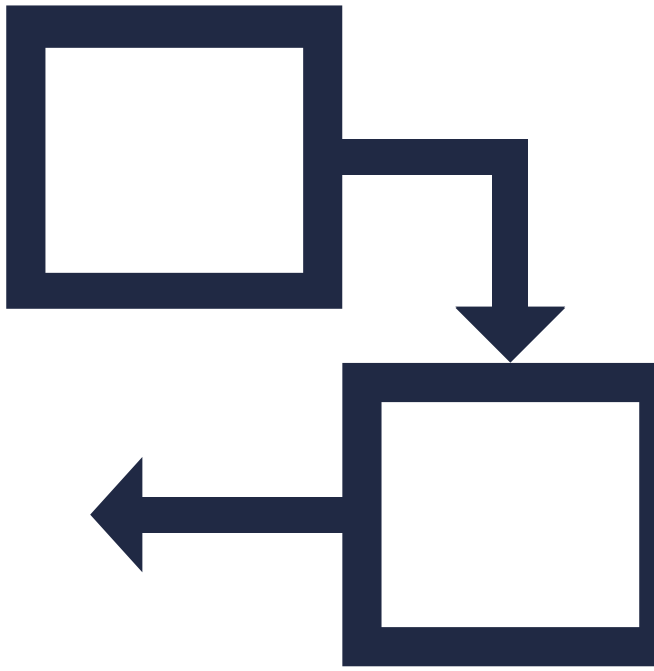
Apteco

Using Automated Alternatives in Apteco PeopleStage™

How highly sophisticated algorithms can crack the code of compelling content and improve campaign performance.

Say goodbye to manual testing

A/B testing has been an important technique in marketing for decades, allowing the best performing creative to be determined statistically from a small-scale test, before being rolled out to the full audience. Each test takes time to run, since a particular volume of communications must be analysed before any conclusions can be considered valid.



Historically, when marketing communications were less frequent and broader in scope (when they involved sending the same product catalogue to all customers once every three months for example), it was feasible for A/B tests to be conducted manually.

Modern marketing involves sending more specialised communications with higher frequency, perhaps targeting specific customer segments with specific product ranges. This makes it difficult to perform A/B tests manually. It might take a few weeks to run a single A/B test, so doing this for a dozen product ranges is a considerable investment of effort.

Segments may respond differently; for one segment A may be best, but for another segment B might be better; and what about trying alternatives C and D? The results of the test might indicate that B is best now, but for how long will this remain true? In a few months' time, the conclusion could be different, especially if there are seasonal effects to consider.

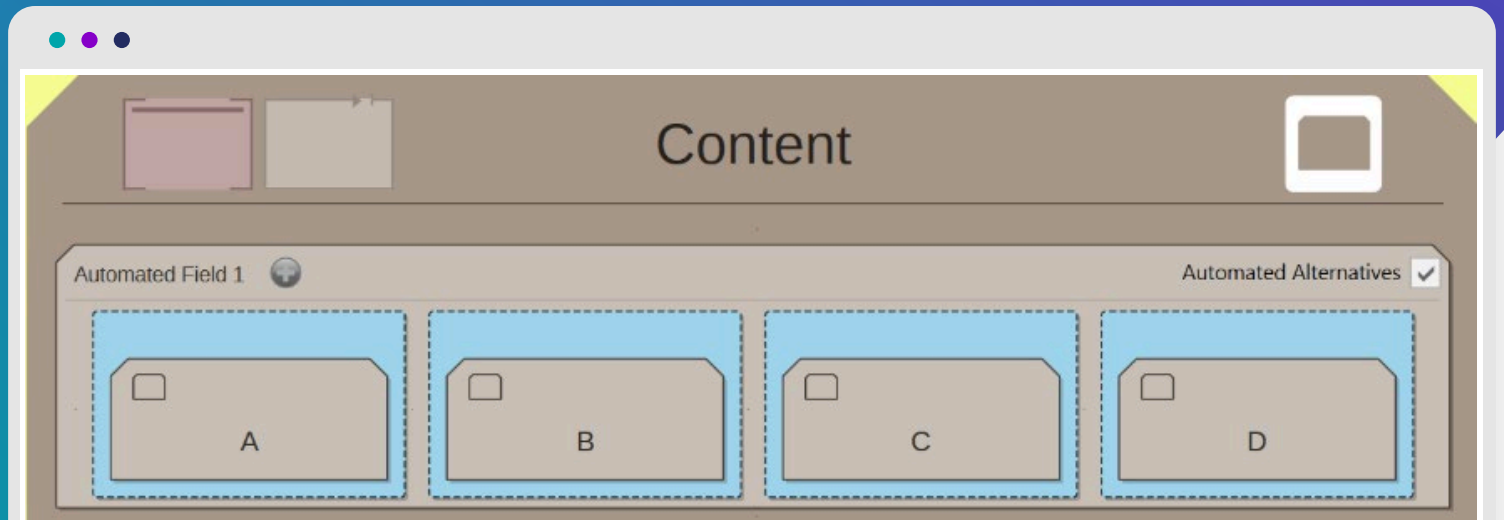
Modern marketing needs to be automated.

A marketing application needs to be able to monitor the performance of the communications being sent and make adjustments accordingly to improve their performance. This "Automated Improvement" could be seen as a form of "Artificial Intelligence".

Apteco PeopleStage™ - Automated Alternatives

Apteco PeopleStage can automatically improve the performance of campaigns through the use of "Automated Alternatives". The user specifies a number of alternative content items (such as subject lines or offers) and PeopleStage then monitors the performance of each over recurring runs of the campaign. The system automatically adjusts the proportion of each alternative in order to increase the overall response rate.

By simply ticking the "Automated Alternatives" option, PeopleStage monitors the performance of each alternative and adjusts the volume of each, using a "Champion-Challenger" algorithm, similar to a winner-stays-on pool competition. The initial run of the message is sent to all the alternatives. The best two become the champion-challenger pair in the first test, while the other alternatives will take it in turn to be the challenger in subsequent tests.



Results

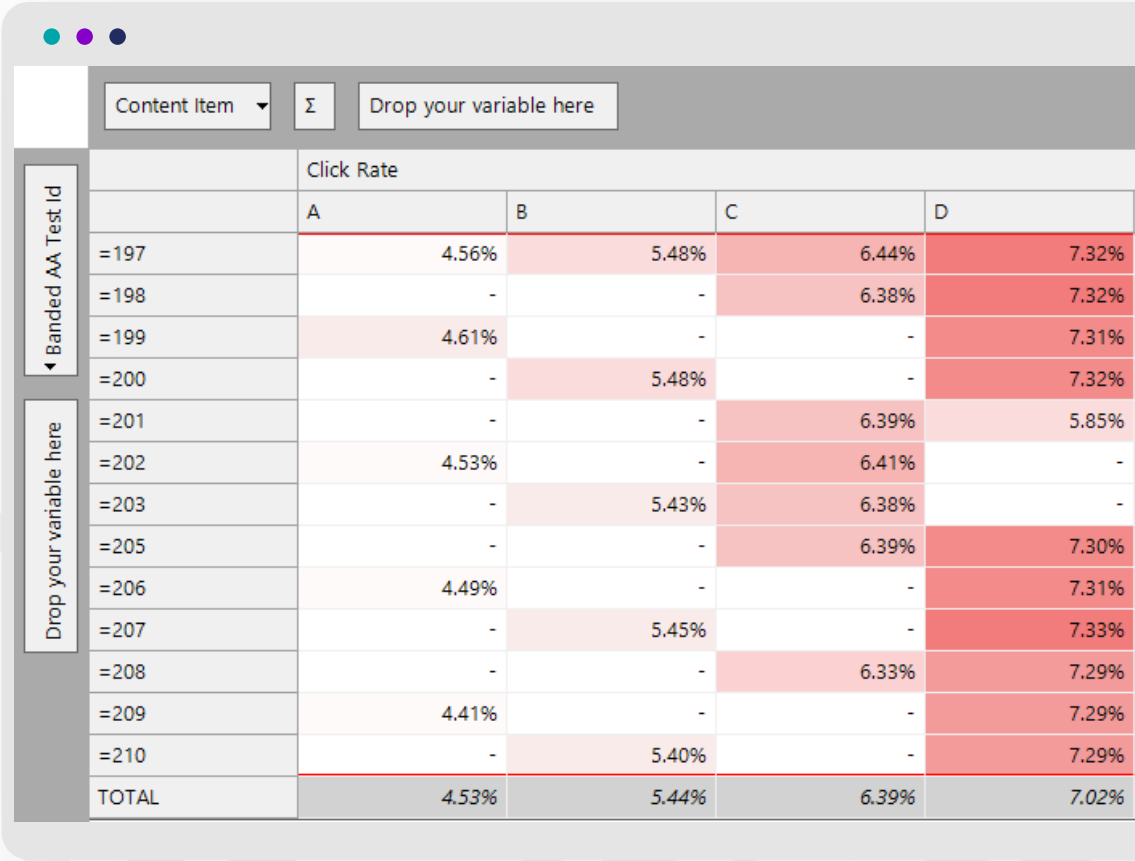
The initial run is – by default – sent to 1,000 people to get a rough indication of the response rates of each alternative. Subsequent tests determine the required volume of communications that need to be sent to make the test statistically significant. Each test sends 80% of communications to the champion and 20% to a single challenger. Once this volume has been sent, possibly over multiple runs of the campaign, the response rates are examined.

Only if the challenger has done statistically better will they become the new champion.

The following analysis in Apteco FastStats® shows the results of a number of successive tests between four alternative pieces of content: A, B, C and D. The standard FastStats design has been extended to allow simple email response history from PeopleStage to be loaded into FastStats.

Each row shows the results from one “AA Test” (Automated Alternatives test). Note that each test could require multiple runs of the campaign in order to achieve the volume of communications required to draw a significant conclusion.

- **The initial test (id 197) is sent to all four alternatives. The response rates for each alternative are shown in the cube with D doing best at 7.32% and C second at 6.44%.**
- **Test 198 is between the first and second placed alternatives from the initial run. D wins this with a response rate of 7.32% and so is used again.**
- **The following tests (199 to 201) cycle through the alternatives in turn, each challenging D.**
- **For some reason D underperforms in test 201 and so C becomes the champion.**
- **Tests 202 to 205 have C as champion. D wins test 205 and so stays on for the following runs being challenged by each alternative in turn.**



	Click Rate			
	A	B	C	D
=197	4.56%	5.48%	6.44%	7.32%
=198	-	-	6.38%	7.32%
=199	4.61%	-	-	7.31%
=200	-	5.48%	-	7.32%
=201	-	-	6.39%	5.85%
=202	4.53%	-	6.41%	-
=203	-	5.43%	6.38%	-
=205	-	-	6.39%	7.30%
=206	4.49%	-	-	7.31%
=207	-	5.45%	-	7.33%
=208	-	-	6.33%	7.29%
=209	4.41%	-	-	7.29%
=210	-	5.40%	-	7.29%
TOTAL	*4.53%*	*5.44%*	*6.39%*	*7.02%*

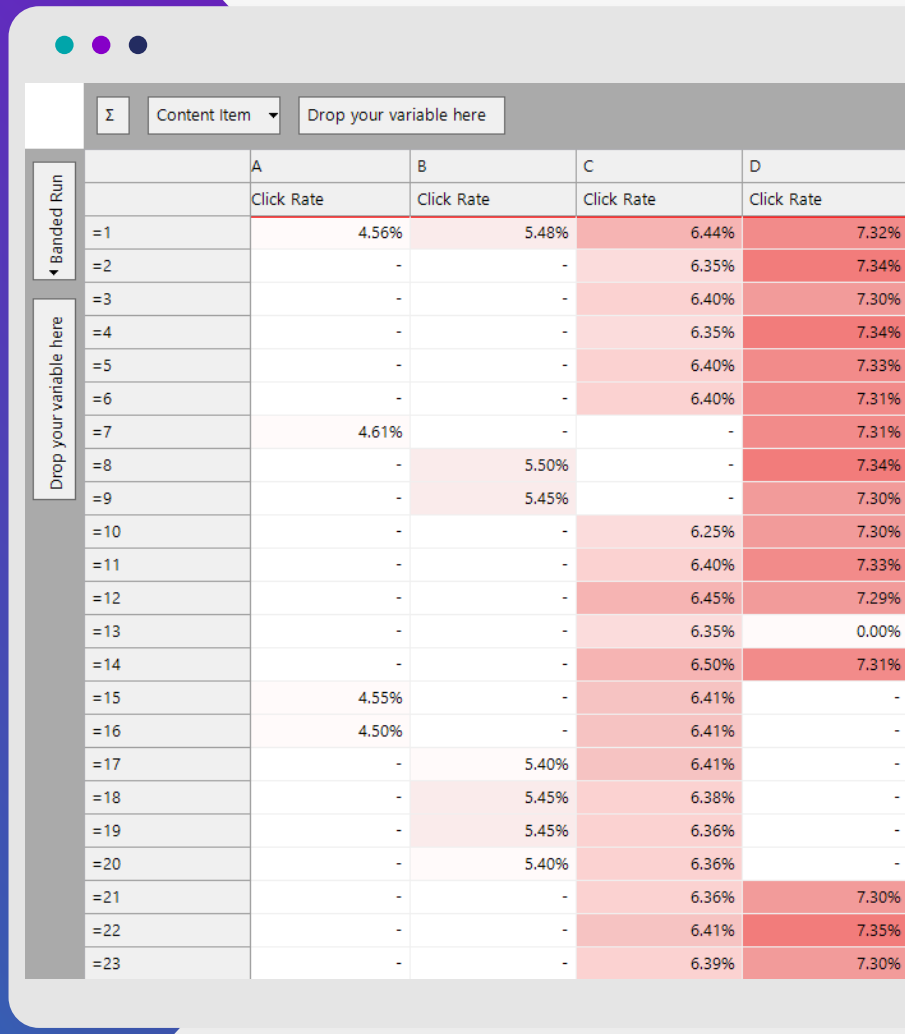
The campaign would continue running with the current best alternative being used for 80% of the communications, and the remaining 20% used to check if any of the other alternatives should be used instead.

Detailed results

The cube right shows the results of the individual runs that went to make up the AA Tests shown previously. Each run of this campaign is for 10,000 communications. The history details are shown on page 10.

Notice the following:

- Only a single run is done for the initial comparison of all the alternatives.
- Runs two to six are needed for the contest between D and C, but there is only one run between D and A. This is because the margin between D and C is smaller and so a larger number of communications is needed to get a significant result (see history: 46,743 vs 5,536).
- Run 13 of the campaign yielded no responses for D (due to some operational error). This is why C manages to becomes champion in Test 201 on page 6.



	A	B	C	D
	Click Rate	Click Rate	Click Rate	Click Rate
=1	4.56%	5.48%	6.44%	7.32%
=2	-	-	6.35%	7.34%
=3	-	-	6.40%	7.30%
=4	-	-	6.35%	7.34%
=5	-	-	6.40%	7.33%
=6	-	-	6.40%	7.31%
=7	4.61%	-	-	7.31%
=8	-	5.50%	-	7.34%
=9	-	5.45%	-	7.30%
=10	-	-	6.25%	7.30%
=11	-	-	6.40%	7.33%
=12	-	-	6.45%	7.29%
=13	-	-	6.35%	0.00%
=14	-	-	6.50%	7.31%
=15	4.55%	-	6.41%	-
=16	4.50%	-	6.41%	-
=17	-	5.40%	6.41%	-
=18	-	5.45%	6.38%	-
=19	-	5.45%	6.36%	-
=20	-	5.40%	6.36%	-
=21	-	-	6.36%	7.30%
=22	-	-	6.41%	7.35%
=23	-	-	6.39%	7.30%

PeopleStage will always wait until the required number of communications have been sent before evaluating the results of the test. At this point it looks at the number of responses from each alternative.

If no responses are received, then PeopleStage will not proceed to the next test and a warning will be displayed in the history as this could indicate a problem with the response gathering mechanism. It will repeat the current combination of alternatives for subsequent runs of the campaign.

Once some responses have been received, if the response rate is below a certain proportion (25% by default) of that expected based on previous history, PeopleStage will again wait before evaluating the test. This could indicate that the next run is being sent too soon without allowing time for the previous run's responses to come in.

PeopleStage will evaluate the test once either sufficient responses have been received or the number of communications sent is above a certain multiple (twice by default) of the required volume.

In some cases, one alternative may have a higher response rate, but it may not be high enough to deem the result as significant. In this case the current champion will stay on. This is determined by the p-Value, which is significant if lower than the value of alpha (0.05).

History

Run 1: Initialising an equal N-way test between ['A', 'B', 'C', 'D'] requiring a volume of 1,000

Run 2: Created new Champion Challenger test ['D'! v 'C'] requiring volume of 46,743
Champion & Challenger from results of N-way test.
('D', 'C', 'B', 'A' = 7.32%, 6.44%, 5.48%, 4.56% from 2,500, 2,500, 2,500, 2,500)
Required volume based on: Expected response rate: 7.32%, Margin: 12.02%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 7.32%
Margin based on historic response rates:
Margin from difference between 0.0732 and 0.0644
Using Challenger's response rate of 6.44%
Previous test was complete (10,000 vs 1,000):
The Equal-N-Way test between ('A', 'B', 'C', 'D') exceeded the minimum volume requirement (10,000 vs 1,000).
The response rate so far (5.95%) is sufficient - over 0.25 of that expected (1.00%).

Run 3: Continuing with the current test ['D'! v 'C'].
Insufficient sent volumes (10,000 vs 46,743):
'D': 8,000 out of 37394 required
(7.34% response rate so far)
'C': 2,000 out of 9349 required
(6.35% response rate so far)
The response rate so far (7.14%) is sufficient - over 0.25 of that expected (7.32%)."

Run 4: Continuing with the current test ['D'! v 'C'].
Insufficient sent volumes (20,000 vs 46,743)...

Run 5: Continuing with the current test ['D'! v 'C'].
Insufficient sent volumes (30,000 vs 46,743)...

Run 6: Continuing with the current test ['D'! v 'C'].
Insufficient sent volumes (40,000 vs 46,743)...

Run 7: Created new Champion Challenger test ['D'! v 'A'] requiring volume of 5,536

Champion 'D' won again: ('D', 'C' = 7.32%, 6.38% from 40,000, 10,000) (p-Value=0.0011)

Required volume based on: Expected response rate: 7.32%, Margin: 37.73%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20

Expected response rate taken from history.

Using Champion's response rate of 7.32%

Margin based on historic response rates:

Margin from difference between 0.0732 and 0.0456

Using Challenger's response rate of 4.56%

Previous test was complete (50,000 vs 46,743):

'D': 40,000 exceeds 37,394 required (7.32% response rate)

'C': 10,000 exceeds 9,349 required (6.38% response rate)

The response rate so far (7.13%) is sufficient - over 0.25 of that expected (7.32%)."

Run 8: Created new Champion Challenger test ['D'! v 'B'] requiring volume of 11,726

Champion 'D' won again: ('D', 'A' = 7.31%, 4.61% from 8893, 1,107) (p-Value=0.0011)

Required volume based on: Expected response rate: 7.31%, Margin: 25.03%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20

Expected response rate taken from history.

Using Champion's response rate of 7.31%

Margin based on historic response rates:

Margin from difference between 0.0731 and 0.0548

Using Challenger's response rate of 5.48%

Previous test was complete (10,000 vs 5,536):

'A': 1,107 exceeds 1,107 required (4.61% response rate)

'D': 8,893 exceeds 4,429 required (7.31% response rate)

The response rate so far (7.01%) is sufficient - over 0.25 of that expected (7.32%).

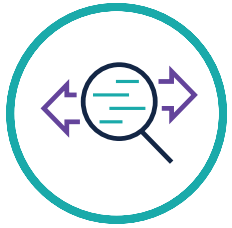
Run 9: Continuing with the current test ['D'! v 'B'].

Insufficient sent volumes (10,000 vs 11,726)...

Sample size considerations

Each “contest” between two alternatives requires a particular number of communications to be sent in order to make the result statistically significant. This “required volume” is worked out in advance based on a number of factors.

Broadly speaking these factors are:



How difficult the differences in response rate are to detect.

How tolerant you are prepared to be of making a mistake (e.g. incorrectly concluding that an alternative was better/ worse than the champion, when in fact it was just due to an atypical set of responses).

The difficulty in identifying differences in response rate will depend on the particular campaign and is based automatically on the actual history that accumulates.

This is characterised by:

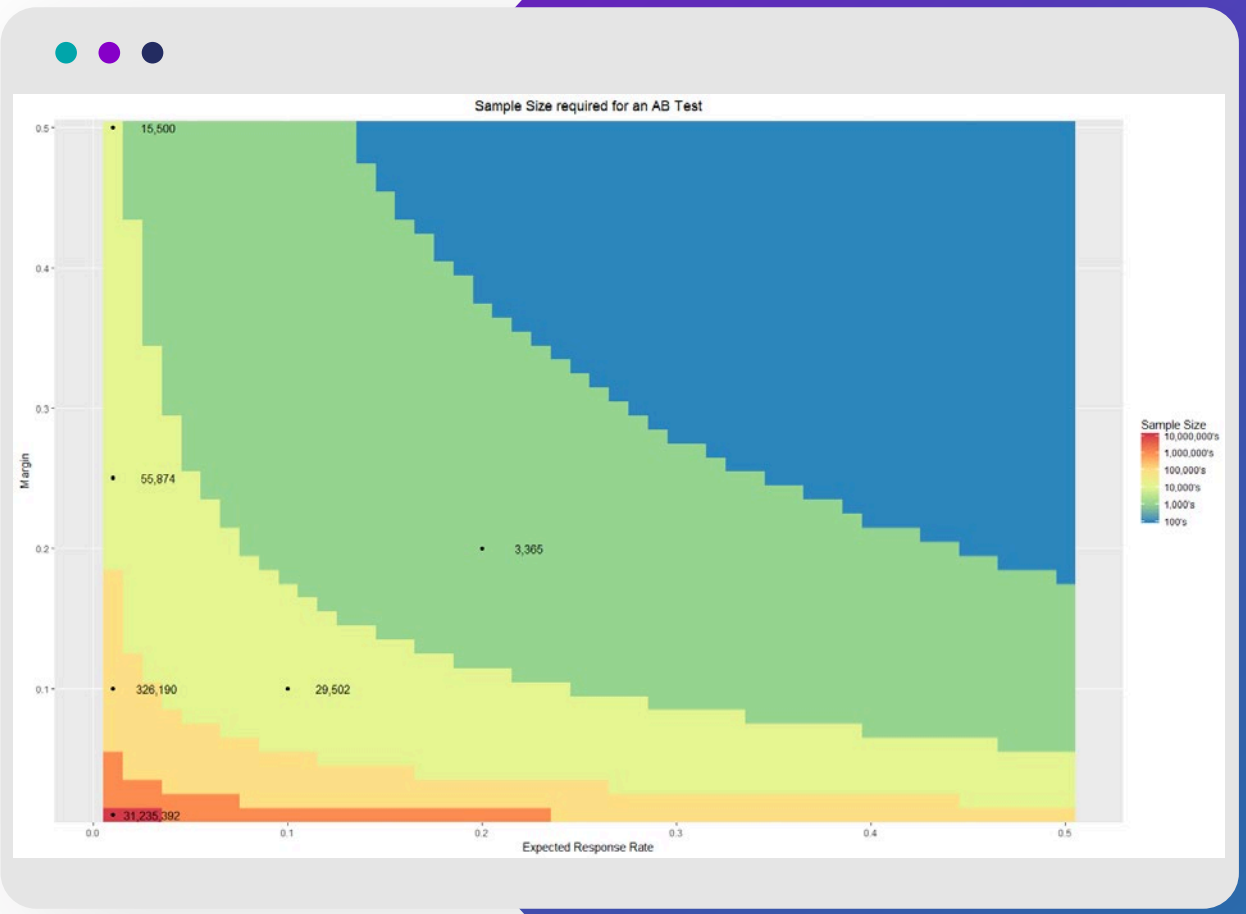
“Expected Response Rate”: the baseline or typical response rate – E.g. 5%, if the champion has historically had a 5% response rate.

“Margin”: the relative difference in response rate which you want to detect – E.g. 10%, would detect a 5.5% response rate for the challenger as being different.

The tolerance of making a mistake is controlled by two statistical parameters:

Alpha: “False positive” rate (default 5%). How prepared you are to conclude that the challenger is better, when in fact it was just the luck of the day.

Beta: “False negative” rate (default 20%). How prepared you are to miss a genuinely better alternative.



The chart to the left shows the required volume for particular combinations of Expected Response Rate and Margin, using default values of alpha and beta. The point values marked show for example that with an expected response rate of 10% and a margin of 10%, the volume required is 29,502. So for a 1% margin on a 1% expected response rate, the volume required is over 30 million!

See here for an online calculator.

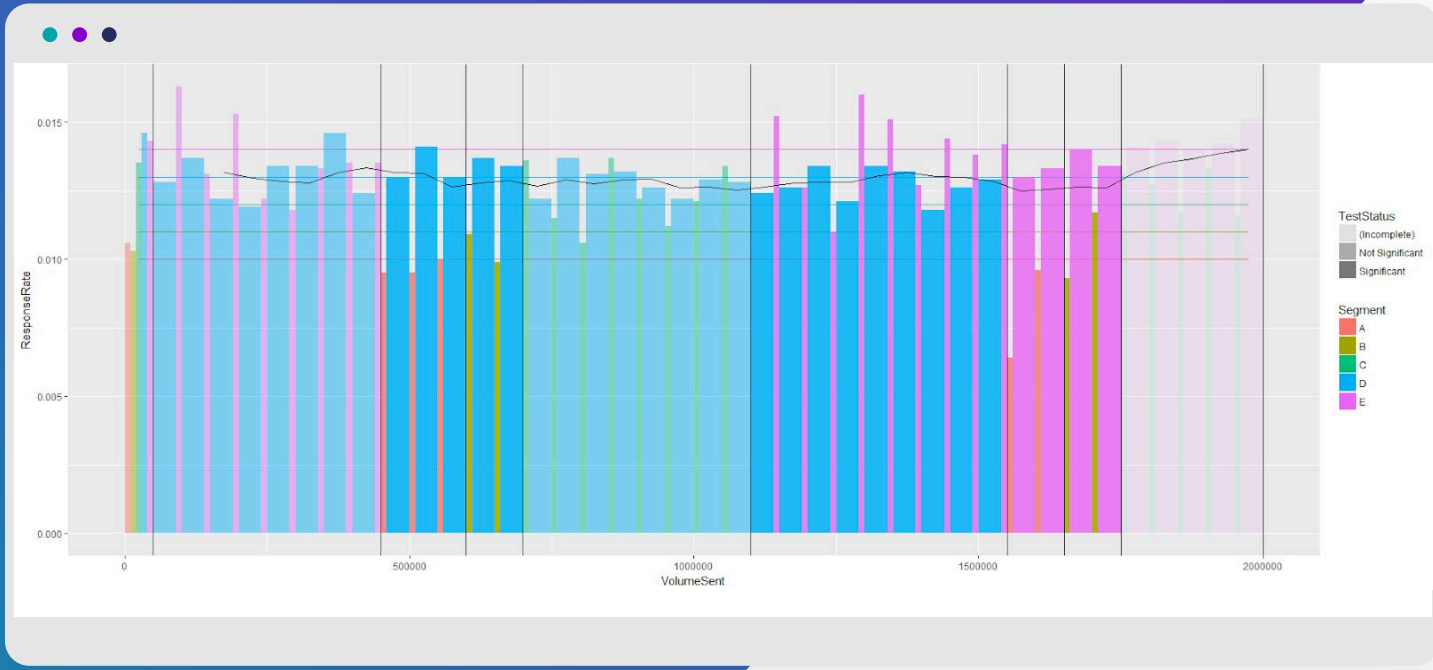
Simulated results

The example below shows what would happen using the standard defaults where a Campaign is sending 50,000 communications per run, with five automated alternatives, A to E. Full details of the history are shown in a later section.

The vertical axis shows the response rate from each alternative during each run. The horizontal coloured lines indicate the expected response rates of each segment used in the simulation (ranging from A with 1% to E with 1.4%).

The x-axis shows the volume sent to each alternative over repeated runs of the campaign.

Each run consists of a pair of alternatives: the width of the bars corresponds to the volume which is in the ratio 80:20; the height of the bar is the response rate for that alternative on that run. The wavy black line indicates a rolling average overall response rate. The vertical black lines separate the individual tests, showing when the volume is sufficient to assess the result. A block of runs are shown faded if that test did not yield a significant result.



- The initial run was sent equally to all five alternatives. By chance, D does slightly better than E and so is the champion in the first test (and so has 80% of the volume of each run).

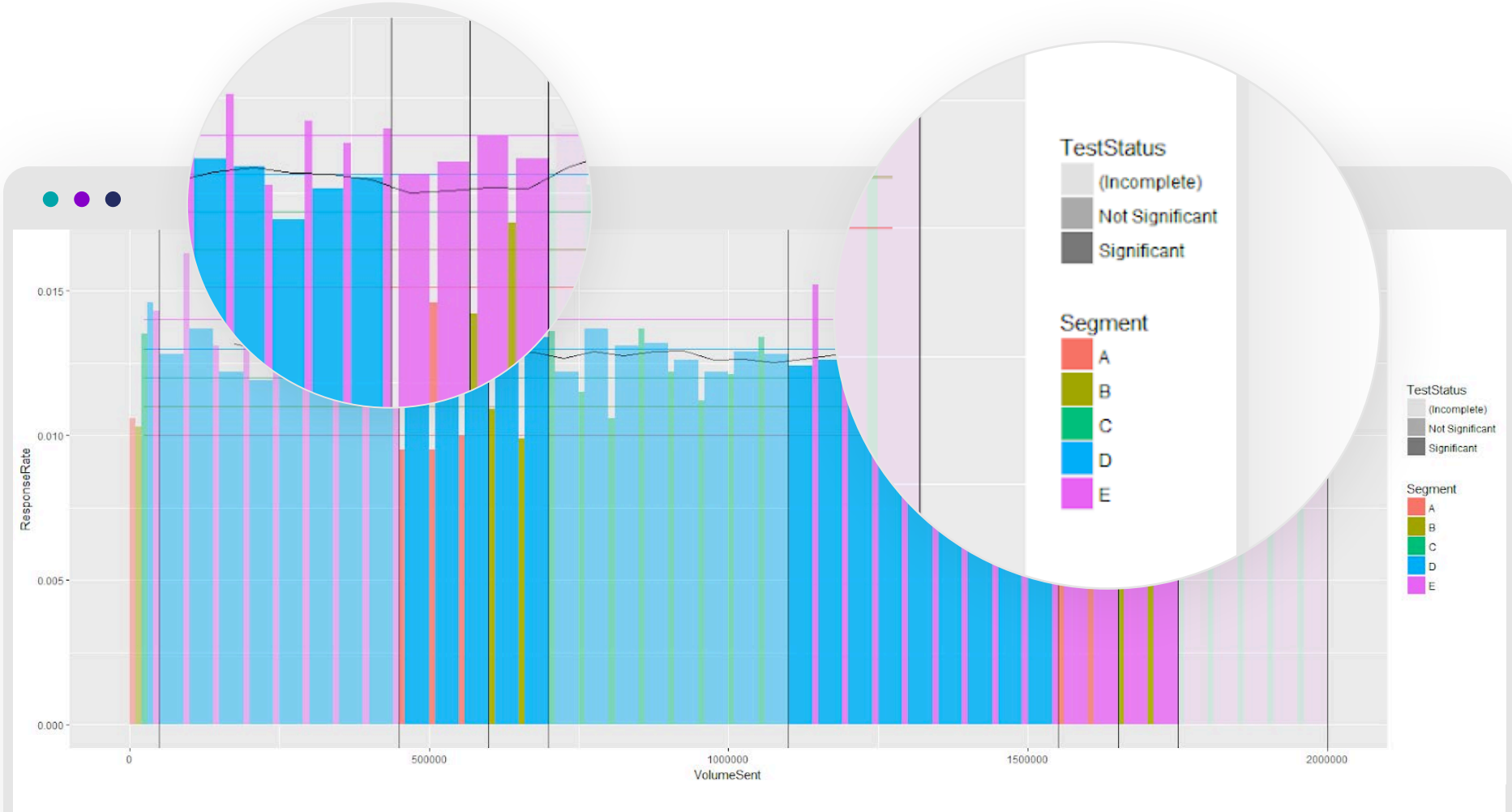
- The next eight runs are needed to evaluate D against E which requires a volume of 357,146. The log shows that although E did have a higher response rate, it was not significantly higher, and so D remains champion.

- Only three runs are needed to evaluate D vs A, since there is a bigger difference in their historical response rates (required volume is 121,085). Similarly only two runs are needed to evaluate D vs B (the historical data actually suggests a bigger difference between D and B, due to B doing unexpectedly worse than A in the initial run).

- The next eight runs are needed to evaluate D vs C – their required volume is higher due to them being closer in historical performance. The statistical test shows that the result is not significant and so D remains champion.

- The volume required to evaluate D vs E is over 400,000, requiring nine runs. This time E does do significantly better and so becomes champion.

- Relatively low volumes are required when E is challenged by A and B.



Settings

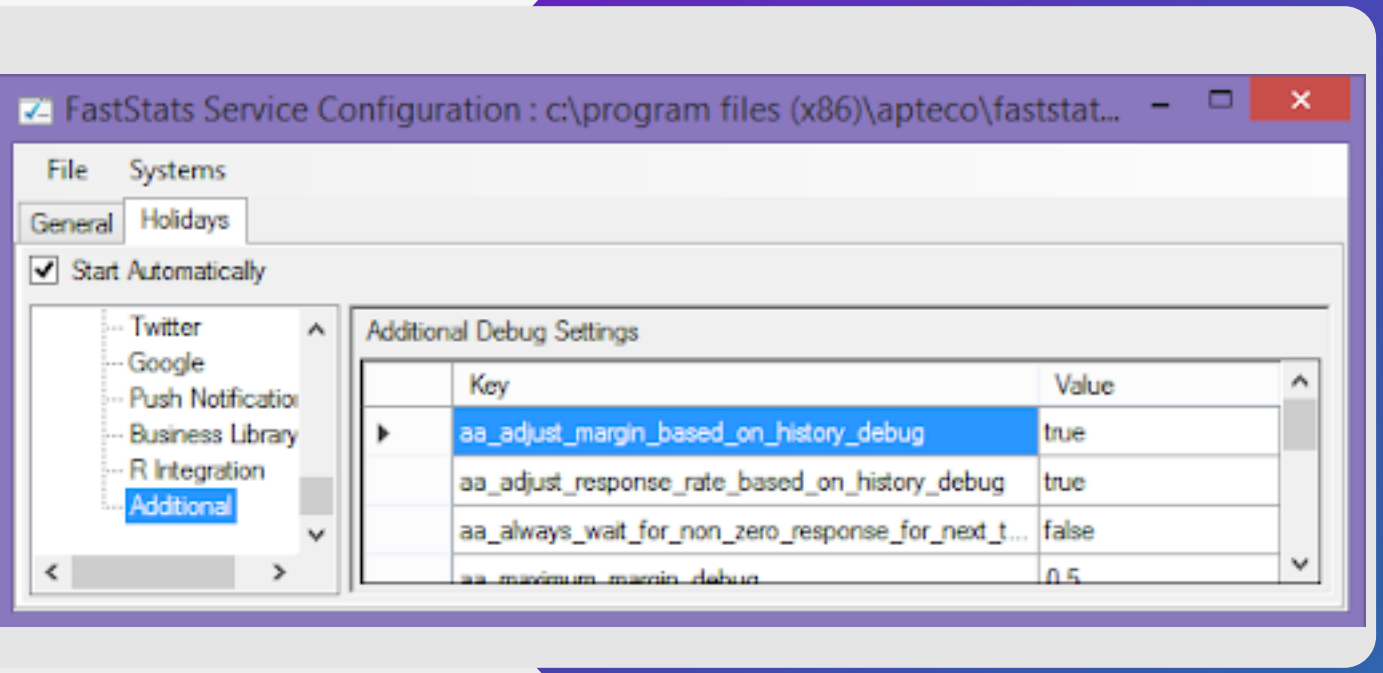
Most settings should not really need to be changed and can only be changed currently via the CLMClient configurator and the FastStats Service configurator.

Setting	Default	Explanation
Maximum Challenger Proportion	0.20 (20%)	Limits the proportion of communications sent to the challenger on each run.
Alpha (false positive)	0.05 (5%)	How prepared you are to incorrectly conclude that the challenger is better.
Beta (false negative)	0.20 (20%)	How prepared you are to miss a genuinely better alternative.
Default Expected Response Rate	0.10 (10%)	Values used in the sample size calculation (see next section). Defaults are only used if the corresponding adjustment option is set to not use the history.
Default Margin	0.25 (25%)	
aa_adjust_response_rate_based_on_history_debug	True	
aa_adjust_margin_based_on_history_debug	True	
aa_minimum_response_rate_debug	0.01 (1%)	Absolute limits on the expected response rate and margin, that are imposed on the historical or default values (depending on which are being used).

Settings

Most settings should not really need to be changed and can only be changed currently via the CLMClient configurator and the FastStats Service configurator.

Setting	Default	Explanation
aa_minimum_margin_debug	0.01 (1%)	Absolute limits on the expected response rate and margin, that are imposed on the historical or default values (depending on which are being used).
aa_maximum_margin_debug	0.5 (50)	
aa_minimum_initial_all_volume_debug	1000	The number of communications required to be sent on the first run.
aa_always_wait_for_non_zero_response_for_next_test_debug	True	If true, will prevent the test progressing to the next alternatives until at least some responses have been obtained.
aa_minimum_response_rate_multiplier_for_next_test_debug	0.25 (25%)	Prevents the test progressing until at least this proportion of the expected responses have been received from the previous test.
aa_maximum_sent_volume_multiplier_for_next_test_debug	2.0	Allows the test to progress even if insufficient responses have been obtained, once the volume sent exceeds this multiple of that required.



The FastStats Service settings are entered in the advanced section by typing in the key and entering a corresponding value. The CLMClient settings are made using the client configurator.

Simulation History

The following history would be displayed within PeopleStage for the R simulation shown previously:

Run 1 = Initialising an equal N-way test between ['A', 'B', 'C', 'D'] requiring a volume of 1,000

Run 2 = Test 2 batch 1
- Created new Champion Challenger test ["D"! v "E"] requiring volume of 357,146
  - Champion & Challenger from results of N-way test. ("D", "E", "C", "A", "B" = 1.46%, 1.43%, 1.35%, 1.06%, 1.03% from 10,000, 10000, 10,000, 10,000, 10,000)
  - Required volume based on: Expected response rate: 1.46%, Margin: 10.00%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
    - Expected response rate taken from history.
      - Using Champion's response rate of 1.46%
    - Margin based on historic response rates:
      - Margin of 2.05% capped at minimum value of 10.00%
      - Margin from difference between 0.0146 and 0.0143
      - Using Challenger's response rate of 1.43%
      - Using Champion's response rate of 1.46%
  - Previous test was complete (50,000 vs 1,000):
    - The Equal-N-Way test between ("A", "B", "C", "D", "E") exceeded the minimum volume requirement (50,000 vs 1,000).
    - The response rate so far (1.27%) is sufficient - over 0.25 of that expected (1.00%).

Run 3 = Test 2 batch 2
- Continuing with the current test ["D"! v "E"].
  - Insufficient sent volumes (50,000 vs 357,146):
    - "D": 40,000 out of 285,717 required (1.29% response rate so far)
    - "E": 10,000 out of 71,429 required (1.63% response rate so far)
    - The response rate so far (1.35%) is sufficient - over 0.25 of that expected (1.46%).

Run 4 = Test 2 batch 3
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (100,000 vs 357,146):
"D": 80,000 out of 285,717 required
(1.33% response rate so far)
"E": 20,000 out of 71,429 required
(1.47% response rate so far)
The response rate so far (1.36%) is sufficient - over 0.25 of that expected (1.46%).

Run 5 = Test 2 batch 4
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (150,000 vs 357,146):
"D": 120,000 out of 285,717 required
(1.29% response rate so far)
"E": 30,000 out of 71,429 required
(1.49% response rate so far)
The response rate so far (1.33%) is sufficient - over 0.25 of that expected (1.46%).

Run 6 = Test 2 batch 5
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (200,000 vs 357,146):
"D": 160,000 out of 285,717 required
(1.26% response rate so far)
"E": 40,000 out of 71,429 required
(1.42% response rate so far)
The response rate so far (1.30%) is sufficient - over 0.25 of that expected (1.46%).

Run 7 = Test 2 batch 6
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (250,000 vs 357,146):
"D": 200,000 out of 285,717 required
(1.28% response rate so far)
"E": 50,000 out of 71,429 required
(1.37% response rate so far)
The response rate so far (1.30%) is sufficient - over 0.25 of that expected (1.46%).

Run 8 = Test 2 batch 7
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (300,000 vs 357,146):
"D": 240,000 out of 285,717 required
(1.29% response rate so far)
"E": 60,000 out of 71,429 required
(1.37% response rate so far)
The response rate so far (1.30%) is sufficient - over 0.25 of that expected (1.46%).

Run 9 = Test 2 batch 8
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (350,000 vs 357,146):
"D": 280,000 out of 285,717 required
(1.31% response rate so far)
"E": 70,000 out of 71,429 required
(1.36% response rate so far)
The response rate so far (1.32%) is sufficient - over 0.25 of that expected (1.46%).

Run 10 = Test 3 batch 1
Created new Champion Challenger test ["D"! v "A"] requiring volume of 121,085
Existing champion stays on: No significant result: ("D", "E" = 1.30%, 1.36% from 320,000, 80,000).
Required volume based on: Expected response rate: 1.30%, Margin: 18.75%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 1.30%
Margin based on historic response rates:
Margin from difference between 0.0130 and 0.0106
Using Challenger's response rate of 1.06%
Previous test was complete (400,000 vs 357,146):
"D": 320,000 exceeds 285,717 required
(1.30% response rate)
"E": 80,000 exceeds 71,429 required
(1.36% response rate)
The response rate so far (1.32%) is sufficient - over 0.25 of that expected (1.46%).

Run 11 = Test 3 batch 2
Continuing with the current test ["D"! v "A"].
Insufficient sent volumes (50,000 vs 121,085):
"D": 40,000 out of 96,868 required
(1.30% response rate so far)
"A": 10,000 out of 24,217 required
(0.95% response rate so far)
The response rate so far (1.23%) is sufficient - over 0.25 of that expected (1.30%).

Run 12 = Test 3 batch 3
Continuing with the current test ["D"! v "A"].
Insufficient sent volumes (100000 vs 121085):
"D": 80,000 out of 96,868 required
(1.35% response rate so far)
"A": 20,000 out of 24,217 required
(0.95% response rate so far)
The response rate so far (1.27%) is sufficient - over 0.25 of that expected (1.30%).

Run 13 = Test 4 batch 1
Created new Champion Challenger test ["D"! v "B"]
requiring volume of 81,574
Champion "D" won again: ("D", "A" = 1.34%, 0.97% from 120,000, 30,000)
Required volume based on: Expected response rate: 1.34%, Margin: 22.89%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 1.34%
Margin based on historic response rates:
Margin from difference between 0.0134 and 0.0103
Using Challenger's response rate of 1.03%
Previous test was complete (150,000 vs 121,085):
"D": 120,000 exceeds 96,868 required (1.34% response rate)
"A": 30,000 exceeds 24,217 required (0.97% response rate)
The response rate so far (1.26%) is sufficient - over 0.25 of that expected (1.30%).

Run 14 = Test 4 batch 2
Continuing with the current test ["D"! v "B"].
Insufficient sent volumes (50,000 vs 81,574):
"D": 40,000 out of 65,259 required (1.37% response rate so far)
"B": 10,000 out of 16,315 required (1.09% response rate so far)
The response rate so far (1.32%) is sufficient - over 0.25 of that expected (1.34%).

Run 15 = Test 5 batch 1
Created new Champion Challenger test ["D"! v "C"]
requiring volume of 384,546
Champion "D" won again:
("D", "B" = 1.36%, 1.04% from 80,000, 20,000)
Required volume based on: Expected response rate: 1.36%, Margin: 10.00%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 1.36%
Margin based on historic response rates:
Margin of 0.55% capped at minimum value of 10.00%
Margin from difference between 0.0136 and 0.0135
Using Challenger's response rate of 1.35%
Using Champion's response rate of 1.36%
Previous test was complete (100,000 vs 81574):

"D": 80,000 exceeds 65,259 required
(1.36% response rate)
"B": 20,000 exceeds 16,315 required
(1.04% response rate)
The response rate so far (1.29%) is sufficient - over 0.25 of that expected (1.34%).

Run 16 = Test 5 batch 2
Continuing with the current test ["D"! v "C"].
Insufficient sent volumes (50,000 vs 384,546):
"D": 40,000 out of 307,637 required
(1.22% response rate so far)
"C": 10,000 out of 76,909 required
(1.36% response rate so far)
The response rate so far (1.24%) is sufficient - over 0.25 of that expected (1.36%)

Run 17 = Test 5 batch 3
Continuing with the current test ["D"! v "C"].
Insufficient sent volumes (100,000 vs 384,546):
"D": 80,000 out of 307,637 required
(1.29% response rate so far)
"C": 20,000 out of 76,909 required
(1.26% response rate so far)
The response rate so far (1.29%) is sufficient - over 0.25 of that expected (1.36%).

Run 18 = Test 5 batch 4
Continuing with the current test ["D"! v "C"].
Insufficient sent volumes (150,000 vs 384,546):
"D": 120,000 out of 307,637 required
(1.30% response rate so far)
"C": 30,000 out of 76,909 required
(1.19% response rate so far)
The response rate so far (1.28%) is sufficient - over 0.25 of that expected (1.36%).

Run 19 = Test 5 batch 5
Continuing with the current test ["D"! v "C"].
Insufficient sent volumes (200,000 vs 384,546):
"D": 160,000 out of 307,637 required
(1.31% response rate so far)
"C": 40,000 out of 76,909 required
(1.24% response rate so far)
The response rate so far (1.29%) is sufficient - over 0.25 of that expected (1.36%).

Run 20 = Test 5 batch 6
Continuing with the current test ["D"! v "C"].
Insufficient sent volumes (250,000 vs 384,546):
"D": 200,000 out of 307,637 required
(1.30% response rate so far)
"C": 50,000 out of 76,909 required
(1.23% response rate so far)
The response rate so far (1.28%) is sufficient - over 0.25 of that expected (1.36%).

Run 21 = Test 5 batch 7
Continuing with the current test ["D"! v "C"].
Insufficient sent volumes (300,000 vs 384,546):
"D": 240,000 out of 307,637 required
(1.28% response rate so far)
"C": 60,000 out of 76,909 required
(1.21% response rate so far)
The response rate so far (1.27%) is sufficient - over 0.25 of that expected (1.36%).

Run 22 = Test 5 batch 8
Continuing with the current test ["D"! v "C"].
Insufficient sent volumes (350,000 vs 384,546):
"D": 280,000 out of 307,637 required
(1.28% response rate so far)
"C": 70,000 out of 76,909 required
(1.21% response rate so far)
The response rate so far (1.27%) is sufficient - over 0.25 of that expected (1.36%).

Run 23 = Test 6 batch 1
Created new Champion Challenger test ["D"! v "E"] requiring volume of 406,967
Existing champion stays on: No significant result: ("C", "D" = 1.23%, 1.28% from 80,000, 320,000).
Required volume based on: Expected response rate: 1.28%, Margin: 10.00%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 1.28%
Margin based on historic response rates:
Margin of 6.72% capped at minimum value of 10.00%
Margin from difference between 0.0128 and 0.0137
Using Challenger's response rate of 1.37%
Using Champion's response rate of 1.28%
Previous test was complete (400,000 vs 384,546):
"D": 320,000 exceeds 307,637 required
(1.28% response rate)

"C": 80,000 exceeds 76,909 required
(1.23% response rate)
The response rate so far (1.27%) is sufficient - over 0.25 of that expected (1.36%).

Run 24 = Test 6 batch 2
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (50,000 vs 406,967):
"D": 40,000 out of 325,574 required
(1.24% response rate so far)
"E": 10,000 out of 81,393 required
(1.52% response rate so far)
The response rate so far (1.29%) is sufficient - over 0.25 of that expected (1.28%).

Run 25 = Test 6 batch 3
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (100,000 vs 406,967):
"D": 80,000 out of 32,5574 required
(1.25% response rate so far)
"E": 20,000 out of 81,393 required
(1.39% response rate so far)
The response rate so far (1.28%) is sufficient - over 0.25 of that expected (1.28%).

Run 26 = Test 6 batch 4
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (150,000 vs 406,967):
"D": 120,000 out of 325,574 required
(1.28% response rate so far)
"E": 30,000 out of 81,393 required
(1.29% response rate so far)
The response rate so far (1.28%) is sufficient - over 0.25 of that expected (1.28%).

Run 27 = Test 6 batch 5
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (200,000 vs 406,967):
"D": 160,000 out of 325,574 required
(1.26% response rate so far)
"E": 40,000 out of 81,393 required
(1.37% response rate so far)
The response rate so far (1.28%) is sufficient - over 0.25 of that expected (1.28%).

Run 28 = Test 6 batch 6
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (250,000 vs 406,967):
"D": 200,000 out of 325,574 required
(1.28% response rate so far)
"E": 50,000 out of 81,393 required
(1.40% response rate so far)
The response rate so far (1.30%) is sufficient - over 0.25 of that expected (1.28%).

Run 29 = Test 6 batch 7
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (300,000 vs 406,967):
"D": 240,000 out of 325,574 required
(1.29% response rate so far)
"E": 60,000 out of 81,393 required
(1.38% response rate so far)
The response rate so far (1.30%) is sufficient - over 0.25 of that expected (1.28%).

Run 30 = Test 6 batch 8
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (350,000 vs 406,967):
"D": 280,000 out of 325,574 required
(1.27% response rate so far)
"E": 70,000 out of 81,393 required
(1.39% response rate so far)
The response rate so far (1.29%) is sufficient - over 0.25 of that expected (1.28%).

Run 31 = Test 6 batch 9
Continuing with the current test ["D"! v "E"].
Insufficient sent volumes (400,000 vs 406,967):
"D": 320,000 out of 325,574 required
(1.27% response rate so far)
"E": 80,000 out of 81,393 required
(1.39% response rate so far)
The response rate so far (1.29%) is sufficient - over 0.25 of that expected (1.28%).

Run 32 = Test 7 batch 1
Created new Champion Challenger test ["E"! v "A"] requiring volume of 54,197
New champion "E":
("E", "D" = 1.39%, 1.27% from 90,000, 360,000)
Required volume based on: Expected response rate: 1.39%, Margin: 28.00%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 1.39%
Margin based on historic response rates:

Margin from difference between 0.0139 and 0.0100
Resetting Challenger's response rate of 0.99%
to minimum allowed value of 1.00%
Previous test was complete (450,000 vs 406,967):
"D": 360,000 exceeds 325,574 required
(1.27% response rate)
"E": 90,000 exceeds 81,393 required
(1.39% response rate)
The response rate so far (1.29%) is sufficient -
over 0.25 of that expected (1.28%).

Run 33 = Test 7 batch 2
Continuing with the current test ["E"! v "A"].
Insufficient sent volumes (50,000 vs 54,197):
"E": 40,000 out of 43,358 required
(1.30% response rate so far)
"A": 10,000 out of 10,839 required
(0.64% response rate so far)
The response rate so far (1.17%) is sufficient -
over 0.25 of that expected (1.39%).

Run 34 = Test 8 batch 1
Created new Champion Challenger test ["E"! v "B"]
requiring volume of 95,864
Champion "E" won again:
("E", "A" = 1.32%, 0.80% from 80,000, 20,000)
Required volume based on: Expected response rate:
1.32%, Margin: 21.17%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 1.32%
Margin based on historic response rates:
Margin from difference between 0.0132 and 0.0104
Using Challenger's response rate of 1.04%
Previous test was complete (100,000 vs 54197):
"E": 80,000 exceeds 43,358 required
(1.32% response rate)
"A": 20,000 exceeds 10,839 required
(0.80% response rate)
The response rate so far (1.21%) is sufficient -
over 0.25 of that expected (1.39%).

Run 35 = Test 8 batch 2
Continuing with the current test ["E"! v "B"].
Insufficient sent volumes (50,000 vs 95,864):
"E": 40,000 out of 76,691 required
(1.40% response rate so far)
"B": 10,000 out of 19,173 required
(0.93% response rate so far)
The response rate so far (1.31%) is sufficient -
over 0.25 of that expected (1.32%).

Run 36 = Test 9 batch 1
Created new Champion Challenger test ["E"! v "C"]
requiring volume of 379,930
Champion "E" won again:
("E", "B" = 1.37%, 1.05% from 80,000, 20,000)
Required volume based on: Expected response rate:
1.37%, Margin: 10.00%, Ratio: 4.0:1, Alpha: 0.05, Beta: 0.20
Expected response rate taken from history.
Using Champion's response rate of 1.37%
Margin based on historic response rates:
Margin of 9.57% capped at minimum value
of 10.00%
Margin from difference between 0.0137
and 0.0124
Using Challenger's response rate of 1.24%
Using Champion's response rate of 1.37%
Previous test was complete (100,000 vs 95,864):
"E": 80,000 exceeds 76,691 required
(1.37% response rate)
"B": 20,000 exceeds 19,173 required
(1.05% response rate)
The response rate so far (1.31%) is sufficient -
over 0.25 of that expected (1.32%).

Run 37 = Test 9 batch 2
Continuing with the current test ["E"! v "C"].
Insufficient sent volumes (50,000 vs 379,930):
"E": 40,000 out of 303,944 required
(1.41% response rate so far)
"C": 10,000 out of 75,986 required
(1.31% response rate so far)
The response rate so far (1.39%) is sufficient -
over 0.25 of that expected (1.37%).

Run 38 = Test 9 batch 3
Continuing with the current test ["E"! v "C"].
Insufficient sent volumes (100,000 vs 379,930):
"E": 80,000 out of 30,3944 required
(1.42% response rate so far)
"C": 20,000 out of 75,986 required
(1.29% response rate so far)
The response rate so far (1.40%) is sufficient -
over 0.25 of that expected (1.37%).

Run 39 = Test 9 batch 4

Continuing with the current test ["E"! v "C"].

Insufficient sent volumes (150,000 vs 379,930):

"E": 120,000 out of 303,944 required

(1.42% response rate so far)

"C": 30,000 out of 75,986 required

(1.25% response rate so far)

The response rate so far (1.38%) is sufficient - over 0.25 of that expected (1.37%).

Run 40 = Test 9 batch 5

Continuing with the current test ["E"! v "C"].

Insufficient sent volumes (200,000 vs 379,930):

"E": 160,000 out of 303,944 required

(1.42% response rate so far)

"C": 40,000 out of 75,986 required

(1.27% response rate so far)

The response rate so far (1.39%) is sufficient - over 0.25 of that expected (1.37%).