Milestone #5:

Estimating Standard ATE

Estee Lauder 1B

01 Code Implementation

Estimating Standard ATE

Code Implementation

```
import pandas as pd
import statsmodels.api as sm

# 1. define treatment and outcome variables
assignment = df["assignment"]        # 1 = treated, 0 = control
revenue = df["revenue (t)"]          # post-experiment revenue

# 2. add a constant term (intercept)
X = sm.add_constant(assignment)
y = revenue

# 3. fit ols regression with robust standard errors
model = sm.OLS(y, X).fit(cov_type="HC0")

# 4. print regression summary
print(model.summary(xname=["const", "treatment (gift sent)"]))

# 5. extract results
ate = model.params["assignment"]
ci_lower, ci_upper = model.conf_int().loc["assignment"]
ci_width = ci_upper - ci_lower
p_value = model.pvalues["assignment"]

print(f"\nestimated ate: ${ate:.2f}")
print(f"95% ci: [{ci_lower:.2f}, {ci_upper:.2f}] (width = {ci_width:.2f})")
print(f"p-value: {p_value:.4f}")
```

Code Results

```
                          OLS Regression Results
==============================================================================
Dep. Variable:            revenue (t)   R-squared:                       0.011
Model:                            OLS   Adj. R-squared:                  0.011
Method:                 Least Squares   F-statistic:                     62.66
Date:                Sat, 25 Oct 2025   Prob (F-statistic):           2.94e-15
Time:                        21:49:29   Log-Likelihood:                -25545.
No. Observations:                5556   AIC:                         5.109e+04
Df Residuals:                    5554   BIC:                         5.111e+04
Df Model:                           1
Covariance Type:                  HC0
=========================================================================================
                          coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------
const                 117.9520      0.453    260.258      0.000     117.064     118.840
treatment (gift sent)   5.1015      0.644      7.916      0.000       3.838       6.365
==============================================================================
Omnibus:                     1111.304   Durbin-Watson:                   2.009
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             2882.493
Skew:                           1.083   Prob(JB):                         0.00
Kurtosis:                       5.785   Cond. No.                         2.62
==============================================================================

Notes:
[1] Standard Errors are heteroscedasticity robust (HC0)

estimated ate: $5.10
95% ci: [3.84, 6.36] (width = 2.53)
p-value: 0.0000
```

02

Reflections

Estimating Standard ATE

1. What does the estimated ATE suggest about the campaign's impact?

The campaign had a **positive** and **significant impact** on post-experiment revenue. In other words, the **gift-sending** campaign **effectively increased customer spending** by about **$5 per person on average**.

2. Is the CI too wide or narrow? Does it include 0? What are the implications of this for business decision-making?

- The **CI is narrow**, which means the **estimate** is more **precise**.
- The interval [3.84, 6.36] so it doesn't include 0. **The treatment had a positive effect on post experiment revenue.**
- Sending the gift **increased the revenue by about $5.10 per customer**. So then it's important to look at the cost of the gift and if its less than $5.10 and profitable.

3. What is the p-value for the treatment coefficient? Does it provide strong evidence to reject the null?

- **Extremely Low P-Value**: Rounded to 0.0000, actual value ≈ 0.0000000000000245 → essentially zero.
- **Implication**: If the null hypothesis were true, the probability of $5.10 revenue increase occurring by random chance is practically nonexistent.
- **Statistical Significance**: $p < 0.001$ → strong evidence to reject the null hypothesis.
- **Treatment Effect**: Gift treatment significantly increases customer revenue.
- **Business Impact**: $5.10 increase per customer is meaningful for overall revenue.

4. How might sample size or variance have influenced your results?

- **Large Sample Size** (5,556 customers): Increases precision of estimated effects by reducing random differences in spending.
- **Narrower Confidence Intervals**: More observations lower standard errors, making estimates more reliable.
- **Higher Statistical Power**: Small effects, like the $5.10 revenue difference, are easier to detect.
- **Robust Standard Errors**: Help account for uneven variance in customer spending between groups.
- **Effect of Sample Size & Variance**: Smaller samples or higher spending variability would widen confidence intervals and reduce significance.

Thank you!

Notebook Link