Analysis of physicians registered under The Tripura State Medical Council

Sikim Chakraborty

I. Introduction

"THE **Tripura State Medical Council" was** established **on** 19th April 2011 with the council members being **nominated by the State Government.** The elected **"The Tripura State Medical Council"** was constituted in December 2014 and the members were elected by an electoral process. The council provides a data of the doctors registered to practice in the state of Tripura from 2012. In this article, an insight is provided towards the state of doctors in the state based on certain parameters viz. age, sex, specialization area, address, and others.

II. Dataset Used For Analysis

The dataset used for evaluating the present condition of physicians is provided by The Tripura State Medical Council. [1]. It comprises doctors registered with the state from a period between 2012 and 2016. The variables in the dataset include Name and Father's Name, Qualification/ Additional Qualification/ College/ University and Year thereof, Reg. No Reg. Date, Date of Birth/Sex, State & Schedule and Address of all the doctors.

III. Approach

The data was scraped from the website and a few additional features were derived for the purpose of the analysis.

A. Scraping

The data was scraped using a list of Python [2] packages viz. BeautifulSoup [3], urllib3 [4] and pandas [5]. The website was scanned for any links indicating the data source of the populated table whether it is a static file viz. .csv or .pdf or it has a backend database running on the server. The website was scanned for 'href' attributes within an 'a' tag containing any known file extensions. Examination revealed a .csv file being the source of the data. The file was downloaded into local and used discussed below.

B. Cleansing

The data was first scanned for any NaN values. An observation was dropped since it contained only the entry for the feature 'Reg. No.' and all other attribute values were 'Cancel'.

The dataset was thereafter cleaned feature wise for removing irrelevant information and for extracting additional variables for analysis. The attribute State & Schedule comprised of same values throughout the dataset and hence was not considered for analysis. The 'Date of Birth/Sex' attribute was next picked up for cleansing. After separating the date of birth and sex using a regular expression, the current age of a physician was computed, and a separate feature 'Age' was appended to the data frame. Some malformed dates were encountered and had to be imputed or altered manually. After that, the registration year of the doctors was extracted from the 'Reg. Date' column of the data frame. Next, the 'Address' variable was used to extract the states from which the doctors originally are. List of cities and India and state and city mapping was extracted from [6]. The representation of the addresses in the dataset is ambiguous. Sometimes a city is mentioned and sometimes the state and sometimes even a district. Few observations were dropped since no mapping could be found between a state or a city and the 'Address'. For ease of analysis, the state has been considered. The 'Qualification/ Additional Qualification/ College/ University and Year thereof' variable was next explored for isolating doctors with specialized focus areas post MBBS. An observation was found wherein the dataset had no record of any basic qualification of the physician but an additional degree from Russia. For ease of analysis, a new feature 'Specialization' was introduced consisting of the following set of categorical variables viz. 'Pathology/ Clinical Pathology', 'Radiology/ Radio Diagnosis/ Radiation Medicine', 'Forensic Medicine/ Forensic Medicine and Toxicology', 'Tropical Medicine', 'Anaesthesiology', 'Dermatology/ Venerology/ Leprosy', 'Psychiatry', 'Opthalmology', 'Community Medicine', 'Sports Medicine', 'Surgery', 'Orthopaedics', 'Physiology', 'Anatomy', 'Paediatrics/ Child Health/ Maternity and Child', 'Pharmacology', 'Microbiology', 'Public Health', 'Gynaecology/Obstetrics', 'Diabetology', 'Biochemistry', 'Otorhinolaryngology/ENT', 'Laboratory Medicine', 'General Medicine' and doctors with no such

Additional plots including all the plots in this article in higher resolution, and source code are uploaded at [7]

mentioned specialization but with an additional qualification were marked as 'Other'. Post cleaning, ~7% of the total data was not considered for analysis. Further details have been provided in [7].

IV. Results and Discussions

The following observations were found while analyzing the dataset. Female doctors practicing in the state of Tripura are found to be much younger (median:30, mean: ~31) than their male (median:32, mean: ~39) counterparts. Fig 1. depicts the ratio of male to female doctors grouped by their ages. Most doctors in Tripura belong to the age group of 30 and below (363), followed by 31-35 (189) and above 60 (76).

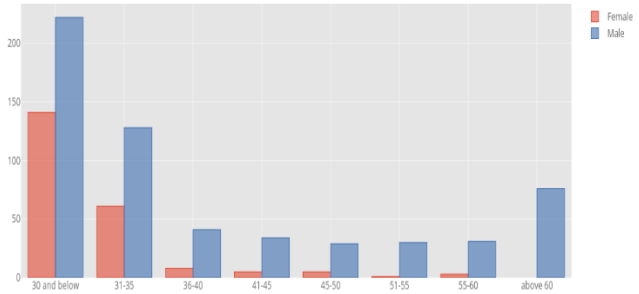

Fig. 1. Distribution of doctors grouped by age group and sex

The average registrations for doctors in Tripura are ~160 (mean:163, median: 160) per year. For the year 2014, there is a sudden surge in the number of doctors being registered (383) (as shown in Fig.7). Fig 2. indicates the doctors in Tripura are predominantly male. Just 224 (~27%) of the doctors are female.

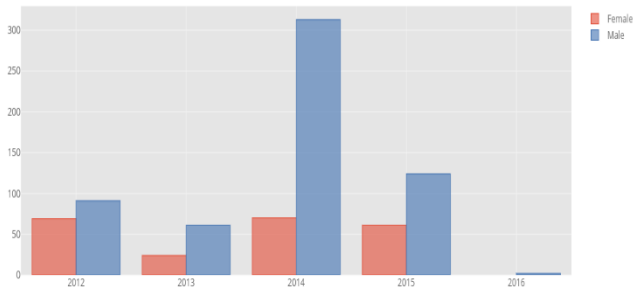

Fig. 2. Registration of doctors from 2012-2016 grouped by sex

The ratio of doctors in Tripura with a specialized focus area to ones with a basic qualification to is ~0.24. Fig. 3. depicts the availability of doctors in Tripura with a specialized focus area with Paediatrics being the most predominant closely followed by Gynaecology/Obstetrics and Surgery whereas Tropical Medicine, Diabetology, and Laboratory Medicine are the least prevalent.

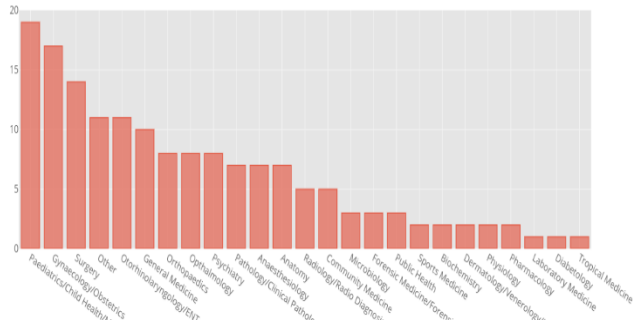

Fig 3. Distribution of doctors with a specialized focus area

Fig. 4. points to the fact that the availability of female doctors with specialized focus areas is limited to a few specific fields viz. Anatomy, Community Medicine, Dermatology, General Medicine, Gynaecology/Obstetrics, Otorhinolaryngology/ ENT, Paediatrics, Pathology, and Pharmacology. The ratio of female to male doctors with an additional qualification is ~0.1 whereas the ratio of female to male doctors without any specialization is ~0.47 (as shown in Fig. 5 and Fig. 6 respectively)

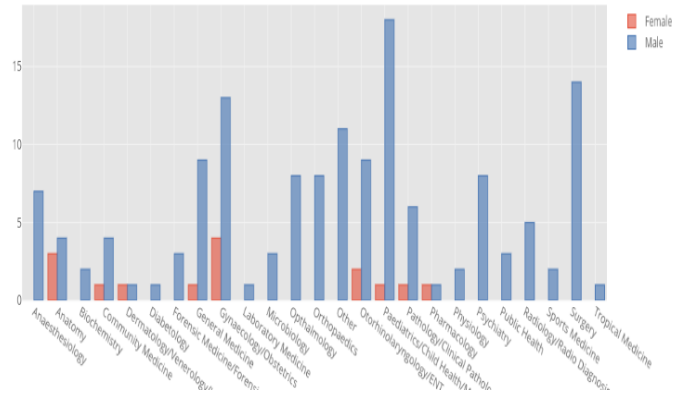

Fig. 4. Sex wise distribution of doctors with specialized focus areas

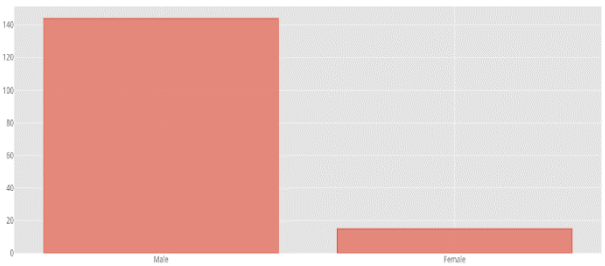

Fig. 5. Sex wise distribution of doctors with any additional qualification

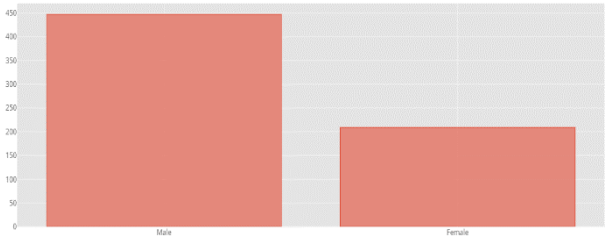

Fig. 6. Sex wise distribution of doctors without any additional qualification

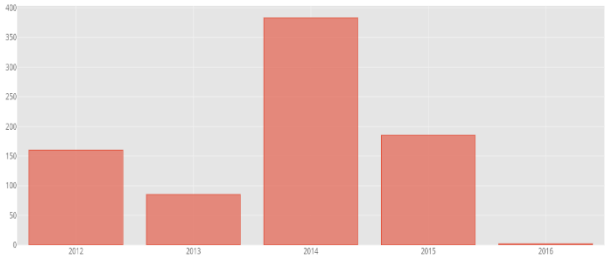

Fig. 7. Year wise registrations of doctors

Table 1. shows that most of the doctors of Tripura are native to the state. Nearby states of West Bengal, Assam and Orissa contribute to more than 50% of the meagre ~13% from the rest of India.

State	Count
Tripura	704
West Bengal	25
Manipur	19
Assam	14
Uttar Pradesh	7
Haryana	6
Rajasthan	5
Kerala	5
Nagaland	5
Madhya Pradesh	4
Delhi	3
Tamil Nadu	3
Meghalaya	2
Mizoram	2
Andhra Pradesh	2
Chandigarh	2
Arunachal Pradesh	2
Jharkhand	1
Orissa	1
Gujarat	1
Bihar	1
Chhattisgarh	1

Table 1. State wise distribution of doctors currently practicing in Tripura

V. Conclusion and Future Work

Tripura having a population of ~36.7 lakhs [8], there is clearly a shortage in the availability of doctors. The above analysis demonstrates that current doctors in Tripura are predominantly male. Since the majority of the doctors are from Tripura itself, this might indicate a gender imbalance in medical education/ profession in the state. The analysis also points to a shortage of doctors with specialization (~0.19%). Fig. 1 suggests that most of the doctors fall within a relatively young age group, thus hinting at the fact that recent graduates are preferring to practice more in Tripura than experienced ones. The next comparable age group is above 60. Further analysis can be done to find the various institutes from which the doctors in Tripura have graduated. A mapping of the district and states of India could be used to gain further insights into the 'Address' variable as discussed in Section III. Also, the sudden surge in registrations from all age groups in 2014 is unexpected and needs to be explored further beyond the scope of this dataset.

References

[1] The Tripura State Medical Council doctors list https://tsmc.tripura.gov.in/doc_list
[2] Python 3.6 https://www.python.org/
[3] BeatuifulSoup 4 https://www.crummy.com/software/BeautifulSoup/
[4] urllib3 https://urllib3.readthedocs.io/en/latest/
[5] pandas https://pandas.pydata.org/
[6] Population Census 2011 https://www.census2011.co.in/
[7] GitHub repository: https://github.com/scrab017/tgdl
[8] https://www.census2011.co.in/census/state/tripura.html