

APPENDIX K - MULTIPLE REGRESSION SCRIPT.R

```
# This script is the main script for running multiple regression on the final set of
variables to come
# out of the trimming process.

# Load this "final" set of variables into the initial multiple regression

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

# Rename the data
multreg1 <- MultReg_22_1

# Run the multiple regression of the cyclingrate on the entire set at once.
Summarize the set to
# view its characteristics.

fit1 <- lm(cyclingrate~., data = multreg1)
summary(fit1)

# Examine the results and remove the least significant variable from the data frame.
Variable with
# the highest P-Value is "emp.density.dev" with a value of 0.7867.

# Remove "emp.density.dev" and create a new data frame for the second iteration of
multiple regression.

multreg2 <- subset(multreg1, select=-c(emp_density.dev))
View(multreg2)

# Check to see the correct column was dropped and nothing else was affected, then
run the regression
# on the cycling rate with the new set of variables.

fit2 <- lm(cyclingrate~., data = multreg2)
summary(fit2)

# In the second iteration of the multiple regression process, the variable with the
highest p-value
# (lowest significance) was "med_sbahndist" with a value of 0.78486.

# Remove "med_sbahndist" and create the next data frame for the next regression
iteration.

multreg3 <- subset(multreg2, select=-c(med_sbahndist))
View(multreg3)

# Check to see the correct column was dropped and then run the regression on this
```

APPENDIX K - MULTIPLE REGRESSION SCRIPT.R

set of variables.

```
fit3 <- lm(cyclingrate~., data = multreg3)
summary(fit3)
```

```
# The process of 1) running regressions, 2) removing the least-significant variable
# from the set, and
# then 3) running the regression on the new set of variables, is repeated until the
# only significant
# variables are left in the set.
```

```
# Iteration 4
```

```
multreg4 <- subset(multreg3, select=-c(autobahn_dev))
View(multreg4)
```

```
fit4 <- lm(cyclingrate~., data = multreg4)
summary(fit4)
```

```
# Iteration 5
```

```
multreg5 <- subset(multreg4, select=-c(job_density.dev))
View(multreg5)
```

```
fit5 <- lm(cyclingrate~., data = multreg5)
summary(fit5)
```

```
# Iteration 6
```

```
multreg6 <- subset(multreg5, select=-c(resofdev))
View(multreg6)
```

```
fit6 <- lm(cyclingrate~., data = multreg6)
summary(fit6)
```

```
# Iteration 7
```

```
multreg7 <- subset(multreg6, select=-c(sidestreets_dev))
View(multreg7)
```

```
fit7 <- lm(cyclingrate~., data = multreg7)
summary(fit7)
```

```
# Iteration 8
```

```
multreg8 <- subset(multreg7, select=-c(med_stammsdist))
View(multreg8)
```

APPENDIX K - MULTIPLE REGRESSION SCRIPT.R

```

fit8 <- lm(cyclingrate~., data = multreg8)
summary(fit8)

# Iteration 9

multreg9 <- subset(multreg8, select=-c(cyclespr_dev))
View(multreg9)

fit9 <- lm(cyclingrate~., data = multreg9)
summary(fit9)

# ALL VARIABLES SIGNIFICANT. Write resulting matrix of variables to a .csv file.
write.csv((multreg9), file = "MultipleRegression_ResultVar.csv")

# THE SUMMARY OF fit9

# Call:
#lm(formula = cyclingrate ~ ., data = multreg9)

# Residuals:
#Min      1Q  Median      3Q      Max
#-1.35419 -0.23350 -0.03099  0.35952  0.94073

#Coefficients:
# Estimate Std. Error t value Pr(>|t|)
#(Intercept)      9.256e+01  8.777e+00  10.546 9.74e-07 ***
# culturalsocial_1km  2.814e-01  4.070e-02   6.915 4.12e-05 ***
# ALG_15to65         4.428e-01  1.946e-01   2.276 0.046120 *
# hh.children       -7.379e-01  7.740e-02  -9.534 2.46e-06 ***
# over65            -4.565e-01  7.470e-02  -6.111 0.000114 ***
# bldgofdev         -1.594e+01  4.790e+00  -3.328 0.007649 **
# busstops_1km      -2.868e-01  6.100e-02  -4.701 0.000840 ***
# tertiary_dev       5.180e+00  5.892e-01   8.791 5.11e-06 ***
# sbahn_density.dev  7.805e+00  1.401e+00   5.573 0.000236 ***
# trunkprimary_dev   1.553e-03  3.139e-04   4.949 0.000580 ***
# med_ubahndist      9.786e-04  2.114e-04   4.629 0.000938 ***
# fahrradstrasse_dev 1.788e-04  4.215e-05   4.242 0.001711 **
# infraofdev        -1.282e+02  1.173e+01  -10.928 7.01e-07 ***
# med_parkdist      -7.332e-02  1.128e-02  -6.502 6.88e-05 ***
# pedway_dev        -4.199e-01  6.996e-02  -6.002 0.000132 ***
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#Residual standard error: 0.7591 on 10 degrees of freedom
#Multiple R-squared:  0.9865, Adjusted R-squared:  0.9676

```

APPENDIX K - MULTIPLE REGRESSION SCRIPT.R

#F-statistic: 52.12 on 14 and 10 DF, p-value: 1.949e-07