

Tests, analyses and optimizations

Tests are the key to improving the effectiveness of direct marketing. Tests should be always performed, and in a systematic manner. Lists are tested, but not only lists: creative initiatives and offers are also tested, thus fine-tuning the aim, until the best possible results are achieved.

In order to compare tests it is necessary for only one variable to differ from a control test; in fact, only in this manner, by comparing redemptions - the percentage of respondents out of what was sent out - is it possible to calculate the "redemption index", which indicates the % of increase or decrease in comparison with a control test.

A control test is normally a customary offer/creative initiative that has been proposed for years. Tests are carried out in addition to these initiatives, simultaneously with them.

Tests are performed to see whether a new subject line of DEM generates more openings, whether a gift in addition to an offer encourages orders and to what extent, whether new listings on the market are more effective, and so forth.

Then results are measured. For example, if a control test set = 100, has given a 1% redemption and a "new

creative" test has given 1.2%, then the redemption index of the "new creative" test is 120. That is, it has improved results by 20%.

Some basic rules for performing tests:

1. Tests should be sent out simultaneously.
2. Addresses are extracted at random from a list in order to establish the sample for a test - that is, by selecting 1 every n addresses.
3. When carrying out de-duplication with other directories and/or extensions, tests have priority 1, meaning that the double in another list/extension is eliminated.
4. The figures to perform a test must not be fewer than 5,000 addresses for mail posting (the ideal amount is 30,000), and fewer than 30,000 for DEM.
5. We stress again that only one test variable must be different from the control test.

We suggest a technique, known as matrix technique, that makes it possible, with a minimum of addresses, to test two variables (in this case, for example, lists and offers) and see their possible connections. Imagine starting off for the first time in the market with a mailing campaign.

We have 2 offers and 2 lists that overlap, as shown in the table below. We obtain 4 combinations. The

column/row with the total will also be analysed, with the advantage of working on a higher figure.

	List A	List B	Total submissions
Offer A	10,000	10,000	20,000
Offer B	10,000	10,000	20,000
Total submissions	20,000	20,000	40,000

Let us imagine hypothetical responses

	List A	List B	Total
Offer A	100	120	220
Offer B	50	77	127
Total	150	197	347

Redemption is calculated thus:

Responses/submissions * 100

	List A	List B	Total
Offer A	1%	1.2%	1.1%
Offer B	0.5%	0.77%	0.64%
Total	0.75%	0.99%	0.87%

Considering that general redemption is 0.87% with the index = 100, then we can calculate the test indexes: for example +38% when using List B in connection with Offer A. This obviously will be the winning

combination for the extension of the next campaign. Offer A significantly shows a 26% increase compared with the overall average and +71% over Offer B.

When the next campaign is launched, and having as foundations (control tests) Offer A and List B, we will test other variables: tests of new lists, new creativity tests, and seasonal tests. In the latter case, the only feature that should differ from the control test is the date of dispatching.

In addition to statistics, one needs to use a bit of craft. It is well known, in fact, that the three weeks before Christmas, and the weeks from July to mid-August are best avoided for direct marketing, and that the best months are September and January, but every product has its specific season, so this is just general advice. Other critical periods are election periods. Areas affected by earthquakes, flooding and so on are to be avoided.

Confidence interval

Excessively small tests widen the confidence interval that predicts the redemption that will be obtained with

an extension. Theoretically, with 30,000 postal submissions and a 1% redemption, the confidence interval of the extension will be between 0.89 and 1.11%. In 95% of cases, this is a valid forecast.

By hypothesis, if a test is repeated 11 times, for example, the 11 results that will be obtained will be distributed on a "normal/Gaussian" curve.

Let us imagine that out of 11 tests, we obtain 1% redemption 5 times, 0.9 and 1.1 twice, and 0.8 and 1.2 once. If we put these redemptions on paper, this draws a Gaussian curve.

This is the theory, but since we will not only perform 1 test, then statistics help us predict results with the following formula:

The interval (Int.) formula is $\pm 1.96 \sqrt{p * q / N}$

Where P is the redemption (1%), Q is non-redemption (99%) and N the size of the sample.

The confidence interval is given by the test redemption \pm Int.

The lower the number of tests the more this confidence interval widens, to the point of making any forecast doubtful. A test on 1,000 cases would mean a confidence interval between 0.4% and 1.6%.

The temptation to save money on sending tests undermines our ultimate goal, which is to predict results, and for these results to be as stable as possible. There are tables that calculate confidence intervals.

Another quick point: the lower the redemption, the more the interval narrows.

Sometimes it can be helpful to consult this table - which, on the basis of the expected redemption and an acceptable interval, shows us how large the sample should be:

In these examples we have worked on a confidence level of 95%, which means accepting 1 mistake every 20 times.

Expected redemption	Size of sample 10,000	Size of sample 20,000	Size of sample 30,000
0.3	+ - 0.11	+ - 0.08	+ - 0.06
0.5	+ - 0.14	+ - 0.10	+ - 0.08
0.7	+ - 0.16	+ - 0.12	+ - 0.10
1	+ - 0.20	+ - 0.14	+ - 0.12

The interval variation to be applied is in yellow. If we expect 0.30% of answers and we want a confidence interval between 0.24 and 0.36% then 30,000 is the sample size that we have to prepare.

When making forecasts one always considers the minimum point of the confidence interval. The reason

is concealed duplication*, which does not exist in tests, but does indeed exist when one extracts a list from a database for the extension. *) This is explained in the previous chapters of the White Paper.

So if a test on 30,000 cases obtains a 1.0% redemption, then the extension estimate will be 0.88% (i.e. $1 - 0.12$).

Due to the fact that one considers only the minimum point of the interval, with 97% confidence - that is, almost certainty - we will obtain this table for projections.

Redemption	10,000	20,000	30,000
0.3	0.19	0.22	0.24
0.5	0.36	0.40	0.42
0.7	0.54	0.58	0.60
1	0.80	0.86	0.88

The minimum point of the interval - that is, the forecast that is made in practice - is in yellow.

The confidence interval, as well as statistics, is the weak point of those who work in this sector; therefore, they are not very well known and are applied little. Too often we see that those who want to save money, those who try to take shortcuts, those who do not follow professional models are then bound to make mistakes.

Direct marketing techniques must be taken seriously.

That being said, let us remember that the entire theory described above must be accompanied by the rigor and transparency of those who set the samples, perform analyses, and then extract extensions, which must comply with the same parameters for test extraction, excluding the names used for the test.

Analyses

A test on external lists of 30,000 submissions makes it possible to perform analyses on one variation, or cross-referenced analyses. This statistical method is more than sufficient when using external lists, which usually have few variables. So, one can break down lists by geographical origin, gender, age and so on.

A precautionary test on 5,000 contacts only says whether to proceed with a more substantial test or to discard a list, given that there are others to choose from.

On a client list the "analysis" issue becomes more interesting, given that the following can be analysed: the date of the last action, the frequency of previous orders, and the amount spent. This approach is called RFM (Recency, Frequency, Monetary).

Other variables can also be analysed, such as the number of promotions received, the type of products purchased, the recruitment channel, the form of payment etc.

An RFM analysis makes it possible to detect highly reactive segments; thanks to analysis, segments with redemption indexes that are double the average can be identified.

When dealing with several variables, one can also perform a statistical analysis, which in jargon is called "scoring". Typically, a logistic function is one that is suitable for the study of redemption, which is a binomial variable based on 2 outcomes: response and non-response. But in practice one can also use a linear model. Let us make an example: Redemption is explained by 3 variables (x), and each variable will have its own weight/coefficient (a).

Here is the function that explains $\text{Redemption} = a_1x_1 + a_2x_2 + a_3x_3$

Firstly, the variables that are more strictly correlated to the answers - that is, the ones that explain the redemption - enter the model. Variables with flat models (indexes tending to an average of 100) do not explain anything and are discarded. If, for example, a variable such as gender shows men with index = 150 and women with index = 50, then in this case the variable enters the formula because it is a key index

ascrivable to two groups of considerable size (half of the sample is male and half female).

Each mode of variable that enters the model will have a weight approximately equal to the redemption index. Each variable will also have its own weight, which is given by to what extent it can explain redemption.

By means of example, let us imagine that three key variables enter the formula: gender - date of last purchase - form of payment, and that all the other variables "are thrown away." At this point, each name in the list will be assigned an extension score, a redemption index.

So Mr. Smith will be given 200 points because he is male, then 170 points because he is a recent customer, and 140 points because he pays by credit card. Mr. Smith's total score is 510. In this way we assign a score to all the other people on the list chosen for the extension.

Once the scores are assigned to all the names in the list, the latter list will be ordered by score, from highest to lowest.

In practice, we have ordered the list from the person with the highest likelihood of making an order to the one with the lowest likelihood.

All this serves to optimize a direct marketing campaign. If we think that the contact cost for each

mailing is about 1 euro, investments can easily be € 50,000-100,000 or more.

Sometimes it can be useful to do the opposite - that is, not to select a list, but to remove the indexes tending to zero.

Why perform scoring? For example, one can decide to send expensive mail only to the 20% of customers with the highest probability of response as indicated by a model, or to send a catalogue to only 50% of customers who score above the average, etc. etc. In brief, scoring is done in order to optimize promotional investments.

We can conclude by saying that the "scoring" model includes some more complex procedures, but for our purposes it is necessary to discuss a quick and intuitive example.

The customer database

Knowledge of customer databases leads to customer segmentation and to identify macro-segments such as:

1. Inactive
2. Dormant
3. Occasional
4. Regular
5. Heavy buyers

6. Advocates

For each category, tailor-made messages will be sent and ad hoc strategies will be adopted.

With inactive customers one can come to the conclusion to send a "last message" saying: "Given that for xx months you have not made any orders, this is the last offer and if you do not respond, you will be deleted from our database."

This issue is connected in some way to the right to be forgotten that it is one of the new rules of the new European law on data protection, which makes it compulsory for companies to delete from databases those who do not respond for a certain amount of time.

At the other end of the scale are the advocates, those who are so excited about our products that they endeavour to recommend and/or sell them to their friends and acquaintances.

One of the goals of those working on client lists is to have such knowledge of the database as to be able to identify the 20% of customers who generate 80% of earnings. Not all customers are worth the same and it is assumed that a company will adopt different strategies and ways of communicating, depending on the type of customer. Summing up, a customer database should be profiled, the RFM model is the most used model, and thereafter there is space for

creativity and for adopting different strategies for different segments.

External databases as support tool

Analyses can also be conducted with the help of data external to the clients database, by for example retrieving data from public records or data from an area, down to the smallest territorial unit, the micro-area.

As regards public records, for the B2B sector the most comprehensive one is undoubtedly the Chamber of Commerce record. In fact, by using a VAT number, up to 30 pieces of information can be gauged: turnover, business, phone, exports.

With regard to the analysis of micro-areas, the following concept applies: tell me where you live and I'll tell you who you are. From the geographical analysis of data, profiles can be gleaned.

Let us take a step back and remember that Italy is divided into regions and provinces and that inside provinces are municipalities (the smallest local government unit).

In Italy there are approximately 8,100 municipalities, and a municipality may include one or more hamlets.

The census defines micro-areas as a group of houses, that is, a set of often neighbouring houses, with 66 households on average.

In the case of a skyscraper a micro-area can correspond to one street and one house number, but in the case of scattered detached house, a micro-area can also be as large as 1 Km.

At the level of micro-area, data are public and thus they never match a specific person.

For Example: in micro-area 2123 we know there are 20 elderly people, 60 adults and 15 young people.

Other details include: professions, nationalities, square meters of houses, the year of construction, and so on.

Those who analyse and sell these data tend to make averages, so it will be said that in micro-zone 2123 the average size is 105 square meters (while in fact, there may be both studio flats and penthouses).

If handled with care, micro-area data can act as a beacon for direct marketing activities whose purpose may be for example, reaching families living in large houses (think gas suppliers).

It can happen that one already has a database available (for example, a publisher who has 2.5 million addresses). If the goal of marketing is to sell a publication for boys, knowing the micro-area where families with children are concentrated can be a guide to select the people who live in certain micro-areas from the 2.5 million records. So, micro-areas data

help focus on areas with a penetration rate of households with children that is higher than the national average. The probability of contacting families within the target is higher in comparison with using a random method.

The postal system, the cornerstone of classic mailing, also needs to be mentioned.

Postcodes are not unique to each municipality. The typical case is postcode 20090, which corresponds to 8 municipalities on the outskirts of Milan.

Moreover, large cities are divided into zones, whereby Milan has around 40 different postcodes.

The key to match files and external databases with territorial indicators are: the postcode, the description of the municipality, the ISTAT Statistics Institute municipality code, the census area code, x/y geo-coordinates. With this latter figure and with the use of mapping it is possible to identify, for example, those who reside up to 16 minutes from a point of sale.

Compared to other countries, such as Spain, it is not customary to indicate the door number in an address. So, it is more than enough to indicate a person's name, street, house number, and, of course, postcode, town and province abbreviation. The rest of the information

that is used in Spain (door, floor, block) is not used in Italy.

These data can be found in the land registry records, but de facto they are bureaucratic data that are never used by people in their daily lives.

Mail to the owner of Door 77 at Via Gulli 40 - 20147 Milan would be returned to sender as "unknown", whereas this practice is normal in Spain.

In conclusion, we mention the land registry, whose purpose is not to allow promotional mailings, but since it is a public record, it makes it possible to extract marketing profiles and/or enhance already existing databases that have been gathered via the explicit consent to receiving advertising material.