

# OPIM 5510 Final Project

Team : **Pratik More**

## (1) Introduction

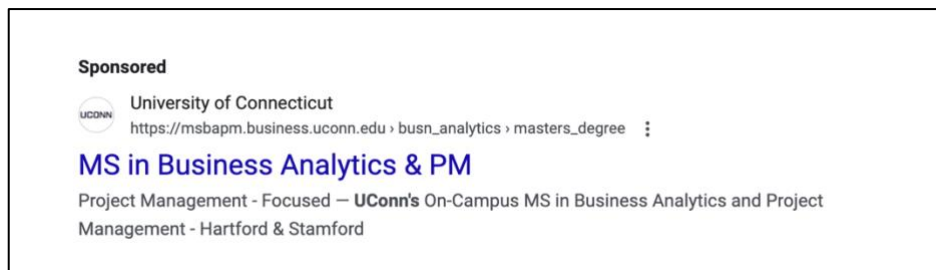
### Business context:

The University of Connecticut's (UConn) MS in Business Analytics and Project Management (MSBAPM) attracts both domestic and international students, each drawn to different aspects of the program. UConn's Google Search Ads are key to driving traffic and applications, and the program aims to test if focusing on analytics over project management can boost clicks and interest from international students while slightly increasing domestic engagement. The goal is to optimize ad messaging for targeted segments, improving marketing efficiency and attracting more applicants. With rising global demand for analytics skills, international students may respond better to analytics-focused ads, with a potential click increase among domestic students as well.

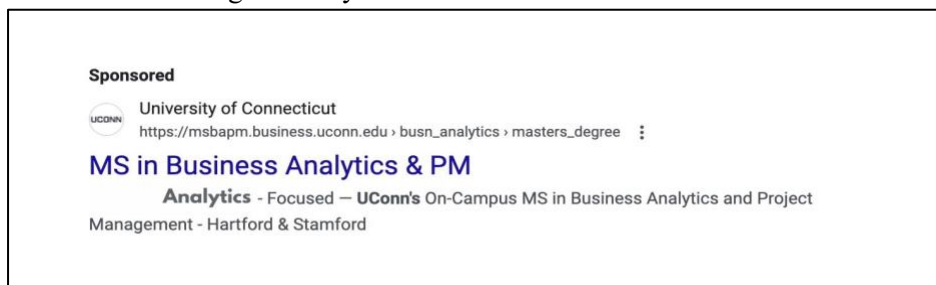
### Web strategy idea:

The University of Connecticut's MS in Business Analytics Program aims to attract global talent through targeted Google Search ads. We propose testing two ad versions—one focused on "Project Management" and the other on "Analytics"—to measure the impact of content focus. The objective is to see if international students, especially from analytics-driven regions, show greater interest in the analytics-focused ad compared to the project management-focused one.

Below is the original ad focusing on Project Management:



Below is the new ad focusing on Analytics:



## (2) Your hypothesis:

We hypothesize that changing the ad content focus from "Project Management" to "Analytics" will increase the number of clicks and interest among foreign prospective students, while keeping domestic interest constant.

- **Null Hypothesis (H<sub>0</sub>):**

There is no difference in the number of clicks between the "Analytics Focused" and "Project Management Focused" ads for both domestic and international students.

- **Alternative Hypothesis (H<sub>1</sub>):**

The number of clicks for domestic as well as international students will be higher with the "Analytics Focused" ad than with the "Project Management Focused" ad.

### (3) Experimental design and data collection

Description of the design:

**Test Setup:**

1. **Ad Variations:**

- **Version A:** Ad focused on Project Management
- **Version B:** Ad focused on Analytics

2. **Target Audience:**

- **Segment A:** Domestic students residing in the USA.
- **Segment B:** International students residing outside of the USA.

3. **Data Collection:**

- We will conduct a randomized control experiment involving 100 students (50 Domestic and 50 International). Each segment will be evenly split:
  - 25 Domestic students exposed to Version A (Project Management focus).
  - 25 Domestic students exposed to Version B (Analytics focus).
  - 25 International students exposed to Version A (Project Management focus).
  - 25 International students exposed to Version B (Analytics focus).
- Each student will be shown the assigned ad version for a one-month period, and their click behavior will be recorded.

**Experiment Design:**

This is an A/B test where we aim to observe how different ad focuses (Project Management vs. Analytics) resonate with Domestic and International students.

- **Metric to Track:** Total ad clicks, segmented by "Student Type" (Domestic or International) and "Ad Version" (Project or Analytics).
- **Expected Outcome:**
  - Higher engagement (more clicks) from International students for the Analytics-focused ad.
  - Similar click behavior across both ad versions for Domestic students.

Description of the data (make clear the unit of analysis):

The data collected in this experiment represents the total number of ad clicks by each individual student over a one-month period.

**Unit of Analysis:**

- **Individual Student:** Each row in the dataset corresponds to one unique student who participated in the experiment.

**Data Columns:**

1. **Clicks:** This column captures the total number of times a student clicked on the ad that was shown within the one-month testing period.
2. **Student Type:** Indicates whether the student is "Domestic" (residing in the USA) or "International" (residing outside of the USA).
3. **Ad Version:** Represents the type of ad shown to the student, either "Project" (focused on project management) or "Analytics" (focused on business analytics).

**Grouping and Segmentation:**

- The data is segmented by **Student Type** (Domestic vs. International) and **Ad Version** (Project vs. Analytics), allowing for comparison across these segments.

Include variable definition and summary statistics in Appendix A

(4) Data analysis, result interpretation and recommendations

Regression model:

$$\text{Clicks} = A + B \text{ Student type} + C \text{ Ad version} + D (\text{Student type} * \text{Ad version}) + \text{Error}$$

Definitions of variables in the regression (make it clear what is your key independent variable and what is the moderator):

- **Clicks (Dependent Variable):** Total clicks on the ad per student over the month.
- **Student Type (Independent Variable):** Domestic (USA) or International (outside USA) wherein 0 being Domestic & 1 being International in the regression model.
- **Ad Version (Key Independent Variable):** Project-focused or Analytics-focused ad wherein 0 represents Project and 1 represents Analytics in the regression model.
- **Student Type \* Ad Version (Moderator):** Interaction term to see if the effect of ad type on clicks varies by student type.

Explain how to interpret the results:

The interaction term (Student Type \* Ad Version) indicates whether the effect of the ad version on clicks differs between Domestic and International students.

After deduction, we can deduce that,

Ad version effect on Clicks = **C** .... (For Domestic Students)

Ad version effect on Clicks = **C+D** .... (For International Students)

Domestic Students: The ad version effect (C) is significant ( $p < 0.05$ ) with a coefficient of 4.920, indicating that the Analytics ad increases clicks by 4.920 compared to the Project ad for Domestic students.

International Students: The combined effect (C + D) is also significant ( $p < 0.05$ ) with an estimated increase of 6.08 clicks, showing that the Analytics ad boosts clicks more for International students compared to the Project ad.

Recommendations:

- Prioritize the **Analytics-focused ad version** across both domestic and international campaigns to maximize engagement.
- **Optimize the landing page** to reinforce the Analytics-focused messaging for a cohesive experience.
- Highlight **analytics coursework details** prominently on the landing page.
- Showcase **partnerships with analytics-driven companies** to increase credibility and appeal.

Include the screenshot of R output in Appendix B.

If you run a linear hypothesis test, explain the test in the table above and provide the screenshot of R output in Appendix C.

## Appendices

- A. Variable definition and summary statistics table (for categorical variables, specify the number of categories and list the category values in their definitions)

Variable	Definition	# of observations	Mean	Standard Deviation	Min	Max
<b>Clicks</b>	Represents the total number of times a student clicked on the ad shown to them during the one-month experiment.	100	7.41	3.98	0	14
<b>Student Type</b>	Indicates whether the student is residing in the USA or outside the USA.  Categories: <ul style="list-style-type: none"> <li>Domestic: Students residing in the USA.</li> <li>International: Students residing outside the USA.</li> </ul>					
<b>Ad Version</b>	Specifies the type of ad shown to the student.  Categories: <ul style="list-style-type: none"> <li>Project: Ad focused on Project Management.</li> <li>Analytics: Ad focused on Business Analytics.</li> </ul>					

- B. Screenshot of the regression result in R

```

> m1 <- lm(Clicks ~ as.factor(StudentType) + as.factor(Adversion) + as.factor(StudentType):as.factor(Adversion), data=Data3)
> summary(m1)

Call:
lm(formula = Clicks ~ as.factor(StudentType) + as.factor(Adversion) +
    as.factor(StudentType):as.factor(Adversion), data = Data3)

Residuals:
    Min       1Q   Median       3Q      Max
-5.52  -2.53   0.48   2.48   4.56

Coefficients:
              (Intercept)              4.8800      0.5777      8.447 3.19e-13 ***
as.factor(StudentType)International -0.4400      0.8170     -0.539  0.591
as.factor(Adversion)Analytics      4.9200      0.8170      6.022 3.16e-08 ***
as.factor(StudentType)International:as.factor(Adversion)Analytics  1.1600      1.1555      1.004  0.318
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.889 on 96 degrees of freedom
Multiple R-squared:  0.4885,    Adjusted R-squared:  0.4726 
F-statistic: 30.57 on 3 and 96 DF,  p-value: 5.866e-14

```

- C. Screenshot of linear hypothesis test result in R if there is any

```
> linearHypothesis(m1,"as.factor(Adversion)Analytics + as.factor(StudentType)International:as.factor(Adversion)Analytics")
```

Linear hypothesis test:  
 $\text{as.factor(Adversion)Analytics} + \text{as.factor(StudentType)International:as.factor(Adversion)Analytics} = 0$

Model 1: restricted model  
 Model 2: Clicks ~ as.factor(StudentType) + as.factor(Adversion) + as.factor(StudentType):as.factor(Adversion)

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	97	1263.12				
2	96	801.04	1	462.08	55.378	4.234e-11 ***

---  
 Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

D. Figures/charts/tables that you referred to in the report, if there are any

E. References if there are any

ChatGPT Prompt that we used to generate human like responses for our dataset

This is an experiment on the search result shown on 'ms in bapm' to which uconn sponsored ad is shown.

There are two ad campaigns. The current one in which project management is highlighted and the new one that is being tested in which analytics is highlighted.

We ran this experiment on total 100 students for a month long. 50 random students residing in the USA which we call it as Domestic students. Out of 50, 25 random students will be shown the current ad and rest 25 students will be shown the new ad for a month. We will note down for total how many times they click on the ad which they were shown for a month.

Similarly, 50 random students living outside of the USA which we call it as the International students were equally divided for current ad and new ad. In same fashion, we will note down the total number of clicks by these 50 students that we observe for a month long.

Finally we get a list of 100 students. 25 domestic and 25 international were shown current ad. Similarly, 25 domestic and 25 international were shown new ad. There will a column next to them representing their respective total number of click on the ad shown.

I want you to generate data for this experiment for which you need to be totally unbiased.

The columns in the table will be, 'Clicks' which shows total number of clicks by a student in the month, 'Student Type' which will either Domestic or International as there are 50 of both, and lastly the 'Ad version' which represent Project ad shown or Analytics.

Please remember one important point.

While generating data go row by row in the sequence I mentioned. When you are generating random clicks value for a student, you will first forgot what you had entered for the previous student and then generate a number for the current row. You need to keep your biases aside and just rely on the information given. While generating the number you need to forgot that this is an experiment and provide real world values and think like actual students. I want human like response.