

# **Data-driven approaches for legacy analysis and prospection**

- Candidates: 38837, 42507, 40781
- LSE supervisor: Tengyao Wang
- Industry partner: LSE PAGE

# Contents

1. Overview of the project

2. Research questions

3. Methodology & plan

**4. Update: Dataset At a Glance**

5. Gantt plot

# Overview of the project

## 1. Legacy marketing programme at LSE

- Focuses on guiding prospective donors from an initial inquiry about leaving a legacy to making a firm commitment to support the institution.
- Aims to ensure that marketing efforts are effectively targeted, enabling PAGE to allocate resources rationally.

## 2. Problems being addressed

- Identify the characteristics of legacy prospects so that PAGE can map them across different stages of the legacy pipeline to adjust their contacting strategies effectively.
- Identifying High-Potential Prospects so that PAGE can enlarge the impact of their efforts.
- Develop a data-driven framework for allocating resources to maximize donor engagement and legacy commitments

# Overview of the project(ctn'd)

## 3.Relevance

- This project is highly relevant to LSE's strategic goals, as legacy giving not only supports immediate financial needs but also strengthens the institution's long-term financial health and ability to fund various events.
- Data-driven approach can benefit with wiser decisions, enhanced efficiency of the legacy marketing programme, and measurable outcomes, which is the growing trend in philanthropy.
- The insights generated through this project will provide a foundation for building stronger relationships with alumni and fostering a culture of giving, ensuring that LSE can continue to thrive and advance its mission of excellence in education and research.

# Research Questions

## Initial questions:

- Based on historical data, what are the main characteristics of donors? How are these characteristics related to whether people become donors or not?
- What is the variability in the characteristics of donor over time? Are they frequently changing so the model needs to adapt to such variations? What are the causes of these changes?
- These questions need to be investigated in 2 aspects: in the legacy pipeline as a whole and at each stage of the legacy pipeline.

# Research Questions

**Hypotheses** - there are several features that could affect whether alumni could join the legacy pipeline:

- Personal wealth: could be reflected on the programmes they studied, the field they have been working in
- Connection to LSE (willingness to get in touch with LSE after graduation): could be reflected on the frequency of participation in LSE alumni events
- Experience in LSE: if they enjoyed their journey at LSE, they are more willing to donate to and support LSE
- Peer impact: if many of their peers, especially who are close friends, join in the legacy pipeline, they are more likely to join as well

# Methodology & plan

Based on the project specification, 3 different types of modelling techniques will be considered.

## **1. Predictive models (to predict whether an alumni is likely to become a donor or not.):**

Logistic regression, decision tree (such as CHAID), and random forest.

## **2. Factor analysis (to detect potential factors influencing alumni participation in the legacy pipeline):**

Methods may include Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA)

## **3. Cluster analysis (to group alumni into groups with different possibility to join the pipeline):**

Approaches may include but not limited to: hierarchical clustering, k-means clustering, model-based clustering, and grid-based clustering.

# Dataset At a Glance

A challenge we are facing is the quality of dataset:

- Small size and limited number of variables
- Mostly categorical instead of numeric
- Up to 40% missing values
- Inner connections among data files unsure
- Variables' meaning ambiguous

Our actions:

- Discussed the situation with supervisor
- Contacted the partner for specification of datasets and variables
- Requested additional data if available (haven't got reply yet)



# Project timeline

