# HW3_boyuj

Boyu Jiang

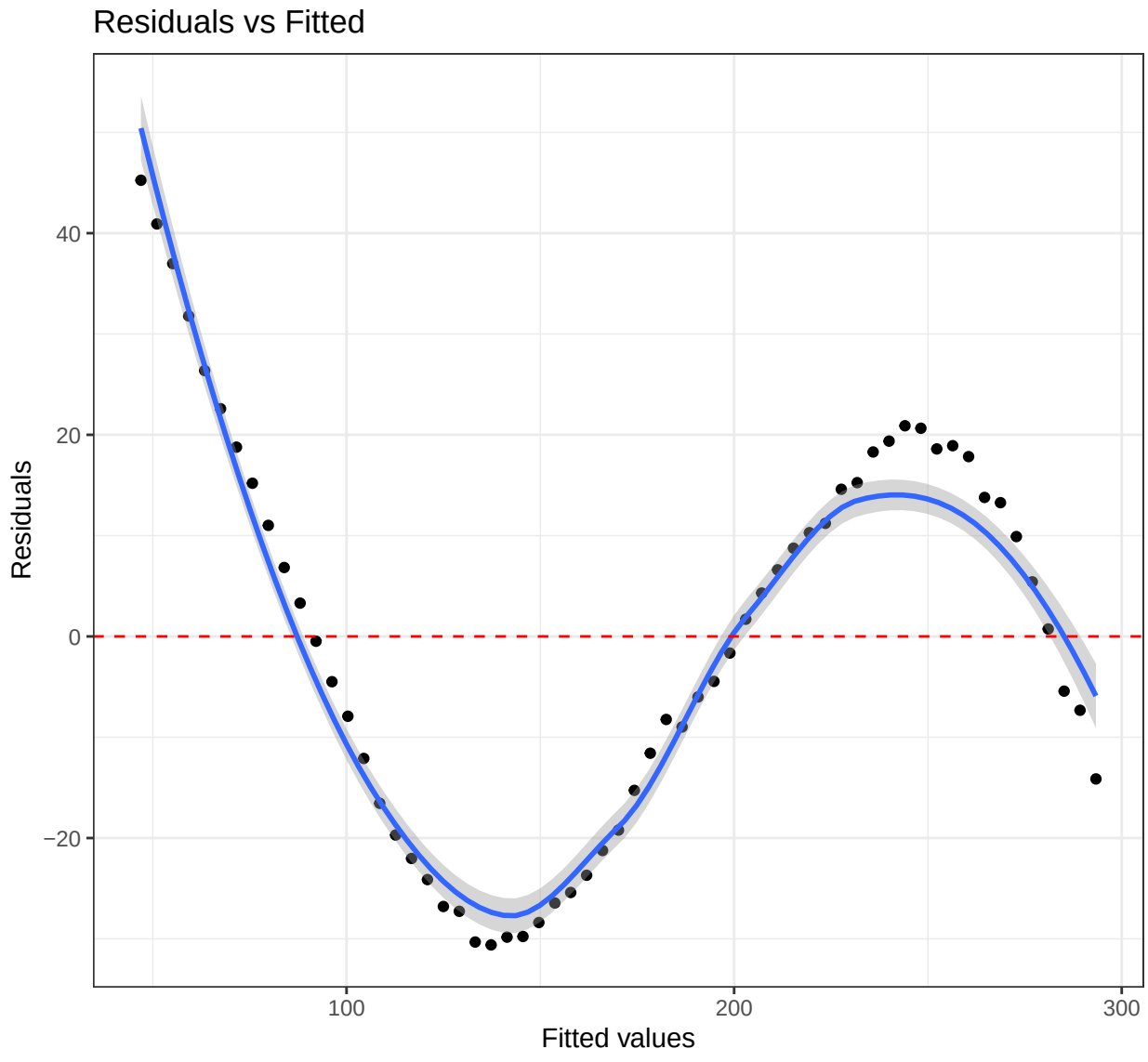
9/28/2021

## Problem 3 and 4

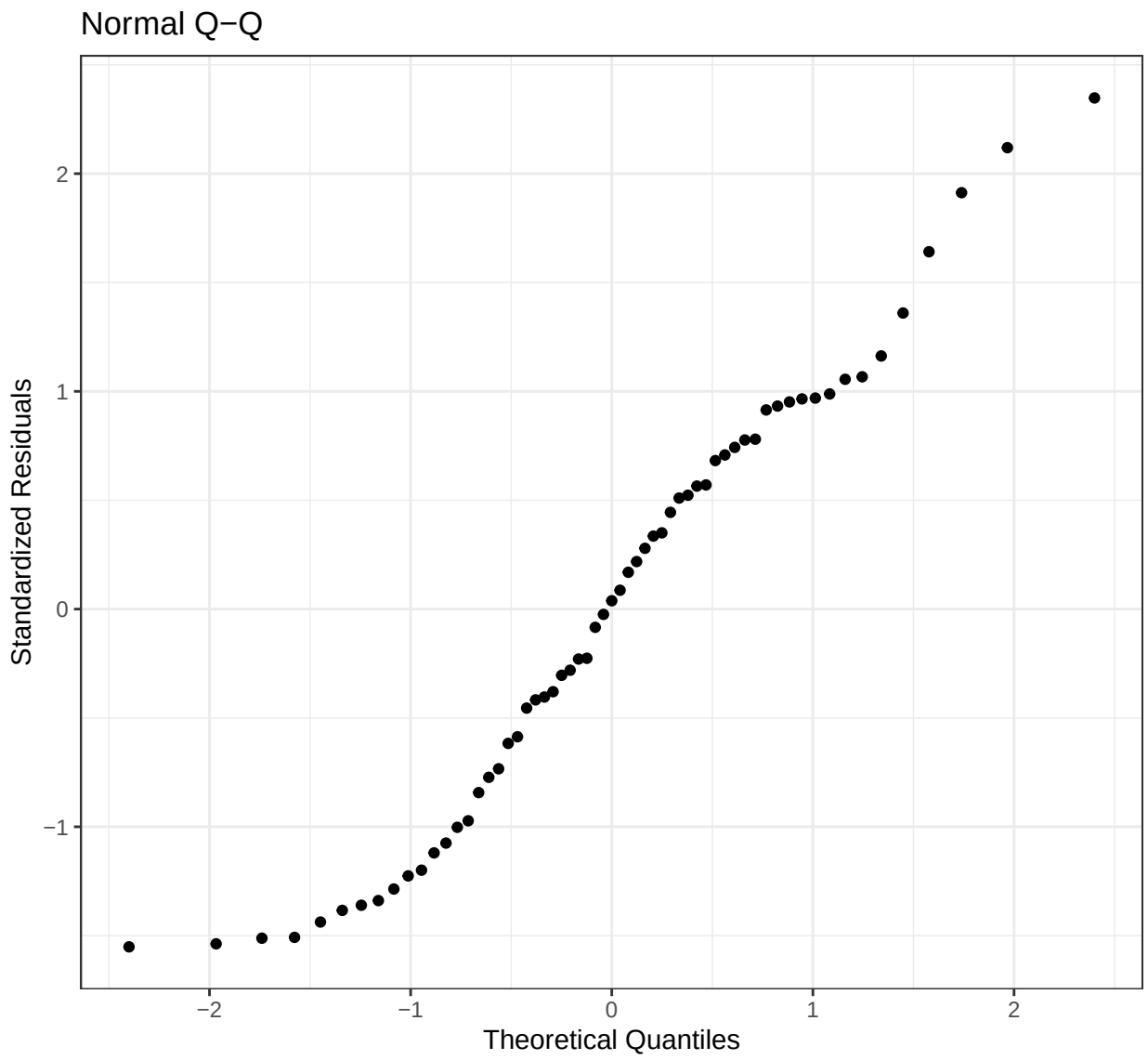
```
library(data.table)
covid_raw <- fread("https://opendata.ecdc.europa.eu/covid19/casedistribution/csv")
us <- covid_raw[covid_raw$countriesAndTerritories == 'United_States_of_America',]
us_filtered <- us[us$month %in% c(6:7),]
us_filtered$index <- rev(1:dim(us_filtered)[1])
fit<-lm(`Cumulative_number_for_14_days_of_COVID-19_cases_per_100000`~index,
        data=us_filtered)

## augment the data as previous
library(broom)
fit.diags <- broom::augment(fit)

library(ggplot2)
library(gridExtra)
plot1 <- ggplot(fit, aes(.fitted, .resid))+
  geom_point()+
  stat_smooth(method="loess")+
  geom_hline(yintercept=0, col="red", linetype="dashed")+
  xlab("Fitted values")+ylab("Residuals")+
  ggtitle("Residuals vs Fitted")+theme_bw()

plot2 <- ggplot(fit, aes(qqnorm(.stdresid, plot.it=FALSE)[[1]], .stdresid))+
  geom_point(na.rm = TRUE)+
  xlab("Theoretical Quantiles")+
  ylab("Standardized Residuals")+
  ggtitle("Normal Q-Q")+theme_bw()

plot3 <- ggplot(fit, aes(.fitted, sqrt(abs(.stdresid))))+
  geom_point(na.rm=TRUE)+
  stat_smooth(method="loess", na.rm = TRUE)+
  xlab("Fitted Value")+
  ylab(expression(sqrt("|Standardized residuals|")))+
  ggtitle("Scale-Location")+theme_bw()

plot4 <- ggplot(fit, aes(.hat, .stdresid))+
  geom_point(aes(size=.cooksd), na.rm=TRUE)+
  stat_smooth(method="loess", na.rm=TRUE)+
  xlab("Leverage")+
```

```
  ylab("Standardized Residuals")+
  ggtitle("Residuals vs Leverage")+
  scale_size_continuous("Cook's Distance", range=c(1,5))+
  theme_bw()+theme(legend.position="bottom")

plot1
```

```
## `geom_smooth()` using formula 'y ~ x'
```

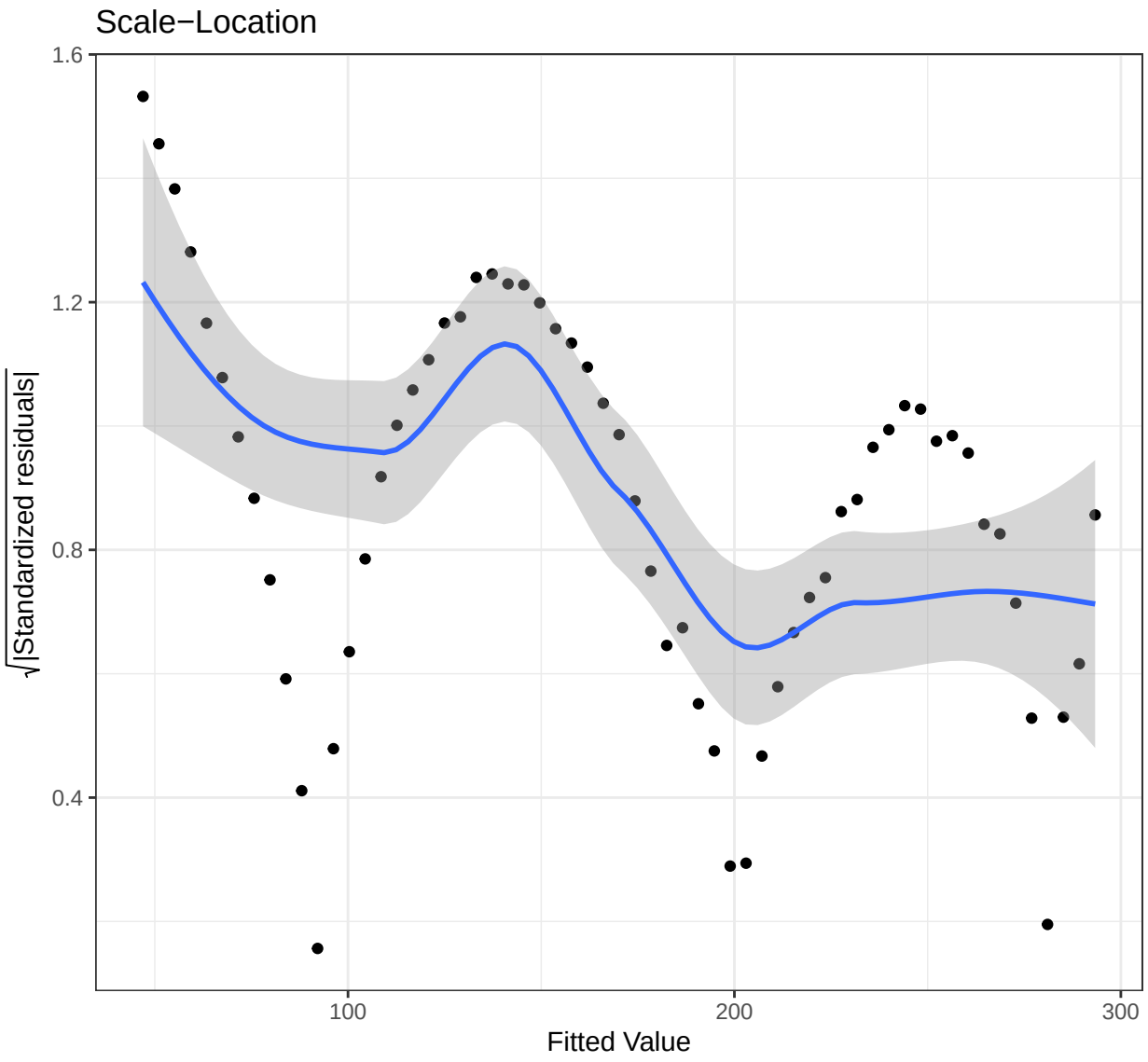



```
plot2
```



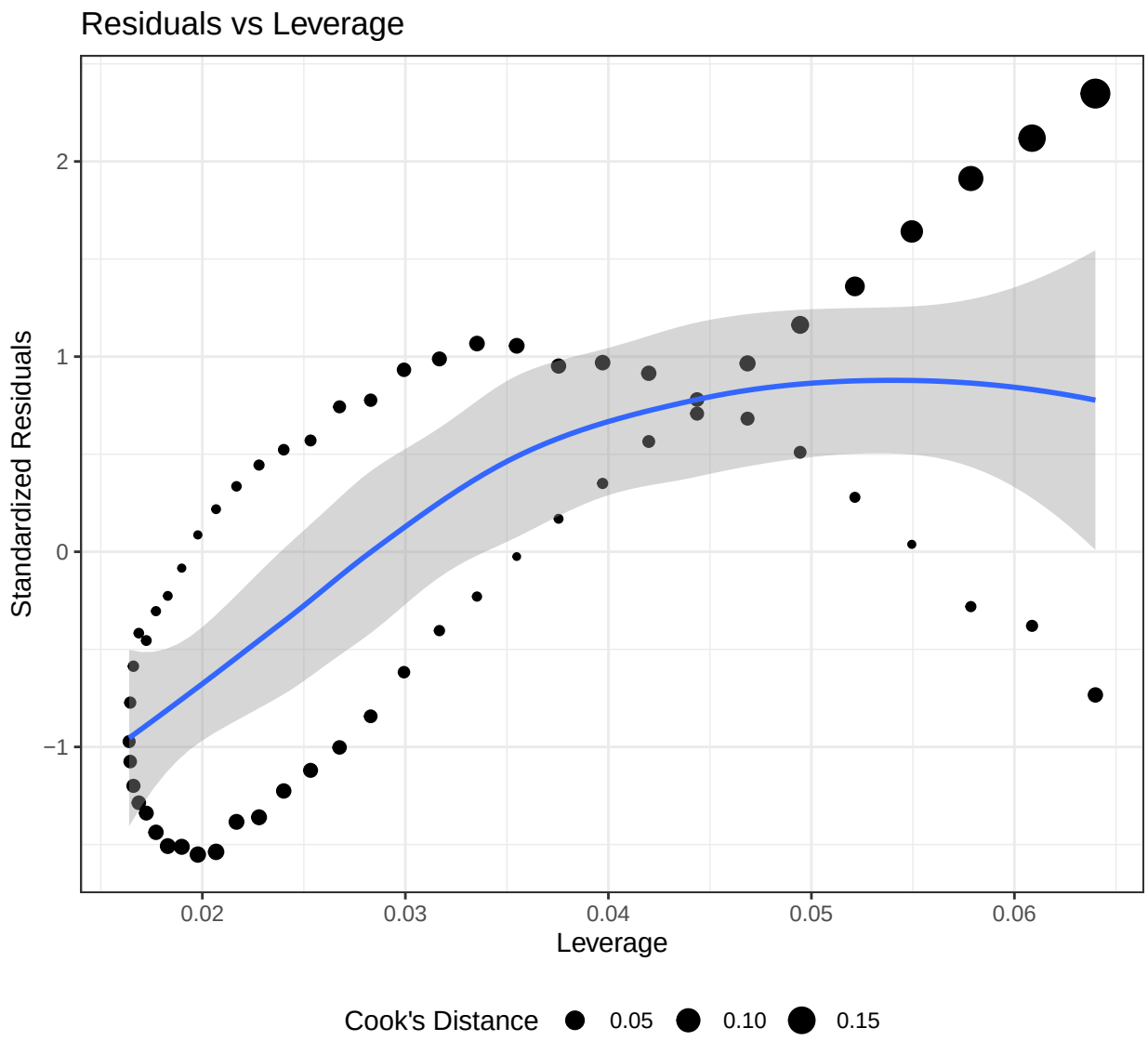

```
plot3
```

```
## `geom_smooth()` using formula 'y ~ x'
```



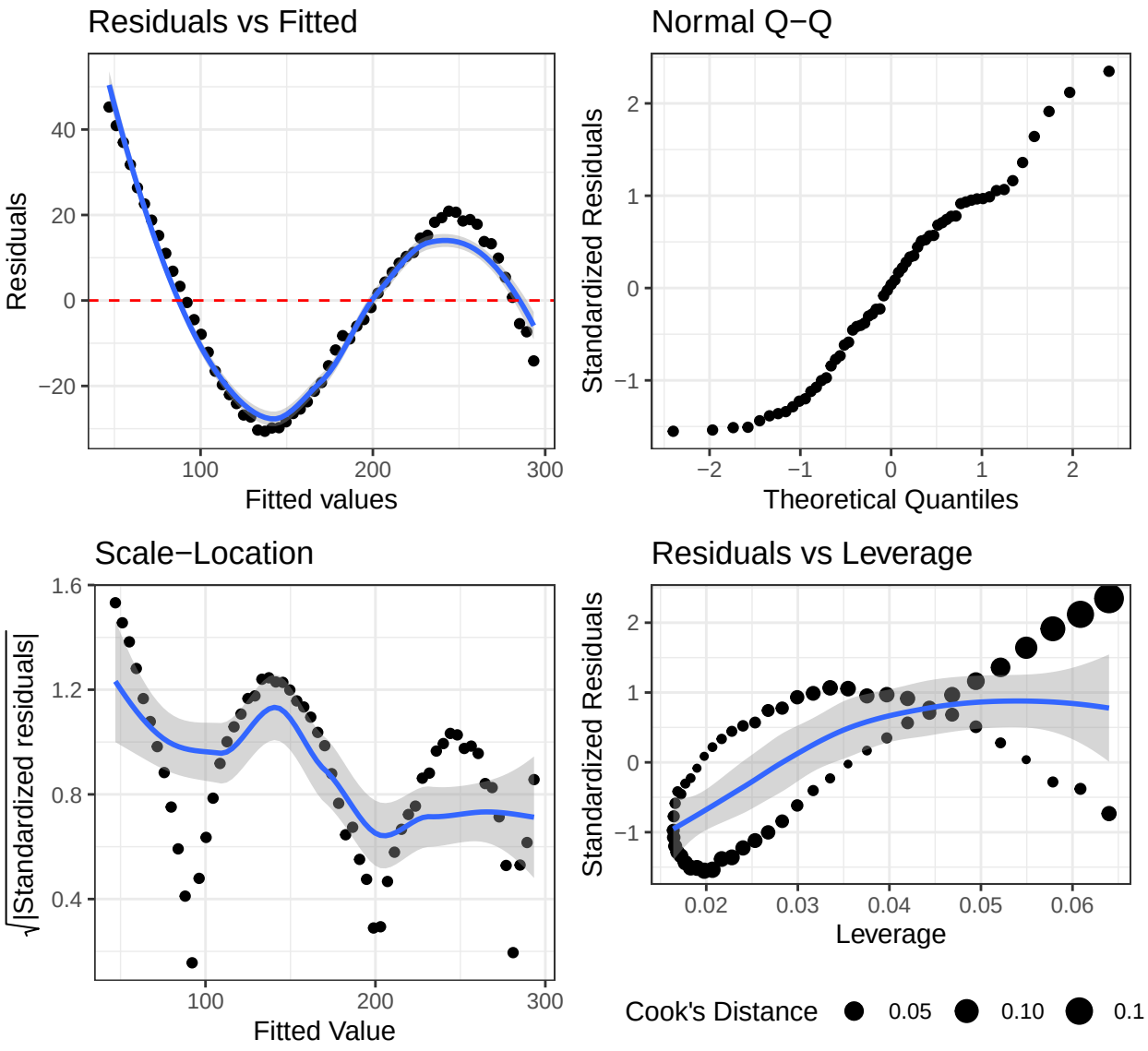

```
plot4
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
grid.arrange(plot1, plot2, plot3, plot4, ncol = 2, nrow = 2)
```

```
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
## `geom_smooth()` using formula 'y ~ x'
```


Residuals vs Fitted
Residuals
Fitted values
Normal Q−Q
Standardized Residuals
Theoretical Quantiles
Scale−Location
√|Standardized residuals|
Fitted Value
Residuals vs Leverage
Standardized Residuals
Leverage
Cook's Distance
0.05
0.10
0.1