**DS 4002** – Instructor: Loreto Alonzi

**Project Topic:** Trends in Admissions Characteristics of Universities from 2014-2023

**Target Audience:** 2nd-year UVA student

**Goal Statement:** The student will produce a **final analysis report** based on a validated logit regression model capable of predicting the use of admissions criteria with an **AUC > 0.75** and identifying up to **three admissions criteria** that exhibit statistically significant ($p<0.05$) shifts in influence following the COVID-19 pandemic, if they exist. From this, the students should be able to properly identify which admissions criteria have shifted because of the Covid-19 pandemic at universities in the United States.

| Spec Category | Task | Criteria for "Meets Spec" |
|---|---|---|
| **I. Project Goal** | The student will replicate the key preprocessing, modeling, and evaluation steps outlined in the case study's technical documentation to determine if admissions criteria importance shifted post-2020. | The student has successfully completed the full analysis pipeline, including data merging, cleaning, feature engineering, and logit regression analysis. |
| **II. Data & Preprocessing** | The student must merge the yearly IPEDS ADM and HD files (2014-2023). They must clean the data, handle missing values, and engineer the **PostCOVID** binary feature (1 for 2020 and later, 0 otherwise). | * Data is successfully merged and cleaned according to the provided script.<br><br>* The **PostCOVID** feature and binary criterion indicators are correctly engineered.<br><br>* The unit of analysis is correctly maintained as the **institution-year**. |
| **III. Modeling & Analysis** | The student will run a series of logit regression models for various admissions criteria (e.g., SAT, GPA, Essay) as the dependent variable. They must include institutional characteristics (CONTROL, REGION) and the **PostCOVID** variable as key independent variables. | * Logit regression models are correctly estimated with **clustered standard errors at the institutional level**.<br><br>* Coefficients for the **PostCOVID** variable are correctly interpreted in terms of odds or probability shifts.<br><br>* Statistical significance is evaluated using a **5% level (alpha = 0.05)**. |
| **IV. Evaluation & Presentation** | The student must evaluate the models using **AUC** and report the pseudo-$R^2$. They must create at least one time-series | * The final model has a reported **AUC**. |

| | | |
|---|---|---|
| | plot and a coefficient plot to visualize the trends and the post-COVID effects. | * At least one admissions criterion with a **statistically significant shift ($p < 0.05$)** post-COVID is identified, if one exists.<br><br>* High-quality visualizations clearly illustrate the shift in admissions criteria importance over the 2014–2023 period. |

Acknowledgements: Special thanks to Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled from Streifer & Palmer (2020). I also want to thank Gianluca Guadagni, who I learned this format of rubric from.