Unwrapping Customer Delight

Optimizing Surprise Gift Strategies with Randomized Controlled Trials and MLRATE

The Estée Lauder Companies

BREAK THROUGH TECH

August 16, 2025

We're excited to be your Challenge Advisors!

Eddy El Chaar (he/him)
The Estée Lauder Companies
Staff Data Scientist
eelchaar@estee.com

Luis Saldaña (he/him)
The Estée Lauder Companies
Staff Data Scientist
lsaldana@estee.com

We're excited to have you as Students!

BREAK THROUGH TECH

- **Name**
- School affiliation/major
- AI-related and/or professional interests
- Fun or boring fact about yourself :)

Company overview

The Estée Lauder Companies (ELC) is a global leader in prestige beauty. We manufacture, market, and sell high-quality skin care, makeup, fragrance, and hair care products, and serve as a steward of consumer-beloved luxury and prestige brands globally.

- **Values: Generosity of spirit, fearless persistence, highest aesthetic standards, respect for the individual, uncompromising quality, and ethics & integrity**
- **ELC is headquartered in NYC, employs 60,000+ people globally, owns 20+ beauty brands, and sells products in ~150 countries and territories**
- **Fun fact: 81% of our workforce is made up of women!**

Business context

- Surprise gift campaigns can be used to enhance customer loyalty and drive sales
- Goal:
  - Understand quantitative impact of surprise gifts on future customer spending given simulated data
- Why?
  - Personalized customer experiences and data-driven marketing may be used to strategize business
  - Quantifying ROI for surprise gift campaigns can inform budget planning and help maximize impact on customer lifetime value
  - Provides insight into optimizing possible future gift-giving strategies

AI Studio Challenge Project Overview

CHALLENGE SUMMARY

This project challenges you to design and analyze a Randomized Controlled Trial (RCT) to determine the causal impact of surprise gifts on subsequent customer spending. You will leverage the Machine Learning Regression-Adjusted Treatment Effect Estimator (MLRATE) technique using Python to analyze the experimental data

YOUR TEAM'S OBJECTIVE

- Design and analyze an RCT to quantify the causal impact of surprise gifts on subsequent customer spending
- Employ the MLRATE technique to enhance statistical power and precisely quantify the gift's impact

DESIRED OUTCOMES

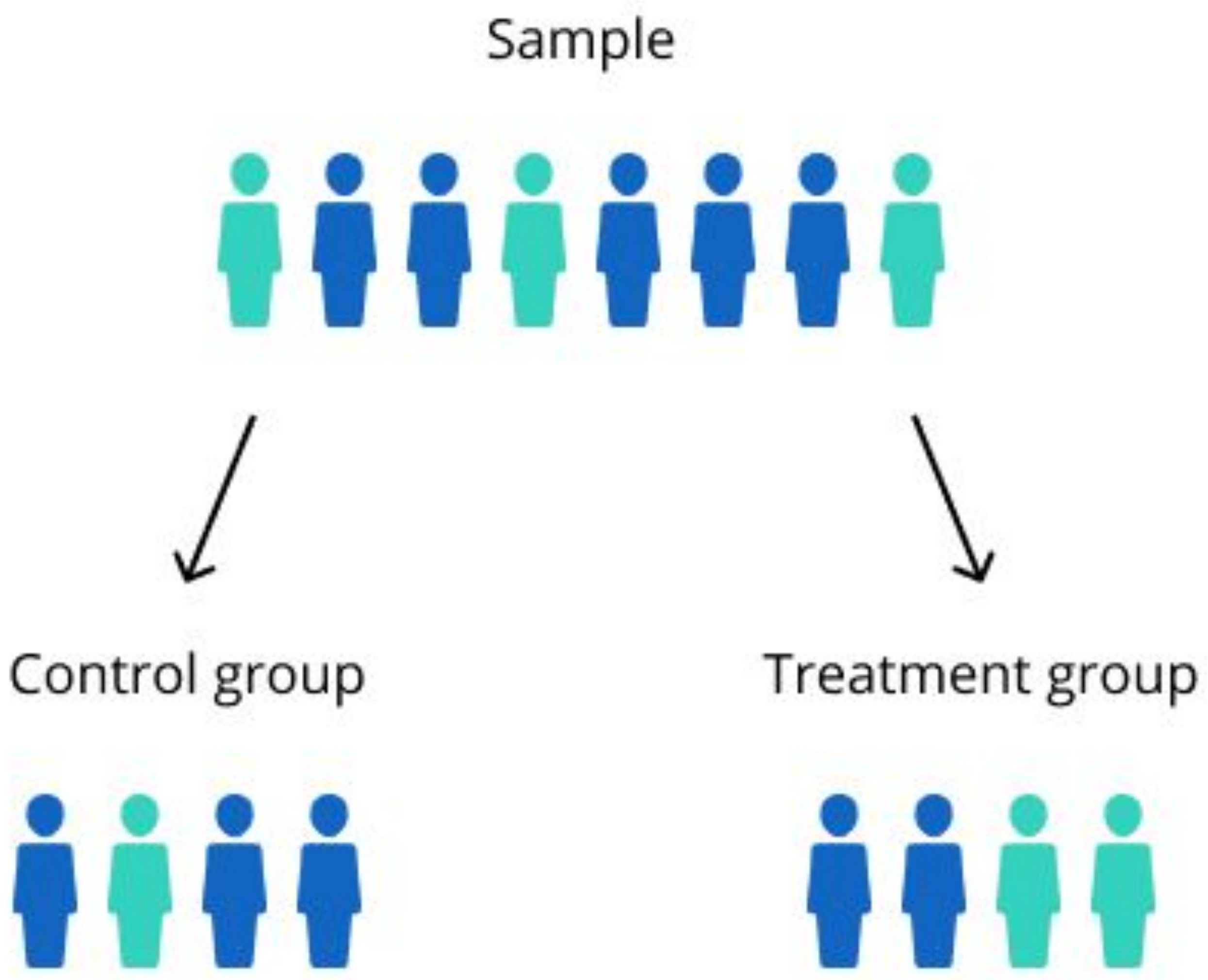
- A robust quantification of the return on investment for surprise gift campaigns
- Insights into optimizing future gift-giving strategies
- Comparison of estimated treatment effect to the simulated true effect

Suggested ML Approach

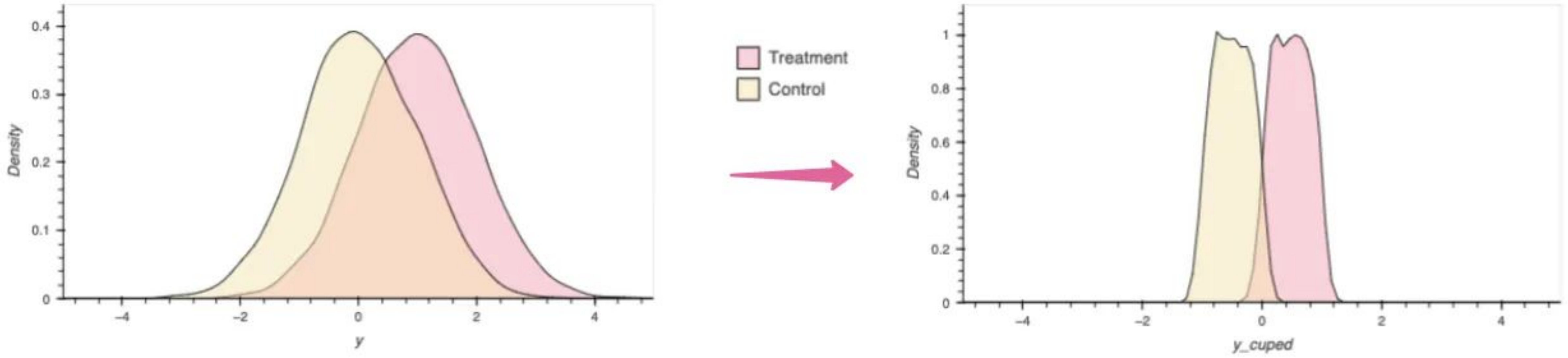
- Utilize a Randomized Controlled Trial (RCT) design
- Employ the MLRATE technique in Python
- Methodology:
  - Train an ML model (e.g., GBTs, ElasticNet) to predict post-intervention customer spending using pre-experiment covariates
  - Employ cross-fitting techniques to generate out-of-sample predictions
  - Use ML-generated predictions as control variates in an Ordinary Least Squares (OLS) regression model to estimate the Average Treatment Effect (ATE) of the surprise gift
- Outcome:
  - Enhanced statistical power of the analysis
  - Precise quantification of the gift's impact
  - Insights into optimizing future gift-giving strategies

Suggested ML Approach

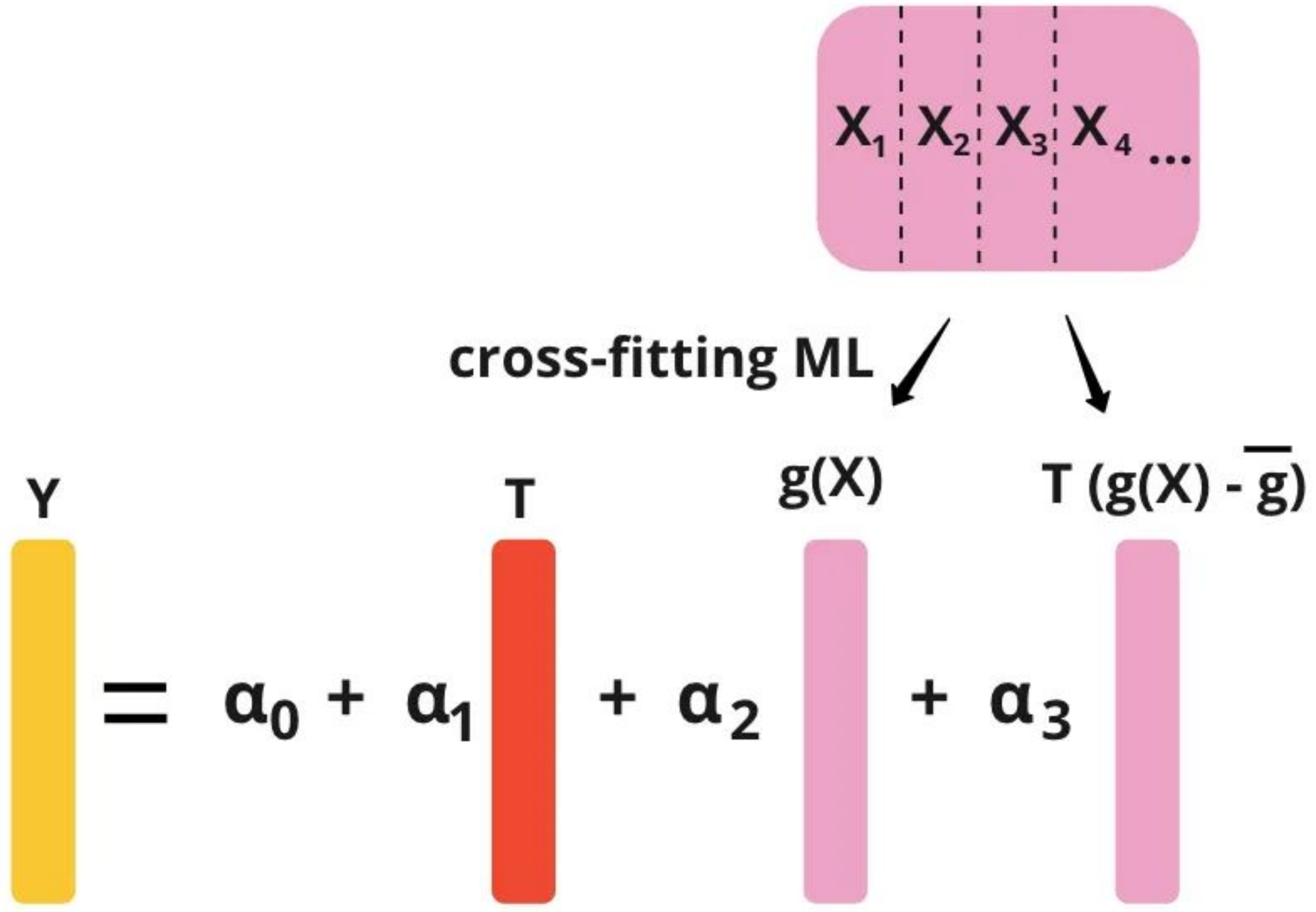
Randomized Control Trials*



Variance Reduction*



MLRATE*



Data overview

- We will provide you with simulated data in a Parquet file
- Type:
  - Pre-Experiment Data (for **Covariates X**):
    - Numerical: Customer ID, Total Prior Spending (e.g., last 12 months), Average Transaction Value, Purchase Frequency, Recency of Last Purchase
    - Categorical: Customer Tenure, Loyalty Program Status
  - Experiment Data:
    - Numerical: Customer ID, Post-Intervention Spending (**outcome Y**)
    - Categorical: Treatment Assignment (binary indicator: 0 for control, 1 for treatment)
- Format: Tabular, provided in one or more data files

Python libraries

- Frequentist regression modeling (OLS, GLS, etc): statsmodels, scipy (optional)
- ML models (for prediction): scikit-learn (for model types like GBTs, ElasticNet) or catboost or xgboost
- Visualization: Matplotlib, Seaborn (optional - has some nice standard statistical plotting tools built in)
- Data wrangling: Pandas, Polars

Suggest Dev and PM tools

- Your choice of personal development tools is welcome!
- Some suggestions:
  - Google Colab: hosted Jupyter notebook service with free access to computing resources (GPU, CPU). This IDE is conducive to team collaboration.
  - Github: dev platform for version control
  - Slack: cloud-based messaging app for team collaboration (chatting, file sharing)
    - Students: please create a team slack channel and add your respective Challenge Advisor!

Helpful resources

The following resources will help you get up to speed on Randomized Controlled Trials (RCTs) and Regression-Adjusted Treatment Effect Estimation (MLRATE). Your first task in this project will be to read these resources to develop an understanding of the concepts and how to implement them.

- Regression:
  - We assume you're familiar with linear regression (link: Chapters 7 & 8 may be helpful for review)
- Hypothesis testing, RCTs & Causal Inference (link: Chapters 16, 18 & 19):
  - Specifically:
    - 16.1 through 16.3
    - 18.1 through 18.4
    - 19.1 through 19.6
- Variance reduction in experiments: this Medium post
- MLRATE:
  - Original paper: here on arXiv (can ignore section 2.4, section 3, and appendices)
  - Others: this Medium post and this Towards Data Science post

Project milestones and timeline

These are the milestones for your Challenge Project. They include the CRISP-DM process steps you learned about in your ML Foundations course. In addition, there is an educational component in the front-end.

MILESTONE 1

- Review material from "Helpful resources" slide
- Q&A with challenge advisors

MILESTONE 2

- EDA on pre-experiment data
- Present EDA findings and data preprocessing

MILESTONE 3

- Power analysis
  - Determine the variance-reduced sample size for the upcoming RCT
  - Compare to standard t-test sample size

MILESTONE 4

- You will have received the experiment data
- Perform EDA on experimental data
- Present EDA findings and data preprocessing

MILESTONE 5

- Estimate ATE using two approaches:
  - First by using the standard t-test
  - Then by controlling for pre-experimental revenue

MILESTONE 6

- Estimate MLRATE
- Compare to:
  - Standard t-test ATE
  - Pre-experiment revenue controlled ATE

MILESTONE 7

- Deliver final project presentation: RCT, MLRATE, key findings
- Present quantified gift impact and campaign recommendations
- Final Q&A and project recap

MILESTONE #1 TASKS

Review Materials from "Helpful Resources"

- **Students must complete:**
  - **Literature Review**: Read/work through material in "Helpful Resources" slide
  - **Prepare Questions for Challenge Advisors**: Bring any questions to us for discussion
- **Meeting will cover:**
  - **Review Slides**: We will present a short summary of the concepts covered in resources. Do not take this as a lecture but rather a concise review, i.e please go over the literature in advance :)
  - **Data Sharing**: We will provide data file to students for exploration

How we'll work together this semester

Check-in meetings	• 1-hour meeting with respective Challenge Advisor, once every 2 weeks • Have everything outlined in milestone task completed
Reporting	• To be shared during check-in meetings, biweekly
Communication	• Use TA resources for technical help (python, libraries, code, collaboration tools) • Use the Break Through Tech Slack team channel to communicate with Challenge Advisors. Expect delay in response up to 48 hours. For technical and immediate help, please contact studio advisor first
Tools and platforms	• See "Suggested Dev and PM tools" slide
Challenge teams:	• **Estée Lauder 1A** ◦ **Angela Wang** ◦ **Selena Ho** ◦ **Nevaeh Clark** ◦ **Sahiti Srikakolapu** ◦ **Abubakar Diallo** ◦ **Kaylee Scanlin** ◦ **Rosina Zhou** • **Estée Lauder 1B** ◦ **Xiaoyan Li** ◦ **Ava Leung** ◦ **Sandy Wu** ◦ **Sehr Abrar** ◦ **Aden Garcia** ◦ **Nidhi Parvathala** ◦ **Lydia Aubourg**

Questions?

What questions do you have?

Anything I can help clarify?

What are you most excited about?

Anything you’re unsure about?