# R analytics Dashboard using Shiny


Name-Tanishq Premi

Roll No-41A

## About the Data Set

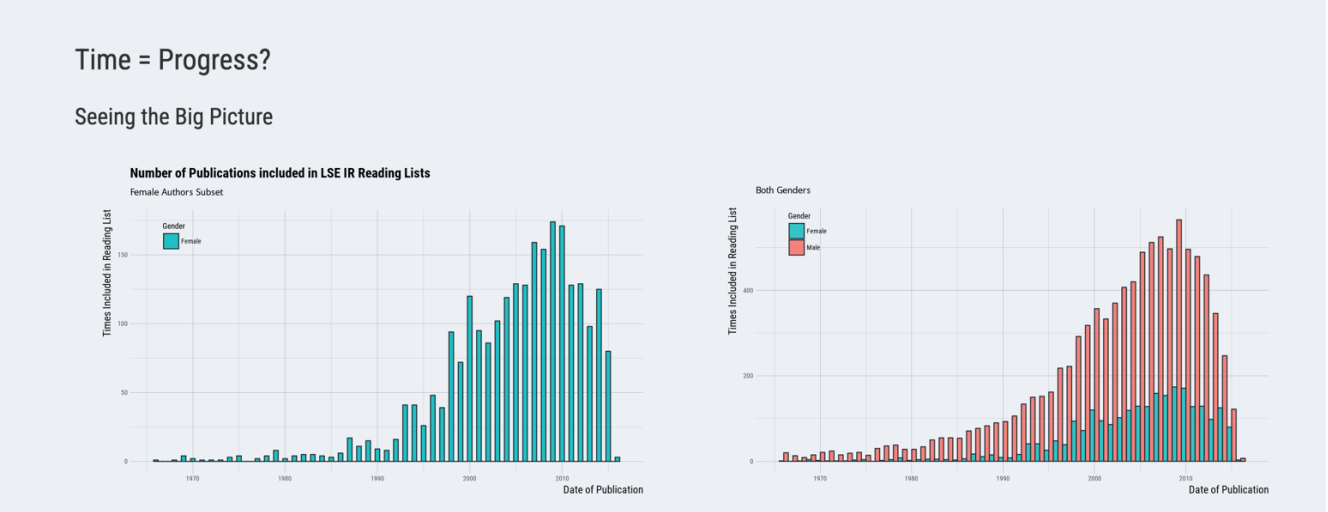
Exploratory analysis of ~13,000 scholarly works included in International Relations (IR) Department syllabi in 2016-2017 at the London School of Economics and Political Science. The main aim of this project is to quantify the publication bias in academia. The findings mirror the US Ivy League patterns of a roughly 80/20 male-to-female author ratio.

## Methodology

The dataset is based on an export of Moodle data containing syllabi for each undergraduate, master's and PhD level IR course on offer at the LSE in the 2015-16 academic year. A total of 43 courses (18 Undergraduate-level, 23 Master's-level, and 2 PhD-level) render 12,358 non-unique (2,574 by female authors) textual sources listed as both essential and background reading material. The analysis focuses on books and articles published between 1960 and 2015. Finally, in order to tackle the gender bias issue as it relates to authorship in academia, sex of the author(s) are coded M/F*. In case of multiple authors, the binary coding indicates at least one female scholar is involved. All coding was done manually by LSE IR PhD candidates on their spare time.

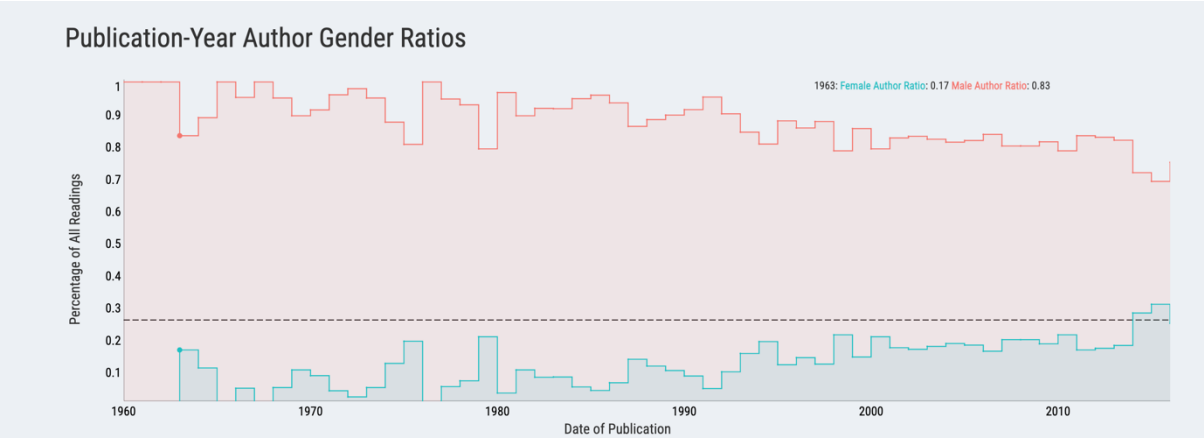
## Introduction

The first plot displays the raw count of publications by female authors included in a reading list by publication year
Within this dataset, we see that works by female authors are more likely to be included as time passes. This is hardly surprising, as prior to 1960's academic authorship was more or less exclusively a male enterprise. However, when plotted side by side with male authors, we find that the trend is universal



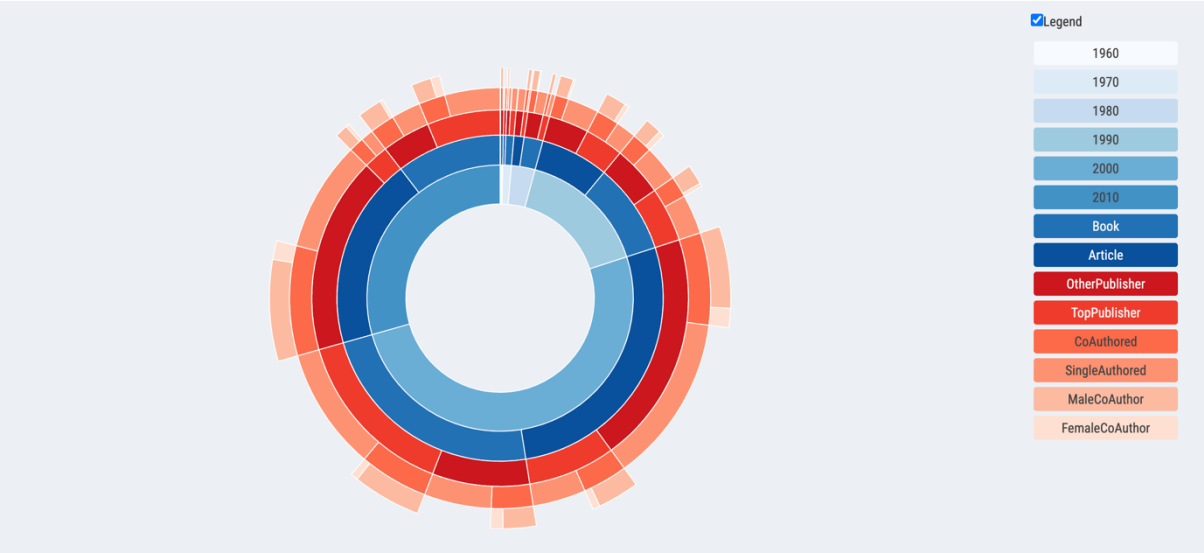
## Time-Series

This interactive time series analogue of the introduction plot shows reading list breakdown based on gender. Female-to-male author ratio, based on their inclusion in LSE IR reading lists, rarely hits 20% in a given publication year. This finding is in line with that of Colgan (2017). Refer to the methodology tab to read about how this study was conducted.




## Publication Pathways

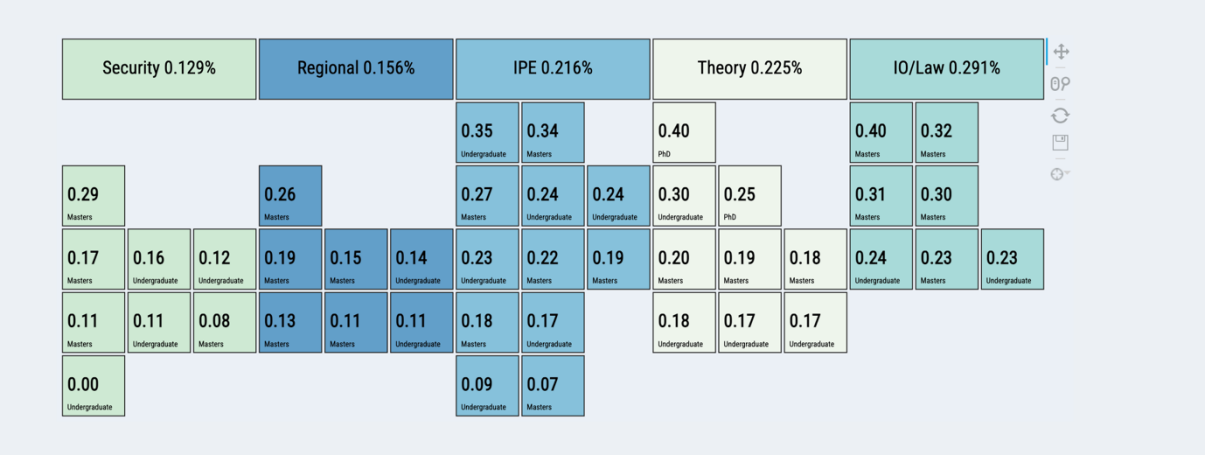
In order to better understand the pathways leading to reading list inclusion, we create and plot sequences involving at least one female author. The numbers inside the dial report the count and overall percentage of the current selection:



Hover on the dial starting from the innermost circle to find out reading list female authorship trends based on the sequence:

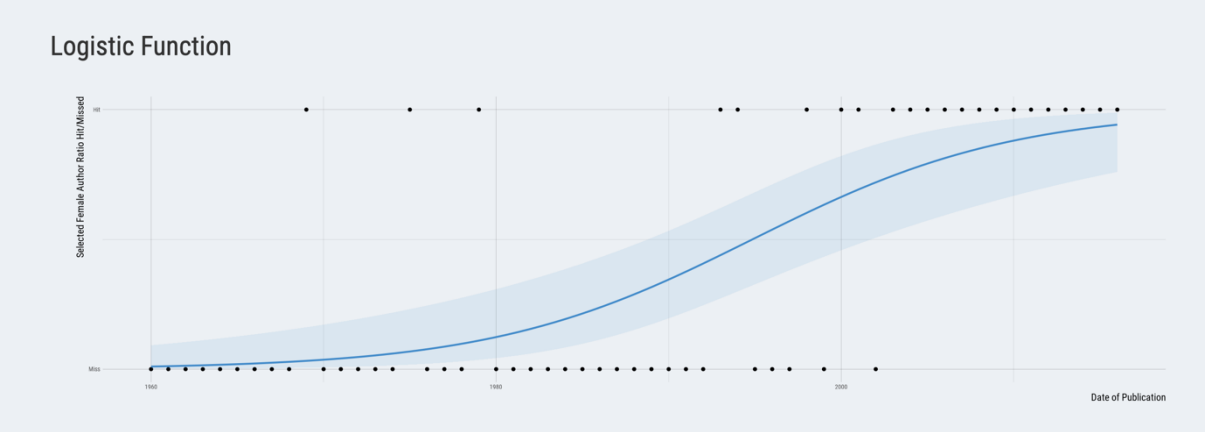
## Course Breakdown

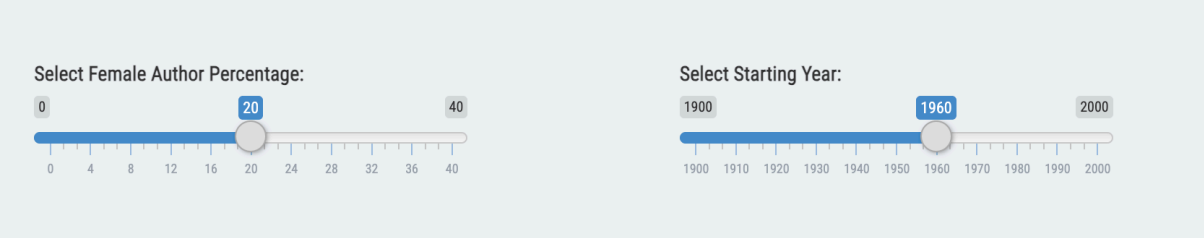
We visualise all 43 courses (18 Undergraduate-level, 23 Master's-level, and 2 PhD-level) included in the dataset. We cluster all courses under five overarching categories: Security/Statecraft Studies, Regional Studies, International Political Economy, Theory, and International Organisations/Law:




## Logistic Link

This graph calculates the logistic function based on selected parameters: female author percentage and starting year. For example, the default settings display which publication-years since 1960 has at least 20% female authors:

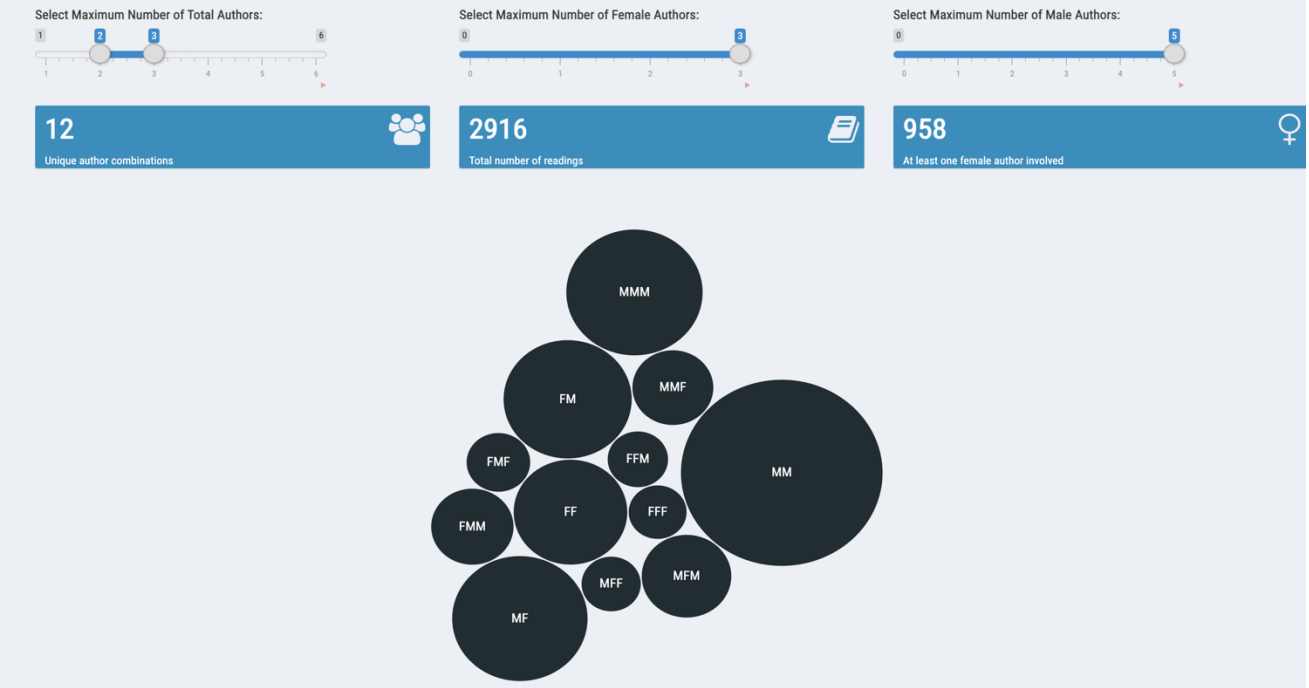





First slider creates a new dummy variable which is set to 0 if it is not met (miss) in a calendar year and 1 (hit) otherwise. The second slider selects the starting year for the data. For example, the default settings produce a logistic link function showing which years had at least 20% female authors since 1960 with a 95% confidence interval.

## Co-Authorship

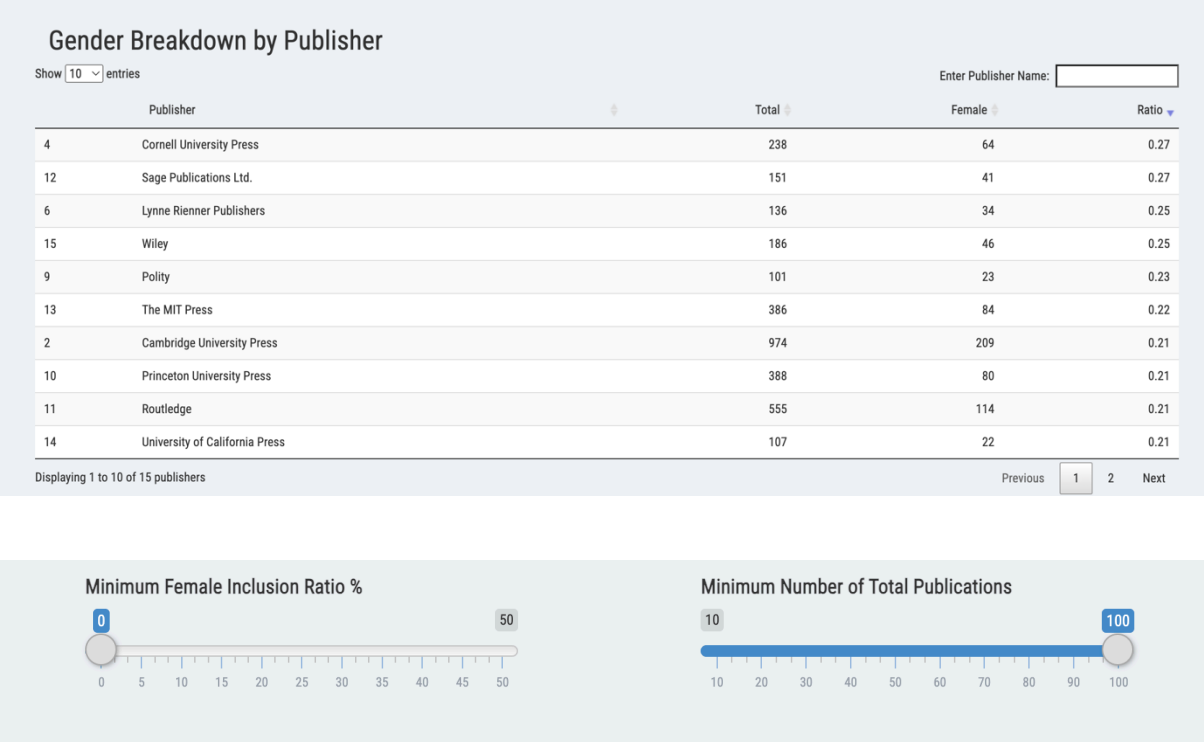
Finally, we visualise co-authorship patters using three parameters: maximum number of authors, maximum number of female authors, ad maximum number of male authors. The bubbles are recalculated each time a slider is updated. The radii are calculated using square root to offset the male single-author dominance.




Explore co-authorship preferences by setting different constraints on author gender and/or total number of authors. The play button under the sliders animates the graph by increasing the selected value by one until it reaches its maximum. The first slider creates a subset based on the value range: The default '2-3' only shows works featuring either 2 or 3 authors, which excludes single-authored pieces by authors from both genders. Using the other sliders, all gender combinations can be analysed. As coding conserved the first-last author sequence, 'FM' is qualitatively different than 'MF'. Note that the bubble circumferences are calculated using square roots: the differences look smaller than they actually are. Otherwise, most female author groups would be invisible.

# Publisher Info

This interactive dataframe allows for searching for specific publishers as well filtering on female author ratio and the total number of publications:




Play around with the sliders to filter publishers. Notice that as publishers get more 'prestigious' (i.e. higher number of total publications included), their F/M ratio goes down.