Exploratory Analysis on Kenya Road Data (Oct. 15)

USC Machine Learning Team

11/5/2017

Data Loading

```
# load data file
if (!exists("Kenya.raw")) {
    Kenya.raw <- read.csv("./kenya_oct_15_data_labeled.csv")
}
print(nrow(Kenya.raw))
```

```
## [1] 15238
```

```
Kenya.73 = Kenya.raw[Kenya.raw$Batch_id == 73, c("time", "label", "velocity", "x", "x_raw")]
Kenya.73$time <- as.POSIXct(Kenya.73$time, format="%Y-%m-%d %H:%M:%OS")
Kenya.73$Speed <- Kenya.73$velocity
Kenya.73$vert_accel <- Kenya.73$x_raw
Kenya.73$vert_jolt <- Kenya.73$x
Kenya.73 <- Kenya.73[, c("time", "Speed", "vert_accel", "vert_jolt", "label")]
```

time: the time stamp of a data point
Speed: traveling speed of the vehicle
vert_accel: vertical acceleration (up and down) of the vehicle without natural gravity
vert_jolt: vertical jolt of the vehicle (incremental change of vertical acceleration)

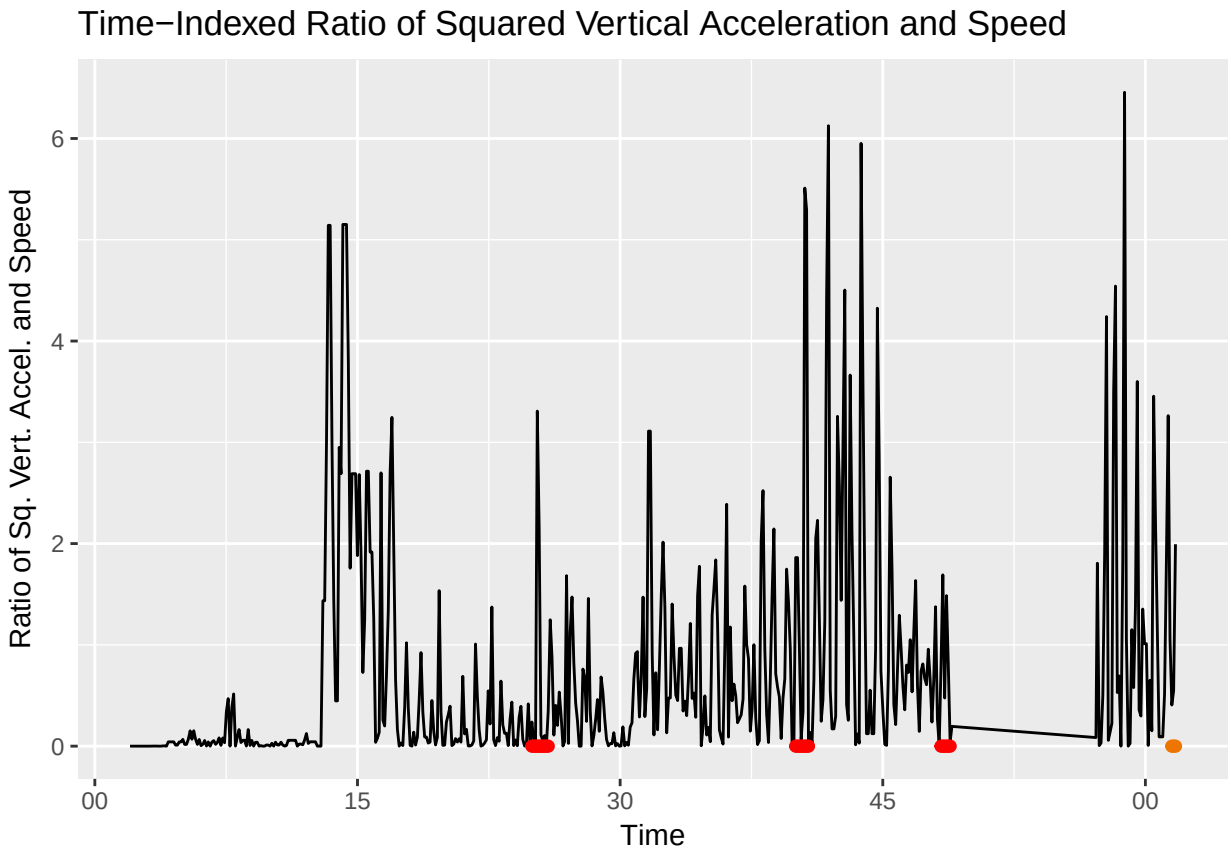
Data Cleaning & Exploratory Analysis

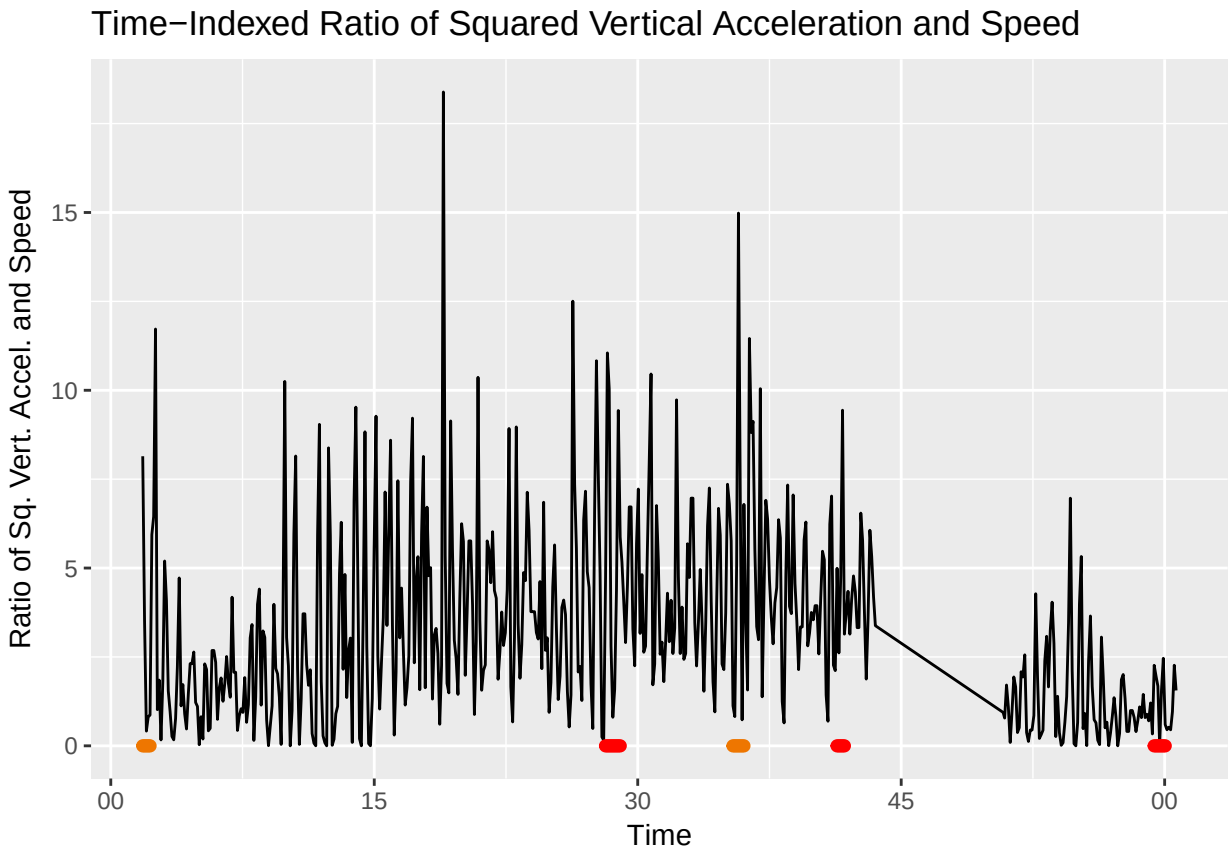
Epoch 1: Display of Ratio betwwen Vertical Acceleration (w/o Natural G) and Speed

sq_vert_accel_ratio_speed: the ratio between squared vertical acceleration (without natual gravity) and traveling speed

```
# calculate ratio between vertical acceleration and speed
for (i in 1:nrow(Kenya.73)) {
    if (Kenya.73$Speed[i] == 0) {
        Kenya.73$sq_vert_accel_ratio_speed[i] = 0
    }
    else {
        Kenya.73$sq_vert_accel_ratio_speed[i] = (Kenya.73$vert_accel[i] * Kenya.73$vert_accel[i]) / Keny
    }
}
```

```
## Loading required package: ggplot2
```

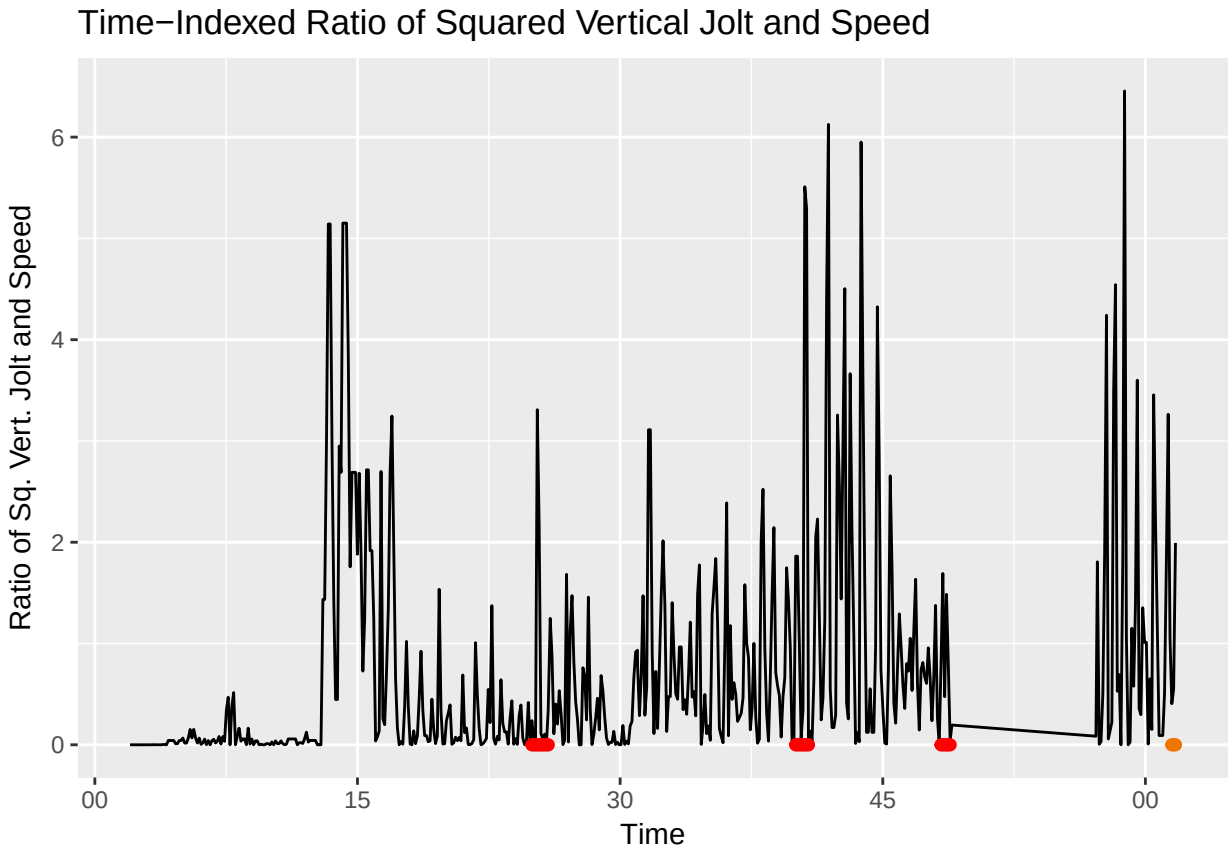


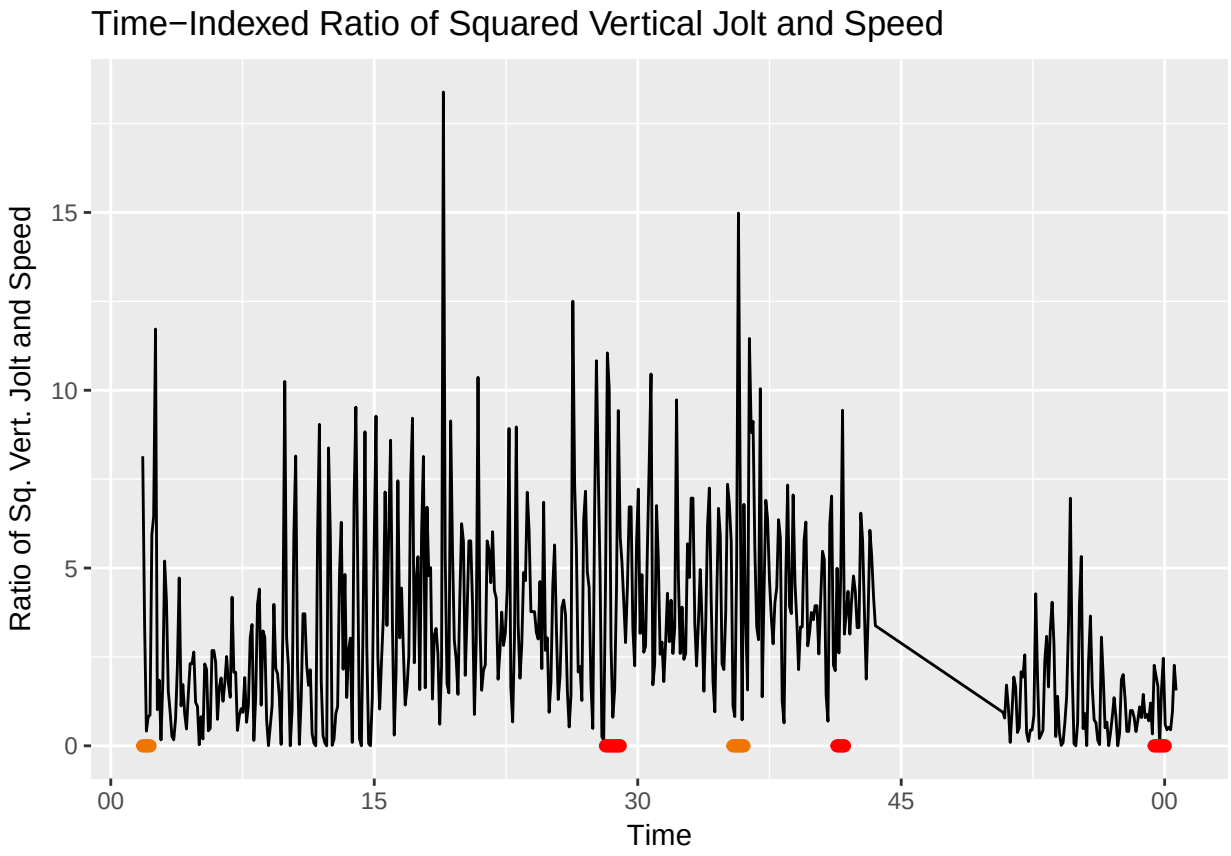

Time−Indexed Ratio of Squared Vertical Acceleration and Speed
Ratio of Sq. Vert. Accel. and Speed
15
10
5
0
00
15
30
45
00
Time

Epoch 2: Display of Ratio betwwen Vertical Jolt and Speed

sq_vert_jolt_ratio_speed: the ratio between squared vertical acceleration (without natual gravity) and traveling speed

```
# calculate ratio between vertical acceleration and speed
for (i in 1:nrow(Kenya.73)) {
    if (Kenya.73$Speed[i] == 0) {
        Kenya.73$sq_vert_jolt_ratio_speed[i] = 0
    }
    else {
        Kenya.73$sq_vert_jolt_ratio_speed[i] = (Kenya.73$vert_accel[i] * Kenya.73$vert_accel[i]) / Kenya
    }
}
```




Time−Indexed Ratio of Squared Vertical Jolt and Speed
Ratio of Sq. Vert. Jolt and Speed
15
10
5
0
00
15
30
45
00
Time

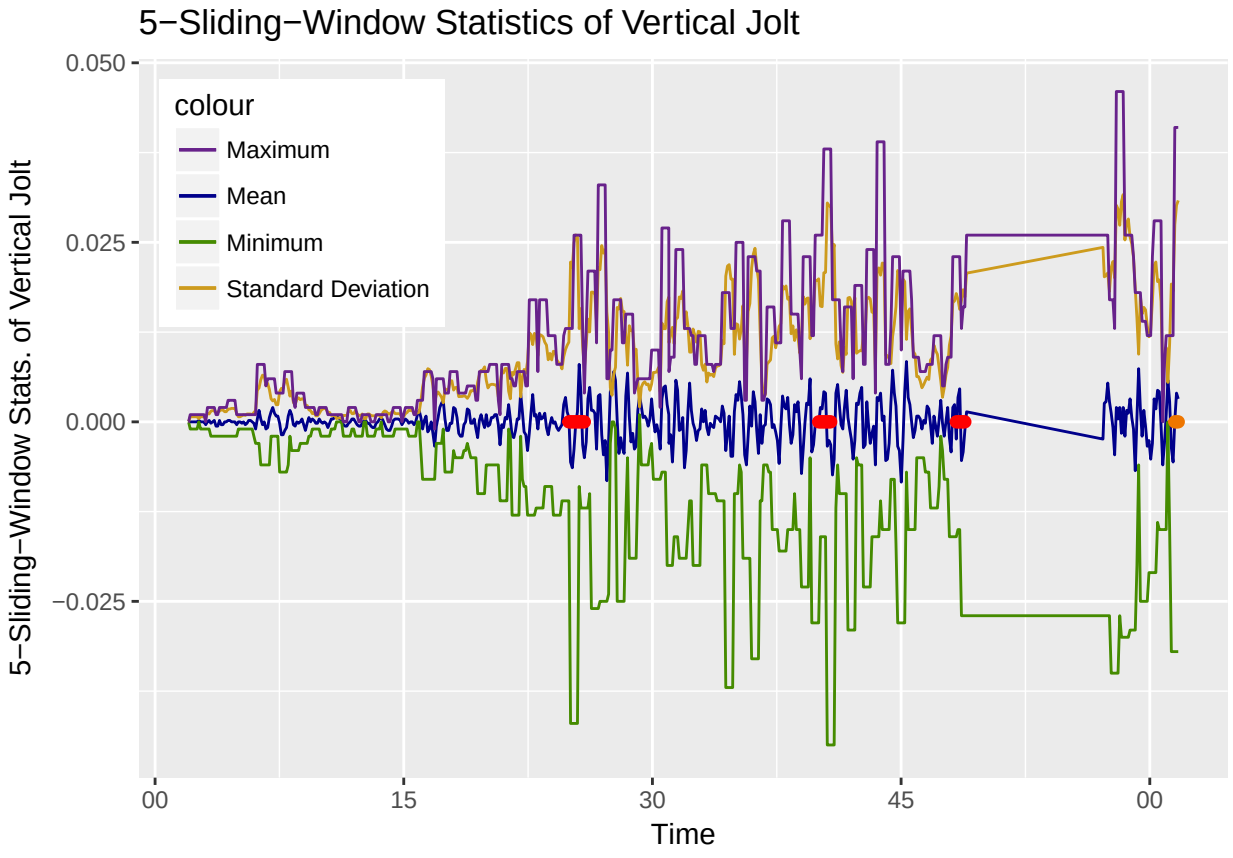
Epoch 3: Display of 5-Sliding-Window Statistics of Vertical Jolt

vert_jolt_mean: 5-sliding-window mean of vertical jolt
vert_jolt_sd: 5-sliding-window standard deviation of vertical jolt
vert_jolt_min: 5-sliding-window minimum of vertical jolt
vert_jolt_max: 5-sliding-window maximum of vertical jolt

```
# calculate 5-sliding-window mean of vertical jolt
for (i in 3:nrow(Kenya.73)-2) {
    Kenya.73$vert_jolt_mean[i] = mean(c(Kenya.73$vert_jolt[i-2],
                                        Kenya.73$vert_jolt[i-1],
                                        Kenya.73$vert_jolt[i],
                                        Kenya.73$vert_jolt[i+1],
                                        Kenya.73$vert_jolt[i+2]))
}
Kenya.73$vert_jolt_mean[1] = 0
Kenya.73$vert_jolt_mean[2] = 0
Kenya.73$vert_jolt_mean[nrow(Kenya.73) - 1] = 0
Kenya.73$vert_jolt_mean[nrow(Kenya.73)] = 0
# calculate 5-sliding-window standard deviation of vertical jolt
for (i in 3:nrow(Kenya.73)-2) {
    Kenya.73$vert_jolt_sd[i] = sd(c(Kenya.73$vert_jolt[i-2],
                                    Kenya.73$vert_jolt[i-1],
                                    Kenya.73$vert_jolt[i],
                                    Kenya.73$vert_jolt[i+1],
                                    Kenya.73$vert_jolt[i+2]))
}
Kenya.73$vert_jolt_sd[1] = 0
Kenya.73$vert_jolt_sd[2] = 0
Kenya.73$vert_jolt_sd[nrow(Kenya.73) - 1] = 0
Kenya.73$vert_jolt_sd[nrow(Kenya.73)] = 0
# calculate 5-sliding-window minimum of vertical jolt
for (i in 3:nrow(Kenya.73)-2) {
    Kenya.73$vert_jolt_min[i] = min(c(Kenya.73$vert_jolt[i-2],
                                      Kenya.73$vert_jolt[i-1],
                                      Kenya.73$vert_jolt[i],
                                      Kenya.73$vert_jolt[i+1],
                                      Kenya.73$vert_jolt[i+2]))
}
Kenya.73$vert_jolt_min[1] = 0
Kenya.73$vert_jolt_min[2] = 0
Kenya.73$vert_jolt_min[nrow(Kenya.73) - 1] = 0
Kenya.73$vert_jolt_min[nrow(Kenya.73)] = 0
# calculate 5-sliding-window maxium of vertical jolt
for (i in 3:nrow(Kenya.73)-2) {
    Kenya.73$vert_jolt_max[i] = max(c(Kenya.73$vert_jolt[i-2],
                                      Kenya.73$vert_jolt[i-1],
                                      Kenya.73$vert_jolt[i],
                                      Kenya.73$vert_jolt[i+1],
                                      Kenya.73$vert_jolt[i+2]))
}
Kenya.73$vert_jolt_max[1] = 0
Kenya.73$vert_jolt_max[2] = 0
```

```
Kenya.73$vert_jolt_max[nrow(Kenya.73) - 1] = 0
Kenya.73$vert_jolt_max[nrow(Kenya.73)] = 0
```

```
require(ggplot2)
Kenya.73.vert_jolt_5.plot <- ggplot(Kenya.73[1:500, ], aes(x = time))
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_mean, colour = "Mea
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_sd, colour = "Stand
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_min, colour = "Mini
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_max, colour = "Maxi
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + scale_colour_manual(values = c("darkorchid4", "
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + theme(legend.position = c(0.15, 0.8))
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + xlab("Time")
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + ylab("5-Sliding-Window Stats. of Vertical Jolt
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + ggtitle("5-Sliding-Window Statistics of Vertica
for (i in 1:500) {
    if (Kenya.73$label[i] == 1) {
         Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot +
             annotate("pointrange", x = Kenya.73$time[i], y = 0, ymin = 0, ymax = 0, colour = "red", siz
    }
    if (Kenya.73$label[i] == 2) {
        Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot +
             annotate("pointrange", x = Kenya.73$time[i], y = 0, ymin = 0, ymax = 0, colour = "darkorang
    }
}
print(Kenya.73.vert_jolt_5.plot)
```



```
require(ggplot2)
Kenya.73.vert_jolt_5.plot <- ggplot(Kenya.73[501:1000, ], aes(x = time))
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_mean, colour = "Mea
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_sd, colour = "Stand
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_min, colour = "Mini
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + geom_line(aes(y = vert_jolt_max, colour = "Maxi
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + scale_colour_manual(values = c("darkorchid4", "
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + theme(legend.position = c(0.8, 0.925))
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + xlab("Time")
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + ylab("5-Sliding-Window Stats. of Vertical Jolt"
Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot + ggtitle("5-Sliding-Window Statistics of Vertica
for (i in 501:1000) {
    if (Kenya.73$label[i] == 1) {
        Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot +
            annotate("pointrange", x = Kenya.73$time[i], y = 0, ymin = 0, ymax = 0, colour = "red", siz
    }
    if (Kenya.73$label[i] == 2) {
        Kenya.73.vert_jolt_5.plot <- Kenya.73.vert_jolt_5.plot +
            annotate("pointrange", x = Kenya.73$time[i], y = 0, ymin = 0, ymax = 0, colour = "darkorang
    }
}
print(Kenya.73.vert_jolt_5.plot)
```