Motivation:

Politicians are the instrument of the government, they play an instrumental role in shaping the laws, policies, and decisions that govern our daily lives. The world has seen many politicians, from conservatives to liberals, dictators to democratic, aggressive to charismatic and efficient to futile, but what determines who becomes a politician? Is it possible that the explanation goes beyond the standard factors of education, party affiliations, and characteristic traits? Is it possible that underlying factors people completely overlook such as birth month have a bearing on who becomes a politician? Various studies have found relationships between birth month and disease risks or likelihood of succeeding in sports. For example, between those born in November and risk of ADHD, or those born earlier in the year and the chance of becoming a professional basketball player. The perks of positive birth month effects relate to a phenomenon called the relative Age effect. This describes the case where older peers have an inherent advantage over those born later due to factors potentially relating to physical/mental maturity and experience,leading to greater opportunities for development. So could this factor i.e birth month effects potentially have a correlation with becoming a politician? Are January babies more likely to rule the world?! Read further to find out …

Data collection

To effectively explore the correlation between birth month and political representativeness that was broadly true for a large set of politicians spanning a range of countries we needed a comprehensive dataset. After sifting across different datasets we decided to use EveryPolitician that collected birth month data for over 5 million politicians for 233 countries across recent history. Using a broad dataset that accounts for a wide range of countries and political systems provided us with a more diverse and representative sample of politicians compared to other datasets such as World Bank's World Development indicators (WDI) dataset that wasn't as comprehensive. Furthermore, because EveryPolitician is open-source and community-driven, the data is regularly updated and enhanced by volunteers, hence enhancing the data quality and completeness of our dataset. Moreover, Every Politician's data is in a structured and standardised format called Popolo where a person's birth date is represented as a string in the format YYYY-MM-DD. This standardised format makes web scraping easier because the birth date is always in the same format regardless of nation or politician, simplifying the data extraction process.

Data trends

After data collection, cleaning the data was necessary as, despite the advantage of using a more broad and complete dataset, there was still the possibility of working with incomplete or missing values which had the potential to confound results if they were clustered around a certain birth month.Following from this, issues with data quality lead to an initial overstatement of the likelihood of being a politician for January which was more than 4 times greater than the rest of the months. Upon deeper analysis, this was attributed to the standardised value of 1st Jan being assigned to politicians for which birth month data was not available. We identified this phenomenon as the January skew. To correct for this we omitted countries that had an unreasonable number of 1 Jan values. The threshold for the omission was computed by excluding countries where the ratio of records on January 1st compared to the total number of birth month records was more than 10 times the expected ratio of records for Jan 1st.. Countries for which this threshold was violated were deemed unreliable and include Syria and Cameroon for which the proportion of people born in January were 98.2% and 26.8% respectively. This alters greatly from the sample average of 8.97% for the month of January alone. Hence such results were removed from our birth month analysis to limit any upward bias and safeguard data quality.

Data trends

To better explore the factors relating to political representativeness across the range of countries we decided to investigate the correlation of political representation across a range of factors with a specific focus on birth month effects. The trends observed were as follows

- **Birth month & relative age effect**

When looking at the chart for the percentage of MPs against month of birth, we can see a slightly higher number of MPs relative to the number of births in each month before June, whereas afterwards the number of MPs is lower relative to births. This perhaps supports the Relative Age Effect by showing that those born before June have a greater chance of becoming an MP. We see this more clearly from the percentage difference graph, clearly illustrating the different trends pre and post June. Overall, the highest proportion of MPs are born in January, supporting the correlation between being born early in the year and having more success.

- **Gender; The correlation between being a politician and being male is relatively robust to different groupings based on HDI,GII or other development indicators. It certainly showcases a heavily positive correlation between being a politician and being male. Although the disparity between being a politician given you are male versus female is significantly more apparent for economically less developed countries. This is shown most clearly by the percentage difference between being a politician given you are female for high GII countries when compared to low ones.**
- **Zodiac;**
- **Day of birth; When compared to population statistics for average number of people born on each day, MP's were more disportionately born on the weekends. This is exhibited by the percentages table and graph which show positive values for those born on the weekends. This is significant considering that MP births are**

much lower compared to population average for all other days in the week shown by negative percentage difference values.

- **University education (UK only)**
- **Age (UK only)**