

## APPENDIX O - FURTHER ANALYSIS MULTIPLE REGRESSION PROCESS SCRIPT.R

```
# Further investigation with set of variables produced by the variable trimming
process
# All variables with a collinearity higher than 0.5 were removed from this set, and
the
# stepwise process was repeated starting with that new set.

# Load the final set of variables which was produced by the variable trimming
process, and
# check for any moderately collinear variables. Create a new .CSV in Excel that does
not
# contain them.

library(readr)
MultReg_22_1 <- read_csv("MultReg_22_1.csv")
View(MultReg_22_1)
attach(MultReg_22_1)

corset_f1 <- cor(MultReg_22_1[,2:23])
write.csv((corset_f1), file = "FurtherReg_CollCheck.csv")

# Import new .CSV with set of variables which are not even moderately collinear.

library(readr)
FurtherReg_NoModColl <- read_csv("FurtherReg_NoModColl.csv")

# Begin the new Stepwise Multiple Regression Process with this set of variables.

MultReg_F1 <- FurtherReg_NoModColl
View(MultReg_F1)
fit_f1 <- lm(cyclingrate~., data = MultReg_F1)
summary(fit_f1)

# Analyze the results, and continue to the next iteration of the process, removing
# the least significant variables from the set.

MultReg_F2 <- subset(MultReg_F1, select=-c(fahrradstrasse_dev))
View(MultReg_F2)
fit_f2 <- lm(cyclingrate~., data = MultReg_F2)
summary(fit_f2)

# This process was repeated until all variables in the set were significant, as in
the original stepwise
# multiple regression process.

# Iteration 3

MultReg_F3 <- subset(MultReg_F2, select=-c(med_parkdist))
View(MultReg_F3)
```

## APPENDIX O - FURTHER ANALYSIS MULTIPLE REGRESSION PROCESS SCRIPT.R

```
fit_f3 <- lm(cyclingrate~., data = MultReg_F3)
summary(fit_f3)

# Iteration 4

MultReg_F4 <- subset(MultReg_F3, select=-c(trunkprimary_dev))
View(MultReg_F4)
fit_f4 <- lm(cyclingrate~., data = MultReg_F4)
summary(fit_f4)

# Iteration 5

MultReg_F5 <- subset(MultReg_F4, select=-c(pedway_dev))
View(MultReg_F5)
fit_f5 <- lm(cyclingrate~., data = MultReg_F5)
summary(fit_f5)

# Iteration 6

MultReg_F6 <- subset(MultReg_F5, select=-c())
View(MultReg_F6)
fit_f6 <- lm(cyclingrate~., data = MultReg_F6)
summary(fit_f6)

# Iteration 7

MultReg_F7 <- subset(MultReg_F6, select=-c(sidestreets_dev))
View(MultReg_F7)
fit_f7 <- lm(cyclingrate~., data = MultReg_F7)
summary(fit_f7)

# At this point, the intercept was the least significant term in the equation,
therefore
# it was removed by adding a "-1" to the regression equation.

# Iteration 8

fit_f8 <- lm(cyclingrate~-1+., data = MultReg_F7)
summary(fit_f8)

# Iteration 9

MultReg_F9 <- subset(MultReg_F7, select=-c(emp_density.dev))
View(MultReg_F9)
fit_f9 <- lm(cyclingrate~-1+., data = MultReg_F9)
summary(fit_f9)

# Iteration 10
```

# APPENDIX 0 - FURTHER ANALYSIS MULTIPLE REGRESSION PROCESS SCRIPT.R

```
MultReg_F10 <- subset(MultReg_F9, select=-c(med_sbahndist))
View(MultReg_F10)
fit_f10 <- lm(cyclingrate~-1+., data = MultReg_F10)
summary(fit_f10)

# After the 10th iteration, all of the remaining variables in the set were
significant (p < 0.05).
# Below is the summary of the 10th iteration in comment form.

# Call:
# lm(formula = cyclingrate ~ -1 + ., data = MultReg_F10)

# Residuals:
#   Min       1Q   Median       3Q      Max
# -5.4976 -1.7877  0.1493  1.3636  4.5941

# Coefficients:      Estimate Std. Error t value Pr(>|t|)
# culturalsocial_1km  0.16659    0.05924   2.812   0.0104 *
# ALG_15to65         -0.84964    0.34481  -2.464   0.0225 *
# bldgofdev          20.80216    3.05460   6.810 9.83e-07 ***
# tertiary_dev        3.24859    1.33469   2.434   0.0239 *
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#
# Residual standard error: 2.828 on 21 degrees of freedom
# Multiple R-squared:  0.9664, Adjusted R-squared:  0.96
# F-statistic: 150.9 on 4 and 21 DF, p-value: 3.767e-15
```