Decomposition Heuristics Experiment

Variables NoofSubprocesses, Size_AvgAct, AvgNodes, AvgCohesion, AvgCNC, AvgDensity

Testing the difference in means of variables "NoofSubprocesses", "Size_AvgAct", "AvgNodes", "AvgCohesion", "AvgCNC", "AvgDensity by *Method* groups.

Results are presented in the following table:

	shapiro.BP	shapiro.DO	Ftest	Test	P.Value	mean.BP	mean.DO
NoofSubprocesses	0.99	0.454	0.035	Welch Two Sample t-test	0.502	16.73	15.14
Size_AvgAct	0	0.001	**	Mann-Whitney test	0.809	6.48	5.34
AvgNodes	0	0.008	**	Mann-Whitney test	0.687	12.46	10.08
AvgCohesion	0.953	0.323	0.374	Two Sample t-test	0.004	0.26	0.38
AvgCNC	0.124	0.004	**	Mann-Whitney test	0.309	1.01	0.96
AvgDensity	0.859	0.006	**	Mann-Whitney test	0.699	0.15	0.16

Figures and outputs of tests

No of Subprocesses

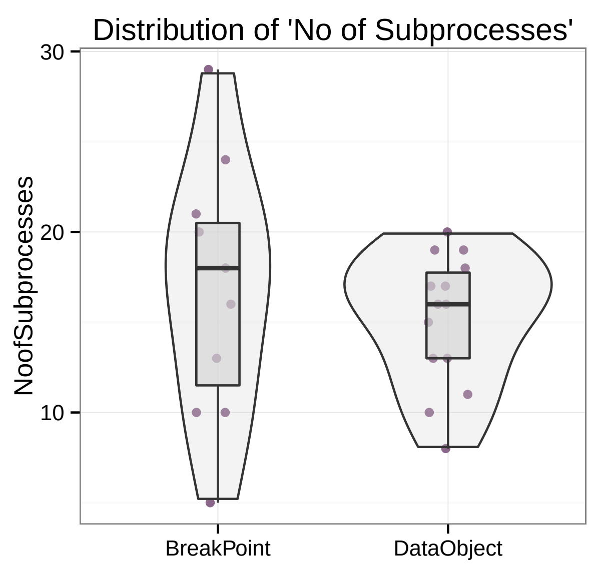

Figure 1: Violoin plot and boxplot of No of Subprocesses

```
##
##  Shapiro-Wilk normality test
##
## data:  data$NoofSubprocesses[data$Method == "BreakPoint"]
## W = 0.9857, p-value = 0.9896
##
##  Shapiro-Wilk normality test
##
## data:  data$NoofSubprocesses[data$Method == "DataObject"]
## W = 0.9427, p-value = 0.4539
##
##  F test to compare two variances
##
## data:  NoofSubprocesses by Method
## F = 3.5785, num df = 10, denom df = 13, p-value = 0.03461
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##   1.101189 12.822444
## sample estimates:
## ratio of variances
##           3.578499
##
##  Welch Two Sample t-test
##
## data:  NoofSubprocesses by Method
## t = 0.6881, df = 14.342, p-value = 0.5024
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -3.342980  6.511811
## sample estimates:
## mean in group BreakPoint mean in group DataObject
##                 16.72727                 15.14286
```

Size AvgAct

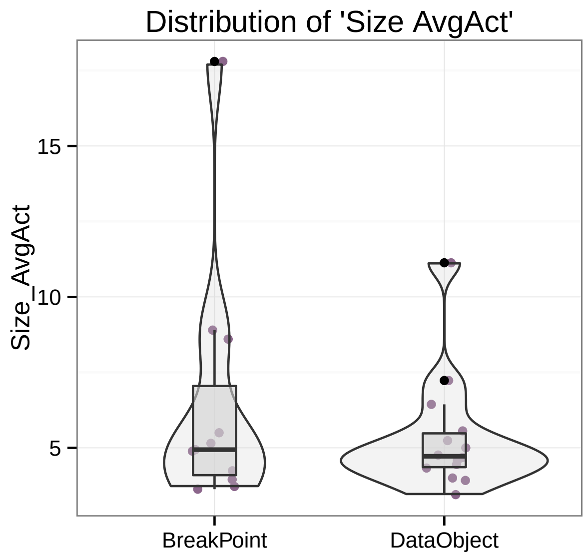

Figure 2: Violoin plot and boxplot of Size AvgAct

```
##
##  Shapiro-Wilk normality test
##
## data:  data$Size_AvgAct[data$Method == "BreakPoint"]
## W = 0.6859, p-value = 0.0003089
##
##  Shapiro-Wilk normality test
##
## data:  data$Size_AvgAct[data$Method == "DataObject"]
## W = 0.7547, p-value = 0.001448
##
##  Wilcoxon rank sum test
##
## data:  Size_AvgAct by Method
## W = 82, p-value = 0.8089
## alternative hypothesis: true location shift is not equal to 0
```

Avg Nodes

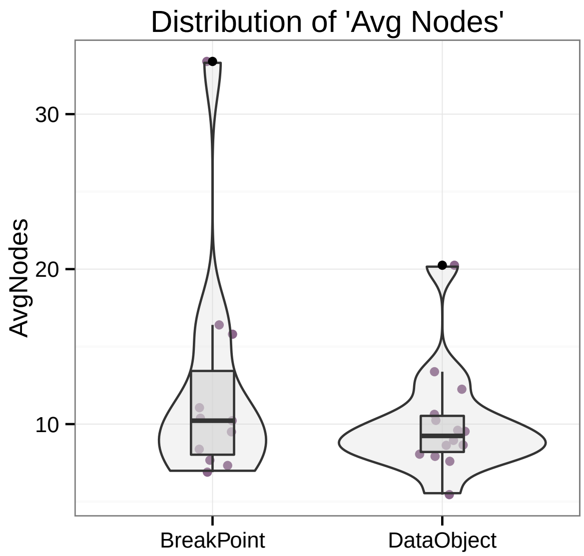

Figure 3: Violoin plot and boxplot of Avg Nodes

```
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgNodes[data$Method == "BreakPoint"]
## W = 0.6962, p-value = 0.0004181
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgNodes[data$Method == "DataObject"]
## W = 0.818, p-value = 0.008408
##
##  Wilcoxon rank sum test
##
## data:  AvgNodes by Method
## W = 85, p-value = 0.6867
## alternative hypothesis: true location shift is not equal to 0
```

Avg Cohesion

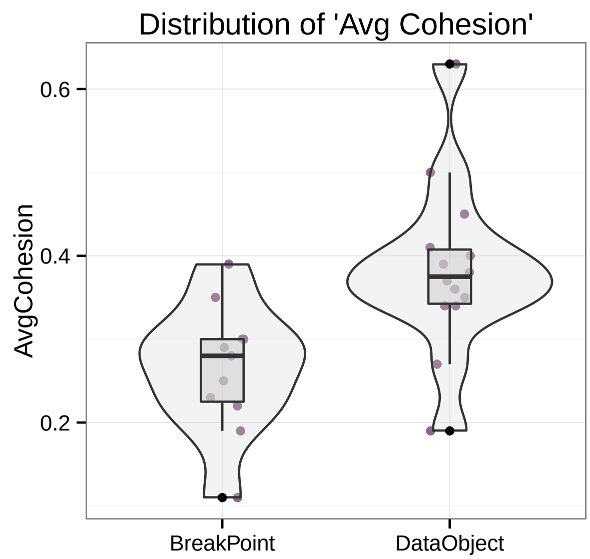

Figure 4: Violoin plot and boxplot of Avg Cohesion

```
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgCohesion[data$Method == "BreakPoint"]
## W = 0.9779, p-value = 0.953
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgCohesion[data$Method == "DataObject"]
## W = 0.9318, p-value = 0.323
##
##  F test to compare two variances
##
## data:  AvgCohesion by Method
## F = 0.5677, num df = 10, denom df = 13, p-value = 0.3744
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1746796 2.0340013
## sample estimates:
## ratio of variances
##          0.5676508
##
##  Two Sample t-test
##
## data:  AvgCohesion by Method
## t = -3.2275, df = 23, p-value = 0.003726
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  -0.19648860 -0.04299192
## sample estimates:
## mean in group BreakPoint mean in group DataObject
##                0.2645455                0.3842857
```

Avg CNC

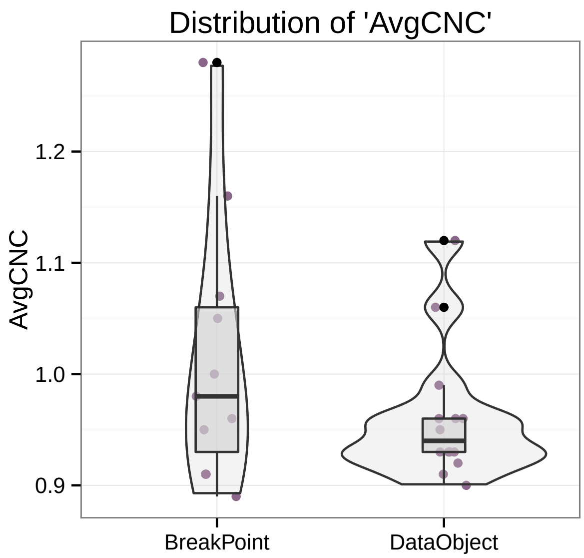

Figure 5: Violoin plot and boxplot of AvgCNC

```
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgCNC[data$Method == "BreakPoint"]
## W = 0.8861, p-value = 0.1244
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgCNC[data$Method == "DataObject"]
## W = 0.7882, p-value = 0.003585

## Warning in wilcox.test.default(x = c(1.28, 0.98, 0.89, 0.91, 0.95, 1, 1.16, :  cannot compute
exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  AvgCNC by Method
## W = 96, p-value = 0.3088
## alternative hypothesis: true location shift is not equal to 0
```

Avg Density

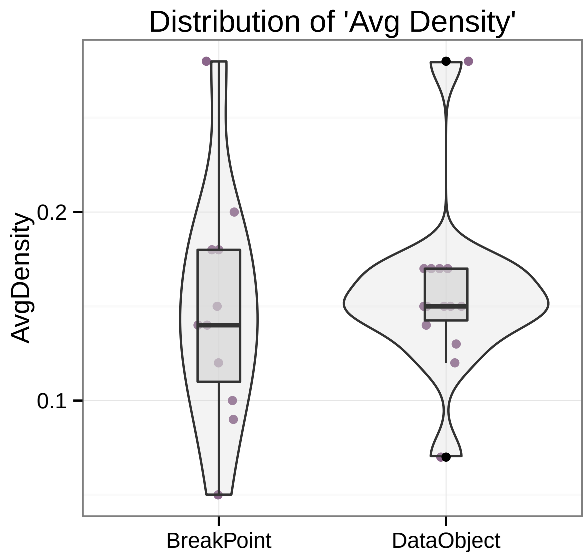

Figure 6: Violoin plot and boxplot of Avg Density

```
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgDensity[data$Method == "BreakPoint"]
## W = 0.9674, p-value = 0.8591
##
##  Shapiro-Wilk normality test
##
## data:  data$AvgDensity[data$Method == "DataObject"]
## W = 0.8054, p-value = 0.005826

## Warning in wilcox.test.default(x = c(0.28, 0.15, 0.18, 0.14, 0.14, 0.2, :  cannot compute
exact p-value with ties

##
##  Wilcoxon rank sum test with continuity correction
##
## data:  AvgDensity by Method
## W = 69.5, p-value = 0.6987
## alternative hypothesis: true location shift is not equal to 0
```

Variables A1-A6, A10, B5-B6

Testing the difference in means of variables "A1_Years", "A2_Number", "A3_Number", "A4_Number", "A5_NoofDays", "A6_NoofDays", "A10_Years", "B5_Hours", "B6_Hours" by *Method* groups.

Results are presented in the following table:

	shapiro.BP	shapiro.DO	Ftest	Test	Pval	mean.BP	mean.DO
A1_Years	0	0	**	Mann-Whitney test	0.679	0.65	0.43
A2_Number	0.116	0.135	0.016	Welch Two Sample t-test	0.678	11.27	12.57
A3_Number	0	0.041	**	Mann-Whitney test	0.164	7.64	7.43
A4_Number	0.011	0.272	**	Mann-Whitney test	0.738	18.64	18.79
A5_NoofDays	0.019	0.001	**	Mann-Whitney test	0.67	3.36	2.67
A6_NoofDays	0.013	0.011	**	Mann-Whitney test	0.598	5.27	3.98
A10_Years	0	0	**	Mann-Whitney test	0.929	0.62	0.28
B5_Hours	0.114	0.062	0.368	Two Sample t-test	0.361	2.05	1.61
B6_Hours	0.129	0.024	**	Mann-Whitney test	0.5	1.18	1.54

Variables A7_Scale_1_7 - A9_Scale_1_7, B1_Scale_1_5 - B4_Scale_1_5

Conducting Fisher's exact test for testing the null of independence of rows (*factor "Methods" levels*) and columns (*variable levels*) in a contingency table with fixed marginals.

Results are presented in the following table:

	Test	P.Value
A7_Scale_1_7	Fisher's Exact Test	0.823
A8_Scale1_7	Fisher's Exact Test	0.730
A9_Scale1_7	Fisher's Exact Test	0.508
B1_Scale_1_5	Fisher's Exact Test	0.584
B2_Scale_1_5	Fisher's Exact Test	0.601
B3_Scale_1_5	Fisher's Exact Test	0.188
B4_Scale_1_5	Fisher's Exact Test	0.366

Contingency tables and test's output for every variable:

A7_Scale_1_7

```
##
##              2 3 4 5 6
##   BreakPoint 3 3 2 2 1
##   DataObject 2 5 4 3 0
##
##                       2          3          4          5          6
##   BreakPoint 0.27272727 0.27272727 0.18181818 0.18181818 0.09090909
##   DataObject 0.14285714 0.35714286 0.28571429 0.21428571 0.00000000
##
##  Fisher's Exact Test for Count Data
##
## data:  data$Method and data$A7_Scale_1_7
## p-value = 0.8225
## alternative hypothesis: two.sided
```

A8_Scale1_7

```
##
##              2 3 4 5 6
##   BreakPoint 4 3 3 0 1
##   DataObject 4 5 3 2 0
##
##                       2          3          4          5          6
##   BreakPoint 0.36363636 0.27272727 0.27272727 0.00000000 0.09090909
##   DataObject 0.28571429 0.35714286 0.21428571 0.14285714 0.00000000
##
##  Fisher's Exact Test for Count Data
##
## data:  data$Method and data$A8_Scale1_7
```

```
## p-value = 0.7305
## alternative hypothesis: two.sided
```

A9_Scale1_7

```
##
##              2 3 4 5 6
##   BreakPoint 2 4 2 1 2
##   DataObject 1 6 4 3 0
##
##                       2          3          4          5          6
##   BreakPoint 0.18181818 0.36363636 0.18181818 0.09090909 0.18181818
##   DataObject 0.07142857 0.42857143 0.28571429 0.21428571 0.00000000
##
##  Fisher's Exact Test for Count Data
##
## data:  data$Method and data$A9_Scale1_7
## p-value = 0.5085
## alternative hypothesis: two.sided
```

B1_Scale_1_5

```
##
##              2 3 4
##   BreakPoint 2 4 5
##   DataObject 4 7 3
##
##                      2         3         4
##   BreakPoint 0.1818182 0.3636364 0.4545455
##   DataObject 0.2857143 0.5000000 0.2142857
##
##  Fisher's Exact Test for Count Data
##
## data:  data$Method and data$B1_Scale_1_5
## p-value = 0.5844
## alternative hypothesis: two.sided
```

B2_Scale_1_5

```
##
##               2  3  4
##   BreakPoint  3  6  2
##   DataObject  3 10  1
##
##                       2          3          4
##   BreakPoint 0.27272727 0.54545455 0.18181818
##   DataObject 0.21428571 0.71428571 0.07142857
```

```
##
##  Fisher's Exact Test for Count Data
##
## data:  data$Method and data$B2_Scale_1_5
## p-value = 0.6006
## alternative hypothesis: two.sided
```

B3_Scale_1_5

```
##
##              2 3 4
##   BreakPoint 3 6 2
##   DataObject 2 4 8
##
##                      2         3         4
##   BreakPoint 0.2727273 0.5454545 0.1818182
##   DataObject 0.1428571 0.2857143 0.5714286
##
##  Fisher's Exact Test for Count Data
##
## data:  data$Method and data$B3_Scale_1_5
## p-value = 0.1882
## alternative hypothesis: two.sided
```

B4_Scale_1_5

```
##
##              2 3 4
##   BreakPoint 3 5 3
##   DataObject 4 3 7
##
##                      2         3         4
##   BreakPoint 0.2727273 0.4545455 0.2727273
##   DataObject 0.2857143 0.2142857 0.5000000
##
##  Fisher's Exact Test for Count Data
##
## data:  data$Method and data$B4_Scale_1_5
## p-value = 0.3661
## alternative hypothesis: two.sided
```