

**FAIRGEN**

**T Mobile**

**BiG VILLAGE**

# How we used augmented samples to optimize local media decisions



Samuel Cohen, PhD  
CEO & Founder @ Fairgen

Antoinette Sandy  
Sr. Manager, Data Science @ T-Mobile

Andy Davidson, PhD  
Chief Data Officer @ Bright Mountain

What is the future of data collection?

Will synthetic replace fieldwork?

Are all synthetic created equal?

# The future of research is hybrid

**AIRGEN**

Description	<b>100% Synthetic</b>  Synthetic respondents entirely sampled from AI trained on past surveys and public data.	<b>AI Boosts</b>  Augmented dataset that uses AI to boost real field collection based on real respondents only
Benefits	Ideal for ideation and brainstorming with cheap and immediate access to data	Allow confident reads of hard-to-reach groups and allows for shorter questionnaires
Technology	Large-Language Models	Statistical imputation models
Risk	Inaccurate insights due to hallucinations and bias; Lack of diversity and seasonality	Minimal risk

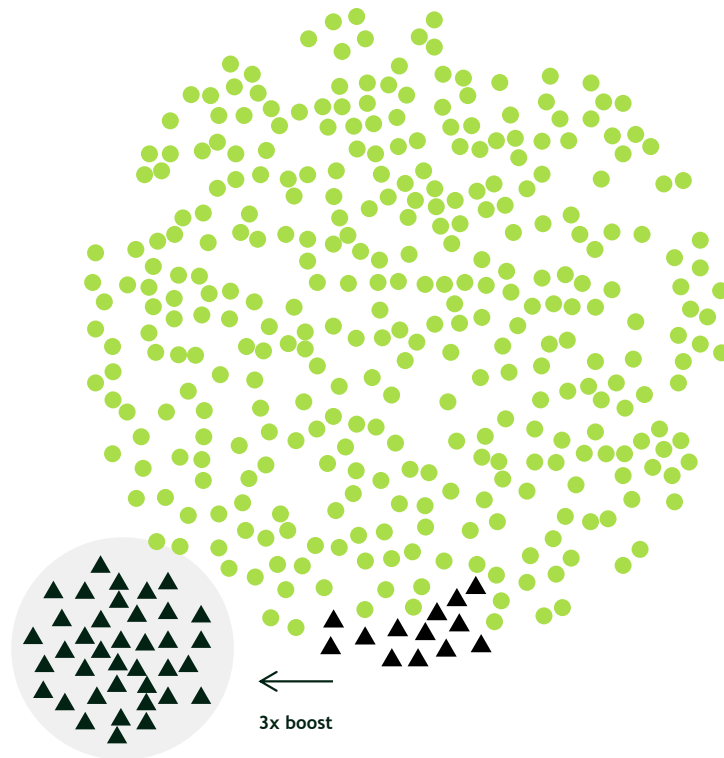
# Getting started with AI boosting

## What can be boosted?

- Any segment under 30% penetration.
- We mostly boost < n=150
- Part of a study with at least 300 respondents overall

## How does it work?

- AI learns patterns from adjacent segments
- Generate predictive responses with same statistical properties as if you had collected 3x more respondents



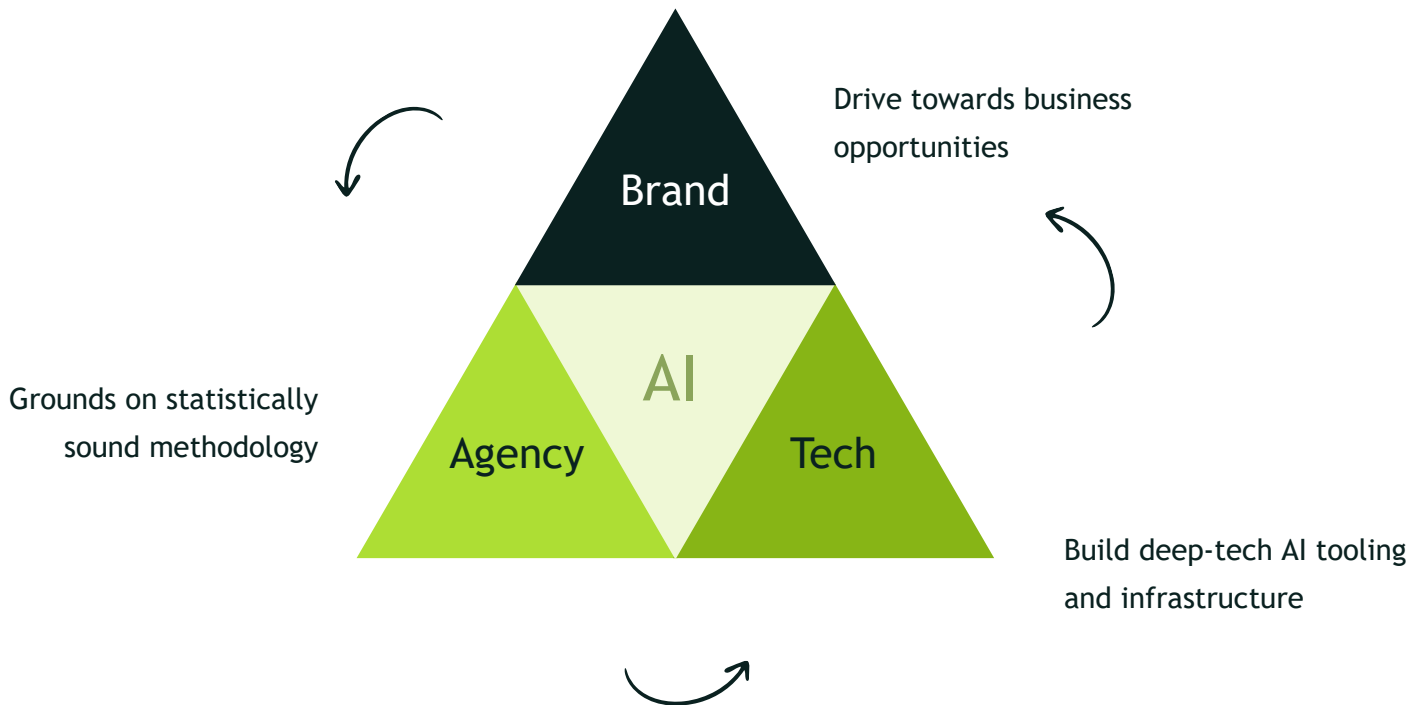
🔍 Improves error margins

👤+ Creates unique respondents

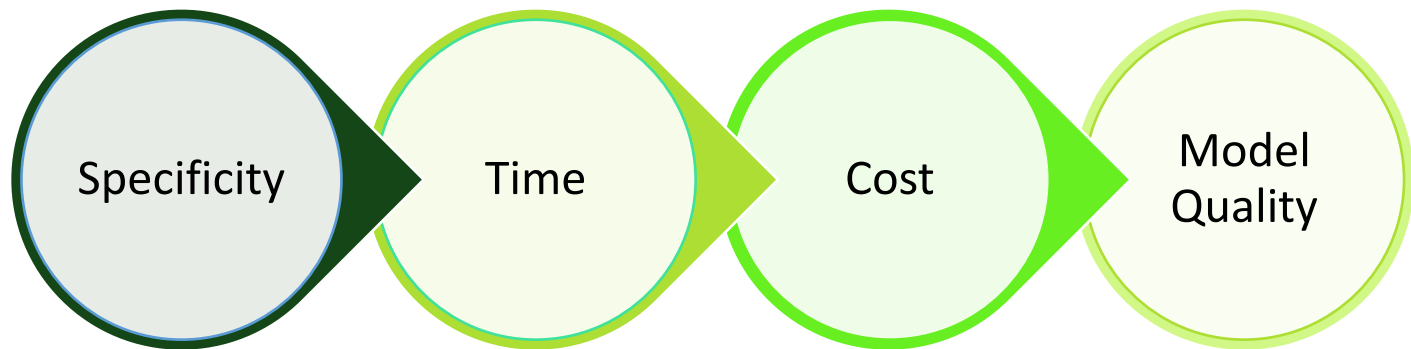
🌟 Statistical modeling

🗨️ Agnostic to use cases

# The triad for impactful AI

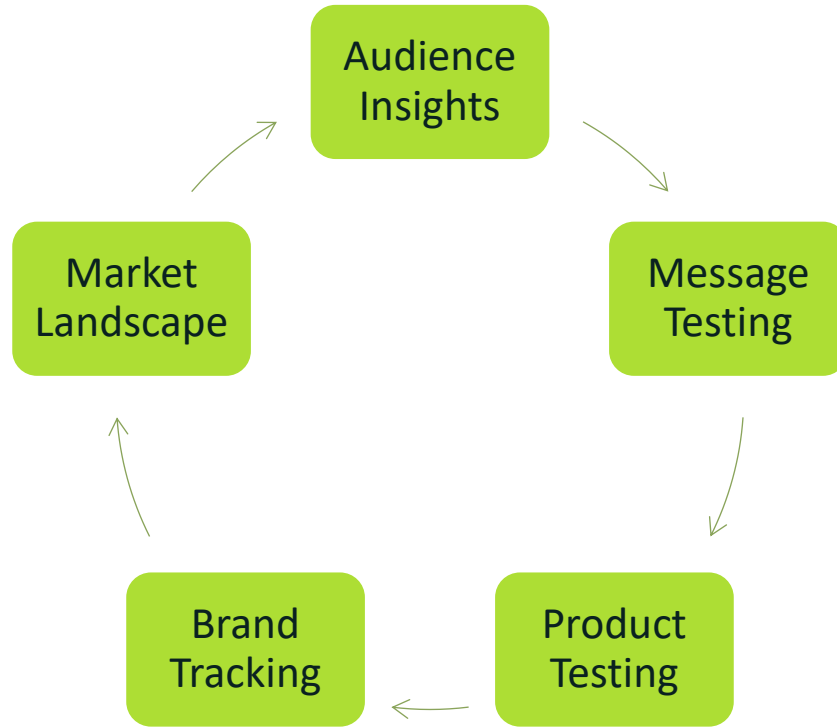


# Synthetic data solves age old problems for our clients



For us, the value is in the  
often *novel insights* a  
synthetic boost can generate

# Use cases continue to expand





# Why we needed a new approach:

## Our local measurement conundrum

### Our internal stakeholders needed:

- Fast, *affordable*, *granular* data
- *Trustworthy* inputs to guide marketing decisions

### Traditional Brand Measurement was:

- Too expensive to measure locally
- Slow and limited to National/Large market reads
- Not actionable for local marketing teams

### Synthetic sample is:

- Affordable to scale
- Faster and allows reads on niche target audiences
- Allows for real measurement and insights for local marketing teams

# “Fake data?” Not exactly...

## From synthetic data:

Initial stakeholder concerns:

- “Isn’t it fake?”
- “How accurate is this?”
- “Can I trust it to spend real money?”

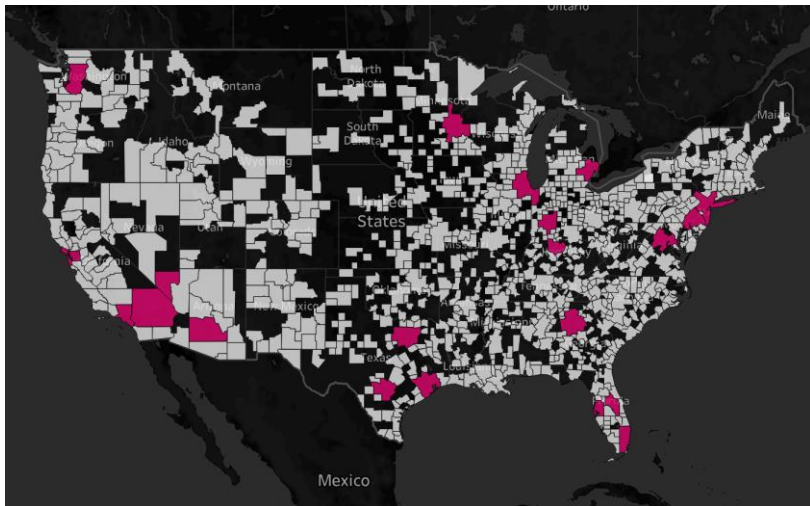
## To augmented sample

How we addressed them:

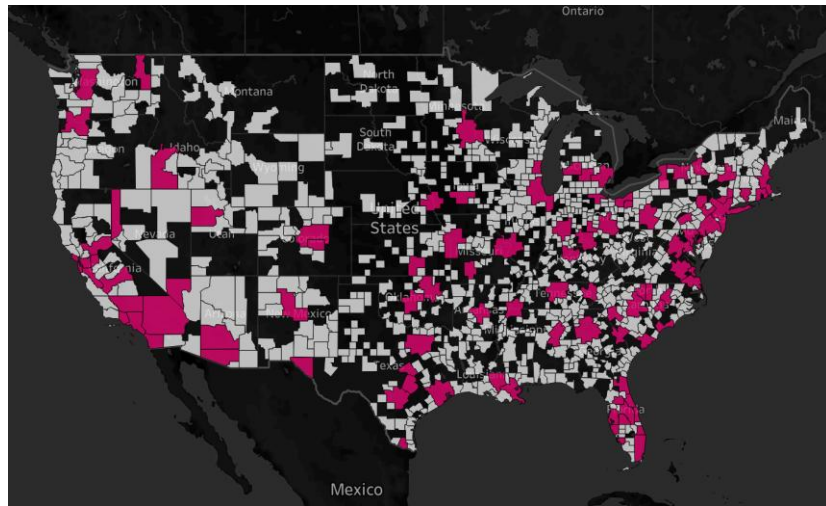
- “Modeled data is augmenting our survey data”
- “You can test the predictive strength of models”
- “We make marketing decisions on statistical models everyday”

# The unlock: what this makes possible

- Local-level Brand lift insights, even from a nationally balanced survey
- Smarter local media optimization
- Ability to run 'test & learns' on niche audiences



Insights on only **21** of the biggest markets  
**35%** of the US 18+ Population



Insights on **98** of the biggest markets  
**70%** of the US 18+ Population

**FAIRGEN**

**T Mobile**

**BIGVILLAGE**

# Thanks! Looking forward to connecting.



Samuel Cohen, PhD  
CEO & Founder @ Fairgen

Antoinette Sandy  
Sr. Manager, Data Science @ T-Mobile

Andy Davidson, PhD  
Chief Data Officer @ Bright Mountain



Scan to connect on **LinkedIn**