  xlab = "Social Isolation",  
  ylab = "Peak Infected",  
  trace.label = "Daily Interactions"  
)
```

Interaction Plot: Social Isolation x Daily Interactions



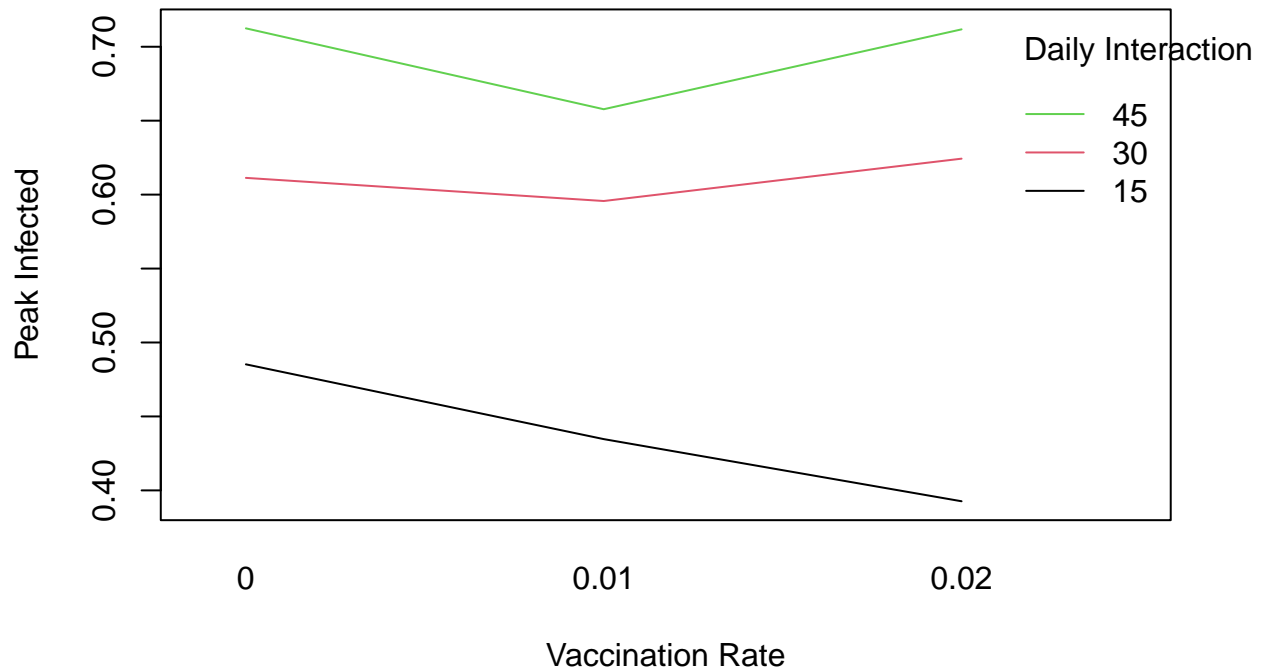
```
interaction.plot(  
  x.factor = data$rate.vac,  
  trace.factor = data$quar.dur,  
  response = data$peak.inf,  
  col = 1:4,  
  lty = 1,  
  main = "Interaction Plot: Vaccination Rate x Quarantine Duration",  
  xlab = "Vaccination Rate",  
  ylab = "Peak Infected",  
  trace.label = "Quarantine Duration"  
)
```

Interaction Plot: Vaccination Rate x Quarantine Duration



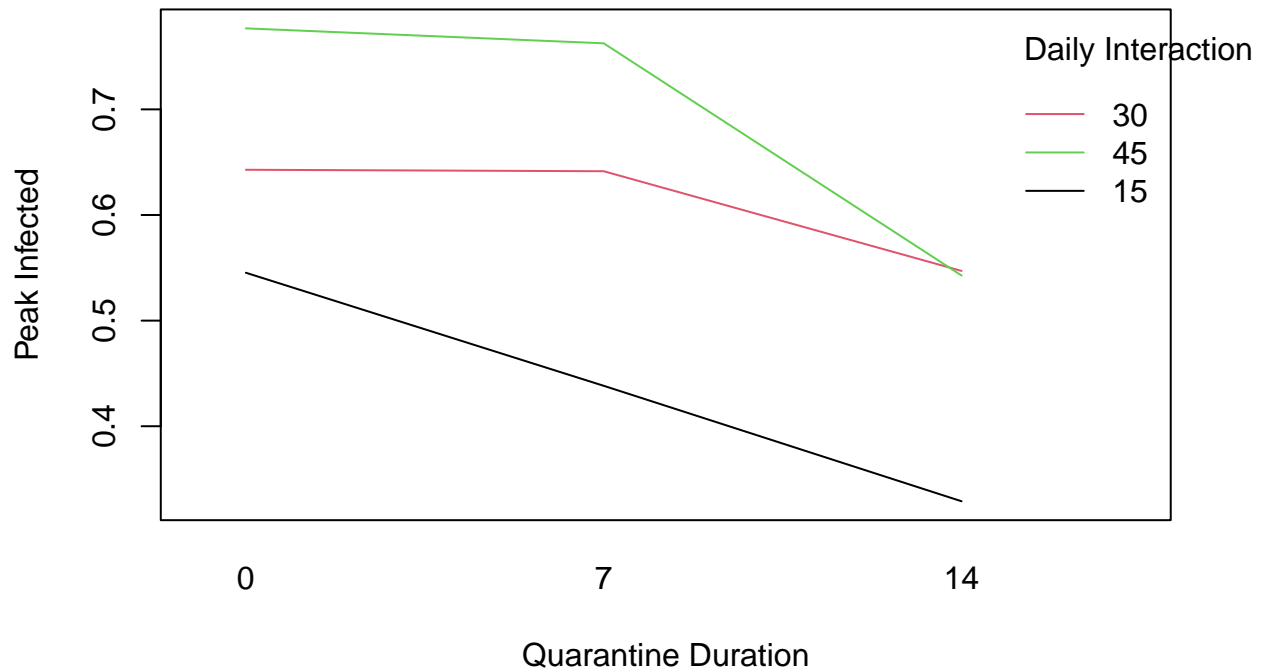
```
interaction.plot(  
  x.factor = data$rate.vac,  
  trace.factor = data$num.daily,  
  response = data$peak.inf,  
  col = 1:4,  
  lty = 1,  
  main = "Interaction Plot: Vaccination Rate x Daily Interactions",  
  xlab = "Vaccination Rate",  
  ylab = "Peak Infected",  
  trace.label = "Daily Interactions"  
)
```


Interaction Plot: Vaccination Rate x Daily Interactions

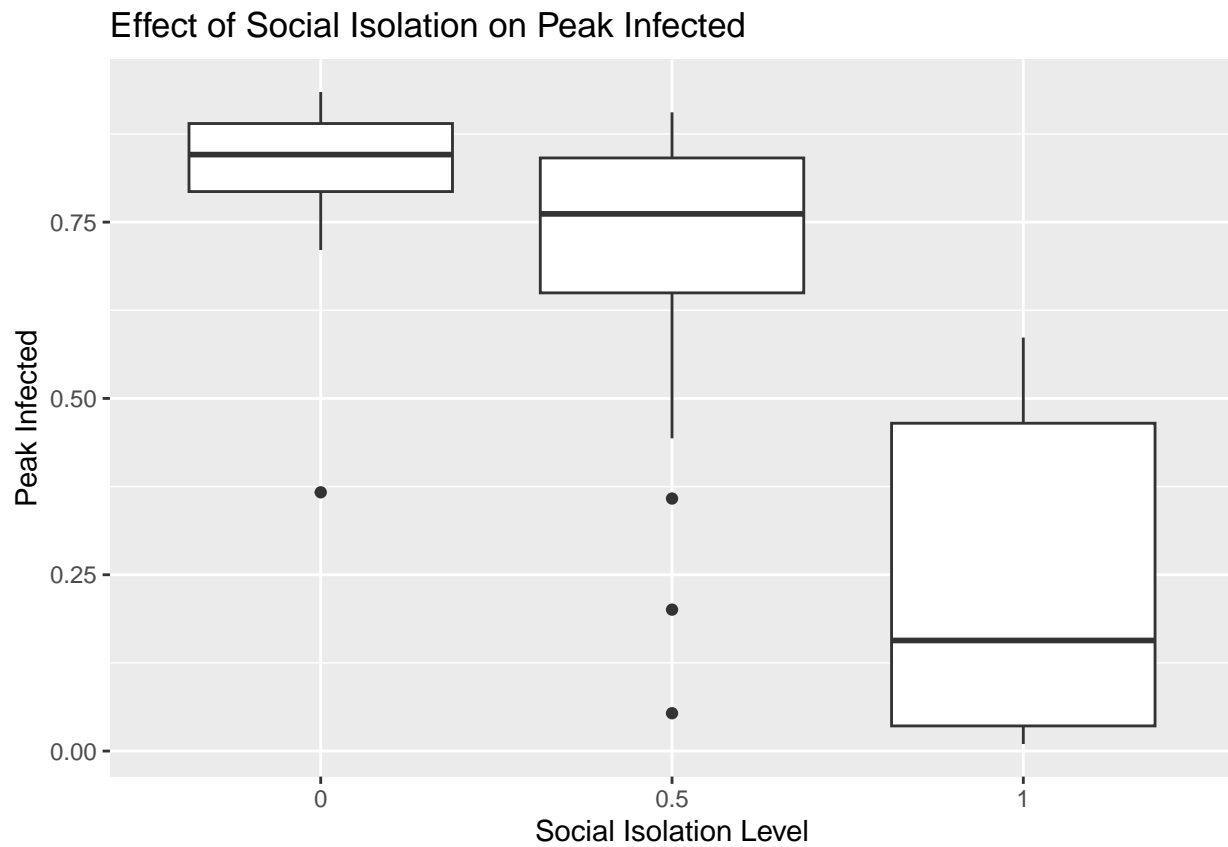


```
interaction.plot(  
  x.factor = data$quar.dur,  
  trace.factor = data$num.daily,  
  response = data$peak.inf,  
  col = 1:4,  
  lty = 1,  
  main = "Interaction Plot: Quarantine Duration x Daily Interactions",  
  xlab = "Quarantine Duration",  
  ylab = "Peak Infected",  
  trace.label = "Daily Interactions"  
)
```

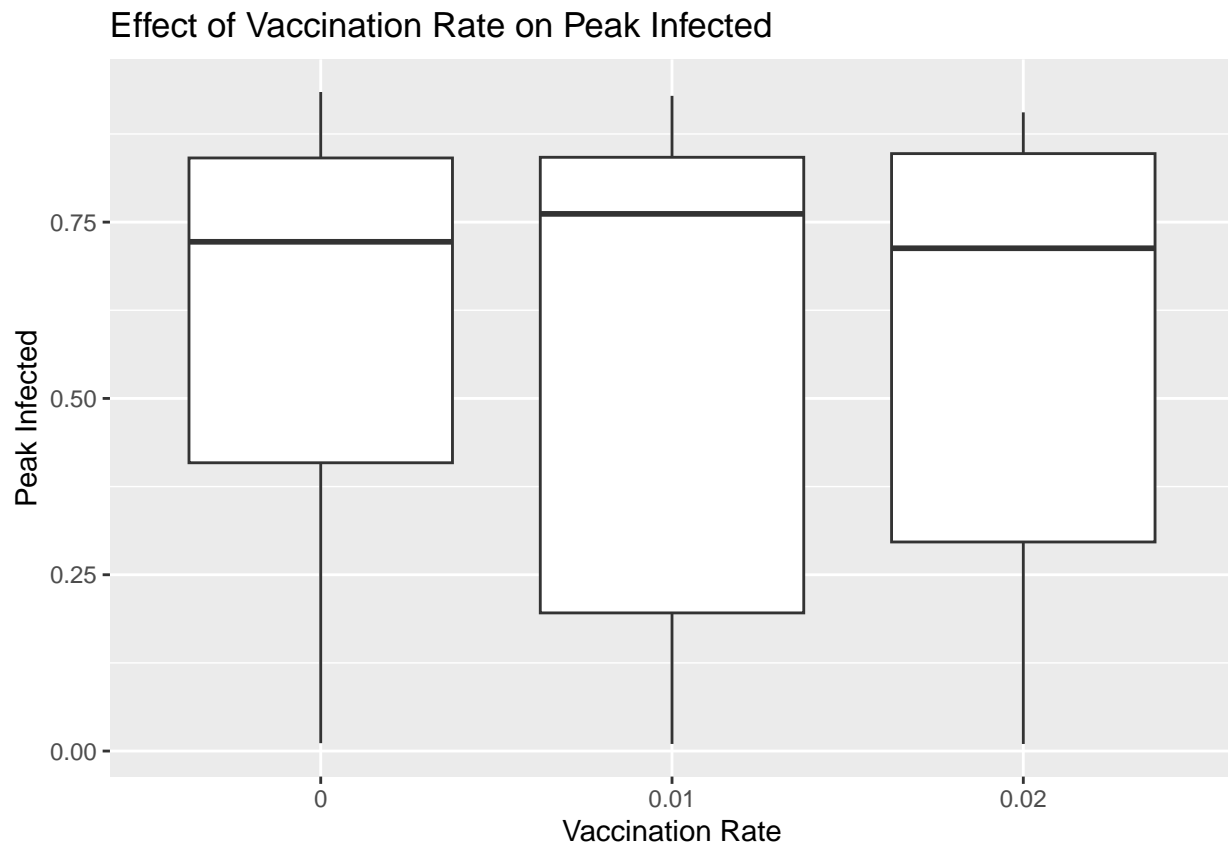
Interaction Plot: Quarantine Duration x Daily Interactions



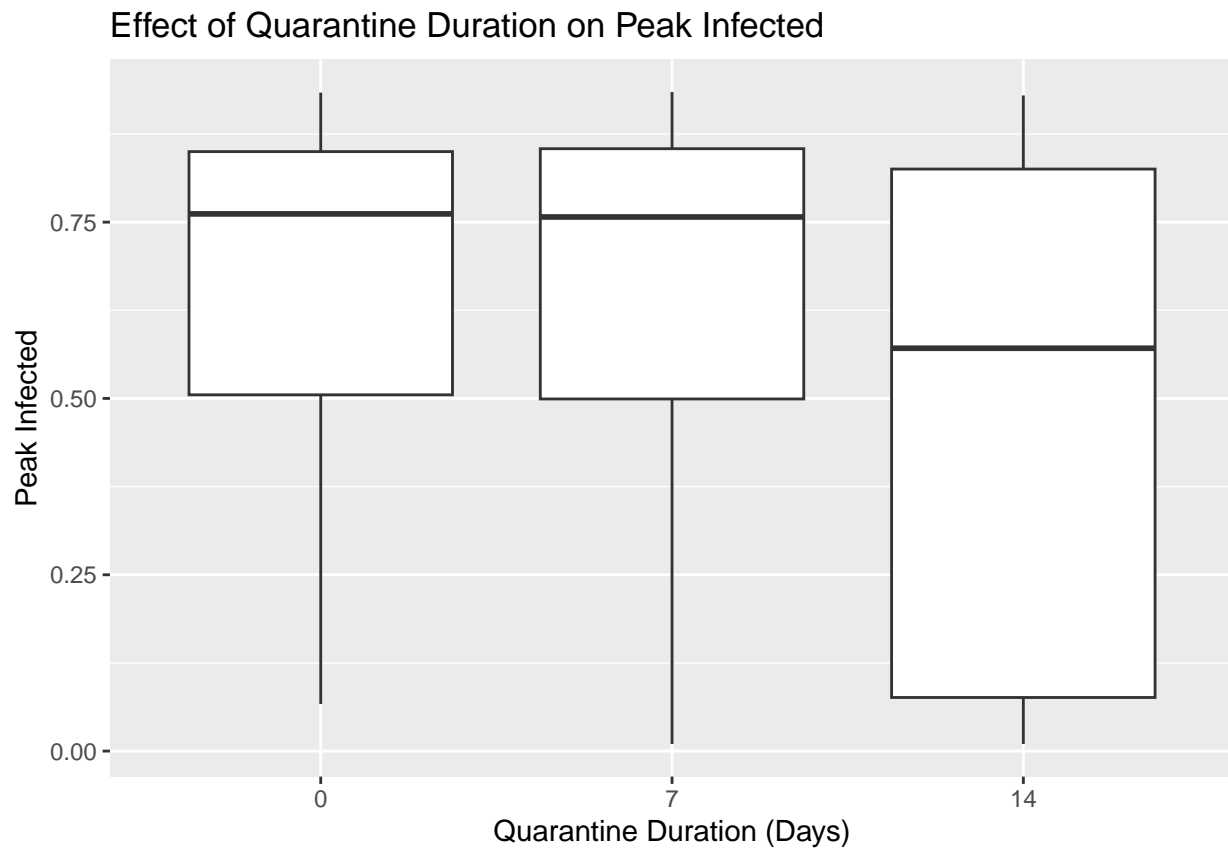
```
# Boxplots
ggplot(data, aes(x = as.factor(soc.iso), y = peak.inf)) +
  geom_boxplot() +
  labs(title = "Effect of Social Isolation on Peak Infected",
        x = "Social Isolation Level", y = "Peak Infected")
```



```
ggplot(data, aes(x = as.factor(rate.vac), y = peak.inf)) +  
  geom_boxplot() +  
  labs(title = "Effect of Vaccination Rate on Peak Infected",  
        x = "Vaccination Rate", y = "Peak Infected")
```

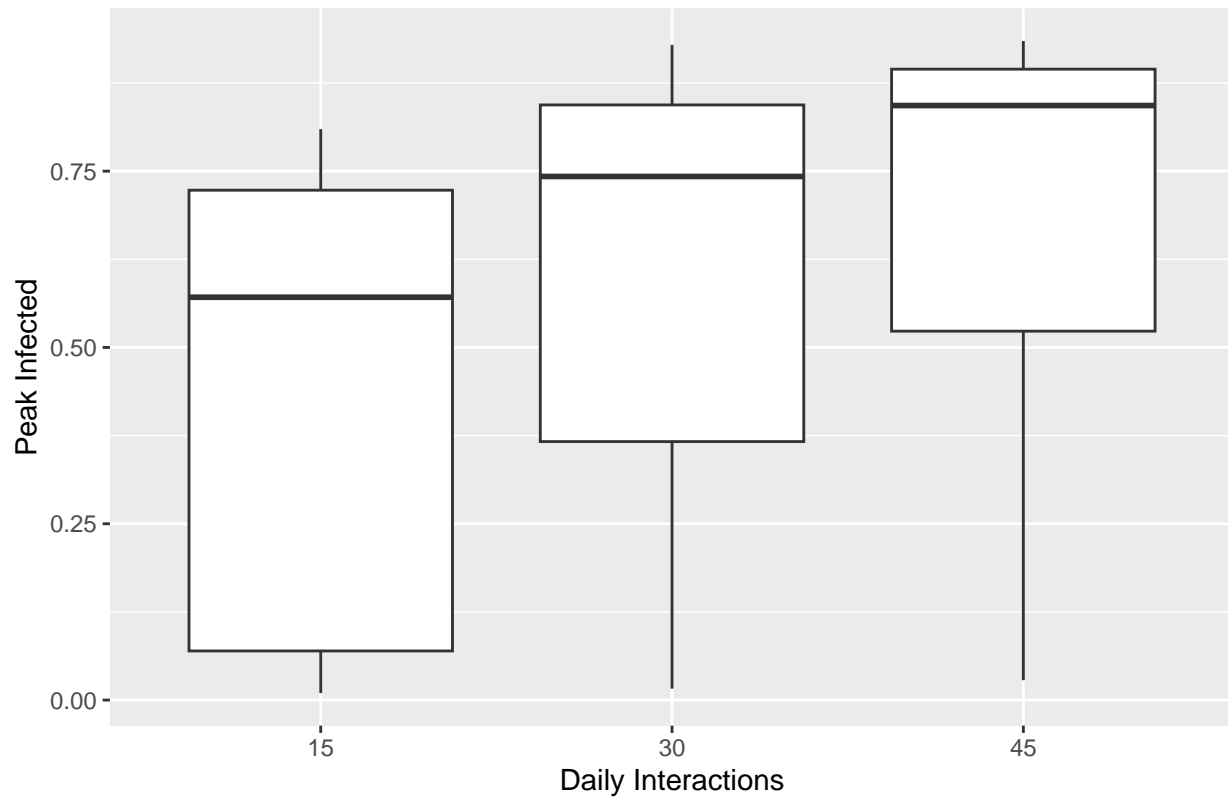


```
ggplot(data, aes(x = as.factor(quar.dur), y = peak.inf)) +  
  geom_boxplot() +  
  labs(title = "Effect of Quarantine Duration on Peak Infected",  
        x = "Quarantine Duration (Days)", y = "Peak Infected")
```



```
ggplot(data, aes(x = as.factor(num.daily), y = peak.inf)) +  
  geom_boxplot() +  
  labs(title = "Effect of Daily Interactions on Peak Infected",  
        x = "Daily Interactions", y = "Peak Infected")
```

Effect of Daily Interactions on Peak Infected



#Optimal Combination

```
aggregated_results <- aggregate(peak.inf ~ soc.iso + rate.vac + quar.dur + num.daily, data = data, mean)
optimal_combination <- aggregated_results[which.min(aggregated_results$peak.inf), ]
optimal_combination
```

```
##      soc.iso rate.vac quar.dur num.daily  peak.inf
## 18         1    0.02         7        15 0.01001624
```