

APPENDIX G:

Visual Dartboarding + Synthetic Data

How to Use This Appendix

This appendix provides a practical walkthrough for using **Visual Dartboarding** in combination with Synthetic Data to drive alignment, improve AI performance, and build stakeholder trust. Whether you're developing a new AI model or exploring data strategies, this framework will help you *show* what's possible early, collaboratively, and clearly. Use it to engage diverse stakeholders, identify hidden risks, and co-create synthetic data strategies that are both technically sound and business-relevant. The following steps are designed to be iterative, visual, and inclusive, giving you a repeatable system for designing, testing, and refining synthetic data approaches with real-world input.

Visual Dartboarding + Synthetic Data: To get stakeholder feedback *on potential applications of synthetic data* and to co-create synthetic data generation strategies that best meet their needs.

Here's a Breakdown of How It Could Work:

- 1. Identify a Problem/Opportunity:** Let's say the problem is "Lack of sufficient real-world data to train an AI model for predicting high-risk insurance claims, particularly in underserved communities."
- 2. Create Initial Visualizations (Rough Drafts):** Instead of just *telling* stakeholders about synthetic data, show them. This could involve several visuals, created as rough drafts:
 - **Potential Synthetic Data Scenarios:** A flowchart showing how synthetic data could be generated to represent different demographic groups, claim types, and risk factors. Think very rough sketches with boxes, arrows, and handwritten labels.
 - **Data Quality/Bias Mitigation Visualization:** A simple chart (again, hand-drawn) showing how synthetic data can be generated to *correct* biases present in the real-world data. For example, if real-world

2 APPENDIX G: VISUAL DARTBOARDING + SYNTHETIC DATA

data under-represents claims from a specific community, the synthetic data generation process could oversample that community.

- **Use Case Visualizations:** Show *examples* of how the AI model, trained on synthetic data, could be used. This could be a user interface mockup (very basic) showing a claims adjuster using the AI to flag potentially high-risk claims. The key here is to visualize the *outcome* of using the AI.
3. **Present to Stakeholders (Visual Dartboarding):** Present these rough drafts to stakeholders (claims adjusters, risk managers, data scientists, community representatives). Follow the Visual Dartboarding principles:
- **Emphasize Roughness:** Make it clear these are just starting points.
 - **Ask Specific Questions:** Don't just ask "What do you think?" Ask:
 - "Are these the *right* scenarios to simulate with synthetic data?"
 - "What biases are *missing* from this visualization?"
 - "How would you *actually use* this AI model in your daily workflow?"
 - **Facilitate Interaction:** Encourage stakeholders to draw on the visualizations, add sticky notes, and suggest changes.
4. **Iterate and Co-Create:** Based on the feedback, refine the visualizations *and* the synthetic data generation strategy. For example:
- Stakeholders might point out that the initial scenarios don't adequately represent certain types of fraud. This would lead to generating more synthetic data focused on those fraud scenarios.
 - Stakeholders might identify new data features that are important for predicting risk. This would lead to incorporating those features into the synthetic data generation process.
 - Stakeholders might raise concerns about the AI model being used to discriminate against certain groups, even with bias mitigation efforts. This could lead to generating synthetic data specifically designed to test for and prevent discriminatory outcomes.
5. **Visualize the Results:** Once the synthetic data is generated and the AI model is trained, *visualize the results* to further build trust and understanding. This could include:
- **Performance Metrics:** Show how the AI model performs on both synthetic and real-world data, highlighting any improvements gained from using synthetic data.

- **Bias Analysis:** Visualize the results of bias detection tests, demonstrating that the AI model is fair and equitable.
- **Impact on Key Metrics:** Show how the AI model, trained on synthetic data, is improving key business metrics (e.g., reducing fraud, improving customer satisfaction).

Benefits of This Approach:

- **Increased Stakeholder Buy-In:** By involving stakeholders in the *design* of the synthetic data generation process and the *visualization* of the results, you increase their sense of ownership and trust.
- **More Relevant Synthetic Data:** Stakeholder input ensures that the synthetic data is relevant to their specific needs and concerns.
- **Better AI Models:** By training AI models on high-quality, representative synthetic data, you can create more accurate and reliable models.
- **Reduced Risk:** Synthetic data allows you to test AI models in a safe and controlled environment, without exposing sensitive real-world data.

By combining Synthetic Data and Visual Dartboarding, you can create a powerful and collaborative approach to AI development that leads to better outcomes and increased stakeholder adoption.

