```
## Loading required package:  Defaults
## Loading required package:  xts
## Loading required package:  zoo
##
## Attaching package:  'zoo'
##
## The following object(s) are masked from 'package:base':
##
## as.Date, as.Date.numeric
##
## Loading required package:  TTR
```

Estimating Traffic Volumes in Athens: TRB data analysis competition

Bhargava R. Chilukuri[a], Adnan Sheikh[a], Gregory S. Macfarlane[a,b,*]

[a]*School of Civil and Environmental Engineering, Georgia Institute of Technology*
790 Atlantic Drive, Atlanta GA 30332-0355
[b]*School of Economics, Georgia Institute of Technology*
221 Bobby Dodd Way, Atlanta, GA 30332

Abstract

Traffic loop detectors are important tools for recording and monitoring vehicle flows along major routes in a region. The reliability of these detectors, however, is such that certain important observations may be missing. In this study, we employ an approach based on traffic flow theory and joined with time series econometrics to impute missing values, and make modest projections, from loop detector data in Athens, Greece.

Keywords: TRB data analysis competition, traffic forecasting

1. Introduction

1.1. Literature

2. Model

```
MyData <- read.csv("./Source_Data/data_additional_April.csv")
lanes <- c(1, 2, 3, 4, 6, 7, 8)



# Convert timestamp
MyData$TIMESTAMP <- as.POSIXct(strptime(MyData$TIMESTAMP, format = "%m/%d/%y %H:%M"))


# change 255 for missing


# critical occupancy dummy
MyData$L101_crit <- ifelse(MyData$L101_occupancy > 20, 1, 0)
```

*Corresponding author. Tel.: +1 801 616 9822
Email addresses: bchilukuri3@gatech.edu (Bhargava R. Chilukuri), asheikh7@gatech.edu (Adnan Sheikh), gregmacfarlane@gatech.edu (Gregory S. Macfarlane)

Table 1: Loop1 models

```
apsrtable(model1lag, model1lagc, model1int, model2lag, Sweave = TRUE,
    coefrows = 1, digits = 4)

## Error:  $ operator is invalid for atomic vectors
```

```
# throw away volume outliers
MyData$L101_volume <- ifelse(MyData$L101_volume > 100, NA, MyData$L101_volume)
```

2.1. Estimation

We estimate an autoregressive model for each lane, and present the coefficient estimates in Table 2

```
model1lag <- lm(L101_volume ~ Lag(L101_volume, k = 1), data = MyData)
model1lagc <- lm(L101_volume ~ Lag(L101_volume, k = 1) + L101_crit,
    data = MyData)
model1int <- lm(L101_volume ~ Lag(L101_volume, k = 1) + L101_crit +
    Lag(L101_volume, k = 1):L101_crit, data = MyData)
model2lag <- lm(L101_volume ~ Lag(L101_volume, k = 1):L101_crit +
    Lag(L101_volume, k = 1) + L101_crit + Lag(L101_volume, k = 2), data = MyData)
```

```
model1 <- lm(L101_volume ~ Lag(L101_volume, k = 1):L101_crit + Lag(L101_volume,
    k = 1) + L101_crit + Lag(L101_volume, k = 960), data = MyData)
model2 <- lm(L102_volume ~ Lag(L102_volume, k = 1), data = MyData)
model3 <- lm(L103_volume ~ Lag(L103_volume, k = 1), data = MyData)
model4 <- lm(L104_volume ~ Lag(L104_volume, k = 1), data = MyData)
model6 <- lm(L106_volume ~ Lag(L106_volume, k = 1), data = MyData)
model7 <- lm(L107_volume ~ Lag(L107_volume, k = 1), data = MyData)
model8 <- lm(L108_volume ~ Lag(L108_volume, k = 1), data = MyData)
AR1models <- list(model1, model2, model3, model4, model6, model7,
    model8)
```

Table 2: Autoregressive Model Coefficients

```
AR1Coefs <- as.table(matrix(NA, ncol = length(AR1models), nrow = 3))
for (i in 1:length(AR1models)) {
    AR1Coefs[1:2, i] <- t(coef(AR1models[[i]]))
    AR1Coefs[3, i] <- summary(AR1models[[i]])$r.squared
    colnames(AR1Coefs)[i] <- paste("Lane ", lanes[i], sep = "")
}
      ## Error:  number of items to replace is not a multiple of replacement length

rownames(AR1Coefs) <- c("Intercept", "Lag", "$R^2$")


coefs.x <- xtable(AR1Coefs)
print(coefs.x, floating = FALSE, sanitize.rownames.function = function(x) {
    x
})
```

	A	B	C	D	E	F	G
Intercept							
Lag							
R^2							

```
model21 <- lm(L101_volume ~ Lag(L101_volume, k = 1) + Lag(L101_volume,
    k = 2), data = MyData)
model22 <- lm(L102_volume ~ Lag(L102_volume, k = 1) + Lag(L102_volume,
    k = 2), data = MyData)
model23 <- lm(L103_volume ~ Lag(L103_volume, k = 1) + Lag(L103_volume,
    k = 2), data = MyData)
model24 <- lm(L104_volume ~ Lag(L104_volume, k = 1) + Lag(L104_volume,
    k = 2), data = MyData)
model26 <- lm(L106_volume ~ Lag(L106_volume, k = 1) + Lag(L106_volume,
    k = 2), data = MyData)
model27 <- lm(L107_volume ~ Lag(L107_volume, k = 1) + Lag(L107_volume,
    k = 2), data = MyData)
model28 <- lm(L108_volume ~ Lag(L108_volume, k = 1) + Lag(L108_volume,
    k = 2), data = MyData)
AR2models <- list(model21, model22, model23, model24, model26, model27,
    model28)
```

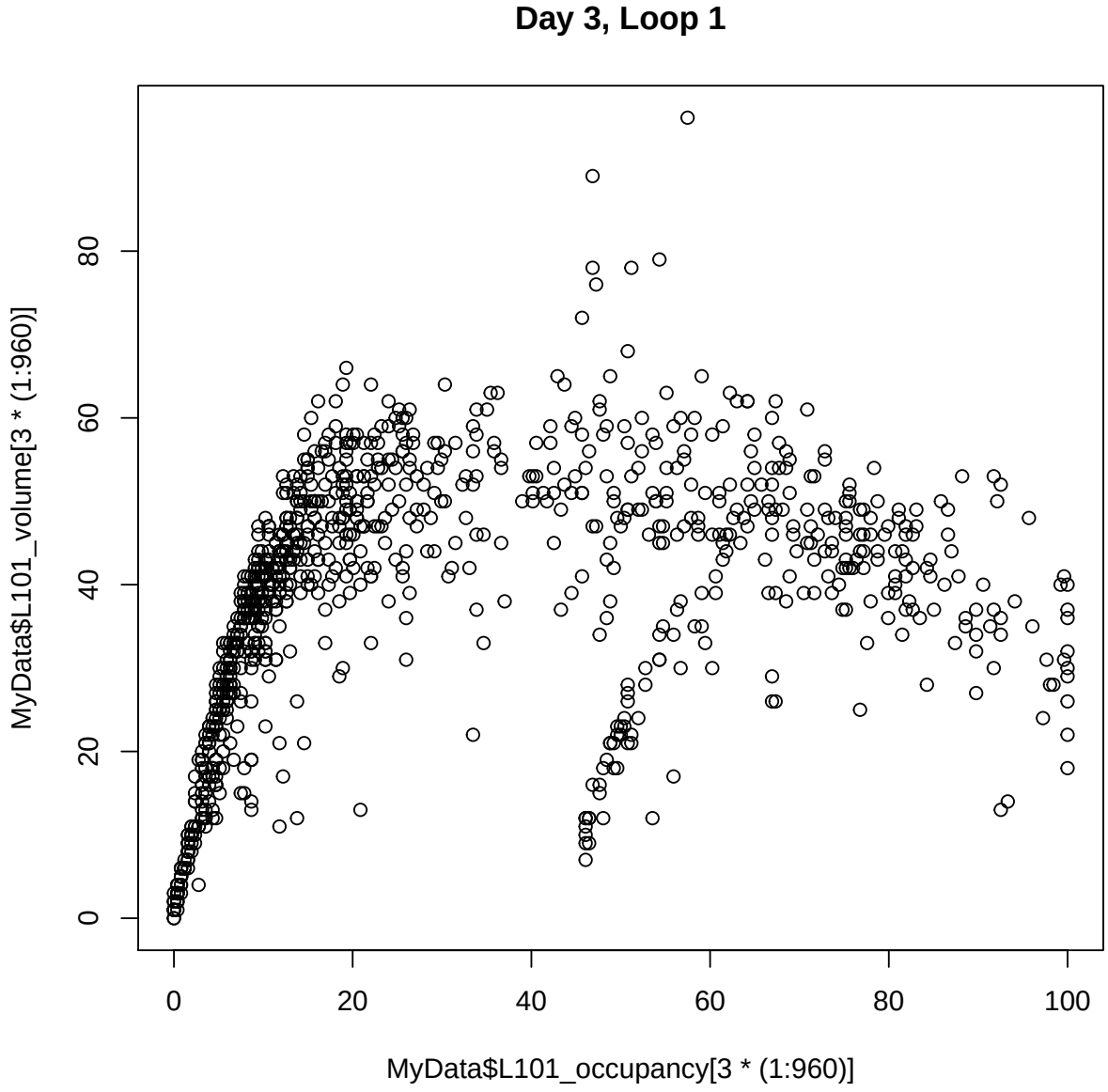
```
plot(MyData$L101_occupancy[3 * (1:960)], MyData$L101_volume[3 * (1:960)],
    main = "Day 3, Loop 1")
```

Table 3: Autoregressive 2 Model Coefficients

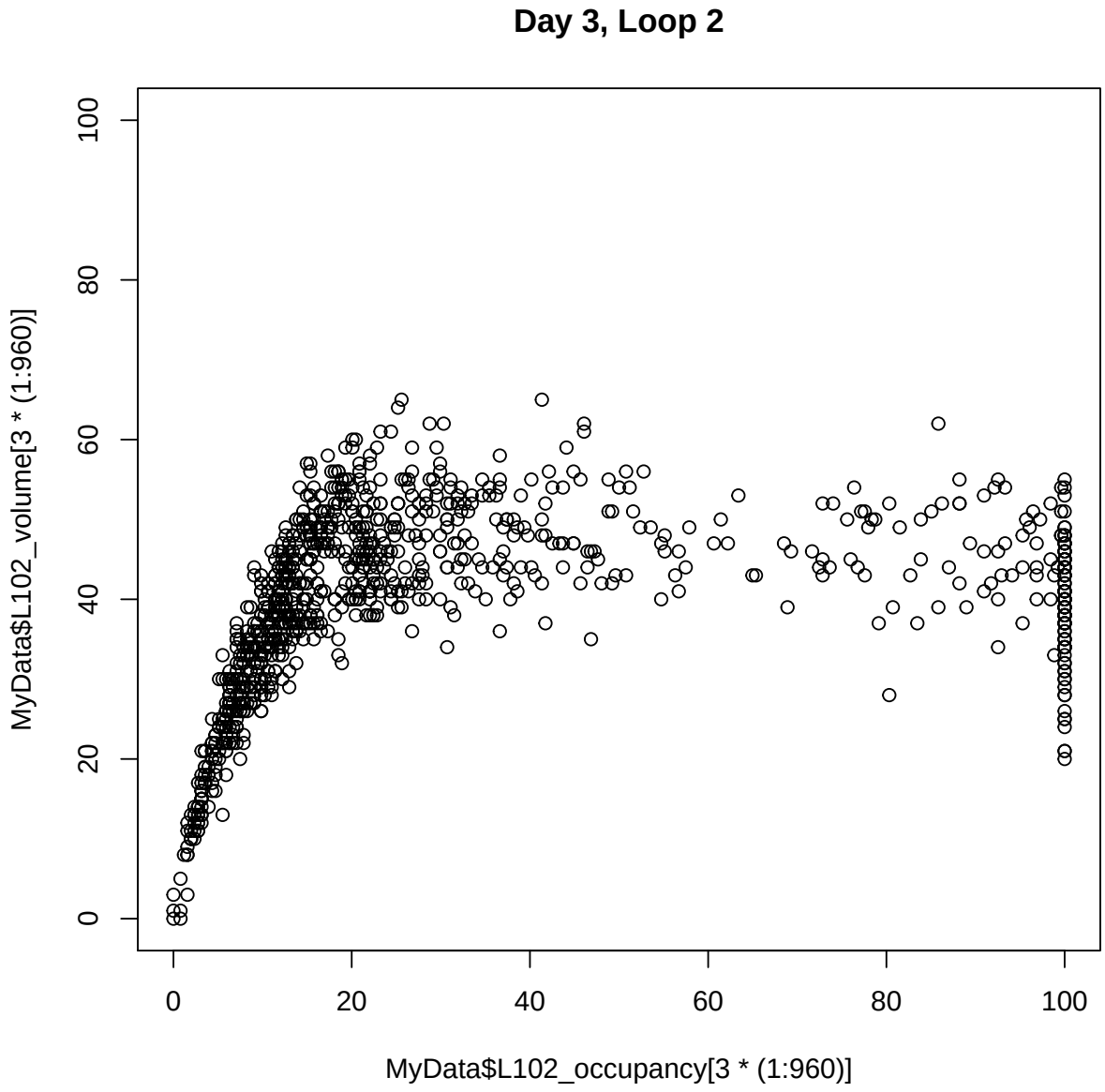
```
AR2Coefs <- as.table(matrix(NA, ncol = length(AR1models), nrow = 4))
for (i in 1:length(AR1models)) {
    AR2Coefs[1:3, i] <- t(coef(AR2models[[i]]))
    AR2Coefs[4, i] <- summary(AR2models[[i]])$r.squared
    colnames(AR1Coefs)[i] <- paste("Lane ", lanes[i], sep = "")
}
rownames(AR2Coefs) <- c("Intercept", "Lag", "Lag-2", "$R^2$")

coefs.x <- xtable(AR2Coefs)
print(coefs.x, floating = FALSE, sanitize.rownames.function = function(x) {
    x
})
```

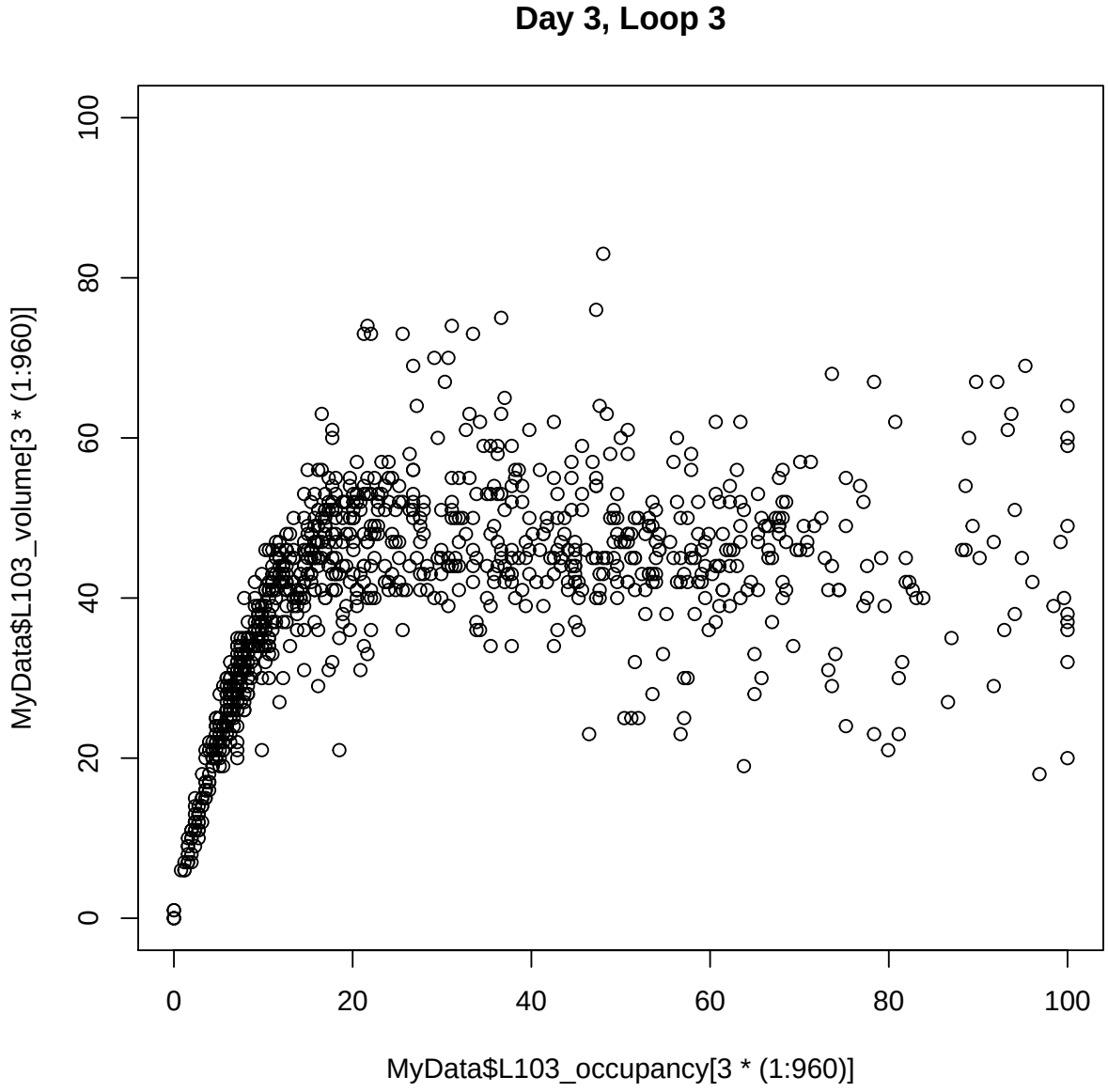
	A	B	C	D	E	F	G
Intercept	4.65	3.55	3.88	7.71	8.12	3.18	3.53
Lag	0.44	0.47	0.57	0.57	0.49	0.49	0.59
Lag-2	0.42	0.44	0.33	0.19	0.33	0.41	0.30
R^2	0.67	0.76	0.75	0.50	0.60	0.76	0.75



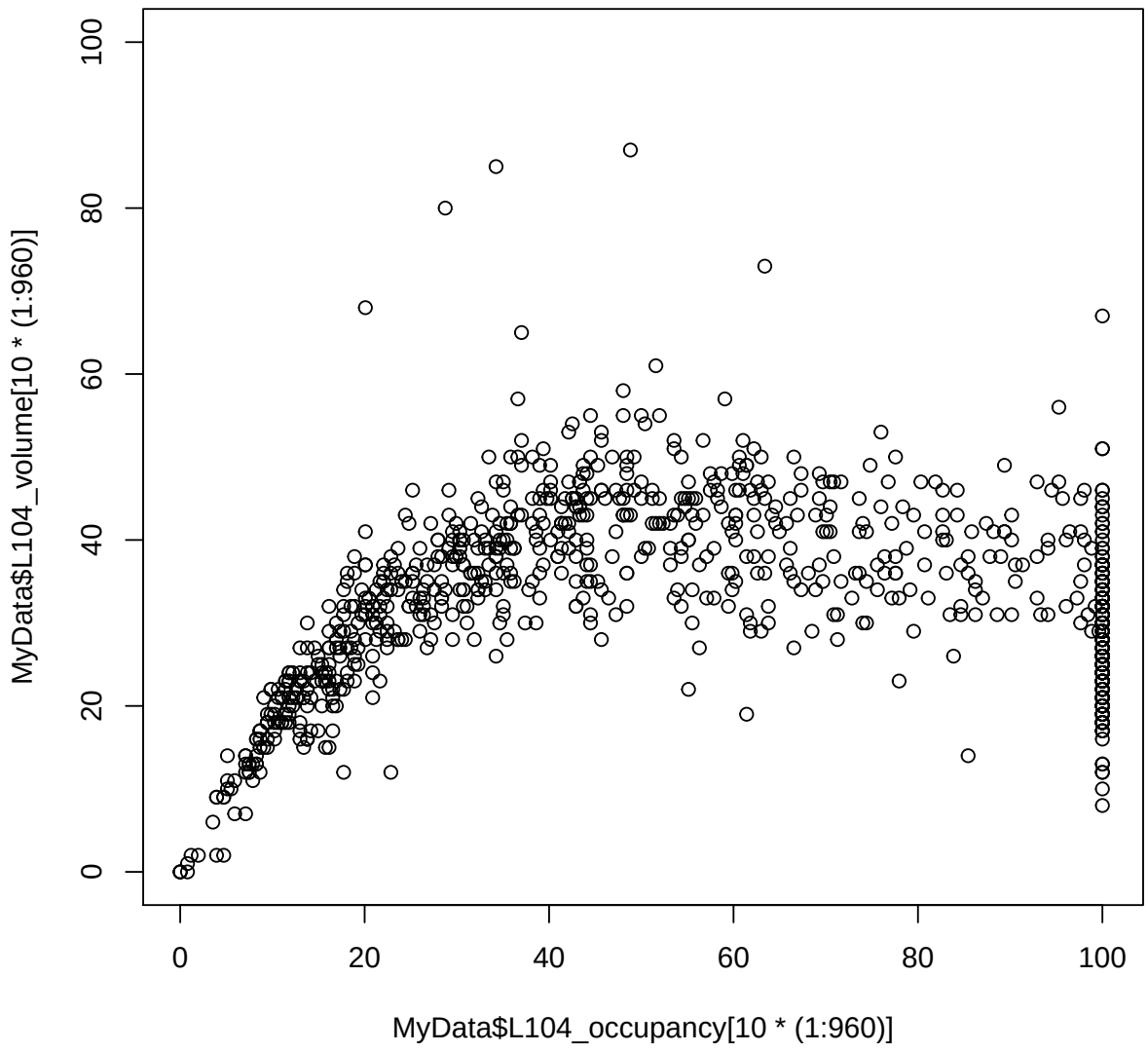

```
plot(MyData$L102_occupancy[3 * (1:960)], MyData$L102_volume[3 * (1:960)],
    main = "Day 3, Loop 2", ylim = c(0, 100))
```



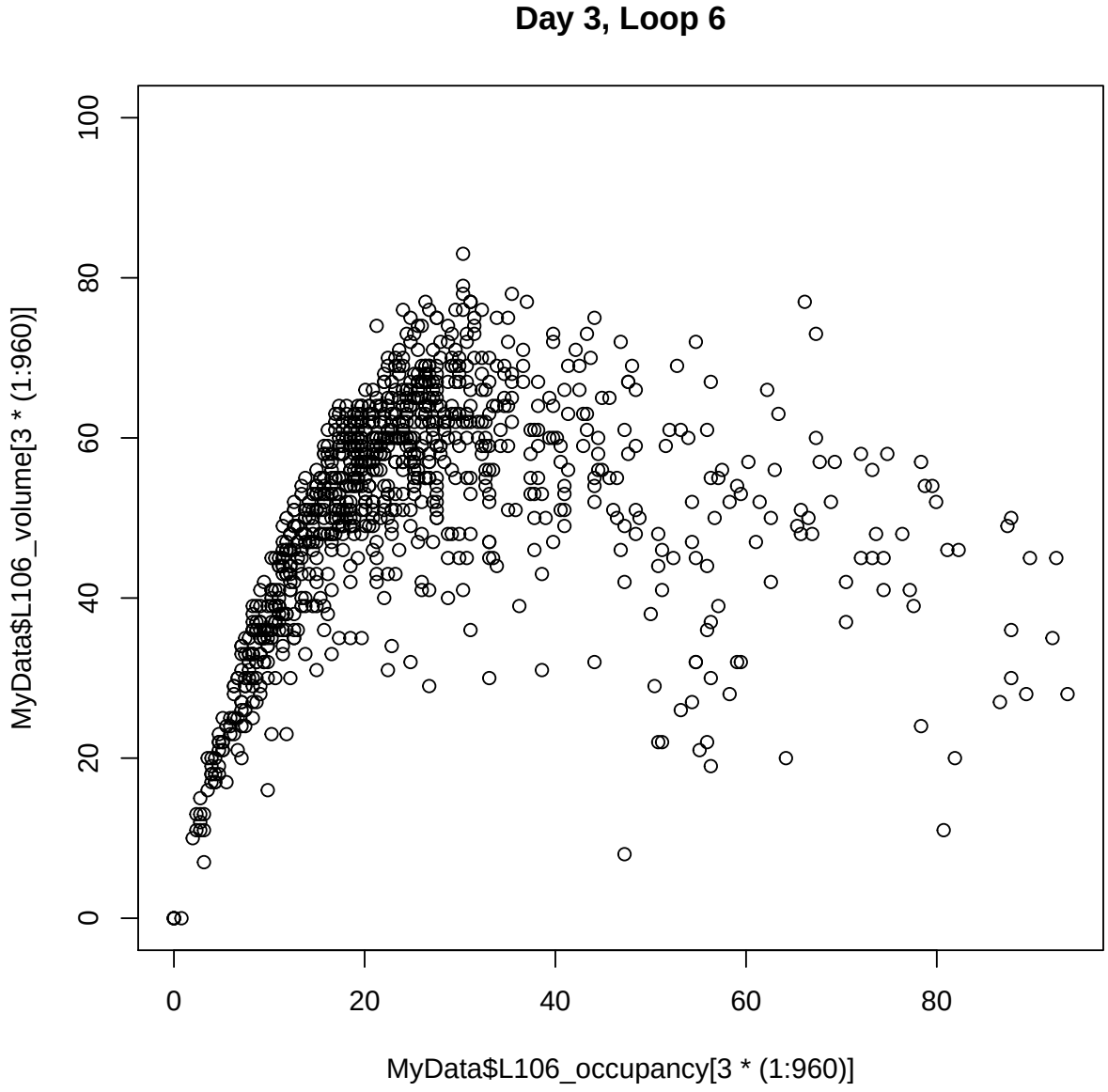

```
plot(MyData$L103_occupancy[3 * (1:960)], MyData$L103_volume[3 * (1:960)],
    main = "Day 3, Loop 3", ylim = c(0, 100))
```



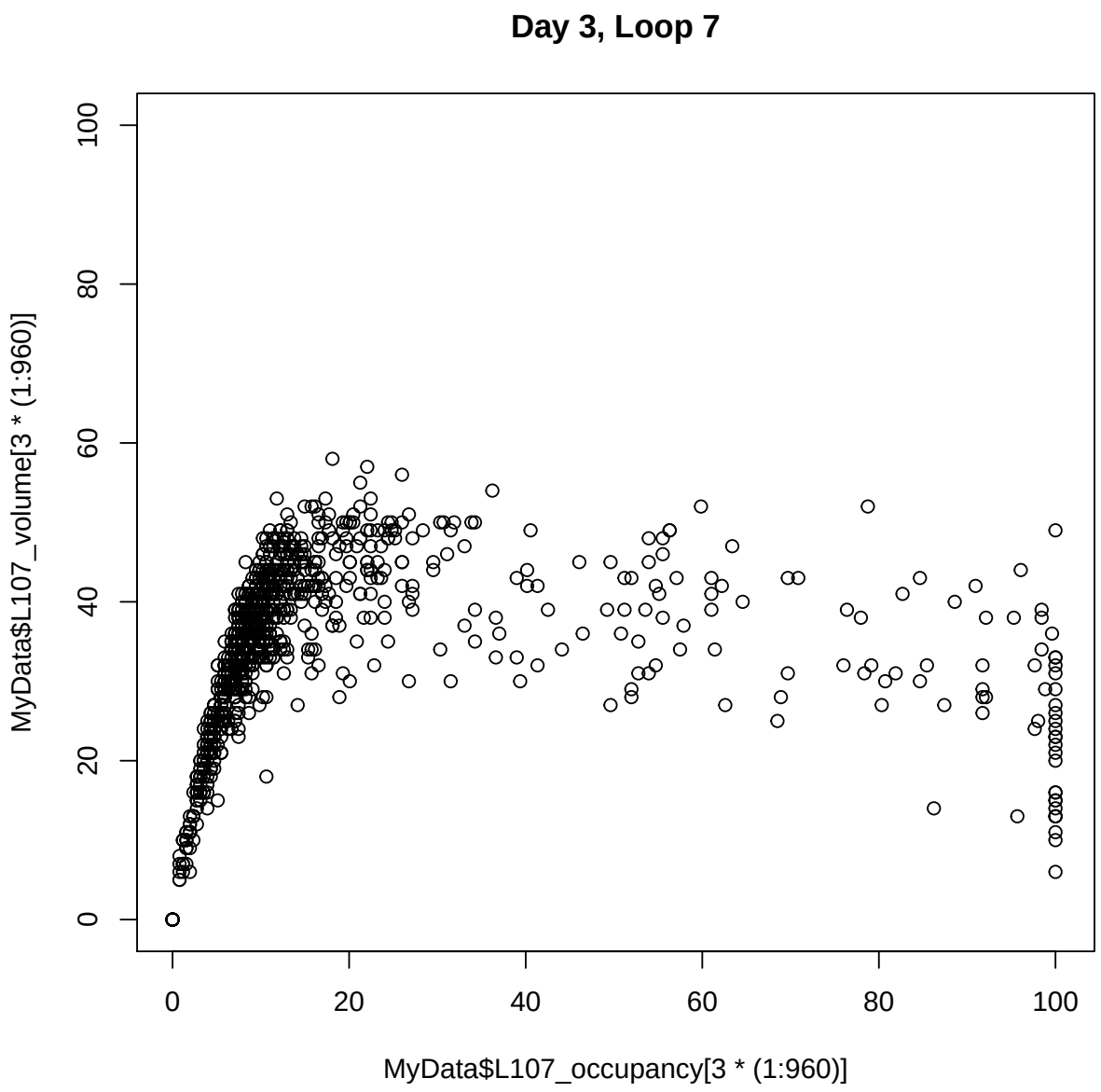

```
plot(MyData$L104_occupancy[10 * (1:960)], MyData$L104_volume[10 *
    (1:960)], main = "Day 3, Loop 4", ylim = c(0, 100))
```



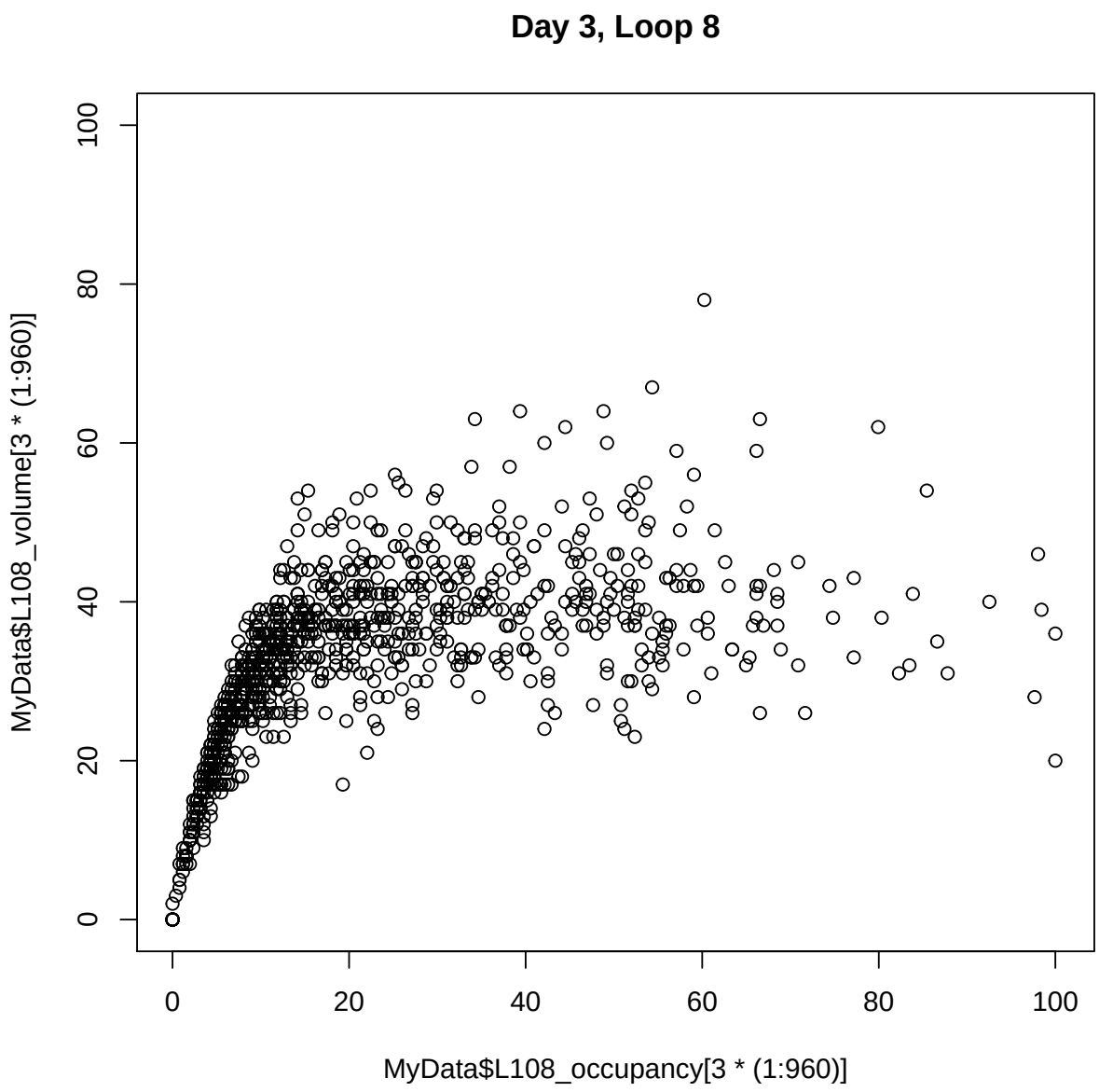

```
plot(MyData$L106_occupancy[3 * (1:960)], MyData$L106_volume[3 * (1:960)],
    main = "Day 3, Loop 6", ylim = c(0, 100))
```




```
plot(MyData$L107_occupancy[3 * (1:960)], MyData$L107_volume[3 * (1:960)],
    main = "Day 3, Loop 7", ylim = c(0, 100))
```




```
plot(MyData$L108_occupancy[3 * (1:960)], MyData$L108_volume[3 * (1:960)],
    main = "Day 3, Loop 8", ylim = c(0, 100))
```



3. Forecasting

4. Conclusion

A word on execution

This project was executed as a training exercise on literate programming using R (R Development Core Team, 2012), `knitr` (Xie, 2012), and LaTeX. The source code is available on GitHub as the `GT_TranspoComp`

project.

References

R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org/`, ISBN 3-900051-07-0, 2012.

Y. Xie, knitr: A general-purpose package for dynamic report generation in R, URL `http://yihui.name/knitr/`, r package version 0.8.5, 2012.