

My title

Firstname Lastname

optional@email.com
My optional Institute
My optional University

Abstract Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua.

Keywords: Lorem, Ipsum

1 Story Line

- Test Time Generation - Directed Random is faster in regards to AVM and AVM-Defaults.
- Coverage - As Directed Random is faster it has the same coverage as the rest of the techniques.
- Mutation Score - Directed Random has the higher or similar mutation score comparing to AVM and AVM-Defaults.

2 RQs

2.1 Test Time Generation

How efficient Directed Random in regard of test time generation comparing to the other data generators?

2.2 Coverage

How do the number of test requirements generated by the coverage criteria differ depending on the data generation technique being used, and how successfully can test cases be automatically generated to satisfy them?

2.3 Fault-Finding Effectiveness

How does fault finding effectiveness vary depending on data generation technique used?

3 What Plots ?

3.1 Test time generation (Efficiency)

3.1.1 Analyse Wilcox Rank and Effect Size for AVM and Directed Random

To statistically analyze the the new technique we conducted tests for significance with the nonparametric Wilcoxon rank-sum test, using the sets of 30 execution times obtained with a specific DBMS and all techniques **Hitchhicker Guide Ref.** A p-value that less than 0.05 is considered significant. To review practical are significance tests we use the nonparametric A12 statistic of Vargha and Delaney **REF** was used to calculate effect sizes. The A12 determine the average probability that one approach beats another, or how superior one technique compared to the other. We followed the guidelines of Vargha and Delaney in that an effect size is considered to be “large” if the value of A12 is < 0.29 or > 0.71 , “medium” if A12 is < 0.36 or > 0.64 and “small” if A12 is < 0.44 or > 0.56 . However, is the values of A12 close to the 0.5 value are viewed as there no effect.

When comparing AVM and Directed Random techniques in regard of time we used Mann-Whitney U-test and the A-hat effect size calculations. As Shown in in the following two tables that there is statistically significant difference between Directed Random and AVM, $p \leq 0.05$. Therefore, we reject the null hypothesis that there is no difference between AVM and Directed Random. As p-value near zero, that directed random is faster than AVM in a statistically significant test. Moreover, the A-12 shows that Directed Random has a large effect size when it comes to test generation timing. Which means that Directed Random is the winner in regard of test generation timing.

p.value	dbms
0	Postgres
0	SQLite
0	HyperSQL

Figure 1 and 2.

3.2 Coverage

Figure 3 and 4.

3.3 Mutation Score

Figure 5, 6 and 7.

- Directed Random has a higher PKColumnA operator kills than AVM-Defaults.
- Primary Key Composite key can be problematic for AV<-Defaults.
- Directed Random has either similar or less (because of coverage) mutation kill compared to AVM.
- AVM has lower coverage than AVM-Defaults however it has a higher mutation score because of empty strings and Composite keys.

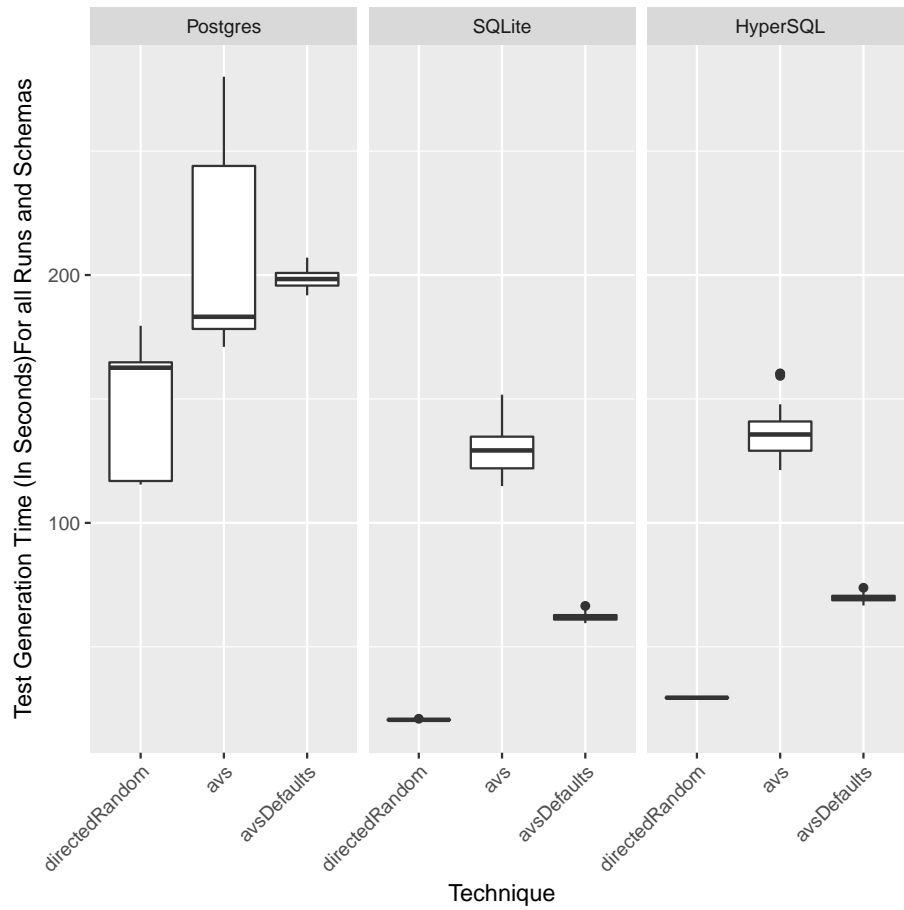


Figure 1. Test generation timing - in seconds. First summing test generation timing for each run then group by data generator and DBMS, then divide test generation timing by 1000 to convert to second.

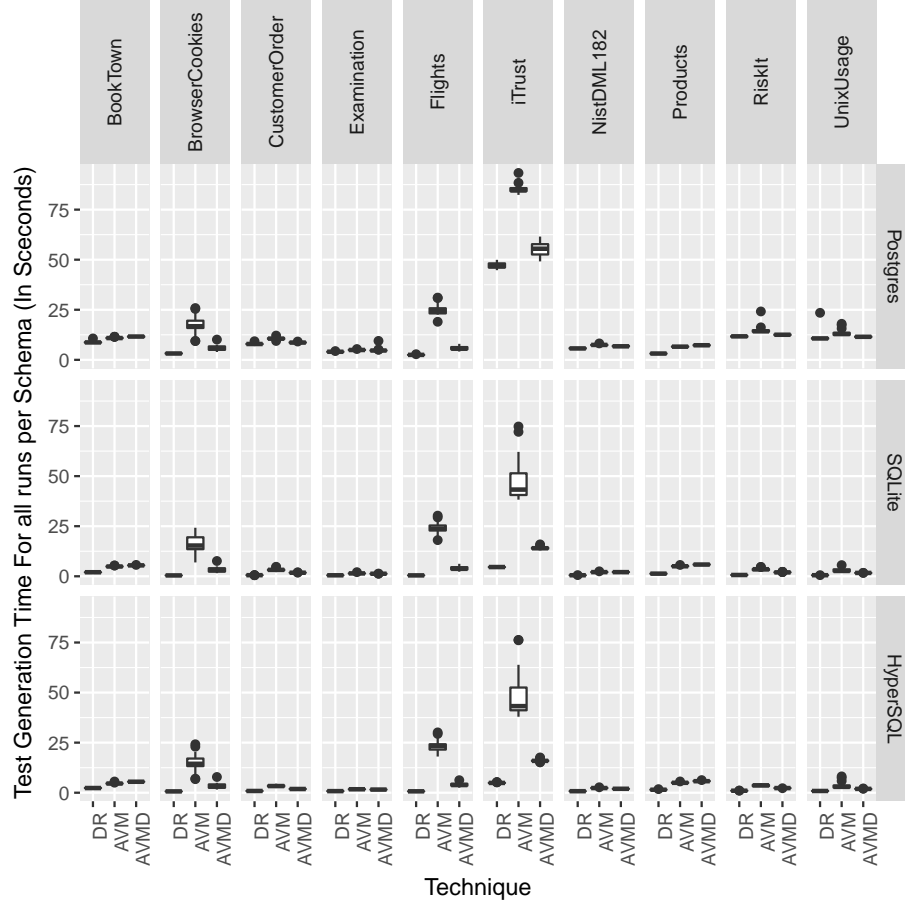


Figure 2. Test generation timing box plot showing some of Game Changer schemas

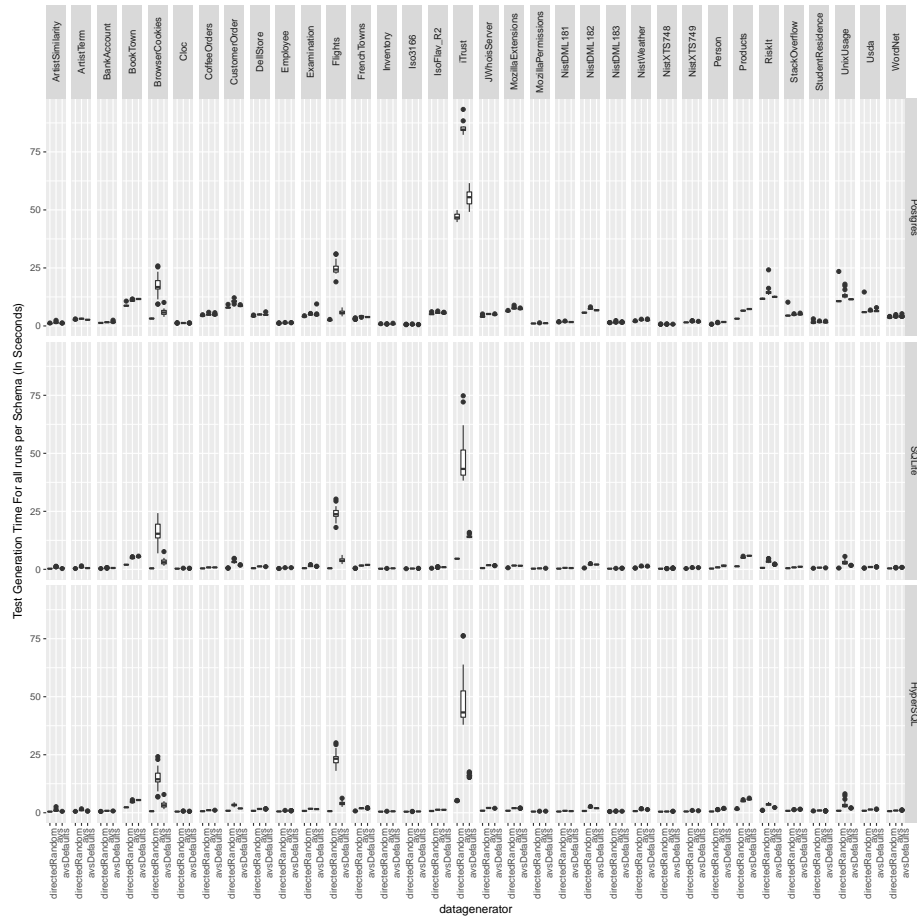


Figure 3. Box plot for Test Generation time for DBMS, techniques and schemas - in percentage. Group by data generator, case study and DBMS, then dividing test generation timing by 1000 to convert to second. No averaging or summing to see the spread of values for all runs per schema

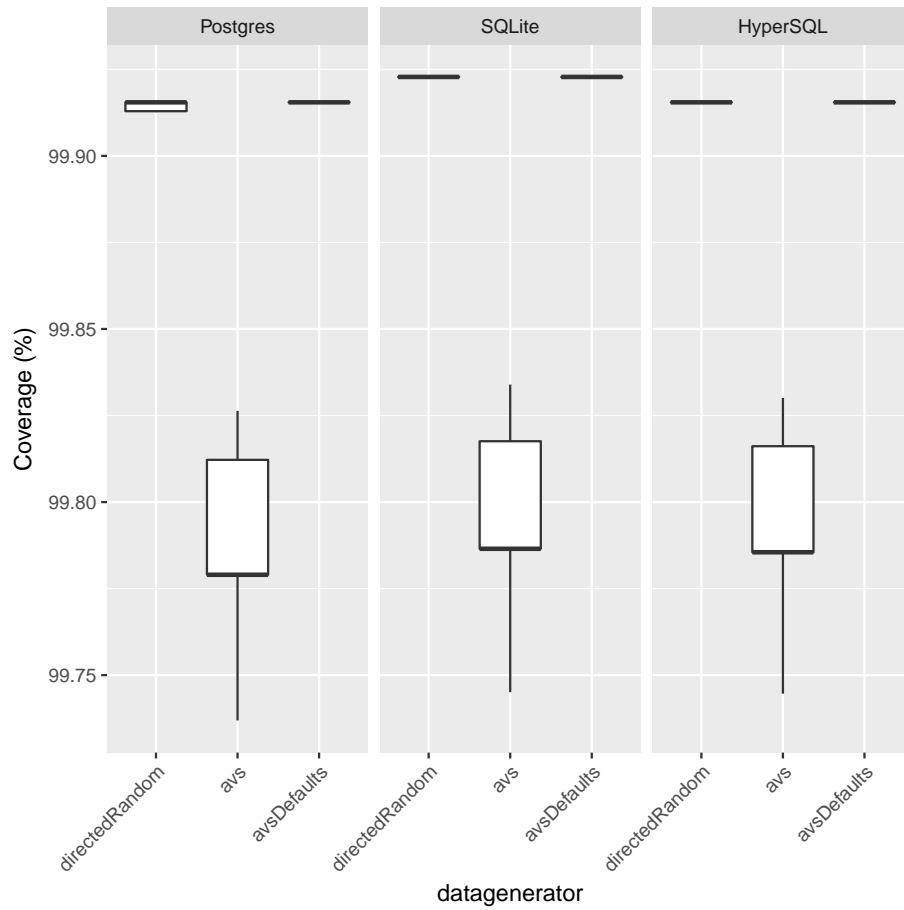


Figure 4. Box plot for Coverage, averaging all schema coverage per run

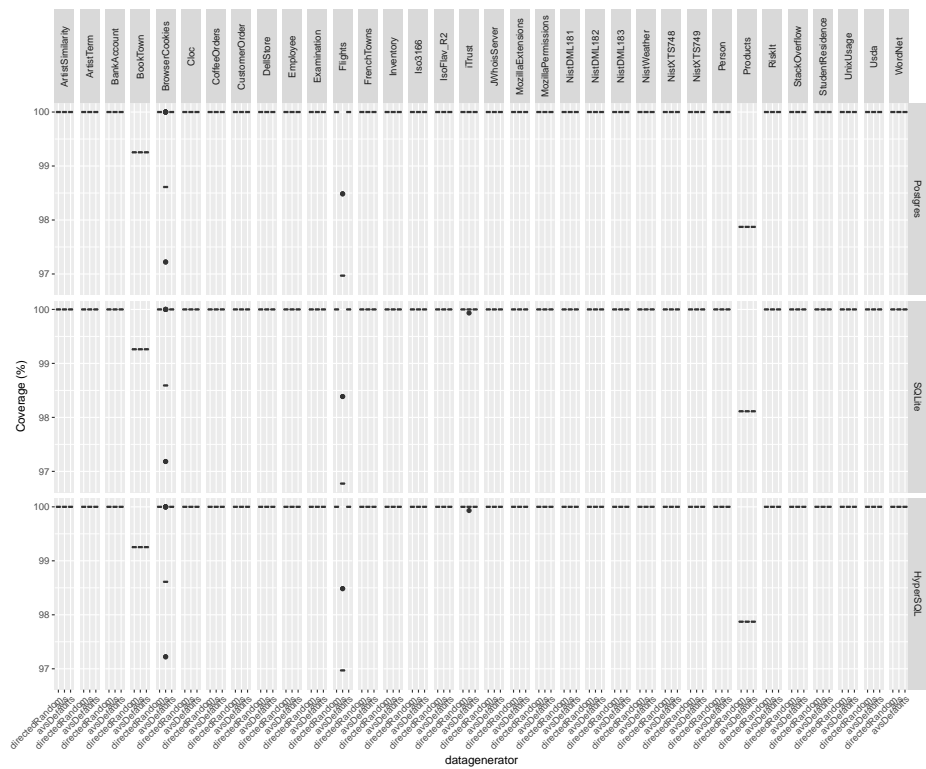


Figure 5. Box plot for Coverage

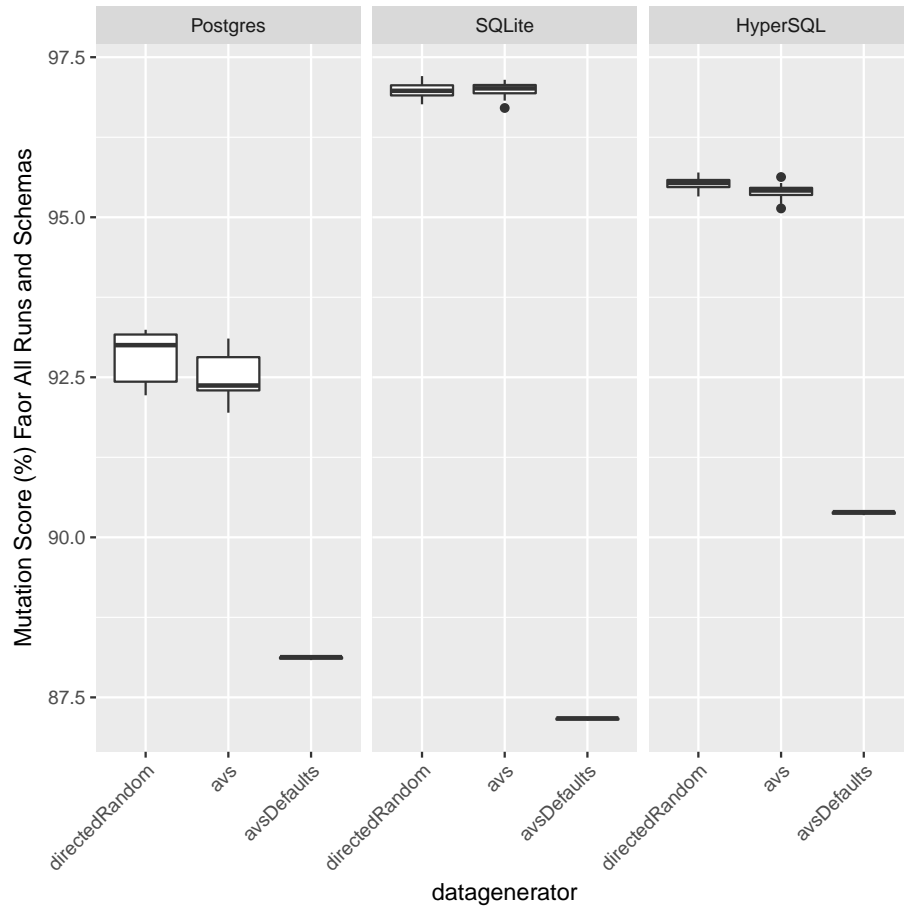


Figure 6. Box plot for mutation scores for DBMSs and techniques - in percentage. First I sum all score numerator and denominator per run per DBMS per data generator, then plotting using group by data generator, DBMS and dividing the score numerator by denominator multiplying by 100

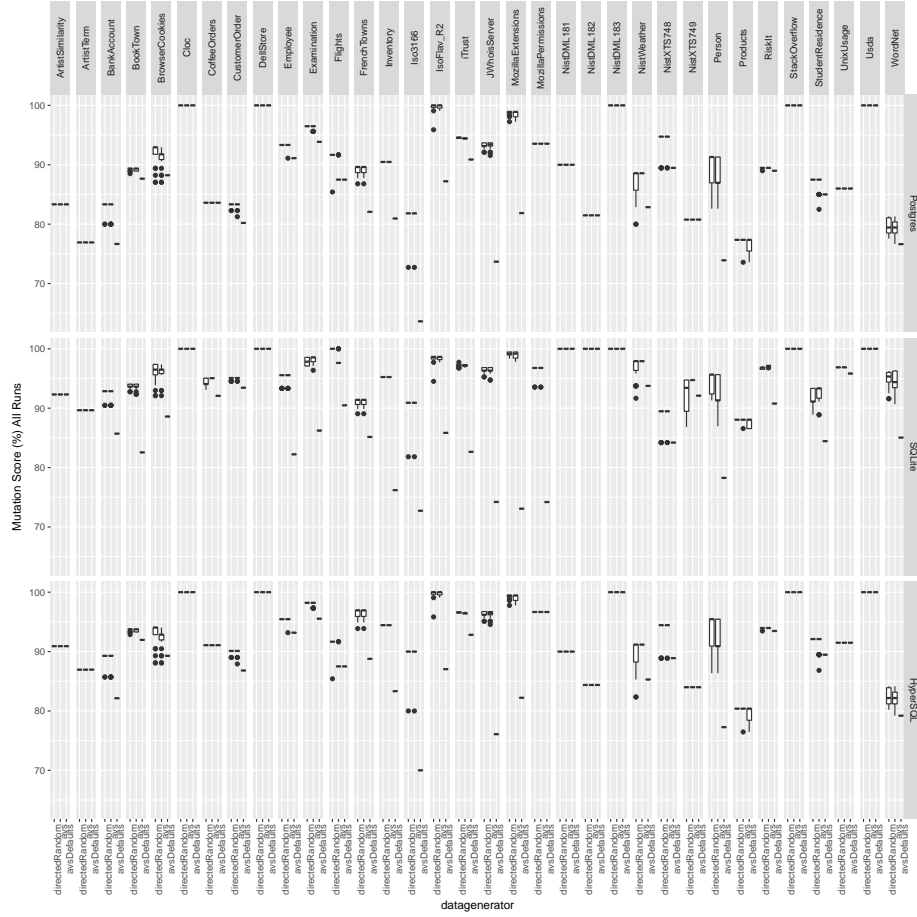


Figure 7. Box plot for mutation scores for DBMS, techniques and schemas - in percentage. Group by data generator, case study and DBMS, then dividing the score numerator by denominator multiplying by 100. No averaging or summing to see the spread of values for all runs per schema

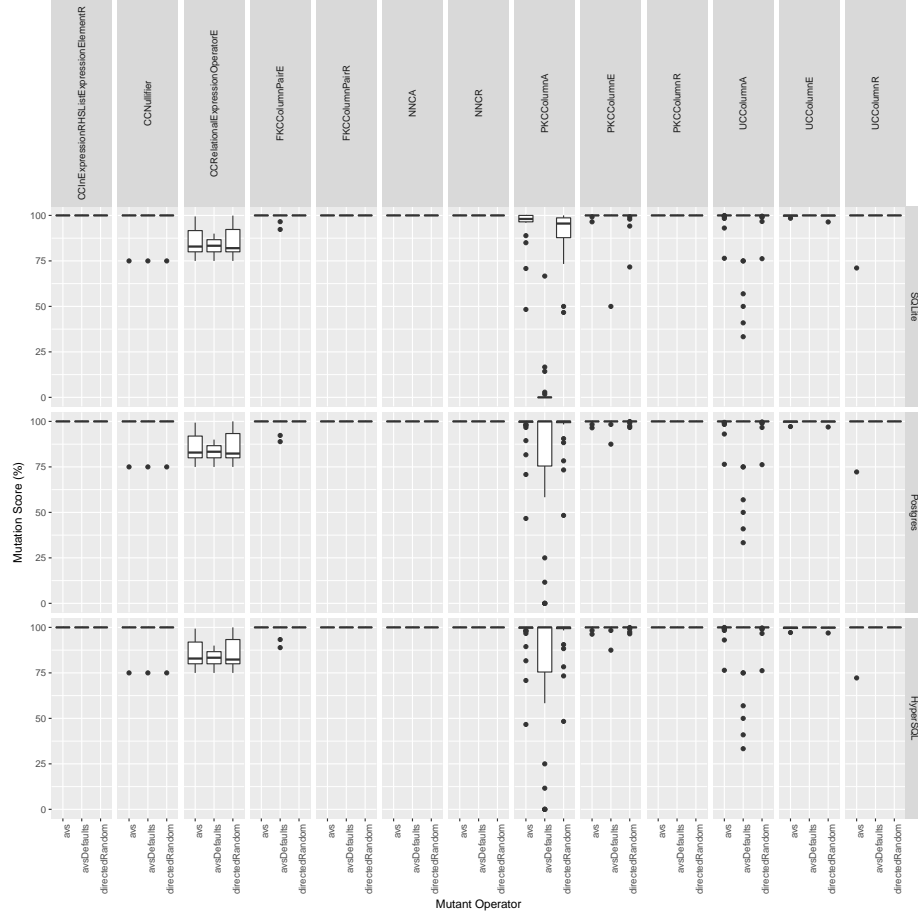


Figure 8. Box Plot of mutant scores in regard of Mutant Operators and DBMSs - in percentage. Using mutanttiming file, I group by generator, DBMS, schema and operator. Then I only select NORMAL mutants, then calculating the percentage of killed mutant by summing all killed divided by (killed mutant plus alive mutant) multiply by 100.

	Postgres			SQLite			HyperSQL			ArtistSimilarity
	DR	AVM	AVMD	DR	AVM	AVMD	DR	AVM	AVMD height	
Case Study	83.33	83.33	***68.33	92.31	92.31	92.31	***64.62	90.91		90.91
90.91	***74.55									
ArtistTerm	76.92	76.92	76.92	***61.79	89.66	89.66	89.66	***62.53		86.96
86.96	86.96	***69.86								
BankAccount	83.11	***82.78	***76.67	***64.56	92.70	***92.46	***85.71	***68.33		89.05
***88.69	***82.14	***69.17								
BookTown	89.30	***89.22	***87.66	***80.51	93.70	***93.63	***82.55	***80.89		93.69
***93.6	***91.96	***84.46								
BrowserCookies	91.92	***91.22	***88.24	***51.18	95.94	***96.05	***88.6	***55.79		93.02
***92.3	***89.29	***51.79								
Cloc	100.00	100	100	***89.78	100.00	100	100	***79.56		100.00
100	100	***89.78								
CoffeeOrders	83.61	83.61	83.61	***38.52	94.29	***95.05	***92.08	***66.63		91.07
91.07	91.07	***41.96								
CustomerOrder	83.26	***83.19	***80.21	***31.87	95.01	***95.05	***93.44	***59.95		90.04
***89.96	***86.81	***34.03								
DellStore	100.00	100	100	***87.12	100.00	100	100	***85.17		100.00
100	100	***87.12								
Employee	93.33	***93.26	***91.11	***85.04	95.11	***95.26	***82.22	***86.44		95.45
***95.38	***93.18	***86.97								
Examination	96.49	***96.32	***93.86	***85.91	97.90	***98.19	***86.23	***78.94		98.21
***98.04	***95.54	***87.44								
Flights	91.46	***87.92	***87.5	***40.76	100.00	***97.86	***90.48	***66.55		91.46
***87.92	***87.5	***40.76								
FrenchTowns	89.09	89.09	***82.08	***21.29	90.96	90.96	***85.16	***32.6		96.36
96.36	***88.78	***23.03								
Inventory	90.48	90.48	***80.95	***74.6	95.24	95.24	***76.19	***79.37		94.44
94.44	***83.33	***79.63								
Iso3166	81.52	81.52	***63.64	***63.33	90.61	90.61	***72.73	***61.52		89.67
89.67	***70	***69.67								
IsoFlav_R2	99.74	***99.85	***87.21	***56.88	98.37	***98.48	***85.84	***49.33		99.74
***99.85	***87.04	***56.28								
iTrust	94.55	***94.44	***90.88	***85.32	97.21	***97.19	***82.65	***85.54		96.57
***96.47	***92.82	***87.14								
JWhoisServer	93.32	***93.26	***73.68	***82.7	96.42	***96.33	***74.21	***85.86		96.36
***96.3	***76.09	***85.4								
MozillaExtensions	98.75	***98.52	***81.87	***59.91	99.15	***98.97	***73.08	***59.74		99.30
***99.06	***82.22	***60.21								
MozillaPermissions	93.55	93.55	93.55	***89.46	96.56	96.56	***74.19	***92.69		96.67
96.67	96.67	***92.44								
NistDML181	90.00	90	90	***64.67	100.00	100	100	***73.23		90.00
90	90	***64.67								
NistDML182	81.48	81.48	81.48	***60.49	100.00	100	100	***89.93		84.38
84.38	84.38	***65.97								
NistDML183	100.00	100	100	100	100.00	100	100	100		100.00
100	100	100								
NistWeather	87.05	***88.57	***82.86	***43.62	96.88	***97.92	***93.75	***72.71		89.61
***91.18	***85.29	***44.9								
NistXTS748	94.04	***94.39	***89.47	***94.56	88.77	***89.12	***84.21	***89.3		93.70
***94.07	***88.89	***94.26								
NistXTS749	80.77	80.77	80.77	***72.31	91.93	***94.74	***92.11	***82.28		84.00
84	84	***75.2								
Person	89.28	***88.12	***73.91	***84.2	94.49	***93.04	***78.26	***88.55		93.33
***92.12	***77.27	***88.03								
Products	77.36	***77.23	***76.67	***53.65	88.06	***88.01	***87.56	***76.92		80.39
***80.26	***79.67	***55.75								
RiskIt	89.46	***89.47	***89	***60.46	96.80	***97.08	***90.78	***76.48		93.95
***93.97	***93.47	***63.5								
StackOverflow	100.00	100	100	***94.75	100.00	100	100	***89.51		100.00
100	100	***94.75								
StudentResidence	87.50	***86.83	***85	***78.92	91.85	***92.59	***84.44	***76.22		92.11
***91.4	***89.47	***83.07								
UnixUsage	86.00	86	86	***42.7	96.88	96.88	***95.83	***69.39		91.49
91.49	91.49	***45.43								
Usda	100.00	100	100	***80.9	100.00	100	100	***75.56		100.00
100	100	***80.9								
WordNet	79.69	***79.53	***76.64	***67.29	94.92	***94.42	***85.05	***76.48		82.44
***82.28	***79.21	***69.83								

Table 1. Table