Journal of Education, Society and Behavioural Science

35(5): 1-14, 2022; Article no.JESBS.85525
ISSN: 2456-981X
(Past name: British Journal of Education, Society & Behavioural Science,
Past ISSN: 2278-0998)

Urban Data Science Education: A Key Actor towards Improving Data-Driven Policy-Making for Solving Urban Problems

S. R. Gandhi [a*] and F. E. Anyiam [b]

[a] Faculty of Technology, CEPT University, Kasturbhai Lalbhai Campus, Gujarat, India.
[b] Centre for Health and Development, University of Port Harcourt, Port Harcourt, Nigeria.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JESBS/2022/v35i530421

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/85525

Original Research Article

Received 04 March 2022
Accepted 08 May 2022
Published 13 May 2022

ABSTRACT

This paper explains the benefits of Urban Data Science Education in improving data driving policy-making and highlights the impact of vital use of available open data using a practical demonstration from the Summer-Winter School (SWS) at CEPT University Ahmedabad India. Good Urban data policies can be an important driver for smart city research and the implementation of good data management practices. Many authors are also becoming interested in the data revolution that is enhancing the way we study and understand cities. It now becomes appropriate to respond to this need in an evidence-informed manner by working with stakeholders to encourage a more resilient approach via education for the younger enthusiast, leading to more transparent, improved, sustainable and safer urban cities. However, the major challenge we foresee is the inability to utilize the large amount of urban data generated daily. With the drive to mobilize more education towards urban data science and analytics, generated data can be effectively utilized to make informed and evidence-based decisions. This paper points to a practical methodology for reproducibility, the key steps taken, as it views the inclusion of an Urban Data Science curriculum as priceless, and a decision that will bring into notice the importance of engaging the younger generation in scholastic learning on data principles and analytics. Data along with spatial enhancement can be the driving factor towards improving data-driven policy-making for solving urban problems in smart cities.

**Corresponding author: E-mail: shaily.gandhi@gmail.com;*

Keywords: *Urban data science; data Analysis; policy-making; open data; smart cities; data science; education; Urban; Cept University; India.*

1. INTRODUCTION

Data is the lifeblood of decision-making and the raw material for accountability [1]. There has never been a more exciting time to be active in data exploitation, as data is transforming today's world on a daily basis. We have large unexplored data with such a capacity and potential to change our living status in many different ways [2]. And, as innovation and technology evolve, particularly in recent developments in big data analytics, a vast amount of this data is collected and evaluated daily, utilizing both classic data analysis methods and a more sophisticated form of data science analytics [3, 4]. Since the end of the Millennium Development Goals (MDGs) in 2000, the world has seen an increased uptake in the use of data and technology to improve people's lives [5,6].

The availability of a large amount of open data in the 21st century has made Data analytics a highly sought-after skill [7–9]. While many data-driven organizations seek data analytic skills at an exponential rate, not many have learned to utilize their potential [6]. Even in many governmental organizations, the promise of a data-driven and evidence-based decision making has failed to materialize because most public institutions lack the capacity to effectively utilize these data, and so the progress in turning data into actionable insights is halted [10,11] Furthermore, in many developing economies, policy-making is still driven by personal sentiment, experience, opinion, or belief, with little empirical input; and only a small number of government institutions are scratching the surface in terms of innovative data and technology use [12, 13].

A data-driven approach entails a methodical and organized process of gathering, compiling, managing, analyzing, interpreting, and applying data in order to generate insight that could aid in the resolution of a social or urban problem [14]. These data can be open source or gathered from a database maintained by a government or a private company. And, if these data can be used effectively and efficiently, training enough data scientists to meet the demand for data utilization becomes crucial.

Every data-driven solution if effectively maximized, can solve varying urban problems. A notable example can be found in the United States, where Pred Pol software uses historical data sets, Data Analytics, and Machine Learning to prevent crimes from occurring, allowing local police departments across cities to deploy their resources [15]. A study conducted by Pokhriyal & Jacques, shows where several data sources are combined for better diagnosis and poverty mapping for improved poverty prediction and mapping [16]. A study conducted in Guatemala and Honduras by Adelman et al. shows how Administrative data predicts school dropout [17].

With the foreseen benefits in Urban Data Science education, the present course we proposed is to complement this drive of data explosion and its benefits in the contemporary world in creating the possibility of bringing data into reality through sound empirical research and analysis, thus minimizing uncertainty towards the best course of action in a particular topic in policy design. The paper, therefore, aims to show that Urban data science education can both help to solve various urban problems by enhancing participants' data skills in other to nourish the culture of data science.

2. THE URBAN DATA REVOLUTION FOR CITIES RESEARCH

In the last decade, there has been an explosion of city and urban life data [1,18]. According to the United Nations, "By some estimates, 90% of the data in the world has been created in the last two years, and it is projected to increase by 40% annually; simply because data is being gathered by inexpensive and numerous information-sensing, mobile devices, which encompasses the open data movement, the rise of crowdsourcing, new ICTs for data collection, and the explosion in the availability of big data, together with the emergence of Artificial Intelligence (AI) and the Internet of Things (IOT)" [19, 20]. In addition, many authors are becoming interested in the data revolution that is enhancing the way we study and understand cities [21–23], often referred to as "Urbanism" [24–26]. These new forms of generated data can range from social media tweets via mobile phones, monitor digital camera feeds (e.g: Video vehicle detection), GPS, transponders, transportation (location/movement), and mapping surveys that monitor city operating systems. These operations enable us to tag locations and coordinate movements using location-based services (LBSs) [27,28], thereby changing how we plan and govern our cities. For example, movement and location Data generated by intelligent transportation systems (ITS) travel can be used to create and refine models and simulations to guide future urban development, assisting in decongesting some regions while improving others to raise land prices [29,30]. This can effectively solve traffic congestion in most cities [31, 32]. This daily transformation enables data-driving planning for the cities. It now becomes appropriate to respond to this need in an evidence-informed manner by working with stakeholders to encourage a more resilient approach via education for the younger enthusiast, leading to more transparent, improved, sustainable and safer urban cities. The significant challenges we foresee are the inability to utilize the large amount of urban data generated daily; and therefore, the need to mobilize more education towards urban data science and analytics to link these data together to make informed and evidence-based decisions under smart city movement.

3. METHODOLOGY

The growing urban cities are increasing demands for evidence and the use of data, making Urban Data Science education a more prevalent approach. Governments and institutional leaders now recognize the need to provide education in data, and interpretation towards sound training on data-driven decision making (DDDM), which is lacking in many developing settings [33,34]. With this need in mind, We organized a Summer School on Urban Data Science to highlight the need, potential, and importance of an institutionally based program in establishing the knowledge and skills necessary to enhance data science practices in governance and policymaking. The Summer Winter School (SWS) at CEPT University, Ahmedabad, India was the primary host of the program.

Two cycles of students have undertaken the course, with the first cycle in 2018 with 22 students and the second cycle in 2019 with 26 students, making 48 students from different disciplines like planning, architecture, civil engineering, and geomatics. Though from different disciplines, they all have one thing in common: learning how data analytics skills can help formulate effective decision-making policies in the various urban sectors. And also, exposing them to the availability of new sources of data and the emergence of new data science technologies will propel an inner drive for making use of such data towards contributing and impacting significantly on public institutions.

Student’s enrollments in the course were both from the bachelor's and master's levels involving technical and non-technical fields. The summer school provided a first-hand practical experience to students to get started.

The school adopted a simple methodology that guided the students in formulating their case studies. The process began with the students learning how to search for a research topic of interest, formulating the topic into a research title, recognizing the relevance of the topic by improving the research objectives, searching for a relevant open-source data in line with their study objectives chosen, understanding the structure, format, and dimensions of the open-source data, and if there are any flaws in the data which will require cleaning the data sets, making graphs and charts for the exploratory phase of the data analysis and then analyzing the data with the aid of the Statistical software package “R”, which was the primary statistical and visualization tool used for the entire duration of the course (https://www.r-project.org/) as shown in Fig. 1.

The geographic information system (GIS) was also incorporated into the study modules to help students formulate a system designed to capture, store, manipulate, analyze, manage, and present spatial and geographic data (http://www.arcgis.com/index.html). Students were also exposed to various open-source software for data cleaning like OpenRefine. An introduction to Git and GitHub was given to students with a proper understanding of the importance of data sharing and various licenses.

The course was taught keeping in mind the heterogeneity of the students. A perfect blend of theory and practical’s were planned with relevance to each topic. The course was intensive and to keep every student up to speed at the same pace in the classwork or exercises, students' feedback was taken daily, which helped enhance the class activity and decisions taken by the tutors for teaching the next concept. The course was modified daily with more group activities and practicals based on students' feedback, as shown in Fig. 2.

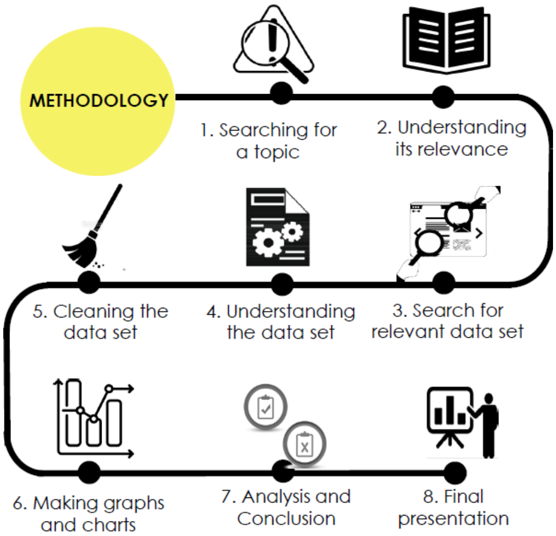

Fig. 1. Methodology adopted by students for their case studies

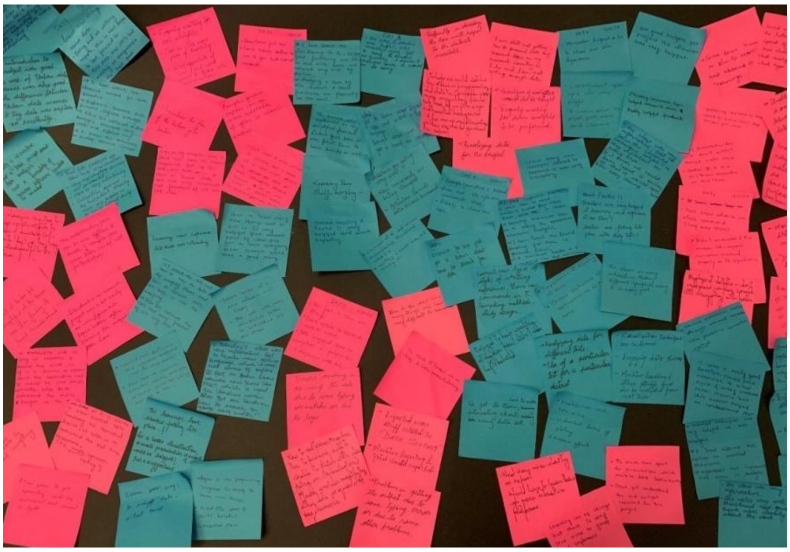

Fig. 2. Student Feedback in Pink and Blue Post-it notes

This process of taking regular feedback from the students and understanding the capability of the students was well appreciated. Two faculties were present in class, which helped give individual attention to the students.

4. RESULTS AND DISCUSSION

General information of the students selected responses to a post feedback survey and students' major strengths of the Summer School were captured in an online survey during the course duration to improve the course methodology from their feedback as shown in Tables 1-3.

As shown in Table 1, the class distribution was fairly equal among male and female students (53.85% vs. 46.15%). The class had more post-graduate than undergraduate students (70.83% vs. 29.19%). The majority of the students (96.15%) used Windows operating system for the class exercises. Around 84.62% of the participants had never used statistical software, while 65.99% of the students had never used programming software. This gave the instructors a good idea about the level of the students. About 30.77% of the students were unsatisfied with their current data management and analytics skills before the summer school while 46.15% were neutral. Only a minor proportion (7.69%)

said they were satisfied with the level of their current data management and analytics skills, and 15.53% were unsure. The most selected reason for participating in the summer school was to learn new skills that can be applied to current work (40.88%) and help get a job or promotion (26.92%).

Table 1. General information of participants (n=26)

Characteristics	Percentage (%)
Participants' level of Education	
Post Graduate	70.83
Under Graduate	29.17
Gender	
Male	53.85
Female	46.15
Operating System used by participants	
Windows	96.15
Apple/Mac OS/UNIX	3.85
How often do participants use a Statistical Software	
Never	84.62
Less than once per year	11.54
Monthly	3.85
How often do participants use a Programming Software	
Never	65.99
Weekly/Daily	14.78
Monthly	11.54
Less than once per year	7.69
Participants satisfaction rating of the level of current data management and analysis skills before summer school	
Unsatisfied	30.77
Neutral	46.15
Satisfied	7.69
Not sure	15.53
Reason for Participating in the Summer School	
To learn new skills that I can apply to my current work	40.88
To learn new skills that will help me get a job or a promotion	26.92
To learn skills that I can apply to my work in the future	24.5
As a requirement for my program	3.85
To refresh or review my skills	3.85

Table 2. Selected responses to the post feedback survey (n=26)

Characteristics	Percentage (%)
I felt comfortable learning in this summer school environment	
Strongly agree	41.67
Agree	50.0
Neutral	8.33
Disagree	0.0
Strongly disagree	0.0
I can immediately apply what I learned at this summer school	
Strongly agree	29.17
Agree	50.0
Neutral	16.67
Disagree	4.17
Strongly disagree	0.0

Characteristics	Percentage (%)
I was able to get clear answers to my questions from the instructors	
Strongly agree	41.67
Agree	50.00
Neutral	4.17
Disagree	4.17
Strongly disagree	0.0
The instructors were enthusiastic about the summer school	
Strongly agree	83.33
Agree	8.33
Neutral	8.33
Disagree	0.0
Strongly disagree	0.0
I felt comfortable interacting with the instructors	
Strongly agree	83.33
Agree	12.50
Neutral	4.17
Disagree	0.0
Strongly disagree	0.0
The instructors were knowledgeable about the material being taught	
Strongly agree	83.33
Agree	16.67
Disagree	0.0
Strongly disagree	0.0

Table 3. Students feedback on the major strength of the Summer School translated verbatim (n=29)

Responses	Percentage (%)
The Big field of data analysis and statistics is explained at the cellular level making it easy to grasp	20.04
The syllabus was broad and the course structure and study materials were well written with access to all the important concepts which can be used further on my own	19.04
Now I can manage, create and analyze larger data-sets in Excel and R	14.28
Impressive depth of knowledge of wide topics covered by the instructors	9.2
Interaction with the instructors by the students and the attention paid by the instructors to the students was great	9.2
The analysis and the capacity of handling the data by the software is good using a step by step instruction process by the instructors	9.2
Learnt programming and statistics that I can implement on my projects as well as job criterion	4.76
Misconceptions about Big Data, Machine Learning and excels were cleared	4.76
Smooth and to the point learning. A lot was taught and I didn't get confused.	4.76
The summer school was designed in a very well planned manner. I am amazed to see how easy it has become for me to work on R Studio. The course gave me so much information about how a paper should be written, how can data be cleaned easily for the first time	4.76

As shown in Table 2, it was very satisfactory to see the results of the post-summer school feedback which observed that 91.67% of the students were comfortable learning in the summer school and 79.17% agreed that they could immediately apply the knowledge gained in this summer school to their respective study. The majority of the students (91.67%) said they were able to get clear answers to their questions from the instructors, and 95.83% said they felt comfortable interacting with the instructors. All of the participants (100.0%) agreed that the instructors were knowledgeable about the material being taught.

Table 3 reflected the students' feedback on the Major strength of the Summer School. The most were that the Big field of data analysis and statistics is explained at the cellular level making it easy to grasp (20.04%) followed by the syllabus was broad and the course structure and study materials were well written with access to all the important concepts which can be used further on their own (19.04%).

The students' feedback made it clear that students understood the open data and data sharing policy, which was reflected in the variety of projects chosen. Also, from the students' feedback, it was concluded that they learned new terminology, which added to a better understanding of the importance of Urban Data Science.

Fig. 3 shows the Word Cloud on the various terms students felt they had heard for the first time in the summer school and would help them in future research.

Fig. 3. Area of Students Learning from the Summer School

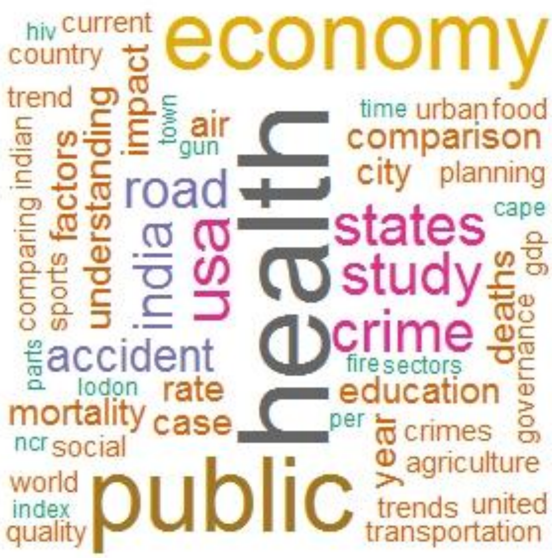

Fig. 4. Topics selected by students for Case Study

5. THE VALUE

The outputs of the case studies were spread in varying domains of Crime, the Economy, Education, Environment, Public Health, Road Accident, Sports, Governance and Planning, etc. The Word Cloud shown below depicts the various topics covered by the students from the 2018 and 2019 batch; addressing a vast domain of urban problems.

One of the interesting studies presented by the summer school student was on the Prevalence of Obesity among Socially Vulnerable Groups in the United States (Fig. 5). Heart disease, linked to obesity, is one of the leading causes of death in the United States [35,36]. Because it's debatable whether or not it's a disease, finding an effective treatment is slow [37,38,39]. The opensource dataset used for the study was the National Health and Nutrition Examination Survey (NHANES) from the year 2013-14 (kaggle.com). The Pearson's correlational coefficient (r) result showed a weak but significant negative correlation between annual household income and BMI (see Fig. 6). From the analysis, as the annual household income increases, the BMI decreases. This could be due to the ability to afford high-quality, nutritious food as well as having access to healthcare.

Cross-tabulations using a statistical test like Chi-Square, Pearsons Correlation, Kendall's rank correlation, ANOVA (Analysis of Variance) were performed on Race, Age, Income, Education with its relation to Obesity. Interesting results were found proofing that Obesity is prevalent among Ethnic/Racial minorities, and that socioeconomic, racial factors influence obesity in children and the elderly. The study's findings supported that people from Low-income households having lower educational levels showed a tendency to be more vulnerable to obesity due to poor dietary options [40].

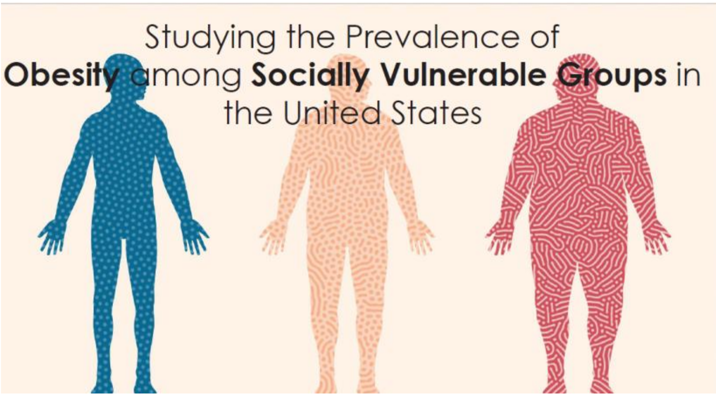

Fig. 5. Project cover page for Prelevance of Obesity among Socially Vulnerable Groups in the United States

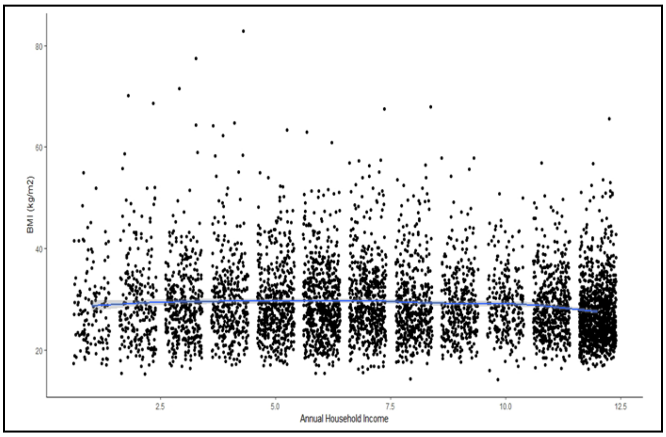

Fig. 6. Scatterplot showing a weak positive correlation between annual household income and BMI

Recommendations were made from this study based on the findings, which include: Food stamps, raw food materials could be provided for economically weaker households; The current school lunch program in the United States is not focused on providing a menu of healthy foods, and could be corrected; Unhealthy fast foods and sugary sodas should be taxed higher to discourage their popularity; Better medical care, especially the public medical care and insurance system in the United States needs to be strengthened and strong socialist reforms are needed in this sector to provide free healthcare for all.

A second student analyzed the performance of Indian states and union territories in terms of Sustainable Development Goals (SDG) for the year 2018. NITI Aayog has been entrusted with the role to co-ordinate 'Transforming our world: the 2030 Agenda for Sustainable Development' at the national level. From the study, evaluation is carried out at the country level as well as for each state and Union territory and a score is given between 0 to 100; data was collected from https://niti.gov.in/. The spatial distribution of the Scores is shown in Figure 7 which rightly indicates that none of the states have been able to achieve 100 scores to date. Analyzing each goal individually helps identify the direction in which the government should take action. "Gender Equality" is still one of the alarming situations in India for all the states.

Fig. 8 shows that for goals 6, 10, and 15 some of the states have reached the Achievers level. It was highlighted that there is an alarming situation for Goal 9 and 11 in most India states. Progress on SDG 12, 13, and 14 could not be measured because relevant state-level data could NOT be consolidated or found. SDG 17 was left out because the goal is focused on international partnerships, being less relevant for domestic level policy actions. After studying the individual states SDG scores, it was observed that Uttar Pradesh is having the lowest score i.e. 42, while Kerala and Himachal Pradesh are having the highest score (69). The range of scores is from 42-69. Hence, there is a tremendous scope of improvement to bring the National SDG Index to 100 before 2030. Various states can adopt successful models from other states to achieve every goal efficiently in less time. The recommendation from the analysis of this study is that political willingness via effective policy decision-making from the study findings could help towards enhancing Indian's sustainable development targets.

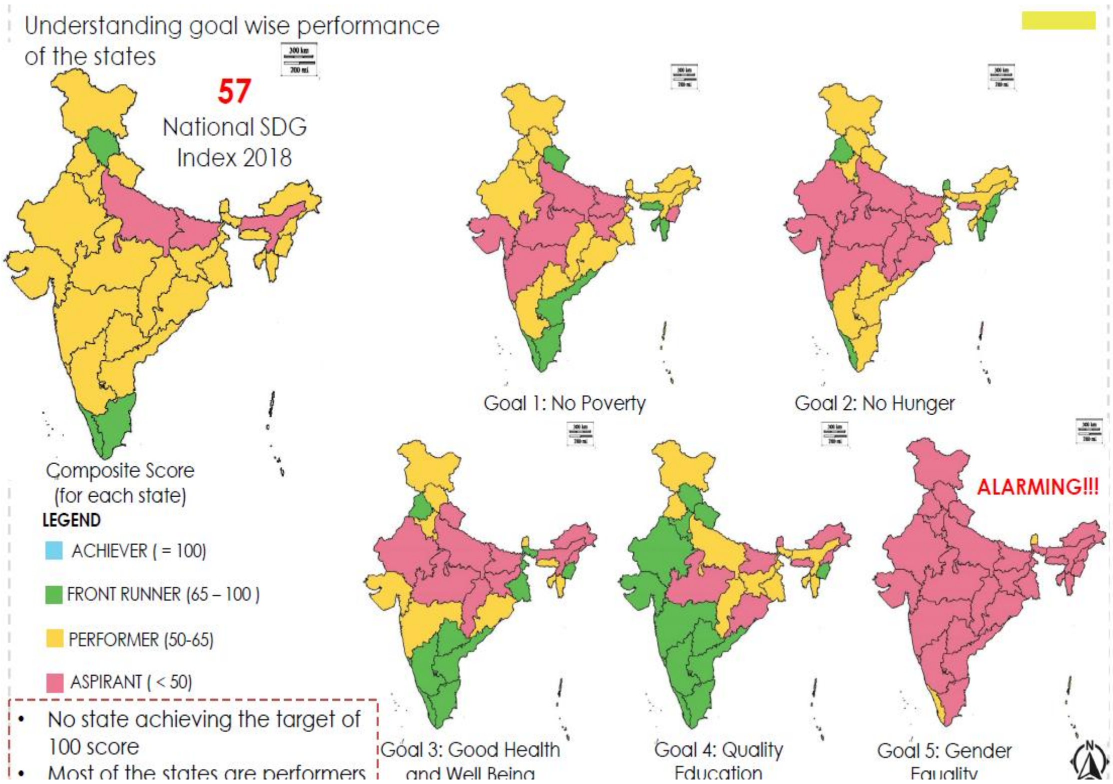

Fig. 7. Understanding the Status of SDG goals 1-5 for the states of India

Another case study from a student accessed the behavioural pattern of Terrorist Attacks in India. The study findings showed a high prevalence of attacks in two Indian cities: Jammu and Kashmir. The case findings can be a catalyst for innovative strategy in India's ministry of defense in preparation for future attacks. (Fig. 9).

The students presented many more interesting case studies, captured in Table 4.

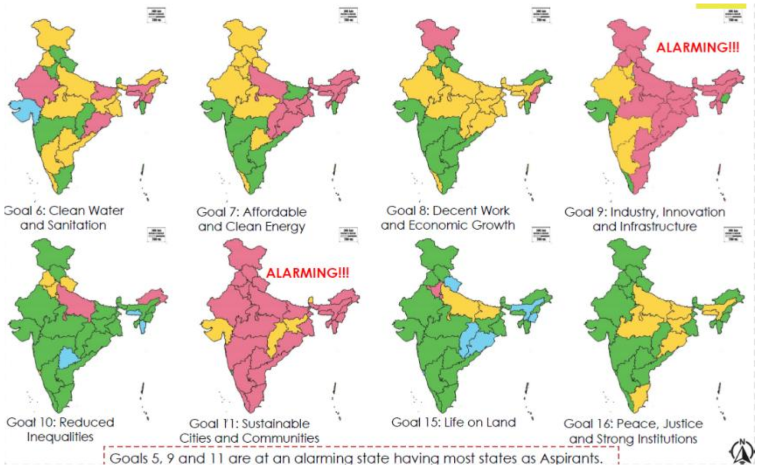

Fig. 8. Understanding the Status of SDG goals 6-16 for the states of India

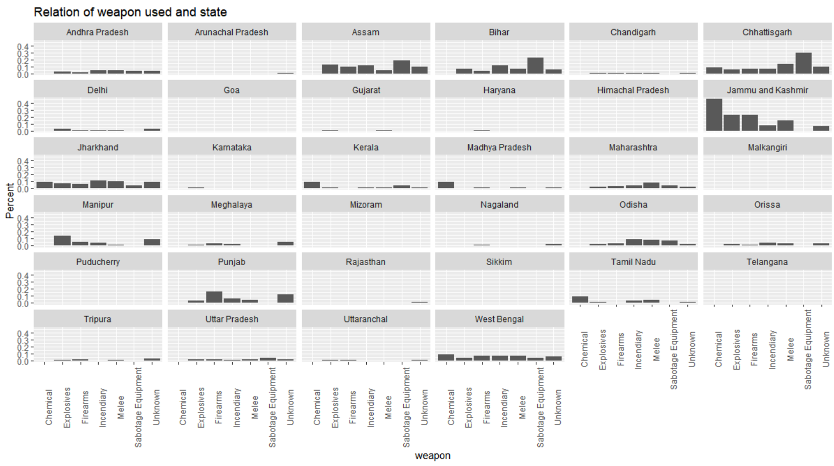

Fig. 9. Relation of the weapon used in the tarries attack with the state

Table 4. Student's Research Titles for the Urban Data Science Summer School 2018 and 2019

Topic	Year	Title
Crime	2018	The crime rate in the City of Chicago
		Analyzing determinants of gun violence In the United States of America
		Understanding and interpreting crime trends in Greater London, England in 2016.
	2019	Crime Against Women in India and its impact on the economy of the Country
		Understanding the pattern of terrorist attacks in India
Disaster	2018	Vegetation fire in Cape Town
Economy	2018	Understanding the economic trend based on different sectors in the USA
		Suicides in India: A Case study
	2019	Impact of Foreign direct investment on the Human Development Index in the past 27 years

Topic	Year	Title
		Impact of Contraction and Expansion of Establishments towards Employment in the US
		Agriculture input versus the urban population and its effects
		Current Health expenditure per capita in the USA
		Food import value and volume changes with respect to time in the USA
		Comparison of Suicide Rate to GDP and HDI
		Comparison of Current health expenditure (CHE) as a percentage of gross domestic product (GDP) (%) - by country
Education	2018	Educational attainment for adults age 25 and older for the U.S., States, and counties, 1970 – 2016
	2019	Contribution of Education in the development of countries across the world
Food	2018	Homemade beer recipes
		Food Accessibility System of US Counties
Governance	2018	Accidental deaths & suicides in India
	2019	Comparing the number of reported confirmed cases to reported deaths
Governance and Planning	2019	Analyzing the performance of Indian states and union territories in terms of sustainable development goals for the year 2018
		Recognizing the disparity between 'planned' and 'unplanned' parts of the city and its effects on people's choices; a case of Ahmedabad
Pollution	2018	Factors degrading air quality level in Gujarat.
		Analyzing Trends of Pollution: A Spatio-Temporal Study of US air pollution
		Air Quality for the Major Urban States of US: Focussed Study On Particulate Matter (PM 10, PM 2.5) from 2008 to 2017
	2019	Accessing policy need for Delhi (NCR): Comparing pollution levels in Delhi and Beijing
Public Health	2018	The relation of Demographic Parameters on mortality of Cancer incidents in the States of USA.
	2019	A study on the prevalence of obesity among Socially Vulnerable groups in the United States
		Correlation and interdependence of diseases across 500 cities in the USA
		Effect of socioeconomic parameters on suicidal deaths
		The Prevalence of HIV in the Year 2002-2011 Globally
		Social Demographic Outcomes Of Death Caused By Drugs In United States 2003-2017
		Understanding the mortality rate and its relation to the health condition of children around the world
Real Estate	2018	Predicting housing prices: a case for Seattle housing crisis, USA
	2019	Factors affecting Real Estate Prices in IOWA city, USA
Road Accident	2018	Mortality from land transport accidents
	2019	Road accident collision analysis and pedestrian safety study in Great Britain(UK)
		An assessment of the association of environmental factors and road fatalities in Britain.
Social Media	2018	YouTube videos
Sports	2018	Indian premier league matches
	2019	Comparison popularity of football player
Transportation	2018	Monitoring the trend of Cross border crossing vehicles in the USA
		The inflow of Foreign Tourists to India
		Traffic Violations in Montgomery County
		Statistical analysis of Annual average daily flow of traffic of east midlands, Britain

6. THE FUTURE

While there are some concerns if the motivation for data science education is commercial rather than intellectual developments according to Donoho [41], we cannot argue the impact it has on data-driven developments in scientific research [42, 43]. In other words, there is a need for further development of such courses in various institutions because opportunities are rapidly expanding and will continue to do so in the coming years [44]. Urban data science is in its formative stages of development and will evolve among professionals and students desiring to shape the future of smart cities [25, 45]. Adopting data science research into a wide range of fields, such as urban planning and public transportation systems, will create a culture of experimentation with machine learning models and cloud infrastructure solutions toward a digitalized world. Advances in this emerging field will make it possible to generate new insights from such data, which will complement government-generated survey data thus adding to the well of information and knowledge on human well-being.

7. CONCLUSION

Based on the findings of the study, We believe that urban data science education can provide the evidence needed to monitor and promote effective development policies, and It can also inform the general public about the effectiveness of education as a critical resource for improving data usage in improving governance through impactful decision making and policies necessary for quality delivery for long-term development. This inclusion should also galvanize a broad effort toward the structured development of an Urban Data Science curriculum, which has the potential to improve data-driven policymaking for solving urban problems in Smart cities.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support from the Summer-Winter School (SWS) of CEPT University, Ahmedabad, India, the school's students, and the Committee on Data for Science and Technology (CODATA). Sincere thanks also go to Dr. Simon Hodson, Executive Director, CODATA.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Kharrazi A, Qin H, Zhang Y. Urban big data and sustainable development goals: Challenges and opportunities. Sustain. 2016;8:1–6.
2. Lee JG, Kang M. Geospatial Big Data: Challenges and Opportunities. Big Data Res. 2015;2: 74–81.
3. Jin X, Wah BW, Cheng X, et al. Significance and Challenges of Big Data Research. Big Data Res. 2015;2:59–64.
4. [Larson D, Chang V. A review and future direction of agile, business intelligence, analytics and data science. Int J Inf Manage. 2016;36:700–710.
5. Archenaa J, Anita EAM. A survey of big data analytics in healthcare and government. Procedia Comput Sci. 2015; 50:408–413.
6. Wang Y, Kung LA, Byrd TA. Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. Technol Forecast Soc Change. 2018;126:3–13.
7. Ahmed F, Capretz LF, Bouktif S, et al. Soft skills and software development: A reflection from software industry. Int J Inf Process Manag. 2013;4:171–191.
8. Börner K, Maltese A, Balliet RN, et al. Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. Inf Vis. 2016;15:198–213.
9. Jang H. Identifying 21st Century STEM Competencies Using Workplace Data. J Sci Educ Technol. 2016;25:284–301.
10. Hastings JS, Howison M, Lawless T, et al. Unlocking data to improve public policy. Commun ACM. 2019;62:48–53.
11. Hoeyer K. Data as promise: Reconfiguring Danish public health through personalized medicine. Soc Stud Sci. 2019;49:531–555.
12. Perl A, Howlett M, Ramesh M. Policy-making and truthiness: Can existing policy models cope with politicized evidence and willful ignorance in a “post-fact” world? Policy Sci. 2018;51:581–600.
13. Ocloo J, Garfield S, Franklin BD, et al. Exploring the theory, barriers and enablers for patient and public involvement across health, social care and patient safety: A systematic review of reviews. Heal Res Policy Syst. 2021;19:1–21.
14. Bibri SE. Data-driven smart sustainable cities of the future: An evidence synthesis

approach to a comprehensive state-of-the-art literature review. Sustain Futur. 2021; 3:1–23.
15. Capotosto J. Data opportunities and risks: the dynamic of public, personal, and commercial interest. J Community Saf Well-Being. 2017;2:1–4.
16. Pokhriyal N, Jacques DC. Combining disparate data sources for improved poverty prediction and mapping. Proc Natl Acad Sci U S A. 2017;114:E9783–E9792.
17. Adelman M, Haimovich F, Ham A, et al. Predicting school dropout with administrative data: new evidence from Guatemala and Honduras. Educ Econ. 2018;26:356–372.
18. Townsend A. Cities of data: Examining the new urban science. Public Cult. 2015;27: 201–212.
19. Nations U. Big Data for Sustainable Development. United Nations 2019;1–2.
20. Hassani H, Huang X, Macfeely S, et al. Big data and the united nations sustainable development goals (UN SDGs) at a glance. Big Data Cogn Comput. 2021;5:1–29.
21. Bannister J, O'Sullivan A. Big Data in the city. Urban Stud. 2021;58:3061–3070.
22. Kitchin R. The ethics of smart cities and urban science. Philos Trans R Soc A Math Phys Eng Sci. 2016;374:1–15.
23. Glaeser EL, Kominers SD, Luca M, et al. Big Data and Big Cities: the Promises and Limitations of Improved Measures of Urban Life. Econ Inq. 2018;56:114–137.
24. Kitchin R. The real-time city? Big data and smart urbanism. GeoJournal. 2014;79:1–14.
25. Williamson B. Educating the smart city: Schooling smart citizens through computational urbanism. Big Data Soc. 2015;2:1–13.
26. Leszczynski A. Speculative futures: Cities, data, and governance beyond smart urbanism. Environ Plan A. 2016;48:1691–1708.
27. Wilson MW. Location-based services, conspicuous mobility, and the location-aware future. Geoforum. 2012;43:1266–1275.
28. Zhou X, Ding Y, Wu C, et al. Measuring the spatial allocation rationality of service facilities of residential areas based on internet map and location-based service data. Sustain. 2019;11: 1–20.
29. Brennand CARL, De Souza AM, Maia G, et al. An intelligent transportation system for detection and control of congested roads in urban centers. In: Proceedings - IEEE Symposium on Computers and Communications. 2016;663–668.
30. Nellore K, Hancke GP. A survey on urban traffic management system using wireless sensor networks. Sensors (Switzerland). 2016;16:157.
31. Singh B, Gupta A. Recent trends in intelligent transportation systems: A review. J Transp Lit. 2015; 9:30–34.
32. De Souza AM, Yokoyama RS, Maia G, et al. Real-time path planning to prevent traffic jam through an intelligent transportation system. In: Proceedings - IEEE Symposium on Computers and Communications. 2016;726–731.
33. Marsh JA, Farrell CC. How leaders can support teachers with data-driven decision making: A framework for understanding capacity building. Educ Manag Adm Leadersh. 2015;43:269–289.
34. Datnow A, Hubbard L. Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. J Educ Chang. 2016;17:7–28.
35. Rao G, Powell-Wiley TM, Ancheta I, et al. Identification of obesity and cardiovascular risk in ethnically and racially diverse populations: A scientific statement from the American Heart Association. Circulation. 2015;132:457–472.
36. Rosenthal RJ, Morton J, Brethauer S, et al. Obesity in America. Surg Obes Relat Dis. 2017; 13:1643–1650.
37. Allison DB, Downey M, Atkinson RL, et al. Obesity as a disease: A white paper on evidence and arguments commissioned by the council of the obesity society. Obesity. 2008;16:1161–1177.
38. Stoner L, Cornwall J. Did the american medical association make the correct decision classifying obesity as a disease? Australas Med J. 2014;7:462–464.
39. Kyle TK, Dhurandhar EJ, Allison DB. Regarding Obesity as a Disease: Evolving Policies and Their Implications. Endocrinol Metab Clin North Am. 2016;45:511–520.
40. French SA, Tangney CC, Crane MM, et al. Nutrition quality of food purchases varies by household income: The SHoPPER study. BMC Public Health. 2019;19: 1–7.
41. Donoho D. 50 Years of Data Science. J Comput Graph Stat. 2017;26:745–766.
42. Clauset A, Larremore DB, Sinatra R. Data-driven predictions in the science of science. Science. 2017;355:477–480.

43. Lytras MD, Raghavan V, Damiani E. Big data and data analytics research: From metaphors to value space for collective wisdom in human decision making and smart machines. Int J Semant Web Inf Syst . 2017;13:1–10.
44. Baumer B. A Data Science Course for Undergraduates: Thinking With Data. Am Stat. 2015; 69:334–342.
45. Acuto M, Parnell S, Seto KC. Building a global urban science. Nat Sustain. 2018;1:2–4.

Peer-review history:
The peer review history for this paper can be accessed here:
https://www.sdiarticle5.com/review-history/85525