File Sizes

	File	Size (KB)
1	Original Data Dump	4,529,194
2	Extracted Desired Data	$1,\!669,\!496$
3	Delta Compressed Data by Int Values	$1,\!664,\!451$
4	Delta Compressed Data by Bytes	$1,\!669,\!496$
5	Variable Byte Compressed Data	$1,\!239,\!426$

Performance Analysis

	File	Time Elapsed (sec)
1	Extracted Desired Data	234.168
2	Delta Compressed Data by Int Values	315.033
3	Delta Compressed Data by Bytes	272.397
4	Variable Byte Compressed Data	296.672
5	Variable Byte Compressed Data (On Delta)	185.977

Analysis Reports

File 1: No Compression

```
1
  Running
2
   Opening json file
3 File opened, loading data from json
4 Main dictionary created
   Cleaning/Deleting the old database
5
6
7
   About to populate database with 10M pairs - batch writing (Asynchronous)
8
9
   ( '
10
        Working . . . Currently at:', 0)
11
   ( )
        Working . . . Currently at:', 999999)
12
   ( )
        Working . . . Currently at:', 1999998)
        Working . . . Currently at:', 2999997)
13
   ( '
        Working . . . Currently at:', 3999996)
14
   ( '
   ( '
15
        Working . . . Currently at:', 4999995)
16
   ( '
        Working . . . Currently at:', 5999994)
17
   ( '
        Working . . . Currently at:', 6999993)
   ( '
18
        Working . . . Currently at:', 7999992)
19
   ( '
        Working . . . Currently at:', 8999991)
20
  ( )
        Working . . . Currently at:', 9999990)
21
   ( '
        Working . . . Currently at:', 10999989)
22 ('
        Working . . . Currently at:', 11999988)
23 ('
        Working . . . Currently at:', 12999987)
24 ('
        Working . . . Currently at:', 13999986)
25
   ( '
        Working . . . Currently at:', 14999985)
26
  ( '
        Working . . . Currently at:', 15999984)
```

Copyright C 2017 Created Using \nvdashT_EX

27	('	Working .		Currently	at:',	16999983)
28	('	Working .		Currently	at:',	17999982)
29	('	Working .		Currently	at:',	18999981)
30	('	Working .		Currently	at:',	19999980)
31	('	Working .		Currently	at:',	20999979)
32	('	Working .		Currently	at:',	21999978)
33	('	Working .		Currently	at:',	22999977)
34	('	Working .		Currently	at:',	23999976)
35	('	Working .		Currently	at:',	24999975)
36	('	Working .		Currently	at:',	25999974)
37	('	Working .		Currently	at:',	26999973)
38	('	Working .		Currently	at:',	27999972)
39	('	Working .		Currently	at:',	28999971)
40	('	Working .		Currently	at:',	29999970)
41	('	Working .		Currently	at:',	30999969)
42	('	Working .		Currently	at:',	31999968)
43	('	Working .		Currently	at:',	32999967)
44	('	Working .		Currently	at:',	33999966)
45	Time	-	atcl	h Writing		
46	Rand	omizing the	ke	ys		
47		-	-	-		
48						
49	Abou	t to read a	11 0	data - Rano	dom Oro	ler
50	('	Reading .		Currently	at:',	0)
51	('	Reading .		Currently	at:',	999999)
52	('	Reading .		Currently	at:',	1999998)
53	('	Reading .		Currently	at:',	2999997)
54	('	Reading .		Currently	at:',	3999996)
55	('	~ · · · ·		Currently	at:',	4999995)
56	('	~ · · · ·		Currently	at:',	5999994)
57	('	~ · · · ·		Currently	at:',	6999993)
58	('			Currently	at:',	7999992)
59	('	Reading .		Currently	at:',	8999991)
60	('	Reading .		Currently	at:',	9999990)
61	('	Reading .		Currently	at:',	10999989)
62	('	Reading .		Currently	at:',	11999988)
63	('	Reading .		C	at:',	12999987)
64	('	Reading .		Currently	at:',	13999986)
65	(,			Currently	at:',	14999985)
66	('	D		Currently	at:',	15999984)
67	('	D		Currently	at:',	16999983)
68	('	Reading .		Currently	at:',	17999982)
69	('	D		Currently	at:',	18999981)
70	('	D		Currently	at:',	19999980)
71	(,	D		Currently	at:',	20999979)
72	(,	D		Currently	at:',	21999978)
73	(,	D		Currently	at:',	22999977)
74	Ċ,	D		Currently	at:',	23999976)
75	Ċ,	D		Currently	at:',	24999975)
76	(,	Reading .		Currently	at:',	25999974)
77	(,	Reading .		Currently	at:',	26999973)
-	-	0			,	- · - •

Copyright \bigodot 2017

78	('	Reading				Currently	at:',	27999972)
79	('	Reading				Currently	at:',	28999971)
80	('	Reading				Currently	at:',	29999970)
81	('	Reading				Currently	at:',	30999969)
82	('	Reading		•		Currently	at:',	31999968)
83	('	Reading		•		Currently	at:',	32999967)
84	('	Reading		•		Currently	at:',	33999966)
85								
86								
87	Datab	ase popu	ıla	ate	d	with batch	n write	es
88		Time Sta	art	ced	l:	1496724754	l.58sed	2
89		Time End	led	l :	14	196730692.1	sec	
90		Time Ela	aps	sed	l:	5937.52049	9994sed	2
91	Datab	ase read	ł 1	car	ndo	omly		
92		Time Sta	art	ced	l:	1496730782	2.03sec	2
93		Time End	lec	1:	14	196731016.1	9sec	
94		Time Ela	aps	sed	l:	234.168370)962sed	2
95								
96	F	Sinished	۲ı	cog	gra	am		

File 2: Delta Compression on Ints

```
1 Running
2 Opening json file
3 File opened, loading data from json
4 Main dictionary created
5
   Cleaning/Deleting the old database
6
7
8
   About to populate database with 10M pairs - batch writing (Asynchronous)
9
   ( '
10
        Working . . . Currently at:', 0)
11
   ( '
        Working . . . Currently at:', 999999)
12
   ( '
        Working . . . Currently at:', 1999998)
        Working . . . Currently at:', 2999997)
13 ('
14 ('
        Working . . . Currently at:', 3999996)
15 ('
        Working . . . Currently at:', 4999995)
16 ('
        Working . . . Currently at:', 5999994)
17
  ( '
        Working . . . Currently at:', 6999993)
18 ('
        Working . . . Currently at:', 7999992)
        Working . . . Currently at:', 8999991)
19 ('
20 ()
        Working . . . Currently at:', 9999990)
21 ('
        Working . . . Currently at:', 10999989)
22 ('
        Working . . . Currently at:', 11999988)
23 ('
        Working . . . Currently at:', 12999987)
24 ('
        Working . . . Currently at:', 13999986)
25 ('
        Working . . . Currently at:', 14999985)
26 ('
        Working . . . Currently at:', 15999984)
        Working . . . Currently at:', 16999983)
27 ('
28 ('
        Working . . . Currently at:', 17999982)
29 ('
        Working . . . Currently at:', 18999981)
```

30	('	Working		•	Currently	at:',	19999980)
31	('	Working			Currently	at:',	20999979)
32	('	Working			Currently	at:',	21999978)
33	('	Working			Currently	at:',	22999977)
34	('	Working			Currently	at:',	23999976)
35	('	Working			Currently	at:',	24999975)
36	('				Currently	at:',	25999974)
37	('				Currently	at:',	26999973)
38	('	Working			Currently	at:',	27999972)
39	('				Currently	at:',	28999971)
40	('	Working				at:',	29999970)
41	('				Currently	at:',	30999969)
42	(,	_				at:',	31999968)
43	(,	•			•	at:',	32999967)
44	<i>(</i> ,	0			5	at:',	
45	Time	0			n Writing	,	
46	Rand	omizing t		ke	•		
47		0		•			
48							
49	Abou	t to read	al	1 0	data - Rano	dom Ore	der
50	(,	Reading				at:',	0)
51	<i>(</i> ,	U U				at:',	-
52	<i>(</i> ,					at:',	
53	ò,	- · · ·			Currently	at:',	
54	,	- · · ·			Currently	at:',	
55	<i>(</i> ,	- · · ·			Currently	at:',	4999995)
56	<i>(</i> ,	- · · ·			Currently	at:',	5999994)
57	<i>(</i> ,	- · · ·			Currently	at:',	6999993)
58	<i>(</i> ,	- · · ·			Currently	at:',	7999992)
59	<i>(</i> ,	- · · ·		•	Currently	at:',	8999991)
60	<i>(</i> ,	- · · ·		•	Currently	at:',	9999990)
61	<i>(</i> ,			•	Currently	at:',	10999989)
62	<i>(</i> ,	D			Currently	at:',	11999988)
63	<i>(</i> ,	Reading		•	Currently	at:',	12999987)
64	(,	D	· ·	•	Currently	at:',	13999986)
65	(,	Reading	· ·	•	~	at:',	14999985)
66	(,	Reading			C	at:',	15999984)
67	(,	Reading	· ·	•	Currently	at:',	16999983)
68	(,	Reading	· ·	•	Currently	at:',	17999982)
69	(,	Reading			Currently	at:',	18999981)
0 <i>3</i> 70	(,	Reading	•••	•	Currently	at:',	19999980)
71	(,	Reading	• •	•	Currently	at:',	20999979)
72	(,	Reading	• •	•	Currently	at:',	21999978)
73	(,	Reading	•••	•	•		22999977)
73 74	(,	Reading	· ·	•	Currently Currently	at:', at:',	23999977)
75	(,	-		•	•	-	
75 76	(,	Reading	•••	•	Currently	at:',	24999975) 25999974)
76 77	(,	Reading Reading	• •	•	Currently Currently	at:',	25999974)
	(,	0	• •	•		at:',	26999973)
$\frac{78}{79}$	(,	Reading	•••	•	Currently	at:',	27999972)
79 80	(,	Reading Reading	• •	•	Currently Currently	at:',	28999971) 29999970)
00	(neauring	• •	·	ourrentry	at:',	233333(V)

Copyright $\bigcirc 2017$

```
81
   ( '
        Reading . . . Currently at:', 30999969)
82 ('
        Reading . . . Currently at:', 31999968)
83 ('
        Reading . . . Currently at:', 32999967)
84
  ( '
        Reading . . . Currently at:', 33999966)
85
86
87 Database populated with batch writes
        Time Started: 1496984931.5sec
88
89
        Time Ended: 1496991614.47 sec
90
        Time Elapsed: 6682.9688251sec
91 Database read randomly
92
        Time Started: 1496991841.62sec
93
        Time Ended: 1496992156.66sec
94
        Time Elapsed: 315.033441067sec
95
96
   --- Finished Program ---
```

File 3: Byte Delta Compression

```
1 Running
2
   Opening json file
3 File opened, loading data from json
4 Main dictionary created
5 Cleaning/Deleting the old database
6
7
8
   About to populate database with 10M pairs - batch writing (Asynchronous)
9
10 ('
        Working . . . Currently at:', 0)
   ( '
        Working . . . Currently at:', 999999)
11
12 ('
        Working . . . Currently at:', 1999998)
13 ('
        Working . . . Currently at:', 2999997)
14 ('
        Working . . . Currently at:', 3999996)
15 ('
        Working . . . Currently at:', 4999995)
        Working . . . Currently at:', 5999994)
16 ('
17 ('
        Working . . . Currently at:', 6999993)
18 ('
        Working . . . Currently at:', 7999992)
19 ('
        Working . . . Currently at:', 8999991)
20 ('
        Working . . . Currently at:', 9999990)
21 ('
        Working . . . Currently at:', 10999989)
22 ('
        Working . . . Currently at:', 11999988)
23 ('
        Working . . . Currently at:', 12999987)
24 ('
        Working . . . Currently at:', 13999986)
25 ('
        Working . . . Currently at:', 14999985)
26 ('
        Working . . . Currently at:', 15999984)
27 ('
        Working . . . Currently at:', 16999983)
28 ('
        Working . . . Currently at:', 17999982)
29 ('
        Working . . . Currently at:', 18999981)
        Working . . . Currently at:', 19999980)
30 (,
31 ('
        Working . . . Currently at:', 20999979)
32 ('
        Working . . . Currently at:', 21999978)
```

33	('	Working .	•	•	Currently	at:',	22999977)
34	('	Working .			Currently	at:',	23999976)
35	('	Working .			Currently	at:',	24999975)
36	('	Working .			Currently	at:',	25999974)
37	('	Working .			Currently	at:',	26999973)
38	('	Working .			Currently	at:',	27999972)
39	('	Working .			Currently	at:',	28999971)
40	('	Working .			Currently	at:',	29999970)
41	('	Working .			Currently	at:',	30999969)
42	('	Working .			Currently	at:',	31999968)
43	(,				~	at:',	32999967)
44	(,					at:',	
45	Time	•			n Writing	,	· · · · · · · · ,
46		omizing th		ce	•		
47				,	, –		
48							
49	Abour	t to read	a11	6	lata - Rand	lom Ord	ler
50	(,	Reading .			~	at:',	0)
51	<i>(</i> ,				a	at:',	999999)
52	(,	Reading .			Currently	at:',	1999998)
53	(,	Reading .			Currently	at:',	2999997)
55 54	(,	Reading .	•		Currently	at:',	3999996)
$55 \\ 55$	(,			·	Currently	at:',	4999995)
56	(,		•	·	Currently	at:',	5999994)
50	(,	0	•	·	•		6999993)
57	(,	Reading .	•	·	Currently	at:',	7999992)
	(,	Reading .	·	•	Currently	at:',	-
59 60	(,	Reading .	·	•	Currently	at:',	8999991)
60 61	(,	Reading .	·	•	Currently	at:',	9999990)
61 62		Reading .	·	•	Currently	at:',	10999989)
62 62	('	Reading .	•	·	Currently	at:',	11999988)
63 64	('	Reading .	·	·	Currently	at:',	12999987)
64 67	('	Reading .	·	·	Currently	at:',	13999986)
65	('	Reading .	•	·	Currently	at:',	14999985)
66	('	Reading .	·	·	5	at:',	15999984)
67	('	Reading .	•	·	5	at:',	16999983)
68	('	Reading .	•	•	5	at:',	17999982)
69	('	Reading .	•	•	Currently	at:',	18999981)
70	('	Reading .	•	•	Currently	at:',	19999980)
71	('	Reading .	•	•	Currently	at:',	20999979)
72	('	Reading .	•	•	Currently	at:',	21999978)
73	('	Reading .	•	•	Currently	at:',	22999977)
74	('	Reading .	•	•	Currently	at:',	23999976)
75	('	Reading .	•	•	Currently	at:',	24999975)
76	('	Reading .			Currently	at:',	25999974)
77	('	Reading .			Currently	at:',	26999973)
78	('	Reading .			Currently	at:',	27999972)
79	('	Reading .			Currently	at:',	28999971)
80	('	Reading .			Currently	at:',	29999970)
81	('	Reading .			Currently	at:',	30999969)
82	('	Reading .			Currently	at:',	31999968)
83	('	Reading .			Currently	at:',	32999967)
		0			5		

Copyright \bigodot 2017

```
( '
        Reading . . . Currently at:', 33999966)
84
85
86
   Database populated with batch writes
87
88
        Time Started: 1496995923.99sec
        Time Ended: 1497001728.99sec
89
90
        Time Elapsed: 5805.00258112sec
91
   Database read randomly
92
        Time Started: 1497001857.33sec
93
        Time Ended: 1497002129.73sec
        Time Elapsed: 272.397305012sec
94
95
96
   --- Finished Program ---
```

File 4: Variable Byte Compression

		· · ·
1	Runn	ing
2	Open	ing json file
3	File	opened, loading data from json
4	Main	dictionary created
5	Clea	ning/Deleting the old database
6		
7		
8	Abou	t to populate database with 10M pairs - batch writing (Asynchronous)
9		
10	('	Working Currently at:', 0)
11	('	Working Currently at:', 999999)
12	('	Working Currently at:', 1999998)
13	('	Working Currently at:', 2999997)
	('	Working Currently at:', 3999996)
	('	Working Currently at:', 4999995)
-	('	Working Currently at:', 5999994)
	('	Working Currently at:', 6999993)
-	('	Working Currently at:', 7999992)
	('	Working Currently at:', 8999991)
	('	Working Currently at:', 9999990)
	('	Working Currently at:', 10999989)
	('	Working Currently at:', 11999988)
	('	Working Currently at:', 12999987)
	('	Working Currently at:', 13999986)
25 26	-	Working Currently at:', 14999985)
	(' ('	Working Currently at:', 15999984) Working Currently at:', 16999983)
$\frac{21}{28}$	(,	Working Currently at:', 10999982)
$\frac{28}{29}$	(,	Working Currently at: , 17999981)
-	(,	Working Currently at: , 19999980)
	(,	Working Currently at: , 20999979)
	(,	Working Currently at: , 2000078)
-	(,	Working Currently at: , 22999977)
	(,	Working Currently at:', 23999976)
35		Working Currently at:', 24999975)
00	`	"of a final section of the section o

Copyright © 2017

36	('	Working			Currently	at:',	25999974)
37	<i>(</i> ,	-				at:',	26999973)
38	<i>(</i> ,	Working				at:',	27999972)
39	<i>(</i> ,	Working				at:',	28999971)
40	<i>(</i> ,	Working				at:',	29999970)
41	<i>(</i> ,	Working				at:',	-
42	<i>(</i> ,	Working				at:',	
43	<i>(</i> ,	Working			•	at:',	
44	(,	Working			-	at:',	
45	•	0			n Writing	ut.,	
46		omizing t			0		
47	manac		, 11 0		, 2		
48							
49	About	to read	1 2 1	1	data - Rano	lom Ord	ler
50^{10}	(,	Reading				at:',	0)
51	(,	Reading					999999)
52	(,	Reading				at:',	1999998)
53	(,	Reading	•••				-
$55 \\ 54$	(,	Reading	•••		Currently		
55	(,	Reading			Currently	at:',	4999995)
56	(,	Reading			Currently	at:',	5999994)
$50 \\ 57$	(,	Reading			Currently	at:',	6999993)
58	(,	Reading			Currently	at:',	7999992)
50	(,	Reading			Currently	at:',	8999991)
60	(,		•••		Currently	at:',	99999990)
61	(,	Reading	•••		Currently	at:',	10999989)
62	(,	Reading			Currently	at:',	11999988)
63	(,	Reading	• •		Currently	at:',	12999987)
64	(,	Reading	• •		Currently	at:',	13999986)
65	(,	Reading		•	Currently		14999985)
66	(,	Reading		•	Currently	at:',	15999984)
67	(,	Reading		•	Currently	at:', at:',	16999983)
68	(,	Reading		•	a	at:',	17999982)
69	(,	Reading		•	~	at:',	18999981)
0 <i>3</i> 70	(,	Reading		•	~	at:',	19999980)
70	(,	Reading		•	Currently	at:',	
$71 \\ 72$	(,	Reading			~	at:',	20999979) 21999978)
73	(,						22999977)
73 74	(,	Reading	• •		Currently Currently	at:',	23999976)
7475	(,	Reading Reading		•	Currently	at:',	24999975)
76	(,	Reading		•	Currently	at:', at:',	25999974)
70 77	(,	Reading	• •		Currently	at:',	26999973)
78	(,	Reading	• •		Currently	at:',	27999972)
79	(,	Reading	• •		Currently		28999971)
80	(,	•	• •			at:',	
80 81	(,	Reading Reading	• •		Currently Currently	at:',	29999970) 30999969)
81 82	(,	Reading	• •		•	at:',	31999968)
82 83	(,	0	• •		Currently Currently	at:',	
84	(,	Reading Reading	• •		Currently	at:',	32999967) 33999966)
$\frac{84}{85}$	(meauring	• •	•	ourrentry	at:',	55333300)
86							
00							

07	Detabase newlated with betch whites
87	Database populated with batch writes
88	Time Started: 1497007421.17sec
89	Time Ended: 1497019697.8sec
90	Time Elapsed: 12276.6317949sec
91	Database read randomly
92	Time Started: 1497019802.41sec
93	Time Ended: 1497020099.08sec
94	Time Elapsed: 296.672063112sec
95	
96	Finished Program

File 5: Variable Byte Compression On Delta

	v i
1	Opening json file
2	File opened, loading data from json
3	Main dictionary created
4	Finished compressing dictionary values, writing to file
5	Making Dict Values to Bytes
6	Finished writing
7	
8	
9	
10	About to populate database with 10M pairs - batch writing (Asynchronous)
11	
12	Working Currently at: O
13	Working Currently at: 999999
14	Working Currently at: 1999998
15	Working Currently at: 2999997
16	Working Currently at: 3999996
17	Working Currently at: 4999995
18	Working Currently at: 5999994
19	Working Currently at: 6999993
20	Working Currently at: 7999992
$21 \\ 22$	Working Currently at: 8999991
$\frac{22}{23}$	Working Currently at: 9999990 Working Currently at: 10999989
$\frac{23}{24}$	Working Currently at: 10999988
$\frac{24}{25}$	Working Currently at: 12999987
$\frac{25}{26}$	Working Currently at: 13999986
$\frac{20}{27}$	Working Currently at: 14999985
$\frac{21}{28}$	Working Currently at: 15999984
$\frac{20}{29}$	Working Currently at: 16999983
$\frac{20}{30}$	Working Currently at: 17999982
31	Working Currently at: 18999981
32	Working Currently at: 19999980
33	Working Currently at: 20999979
34	Working Currently at: 21999978
35	Working Currently at: 22999977
36	Working Currently at: 23999976
37	Working Currently at: 24999975
38	Working Currently at: 25999974

39	Working .		Currently	at:	26999973
40	Working .		Currently	at:	27999972
41	Working .		Currently	at:	28999971
42	Working .		Currently		29999970
43	Working .		Currently		30999969
44	Working .		Currently		31999968
45	•		Currently		32999967
46	-		Currently		33999966
47	Timer Ended -		•		
48	Randomizing t		-		
49	0		·		
50					
51	About to read	l all	data - Ra	ndom	Order
52	Reading .		Currently	at:	0
53	•		Currently		
54	Reading .		Currently		1999998
55	Reading .		Currently		
56	Reading .		Currently		3999996
57	Reading .		Currently		4999995
58	Reading .		Currently		5999994
59	Reading .		Currently		6999993
60	Reading .		Currently		7999992
61	Reading .		Currently		8999991
62	Reading .		Currently		9999990
63	Reading .		Currently		10999989
64	Reading .		Currently		11999988
65	Reading .		Currently		12999987
66	Reading .		Currently		13999986
67	Reading .		Currently		14999985
68	Reading .		Currently		15999984
69	Reading .		Currently		16999983
70	Reading .		Currently		17999982
71	Reading .		Currently	at:	18999981
72	Reading .		Currently	at:	19999980
73	Reading .		Currently	at:	20999979
74	•		Currently		21999978
75	Reading .		Currently	at:	22999977
76	Reading .		Currently	at:	23999976
77	Reading .		Currently		24999975
78	Reading .		Currently	at:	25999974
79	Reading .		Currently		26999973
80	Reading .		Currently	at:	27999972
81	Reading .		Currently	at:	28999971
82	Reading .		Currently	at:	29999970
83	Reading .		Currently	at:	30999969
84	Reading .		Currently		31999968
85	Reading .		Currently		32999967
86	Reading .		Currently		33999966
87	Time Elapsed		•		: 71.28672200000437
88	-				
89					

Research	Inverted Index Searching	Danish Vaid
Spring 2017	Compression and Read Perf.	danishvaid@umail.ucsb.edu

90	Database populated with batch writes
30	Database populated with Datch willes
91	Time Started: 85876.857805sec
92	Time Ended: 86098.197818sec
93	Time Elapsed: 221.34001299999363sec
94	Database read randomly
95	Time Started: 86156.352479sec
96	Time Ended: 86342.329792sec
97	Time Elapsed: 185.97731300001033sec
98	
99	Finished Program