ESOMAR Events

Synthetic Data in Marketing Studies

Exploring the promise of generative AI and synthetic data

Samuel Cohen
Thomas Duhard

ESOMAR
Office address:
Burgemeester Stramanweg 105
1101 AA, Amsterdam
Phone: +31 20 664 21 41

Email: info@esomar.org
Website: www.esomar.org

Publication Date: September 2024
ESOMAR Publication Series Volume S416 Congress 2024
ISBN 92-831-0328-9

CONGRESS 2024

Copyright

Published by ESOMAR, Amsterdam,
The Netherlands
Edited by: Bel Kerkhoff-Parnell, PhD

About ESOMAR

ESOMAR is the champion of the insights and analytics sector. It is the business community for every sector professional. Founded in 1947, the global membership association is a network reaching over 40,000 professionals and 750+ companies in 130+ countries. We support individual and corporate members supplying or using insights by helping them raise ethical standards, grow the demand for insights, and improve its uses and applications by all decision-makers.

About ESOMAR Membership

ESOMAR is open to everyone, all over the world, who believes that high quality research improves the way businesses make decisions. Our members are active in a wide range of industries and come from a variety of professional backgrounds, including research, marketing, advertising and media.
Membership benefits include the right to be listed in the ESOMAR Directories of Research Organisations and to use the ESOMAR Membership mark, plus access to a range of publications (either free of charge or with discount) and registration to all standard events, including the Annual Congress, at preferential Members' rates.
Members have the opportunity to attend and speak at conferences or take part in workshops. At all events the emphasis is on exchanging ideas, learning about latest developments and best practice and networking with other professionals in marketing, advertising and research. CONGRESS is our flagship event, attracting over 1,000 people, with a full programme of original papers and keynote speakers, plus a highly successful trade exhibition. Full details on latest membership are available online at www.esomar.org.

Contact us

ESOMAR
ESOMAR Office:
Burgemeester Stramanweg 105
1101 AA, Amsterdam
The Netherlands
Tel.: +31 20 664 21 41

Email: info@esomar.org
Website: www.esomar.org

ESOMAR Events

Synthetic Data in Marketing Studies

Exploring the promise of generative AI and synthetic data

Samuel Cohen
Thomas Duhard

Introduction

With the rise of AI, quantitative research is on the brink of an exciting revolution. The initial tech-driven revolution in our industry was phone sampling in the 1980s. Later, in the early 2000s, the internet enabled the surge of digital sampling companies, further democratising research. Similar to previous breakthroughs, AI may revolutionise and accelerate the data collection process. This paper will explore the impact of AI technology on our industry, driven by AI-generated responses, also known as synthetic samples. The paper will show if synthetic data is already a viable alternative to boosting underrepresented samples, given the statistical guarantees it provides. Our entire industry is now being challenged to deploy AI solutions responsibly, mapping their limitations and use cases. We suggest a validation framework for doing exactly that when employing synthetic samples in quantitative research.

The paper will present the results of over 7,000 parallel tests using publicly available datasets from the Pew Research Center. It will also showcase a real use case of synthetic data: boosting a niche of swing voters for the European elections rolling study conducted by Ifop. These findings result from a year-long research project done in partnership with the research institute IFOP and the start-up Fairgen, managed by the authors and under the supervision of Professor Emmanuel Candes, the Barnum-Simons Chair in Mathematics and Statistics at Stanford University. While examining the validation process and careful deployment in real use cases, a few of the questions burning in everyone's minds will be answered:

- What are the limitations of synthetic data?
- Does synthetic data pose a reliability risk?
- Is synthetic data the latest breakthrough in data collection?

The promise

Our goal was ambitious: starting from real study data, we aimed to increase samples of low-incidence populations to conduct analyses previously deemed impossible. A multidisciplinary team was formed to evaluate and iterate until satisfactory performance was achieved continuously. The team comprised:

- IFOP Group: A renowned research institute ensuring test execution following best market research practices, contributing its proven experience in conducting high-quality studies and connecting these advancements to real customer use cases.
- Fairgen: A team of leading innovators in high-tech, responsible for developing the models behind synthetic sample generation and acting as the technological engine propelling our exploration into new frontiers.
- Professor Emmanuel Candès, Stanford University: Ensuring the statistical correctness of our approach, Professor Candès brought essential academic dimension and mathematical rigour to our exploration.

This team has worked tirelessly to evaluate and determine whether synthetic data can be used confidently by research providers. In the upcoming sections, you will understand this approach's limitations, challenges and benefits.

Laying the foundations for integrating synthetic data

This section of our paper explores the current challenges with data collection and the potential benefits of incorporating synthetic data into research methodologies. It also outlines synthetic samples' limitations and validation frameworks aimed at ensuring quality and reliability.

Challenges associated with quota sampling

Today's standard data collection practice relies on quota sampling, which uses various indicators to verify the representativeness of the data collected. Statistical tests, such as the chi-square test or t-test, are commonly used to quantify the accuracy of results, evaluate relationships between different variables and identify the most significant outcomes. These analysis tools are essential for generating reliable insights from surveys. These statistical tests measure the accuracy of results or the significance of differences, considering the sample size and the dispersion of responses. They are based on the assumption that the samples used are randomly selected and large enough to represent the studied population.

Finding the balance: Representativity versus random selection

The challenge of consumer research has always been a question of balance: How to survey enough people to represent the population while keeping collection costs and delivery times low? The quota method quickly established itself as the simplest and most effective solution to implement. It aims to replicate the demographic structure of the target population by setting quotas for certain characteristics (age, gender, etc.). Thus, instead of recruiting tens of thousands of respondents to be statistically representative of a national population like France, this method makes it possible to limit to a thousand respondents.

The quota method reduces the time and cost of data collection while obtaining results that can reliably be exploited to guide marketing strategies. Conversely, the collected samples are "biased" by the quota method and, by definition, are not randomly selected. From a purely theoretical, scientific standpoint, the statistical tests mentioned beforehand should not be used to violate the so-called i.i.d. assumption all the tests are based on. However, for decades, the marketing research industry has accepted the empirical rule that the statistical tests mentioned previously are valid on samples selected by the quota method. Many quantifiable results (e.g., accurately predicted election results and increased sales following an advertising campaign) have demonstrated over the years that this empirical approach is valid, and its use is now widely accepted. In summary, quota sampling must be used with caution to avoid selection biases: it is necessary to ensure that the recruited individuals do not present anomalies that affect the quality of their responses.

Data quality issues, ranging from fraud to poorly designed surveys, as previously addressed by ESOMAR, are critical when feeding AI models. Ensuring original, high-quality data is essential, as the accuracy and reliability of AI-generated responses are directly dependent on the integrity and quality of the input data.

It's hard to deliver granular insights with traditional methods

The design of a marketing study largely depends on balancing the number of people research providers can afford to survey (for economic or technical reasons) and the precision and/or granularity of the results that are wished to be delivered. Quite often, generating meaningful analysis of underrepresented subsegments poses a real challenge. So far, the industry has met this challenge mainly through oversampling or boosts. To focus on a subpopulation, the solution would be to survey only an additional sample of respondents from this subpopulation. However, this method has its drawbacks: it takes time, it is costly and it is sometimes impossible to implement because it can be difficult to find enough respondents for certain rare targets.

Benefits of synthetic samples: Economic gain and flexibility

In this context, generative AI technology (on paper) can offer a simple solution to several of the challenges inherent to quota sampling and open up new perspectives. By quickly generating additional samples, the problem of the cost and time of oversampling is addressed. It provides researchers with flexibility. All targets, even small ones, can be analysed without identifying them beforehand. New targets of interest that arise during the analysis or after a client request can be easily integrated. Finally, it can also enrich past data in accordance with the answers of the original collection date, which is impossible to achieve with conventional methods. However, like any new AI implementation, this method raises many questions about its credibility. How can we ensure the relevance and reliability of a synthetic data technology whose driving logic is difficult to evaluate?

Challenges working with synthetic samples

Integrating synthetic data into research workflows requires a comprehensive understanding of its reliability and applicability. In retrospect, over the course of integrating our synthetic data solution, several critical questions arose:

- How can the reliability of synthetic samples be measured? This involves formulating a validation framework that rigorously compares synthetic and real data. Key metrics must be monitored to ensure synthetic samples accurately reflect real-world data.
- How to manage and analyse these samples, objectively extracting insights, despite the invalidity of some traditional statistical methods? We know that some traditional statistical techniques are used empirically. How to integrate this additional layer that operates differently to purely random and sufficiently large samples demanded by statistical science?
- How can this technology be integrated into the market research value chain, replacing existing techniques or paving the way for new applications? Once the first two challenges are overcome, how can this technique be implemented at scale?
- What are the limits of this technology, and which best practices should be adopted? Determining the boundaries of synthetic data usage is crucial to preventing misuse and ensuring the technology's effective application.

Defining usage guidelines based on solid experience is finally the *sine qua non* condition for the correct appropriation of the tool by industry participants and clients.

Innovation and continuous improvement

Artificial intelligence has transformed all aspects of our professional and personal lives over the past year, with developments following at an astounding pace. We can be excited and optimistic or hesitant and pessimistic about its purpose, but as professionals, we cannot remain passive. As we often say, the tricky point with AI is that we have to build the plane while it flies. We strive to do this with AI-generated synthetic data in our industry.

Scope, parameters, and thresholds of synthetic boosters

We set guidelines derived through extensive quantitative and qualitative benchmarking to run synthetic boosts with quality guarantees. Firstly, we recommend boosting survey fields of 300 respondents and above. This allows the model to have enough data to understand patterns and correlations that will be leveraged to boost segments. Secondly, we recommend only boosting segments of 15% penetration and less. This allows the model to extract patterns from the rest of the population that will strengthen its understanding of the segment itself. For instance, when boosting 25 to 34 year olds, this ensures the model will be able to learn from other age segments such as 18 to 24 year olds and 34 to 39 year olds. Finally, even under these conditions, we recommend users to always leverage their prior knowledge about the study, questionnaire and audience itself to figure out whether a segment will be “boostable”. For instance, if a segment (part of the population of the study) is extremely different from the rest of the population, it will be hard for the model to leverage the other respondents to reduce its uncertainty about the segment itself. This is where agency partners have a role to play, due to their extensive experience in running and analysing fields.

Limitations and validation framework

There are two main methods to evaluate synthetic samples. The first method involves parallel testing, where historical data and samples are used to compare synthetic boosters with real boosters (additional real-world sample sets). The second method involves a qualitative evaluation of these samples within a project, along with careful statistical testing and confidence interval estimation to accurately assess data quality. The next section presents the results of an extensive series of parallel tests, followed by descriptions of other qualitative analysis methods. Upon review, after both qualitative and quantitative analyses of different synthetic datasets, we established the following limitations:

- The AI model must be trained on a minimum original field of 300 respondents;
- The segments being boosted should constitute less than 15% of the overall field;
- The performance benchmarks that follow adhere to these constraints.

Quantitative benchmarking: Over 7,000 parallel tests on the Pew corpus

Let’s explore the strengths and limitations of synthetic data by leveraging extensive benchmarks. Fairgen evaluated thousands of synthetic boosts on 40 Pew American trends panel datasets, with more than 10,000 samples, boosting categorical demographic columns like age and race across 7,316 segments (average size 48). Using 1,000 training samples per dataset, Fairboost achieved a mean 2.855 effective sample size (more details about this metric follow below), with smaller segments seeing up to 3.5 times quality boost over ground truth. This showcases synthetic data's value in low-data scenarios. These metrics show the predictive power of AI-generated respondents when used to augment real data, and represent the major breakthrough. At the aggregate level, AI can reliably predict human responses, subject to the limitations discussed below.

Good-quality synthetic samples must perform well on aggregated levels and individually, as answers must respect the questionnaire structure. These challenges are addressed in upcoming sections of this paper.

Introduction to parallel testing, the essential validation method

For synthetic boosters targeting hard-to-reach segments, it's recommended to collect responses using traditional methods including real-world boosts, which involve targeting specific segments for additional data collection. In parallel testing, the real-world boost is set aside while the generative model is trained on the main data collection. Synthetic data is then generated from this model and compared to the real boost to determine if they are statistically equivalent.

Determining the average boost factor

What is a boost factor? If, at aggregate levels, a synthetic boost is statistically equivalent to three times the amount of real data, we call it “3x”. Also, how do we measure this equivalency? It’s all about comparing error distributions. To make accurate inferences from data, we want it to correspond to the true population distribution as much as possible. Naturally, the data sampled when polling is a noisy estimate of the population distribution. The more samples we have, the more accurate our estimate of the population distribution. We measure the error of our data as the distance from a large holdout set, the latter serving as a proxy for the actual, unmeasurable population distribution. We refer to this holdout set as our ground truth. We then compare this error to a sequence of errors obtained from training sets of different lengths, which serve as a yardstick to indicate the value of the synthetic sample set in terms of real dataset size. The factor by which we must increase (or decrease) the sample size to achieve the same error as a given set of samples is called effective sample size (ESS). Hence, a subset with the same error achieved by the original sample has an ESS of one (1).

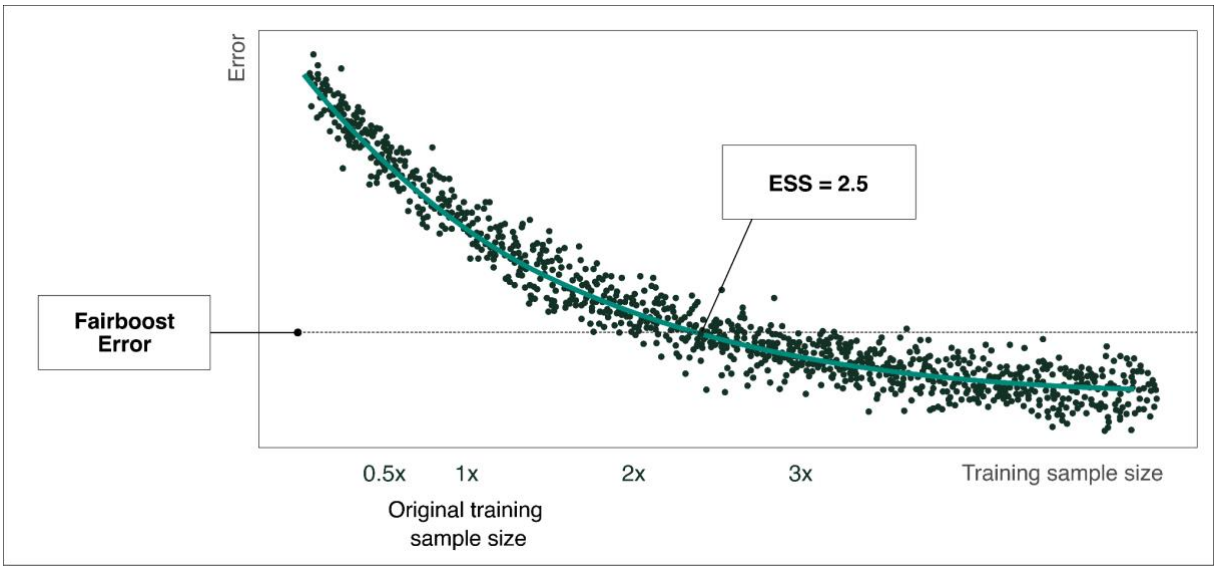


Figure 1

Fairgen averages the ESS for many segments of the data to get the ESS of a dataset. For our evaluation, we selected multiple columns containing demographic information (e.g., age, gender, etc.) to define the segments. Each segment is a subset of the training data for which a subset of the demographic columns

take on a specific set of values (e.g., women between the ages of 30 and 50). To constrain the total number of segments used in the evaluation, we limit our evaluation to segments defined by one or two demographic columns. Segments with extremely low support in the training data cannot be boosted. On the other hand, very large segments likely provide a good estimate of the ground truth and, therefore, do not require more respondents. Therefore, we only boost segments that comprise between 1% and 15% of the data. This intermediate range of segment size corresponds to situations in which a researcher would need to collect additional data.

For a robust quantitative evaluation, we use all the publicly available datasets from the Pew Research Center's American trends panel, which have at least 10,000 samples and include multiple demographic columns. This dataset corpus highly represents the market research use cases in which Fairgen can provide significant value. In total, we include 40 datasets in our evaluation.

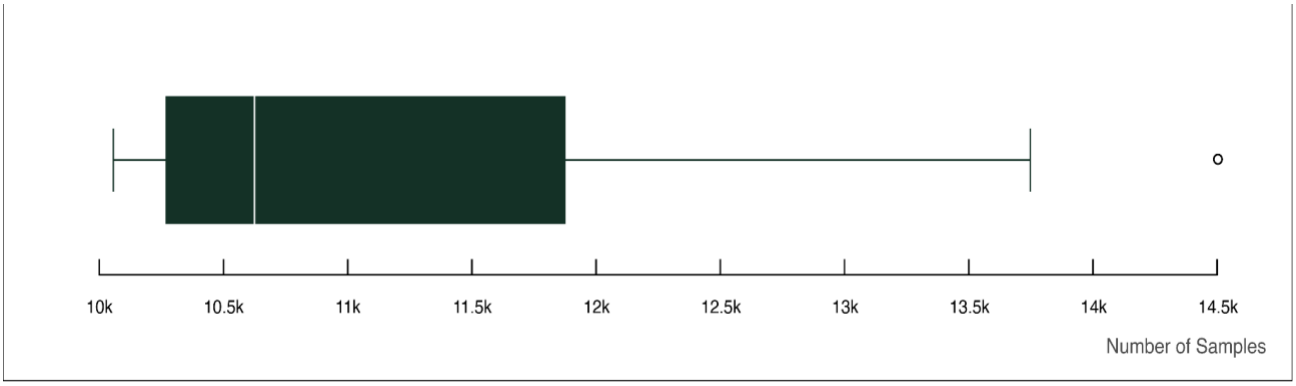


Figure 2

Fairgen's technology works with a wide range of column types. However, some columns should not (e.g., constant value columns) or cannot be boosted (e.g., ID columns). Fairboost, therefore, applies customised logic to these columns. As is often the case with survey data, most of the columns in the datasets are categorical. The demographics that define segments were manually selected based on the datasets' documentation. We selected between three and seven demographic columns for each dataset. Some examples of demographic columns are age, race, religion, gender and political party affiliation. In our evaluation, we used 7,316 segments with sizes in the range [10,150] with a mean of 48.26 and a standard deviation of 37.93. A histogram of the segment sizes is shown in Figure 3. Overall, on average, Fairgen boost can parallel the performance of boosts of between 3.5 times (for smaller segments) and 2.5 times (for larger segments). A Fairgen boost is almost always worth a boost of at least two times.

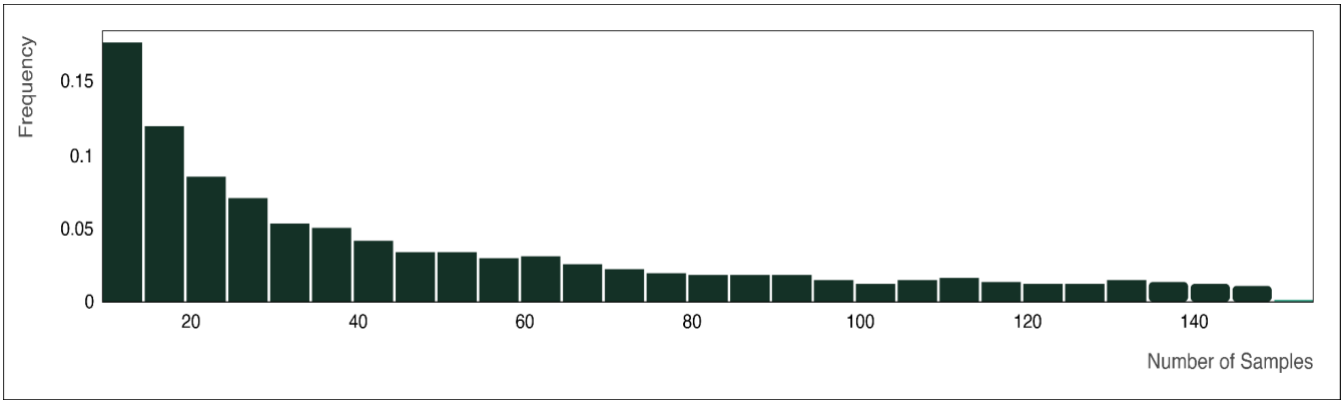


Figure 3

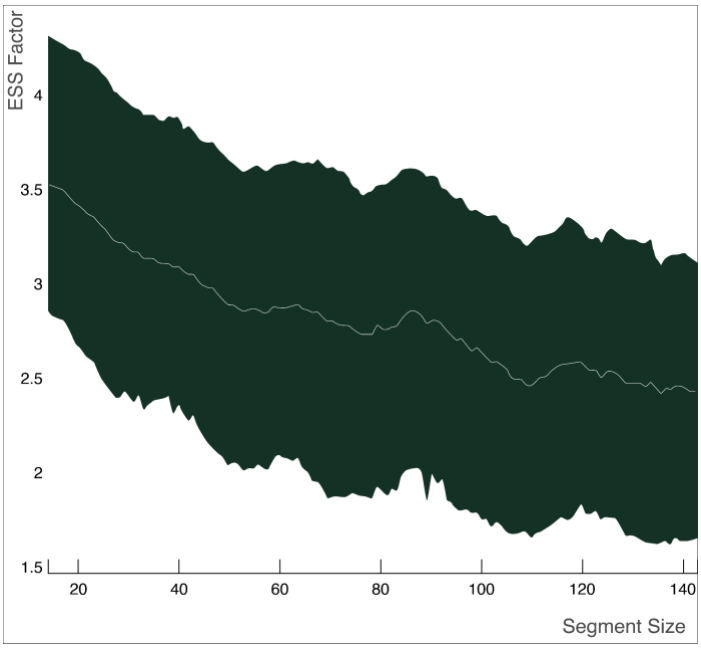


Figure 4

To build intuition around the opportunities Fairboost opens, we now zoom in on a specific Pew dataset where FairBoost's synthetic samples more closely resemble the ground truth distribution than the original training sample. We show this using a social media usage poll (Wave 112) conducted by the Pew Research Center, with 12,147 responses. The demographics selected for our evaluation on this dataset were race, religion, marital status, political party affiliation, education level and whether the respondent uses social media. The mean ESS for the Wave 112 dataset is 2.79, equivalent to increasing the number of respondents from 1,000 to 2,790. In the plots below (see Appendix A), we compare the distributions of the training sample, the FairBoost and the ground truth. We show our results in a grid where each column corresponds to a specific segment and each row corresponds to a specific question (e.g., the top right sub-plot shows the distributions of the question "Which do you prefer for getting the news?" for the segment of high school graduates with the marital status "divorced"). As seen below, the FairBoost often better approximates the ground truth distribution.

Qualitative benchmarking: Boosting a swing niche group for the European elections

Fairgen and Ifop have, for the first time, published election estimates using synthetic data in the scope of European elections. The rolling political study, conducted by Ifop, used AI to boost an underrepresented group of teachers considered a “swing group”. The survey was conducted with a representative sample of 8,000 people of the French population aged 18 and over, including the statistical equivalent of 580 secondary school teachers (derived from 116 interviews extrapolated using synthetic data). Estimating voting intentions requires a substantial number of interviews due to the high level of precision needed. Typically, we limit ourselves to conducting a national estimate of voting intention, from which we then attempt to deduce an estimate for sub-categories by simply filtering the national result for the sub-category (116 interviews out of 8,000 in our example). This method results in an estimate that is less precise as the size of the sub-category decreases.

In our example, we selected a sub-category (secondary school teachers), and boosted it using generative AI to achieve a precise estimate directly for this sub-category, similar to how we would conduct a national estimate of voting intention (see Appendix B). The result was then compared by our political scientists to:

- Previous publications concerning this population, which confirmed the sociological plausibility of the obtained results;
- The results of the sub-category before the generation of synthetic data, which showed that the inconsistencies in the raw results, were corrected by the AI-enhanced results;
- This latter point is crucial and reproducible in many cases: experience shows that generative AI often "corrects" observed inconsistencies in the initial samples.

Statistical benchmarking: Measuring confidence intervals and other qualitative analysis

IFOP has also implemented Fairgen’s technology to boost low-penetration products in an ongoing brand tracker for a consumer brand in the beauty space. The challenge was to increase confidence levels across the entire product portfolio, without recurring to overcostly bills per wave and reducing friction with its customer base. Once that project ran, it was of utmost importance to be able to run statistical tests validating the quality of insights.

Applying confidence intervals

However, a limitation of synthetic data is that statistical assumptions under t-tests and similar tests do not naturally apply; hence, they need to be done with care. We recommend two approaches. Either to be conservative in the N, i.e., the synthetic sample size, which will lead to conservative confidence intervals balancing with the fact that synthetic samples do not satisfy assumptions required for the central limit theorem to apply. Otherwise, Fairgen, in collaboration with IFOP as a design partner, has devised a methodology for estimating accurate confidence intervals from synthetic samples. This methodology consists of training K models, with resamples of the training data, and then computing confidence intervals by bootstrapping and using quantiles of the target estimate (means, difference between means coming from two segments or more). The key question for this methodology is whether a 90% confidence interval really covers true estimates of the quantity of interest 90% of the time (and similar for 95% and 99% intervals). To validate this, Fairgen ran the analysis on thousands of boosts, validating that 90% intervals indeed led to a 90% coverage of the ground truth estimates as expected.

What are the risks of synthetic data?

- Misinterpretation and misuse: Users of synthetic data must be well-versed in its limitations and appropriate applications. Synthetic data could be misused, leading to flawed insights and decisions. Clear guidelines and thorough training are essential to mitigate this risk.
- Integration challenges: Integrating synthetic data into existing research frameworks can be complex. Ensuring that synthetic data complements rather than conflicts with traditional quota sampling methods requires careful planning and execution, especially during post-processing. Failure to integrate properly can result in inconsistencies and reduce the overall quality of its predictions.
- Ethical and privacy issues: The generation and use of synthetic data must adhere to strict ethical guidelines. It should protect individual privacy and be communicated with transparency as part of the methodology.
- How to use this technology: A three-step method (with video feedback from L'Oréal).

We have just reviewed this technology's main principles, including its strengths and limitations. Now, we can propose a modus operandi for brands interested in using it.

Step 1: Choose the right study

As discussed, generative AI can be used for both recent and past studies, provided they include at least 300 high-quality interviews. For past studies, keep in mind that synthetic respondents are generated to "behave like" those who were interviewed at the time of the study (any exogenous events occurring since the study was conducted are not taken into account).

Step 2: Choose the right segment

Generative AI can enhance the statistical quality of segments representing 1% to 15% of the total sample. However, the feasibility of the analysis depends on several factors beyond segment size: the total number of interviews, the overall size of the study (number of questions asked) and the dispersion of responses (it will be more challenging to improve statistical quality if responses are too uniform). Moreover, it is not possible to boost questions posed solely to the chosen segment, as generative AI cannot rely on the responses of other segments to extrapolate results.

Step 3: Choose the right partners

It is crucial to rely on trusted partners to validate the relevance of using generative AI. As highlighted in Step 2, a thorough understanding of the study is essential from the client's side. However, the support of a marketing research professional is invaluable in overcoming most obstacles. Beyond technical efficiency, the choice of a technological partner should also consider ethical considerations: each study should result in the creation of a single generative AI model (no overlap between studies), the process should rely solely on real data (no use of exogenous data), data confidentiality must be guaranteed and no fees should be charged if the analysis proves to be unfeasible.

The formation of the trio "informed client" / "competent marketing research professional in AI" / "leading technological partner" will ensure a smooth process from conception to the final analysis of results, delivering reliable and impactful insights.

Concluding remarks

In this paper, we explored the ins and outs of synthetic data, specifically in the context of augmenting real data, and its potential to transform quantitative research. Similar to other technological breakthroughs, AI-generated responses have the potential to streamline data collection when it comes to granular insights. Through thorough validation and practical applications, we have demonstrated the validity of AI-generated responses when compared to real data, as long as done within the limitations of the technology and with close supervision to ensure a high-quality output. Fairgen's extensive benchmarking, involving over 7,000 parallel tests on the Pew corpus, has demonstrated that synthetic data can reliably augment real data, especially in low-data scenarios. The ESS metric we developed has shown that synthetic data can achieve boost factors ranging from 2.5 times to 3.5 times compared to responses collected via traditional methods. In practical applications, such as the Ifop study for the European election poll, synthetic data was used to study swing groups like secondary school teachers better. Confidence intervals were calculated for the first time for an ongoing brand tracker for a retail brand, and boosts were used to better understand low-penetration products without incurring prohibitive costs.

These findings reaffirm the viability of synthetic data in delivering accurate data at a fraction of the cost and time associated with conventional oversampling techniques. However, integrating synthetic data is not without its challenges. Ensuring the reliability of synthetic samples requires robust validation frameworks and continuous monitoring. Also, our extensive benchmarking showed that synthetic data is, for the moment, only useful to boost segments rather than augment entire fields. The intuition behind that is that the model leverages adjacent segments to boost the segment of interest, but it cannot "double" entire fields out of thin air. In the future, models trained on gazillions of surveys will potentially allow us to expand the scope from boosting segments to boosting entire fields.

The collaboration with Fairgen and IFOP has demonstrated that synthetic data is a viable and powerful tool in the arsenal of modern quantitative research teams. By addressing its limitations and leveraging its benefits, we are confident that synthetic data can unlock new levels of insight and drive the industry forward. How do you think data will be collected 10 years from now?

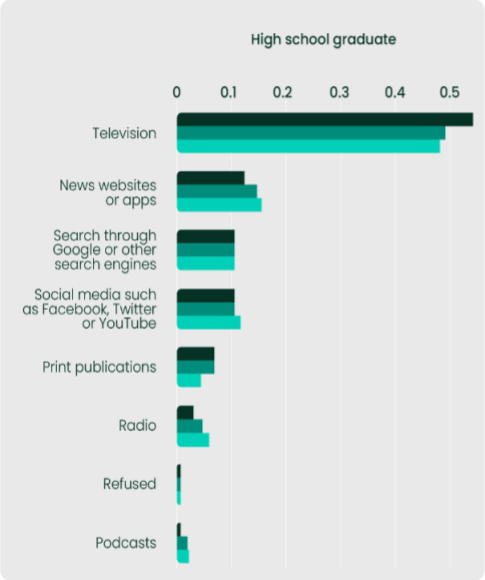
About the authors

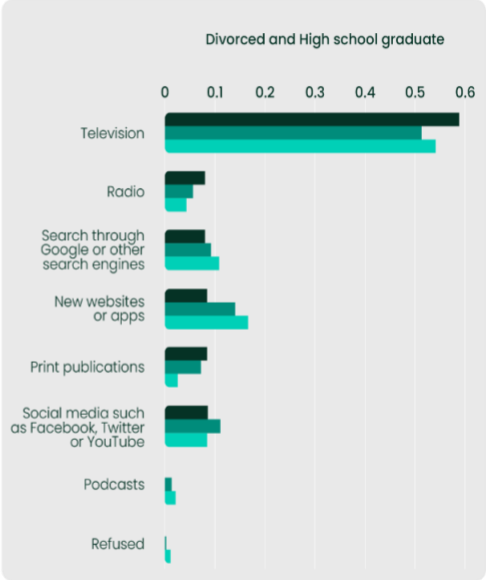
Samuel Cohen, CEO, Fairgen, Israel.
Thomas Duhard, Head of Data Projects, IFOP, France.

Appendix A

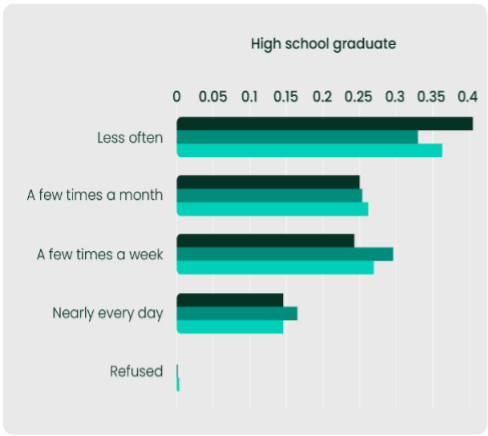
Training Sample FairBoost Ground Truth

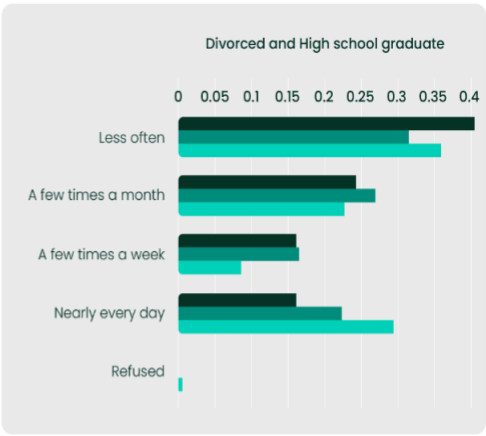
Which do you prefer for getting the news?



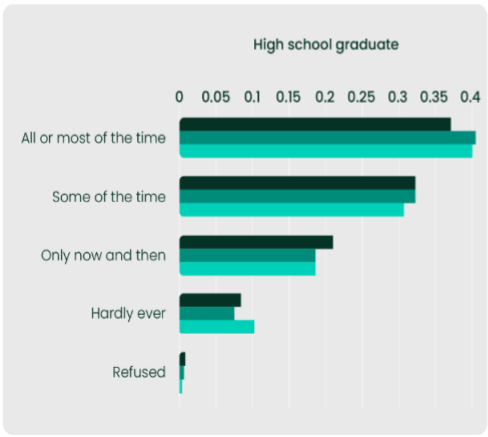


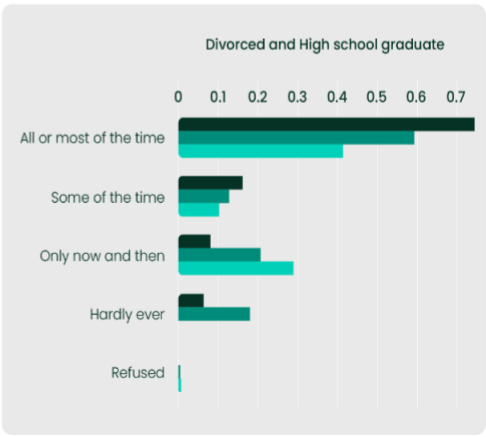
How often do you discuss government and politics with others?





Would you say you follow the news...





Appendix B

Comparison of European election voting intentions among teachers. Source: OpinionWay versus Ifop (May to June 2024)

Filtre : aux professeurs de collège et de lycée	Comparatif OpinionWay 15 au 17 mai[2] 2024 (%)	Ensemble des électeurs 24 mai au 4 juin 2024 (%)
La liste de LO conduite par Nathalie Arthaud	1	0,5
La liste du NPA-Révolutionnaires conduite par Selma Labib	1	0,5
La liste de la France insoumise conduite par Manon Aubry	9	7
La liste de la Gauche Unie du Parti communiste et du MRC conduite par Léon Deffontaines	2	2
La liste du Parti socialiste et de Place Publique conduite par Raphaël Glucksmann	25	32
La liste du Parti Radical de Gauche, de Régions et Peuples Solidaires et de Volt conduite par Guillaume Lacroix	-	-
La liste « Changer l'Europe » de Nouvelle Donne conduite par Pierre Larrouturou	-	-
La liste des Écologistes (ex-Europe Écologie Les Verts) conduite par Marie Toussaint	17	11
La liste « Écologiste au centre » conduite par Jean-Marc Governatori	3	-
La liste du Parti animaliste conduite par Hélène Thouy	2	1
La liste Ecologie positive et territoires conduite par Yann Wehrling	-	-
La liste de Renaissance, du Modem, d'Horizons et de l'UDI conduite par Valérie Hayer	11	13
La liste des Républicains conduite par François-Xavier Bellamy	5	8
La liste de l'Alliance rurale conduite par Jean Lassalle	1	-
La liste de Reconquête conduite par Marion Maréchal	5	5
La liste du Rassemblement National conduite par Jordan Bardella	15	20
La liste de l'UPR conduite par François Asselineau	-	-
La liste des Patriotes et de la Voie du Peuple conduite par Florian Philippot	-	-
La liste « Free Palestine » soutenue par l'Union des démocrates musulmans français et conduite par Nagib Azergui	-	-
Une autre liste	3	-
TOTAL	100	100