

## Data Provenance & Ethics Statement

The data used in this project is the famous "GaltonFamilies" dataset from 1886. It is a public and historical dataset that we got using the statsmodels tool in Python.

Because the data is public, old, and does not contain any private information like names or addresses, we did not need special permission to use it. All the information was already anonymous, so there were no privacy concerns.

It is important to use this model responsibly. The data is from a small group of people in London over 100 years ago. This means the model would not be accurate for people today or people from different parts of the world. It was created for learning purposes only and should not be used to make real-life decisions about anyone's health or growth.