

 **Editorial**



**Dr Asha Subramanian**  
Head – Centre for Open Data Research (CODR)

While data science and big data applications have shown great promise in dealing with social network datasets, electronic commerce and IOT applications using sensor generated data, data science applications have had little impact in governance applications in India especially

integrating data and deriving insights from diverse domains. Governance data includes open government data published by governments across diverse domains, closed data that is generated within each government function and tons of data collected as primary data through surveys by independent organisations. Addressing a governance problem typically involves analysing data across domains since the most of the manifestations are closely interlinked. Data science applications in this domain are particularly hard as data is not usually available a readily use-able form and much of the effort is spent in contextualising and preparing the data for analytics. Governance data when combined with big data sources such as geo-spatial data, telecom data, climate data etc. provide huge potential for insights and inputs to data enabled making.

Public Affairs Centre(PAC)has established the a Centre for Open Data Research(CODR) which focuses on applying data science techniques and innovative research to make data enabled decisions to solve governance and development issues. In this context CODR is in the process of developing two products Semantic Innovation and Harmonisation of Sustainable Development Data in India(SIDDHI) and Data Integration and Insights powered by Semantics (DRSHTI) that help integrate diverse datasets along various themes (spatial, temporal and contextual) and creating an ambitious knowledge graph of interconnected governance data. This self-learning knowledge graph apart from providing comprehensive picture of governance data across many inter-related themes helps facilitate rich insights and inferences. CODR recently acquired its first data science project from the office of the Principal Accountant General(GSSA). The project aims to make their audit process data-led with a potential to assist the office in planning future audits. This project is one step closer to the ambitious knowledge graph of governance data that CODR hopes to build in the upcoming years. Signing off on exciting times ahead!

 **Interview**



**E P Nivedita**  
Principal Accountant General (General & Social Sector Audit), Karnataka

E P Nivedita is the Principal Accountant General (General & Social Sector Audit), Karnataka. She has served in various capacities in different Government departments. Public Affairs Centre interviewed her for this issue of Public Eye.

***How is data used in the policy making space?***

Governments all over the globe including India are harnessing technology and digital platforms to deliver government services. This results in huge volumes of digital data. Apart from this, government agencies like National Sample Survey Office NSSO and other Non- Governmental Organisations (NGOs) collect data. Focus areas for audit are determined by Risk analysis and having access to such enormous amounts of data is a huge opportunity for improving the manner in which we conduct risk analysis. The Indian Audit & Accounts Department has set up the Centre for Data Management and Analytics (CDMA) in 2016 and the Centre for Digital Audit of Revenues in order to integrate big data with the audit process.

***How does your department plan to use the rich data you have to enhance the efficiency of the department?***

With increasing availability of databases the department is using data analytics and visualisation software such as KNIME, IDEA, Tableau and Microsoft Excel to analyse data and look for trends, patterns, correlations and deviations to generate insights for use in audit planning and execution. We have created a pool of staff conversant with different databases. Training programmes are conducted regularly to improve the depth and breadth of the skills of our staff. Systematic data collection and analysis aided by defined processes and capacity building are the measures that we are taking to harness the power of data with a view to enhancing the efficiency of our department.

***How do you think that government departments and organisations like PAC can leverage in the area of data analytics?***

Regular failure of schemes and programmes to address certain deficiencies/problems in the country could stem from causes which cannot be explained with data available from the implementing agency/department of government alone. Data analytics, can help in identifying reasons/correlations for certain phenomena which cannot be explained in isolation. Analytics can help Government in better estimation of various factors of interest and provide inputs on where and how much to steer to cause desired results. Organisations such as PAC are unique since they have rich experience in working with Government agencies and also have the flexibility to hire quality talent from the market.





## Case Studies



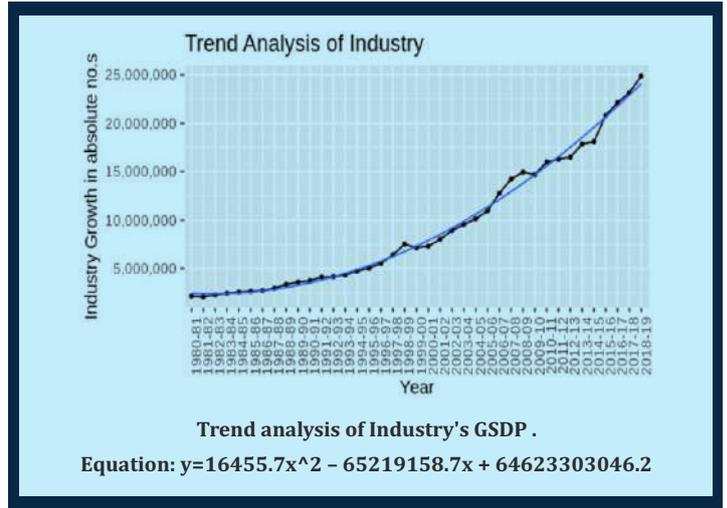
**Mohanapreethi Attuluri**  
Data Engineer

Data driven identification of factors and patterns of natural agglomerations of people over time and analysis of educational, occupational and economic patterns that majorly contribute to the Growth Model for

Karnataka under the conditions of globalisation. This work was done as part of the Sustainable Development Goal (SDG)-8 committee setup by the Government of Karnataka (GoK) to provide a formal framework for definition, collection and dissemination of sustainable development indicators for the Goal-8 namely Decent Work and Economic Growth.

### Approach

The 2001 and 2011 census data was downloaded for initial exploratory data analysis and to identify natural agglomerations of population. Further educational, occupational, tax-payers' data and the Gross State Domestic Product (GSDP)[A1] data were analysed and superimposed with census data to get the trends and patterns of the shifts in population and the corresponding changes in labour distribution across occupations and the related education levels. Finally, a time series analysis and non-linear regression was applied on the yearly GSDP and per capita data at different levels in the industry categories to arrive at predicted targets for 2022 and 2030.



**Mrinalini Kabbur**  
Senior Data Analyst

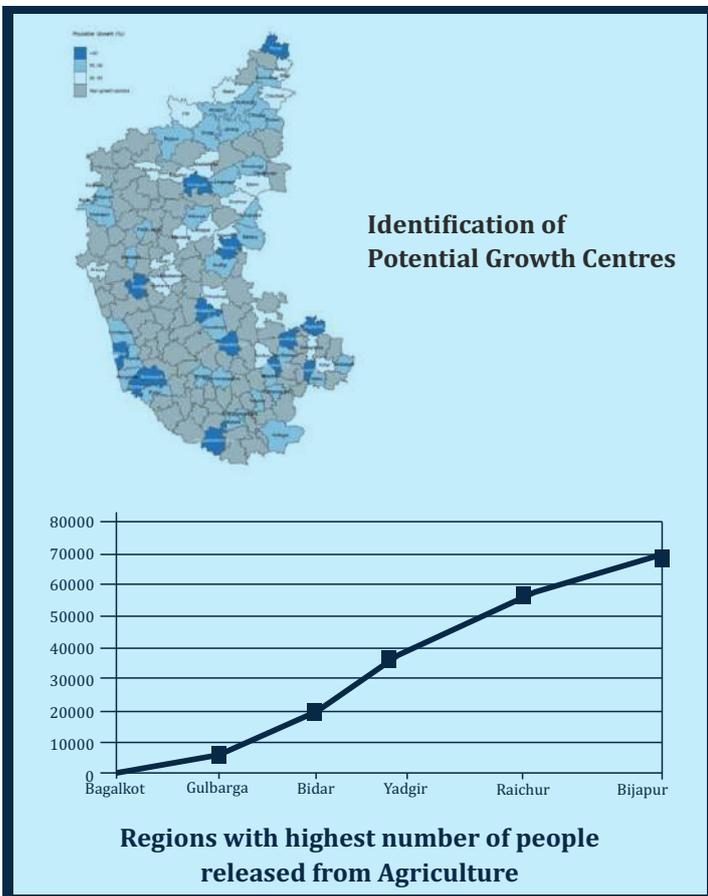
Public Affairs Centre is a part of the Southern Voice on the State of the SDGs (SVSS) initiative and this project tried to understand the enablers and barriers to sustainable livelihoods for women in India.

### Approach

A total of 2503 female respondents selected from the districts of Saharanpur in Uttar Pradesh, Mandi in Himachal Pradesh and Coimbatore in Tamil Nadu and 6 skill centres implementing the Pradhan Mantri Kaushal Vikas Yojana (PMKVY) programme formed the primary data. The respondents answered an exhaustive questionnaire (close to 350 questions) covering all the important stages starting from early life to adulthood including socio-economic, family, skilling, education and livelihood areas to cover enablers and barriers to livelihood participation. First, an Exploratory Factor Analysis (EFA) was used to reduce this large list of demand-and supply-side variables to represent a smaller combination of latent factors or mental models. The mental models suggested by EFA were analysed by splitting the data into two groups based on woman's current working status. Confirmatory Factor Analysis (CFA) was performed separately on each group to understand the characteristics of labour force participation for women.

### Analysis

1. The profile of Working and Non-working women differs largely in terms of the mental models Industry Knowledge and Barriers to Livelihood respectively. Industry Knowledge refers to awareness regarding major sectors providing employment in the district and seems a likely factor in enabling female labour force participation. Barriers to livelihood refers specifically to circumstances from increased responsibilities owing to marriage/children/child rearing, shifting residences, quitting jobs under family influence, lack of agency (voluntary decision to not work or quit work) and lack of appropriate work conditions.



2. Emergence of a considerable impact (inferred from the corresponding loadings of the factors) relating to the indicators associated with the mental model Livelihood and Workplace for both the groups. The significant indicators differentiating the two groups relate to workplace safety/harassment at workplace closely followed by the lack of skills training and reaching place of work.

The analysis clearly demarcated the profile of a working woman vs. the non-working woman, providing cues to potential interventions that could be devised to improve the labour force participation rates in India.



**Manikanta Vikkurthi**  
Data Scientist

### Importance of Data Science Applications in Human Development

Project DRSHTI-Stunting aims to understand the factors influencing stunting in children across the districts in Karnataka. The potential factors impacting stunting can be classified into socio-economic variables such as Households with Clean Fuel, Women married before 18 years of age, Households using improved Sanitation Facility, Literacy of Women etc. or climate change variables such as Rainfall Departures, Drought Severity etc. The data around the variables mentioned were collected from a variety of sources. Monthly Progress Reports (MPRs) containing data regarding the implementation of Integrated Child Development Services (ICDS) provided structural and supply side information such as staff deployed, number of beneficiaries and progress indicators such as live births, deaths, normal, underweight children in numbers. NFHS-4 data provided information regarding stunting percentages along with nutrition and socio-economic related data. Climate datasets from various portals provided information on status of rainfall, drought hit regions and rainfall departures.

Using DRSHTI, variables related to the research question (stunting in this case) were superimposed with variables from multiple datasets across different data sources for similar geospatial locations and for similar time periods to derive insights. DRSHTI uses the underlying graph to identify associated variables for the research question (Stunting) by identifying variables that have close correlations in similar contexts. Early results from our prototypes have been able to detect interesting exclusion patterns. The Figure below illustrates the underlying approach using climate change data and related datasets in unearthing patterns of exclusion across the Agriculture, Nutrition and Livelihoods domains. The initial patterns that emerged include a) Districts with high literacy in women and lower incidence of marriage under the age of 18 directly attributes to significantly low stunting rates. b) Drought severity and rainfall departures play a moderate role in stunting rates for the affected regions. This is the first step towards building a comprehensive model to support data-based decision making in governance.



**Research Papers**

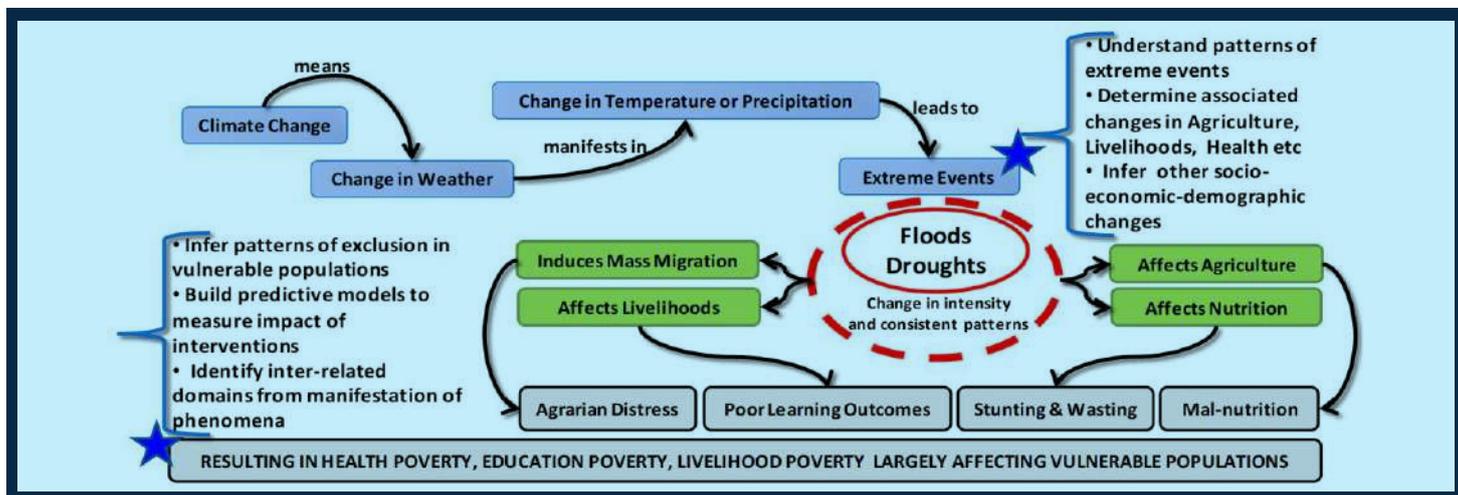
**Semantic Harmonisation of Numeric Data from Open Government Data**

ACM DL DIGITAL LIBRARY

Dr. Asha Subramanian, Pavan Kumar RR  
Manikanta Vikkurthi, Pier Luigi Buttigieg

**Semantic Interpretation and Integration of Open Data Tables**

Dr. Asha Subramanian & S. Srinivasa



 Stakeholders



Mr Shoury Reddy, Executive Director, Balavikasa



Mr Sujit Kumar Chowdhury, Director, Fiscal Policy Institute  
Finance Department Govt. of Karnataka

 Visitors

Visiting PAC after a long gap was like a pilgrim. The energy, enthusiasm and desire for collaboration and synergy building is great!

*Mr Jagadananda, Mentor & Co-Founder Centre for Youth and Social Development (CYSD), Bhubaneswar 16.01.2019*

Excellent conversation with brilliant minds!

*Mrs Yashashree Gurjar, Chief Executive, Sulochana Thapar Foundation 18.01.2019*

It was a pleasure interacting with the bright and experienced team of PAC. The discussions were insightful and the organisation has provided a warm welcome and a potential possibility for future collaboration. Honoured to have visited PAC today.

*Dr. Ravi Kanth M, Consultant, UNICEF 27.02.2019*

Great to have meet passionate team with laser focus purpose of making good and empowering communities. Glad to have a partnership to make our work even more impactful.

*Chandra Shekhar Gowda, COO, STF 01.03.2019*

Inspired with the organisational commitment. Excited for our partnership

*Nandita Bhatt, Martha Farrell Foundation, Delhi 01.03.2019*

 Internspeak



“PAC has taught me the importance of data and the fact that it is crucial to read between the lines to understand the working of the government in a complex democracy like ours”.

*Anoushka Roy, National Law School, Bengaluru*



“This is my first internship ever, and I got lucky that I got one in PAC, so the work environment, everyone and the energy here is fantastic, you get to learn so much”.

*Vidushi Pant, Symbiosis College, Pune*



“Working in PAC has been a wonderful experience. A lot of my mentors here have been super supportive and very patient with me. They have heard out all my doubts and patiently cleared them all”.

*Avani Darne, Symbiosis College, Pune*

To watch more interviews, log on to our YouTube channel <https://www.youtube.com/user/pacindiavideos>

## Community Voices

"I am regularly visiting FPS to check hygiene condition, quality of grains and to address any problems raised in the community."

*Mangala, Citizen Monitoring and Action Team (CMAT) member (also GP member), Mandya*

"I am very happy being a member of CMAT. I learned about my rights on food security and I educated and guided my village people to get their right entitlements."

*Hemavati, Citizen Monitoring and Action Team (CMAT) member, Chickballapur*

## Outreach and Stakeholder Engagement



G Gurucharan, Director, PAC & Dr Annapoorna Ravichander, Head Policy Engagement & Communication, Resource Mobilisation & Training at the On Think Tanks conference in Geneva



G Gurucharan, Director PAC as a panelist at the Round Table on *Governing Diversity in the Cities of the Global South: Policy and Practice Insights* at the European University Institute, Florence.



Dr Asha Subramanian, Head Centre for Open Data Research at the Artificial Intelligence Think Tank forum at the University of Pennsylvania discussing how PAC is using the power of technology and data insights to fight livelihood, health and education poverties.



Dr Annapoorna Ravichander, Head Policy Engagement & Communication, Resource Mobilisation & Training, at the Strengthening Policy Research book launch and round table discussion on organisational challenges and sustainability in Nepal.



Varsha Pillai, Programme Manager, Policy Engagement and Communication was selected as one of the 10 fellows under the On Think Tanks and Think Tank Initiative fellowship to attend the Winterschool for Thinktankers in Geneva in January 2019.



Dr Asha Subramanian, Head Centre for Open Data Research at the technical paper presentation at the ACM India Joint International conference on #DataScience and Management of Data CODSCOMAD-2019, Kolkata.



## Trainings



G Gurucharan, Director Public Affairs Centre PAC and Dr Asha Subramanian, Head Centre for Open Data Research at the Workshop for Development at IIIT, Bangalore on “Data Science: A necessary condition for inclusive development in India”.



G Gurucharan, Director, PAC & Dr C K Mathew, Senior Fellow, PAC delivered a module on Urban Local Governance: Financial Sustainability and Social Accountability at the Executive Development Programme for Officers of the Comptroller and Auditor General of India at the Indian Institute of Management, Bangalore.



Dr CK Mathew was one of the expert commentators at one day workshop on deliberation and validation of the outline of Karnataka Municipalities Bill, 2019 held at the National Law School of India University, Bengaluru.



Dr Annapoorna Ravichander, Head Policy Engagement & Communication, Resource Mobilisation & Training held a Policy Engagement and Communication workshop at the Indian Institute for Dalit Studies, Delhi.



G Gurucharan, Director, PAC took a session on Ethics in government & the political economy of development at Administrative Training Institute, Mysore as part of the Common Foundation Course for a cohort of 50 direct recruits of the Government of Karnataka in departments of health and agriculture.



G Gurucharan, Director PAC and Dr Meena Nair, Head of Research, PAC delivered a training program on Social Accountability Tools(SATs), their relevance & effectiveness to the Trainee officers at the Lal Bahadur Shastri National Academy of Administration, Mussoorie.

### From the Editor's Desk

Public Eye will be published in January, May and September. Do send us your comments/feedback to [pec@pacindia.org](mailto:pec@pacindia.org). If you wish to unsubscribe please click [here](#)