# Effect of dataset preparation steps (section 4.1.2)

## Data set

> prepStepsData

AmaLgam1 AmaLgam2 BLUiRPlus1 BLUiRPlus2 BLUiR1 BLUiR2

akka.net 0.250 0.124 0.291 0.160 0.295 0.168

AutoMapper 0.458 0.132 0.208 0.099 0.208 0.094

CefSharp 0.359 0.565 0.457 0.442 0.569 0.413

corefx 0.198 0.000 0.217 0.050 0.197 0.050

EntityFramework 0.125 0.044 0.125 0.048 0.112 0.047

gitextensions 0.338 0.306 0.339 0.295 0.339 0.270

Glimpse 0.213 0.213 0.257 0.232 0.261 0.183

Hearthstone-Deck-Tracker 0.259 0.083 0.281 0.101 0.209 0.082

ILSpy 0.515 0.190 0.451 0.173 0.445 0.165

MahApps.Metro 0.747 0.481 0.767 0.493 0.770 0.499

Mvc 0.120 0.128 0.139 0.139 0.103 0.131

Nancy 0.158 0.397 0.176 0.370 0.175 0.293

NLog 0.367 0.164 0.350 0.177 0.350 0.179

OpenRA 0.356 0.165 0.370 0.185 0.370 0.186

orleans 0.196 0.139 0.204 0.113 0.188 0.113

roslyn 0.383 0.083 0.400 0.053 0.392 0.020

ShareX 0.173 0.109 0.147 0.101 0.120 0.103

SparkleShare 0.189 0.404 0.183 0.418 0.170 0.407

VsVim 0.604 0.075 0.596 0.051 0.596 0.050

Wox 0.238 0.313 0.176 0.253 0.176 0.198

## Descriptive statistics

> uniNorm(prepStepsData)

$`Descriptive Statistics`

n Mean Std.Dev Median Min Max 25th 75th Skew Kurtosis

AmaLgam1 20 0.312 0.167 0.254 0.120 0.747 0.194 0.371 0.996 0.195

AmaLgam2 20 0.206 0.155 0.152 0.000 0.565 0.103 0.308 0.848 -0.460

BLUiRPlus1 20 0.307 0.166 0.269 0.125 0.767 0.181 0.377 1.154 0.741

BLUiRPlus2 20 0.198 0.139 0.166 0.048 0.493 0.100 0.264 0.761 -0.761

BLUiR1 20 0.302 0.180 0.235 0.103 0.770 0.176 0.375 1.023 0.156

BLUiR2 20 0.183 0.133 0.166 0.020 0.499 0.091 0.216 0.917 -0.204

$`Shapiro-Wilk's Normality Test`

Variable Statistic p-value Normality

1 AmaLgam1 0.8960 0.0347 NO

2 AmaLgam2 0.8969 0.0360 NO

3 BLUiRPlus1 0.8795 0.0173 NO

4 BLUiRPlus2 0.8855 0.0223 NO

5 BLUiR1 0.8831 0.0201 NO

6 BLUiR2 0.8921 0.0294 NO

## Wilcoxon signed-rank test

> wilcox.test(prepStepsData[, 2], prepStepsData[, 1], paired = TRUE, alternative = "less")

Wilcoxon signed rank test with continuity correction

data: prepStepsData[, 2] and prepStepsData[, 1]

V = 45, p-value = 0.02319

alternative hypothesis: true location shift is less than 0

> wilcox.test(prepStepsData[, 4], prepStepsData[, 3], paired = TRUE, alternative = "less")

Wilcoxon signed rank test with continuity correction

data: prepStepsData[, 4] and prepStepsData[, 3]

V = 35, p-value = 0.008324

alternative hypothesis: true location shift is less than 0

> wilcox.test(prepStepsData[, 6], prepStepsData[, 5], paired = TRUE, alternative = "less")

Wilcoxon signed rank test

data: prepStepsData[, 6] and prepStepsData[, 5]

V = 30, p-value = 0.001827

alternative hypothesis: true location shift is less than 0

# Usage of More Constructs (Section 4.2)

## Data set

> modeData

Default Complete Mixed

akka.net 0.124 0.099 0.106

AutoMapper 0.132 0.194 0.149

CefSharp 0.565 0.573 0.503

corefx 0.000 0.166 0.000

EntityFramework 0.044 0.059 0.067

gitextensions 0.393 0.431 0.402

Glimpse 0.213 0.202 0.185

Hearthstone-Deck-Tracker 0.083 0.109 0.104

ILSpy 0.190 0.213 0.203

MahApps.Metro 0.481 0.573 0.506

Mvc 0.128 0.153 0.130

Nancy 0.397 0.385 0.304

NLog 0.164 0.257 0.176

OpenRA 0.165 0.145 0.143

orleans 0.139 0.167 0.266

roslyn 0.084 0.249 0.247

ShareX 0.109 0.144 0.146

SparkleShare 0.404 0.441 0.430

VsVim 0.075 0.055 0.058

Wox 0.313 0.349 0.396

## Descriptive statistics

> library(MVN)

> uniNorm(modeData)

$`Descriptive Statistics`

n Mean Std.Dev Median Min Max 25th 75th Skew Kurtosis

Default 20 0.210 0.159 0.152 0.000 0.565 0.103 0.333 0.775 -0.715

Complete 20 0.248 0.158 0.198 0.055 0.573 0.145 0.358 0.769 -0.671

Mixed 20 0.226 0.151 0.180 0.000 0.506 0.124 0.327 0.526 -1.036

$`Shapiro-Wilk's Normality Test`

Variable Statistic p-value Normality

1 Default 0.8896 0.0265 NO

2 Complete 0.8926 0.0301 NO

3 Mixed 0.9246 0.1215 YES

## Wilcoxon signed-rank test

> wilcox.test(modeData[,2], modeData[,1], paired=TRUE, alternative="greater")

Wilcoxon signed rank test with continuity correction

data: modeData[, 2] and modeData[, 1]

V = 185.5, p-value = 0.001409

alternative hypothesis: true location shift is greater than 0

> wilcox.test(modeData[,3], modeData[,1], paired=TRUE, alternative="greater")

Wilcoxon signed rank test with continuity correction

data: modeData[, 3] and modeData[, 1]

V = 123, p-value = 0.1342

alternative hypothesis: true location shift is greater than 0