

Training on

# Data Analysis for Governance

(Report)

January 10-11, 2022

*Delivered to*



DEPARTMENT  
OF PUBLIC  
ENTERPRISES

*Conducted by*



**PUBLIC AFFAIRS FOUNDATION**  
Partnership for Better Governance



## Contents

Introduction.....	1
Objectives of the Training .....	2
Structure of the Programme and Methodology .....	2
Participants.....	3
Case Studies .....	4
Questions and Answers .....	4
Annexures .....	i
Annexure 1: List of Participants .....	i
Annexure 2: Snapshots.....	iv



## Introduction

The volume of data in the world is increasing exponentially. Data is becoming a vital part of decision-making. Today, in the private sector, analysis of big data is commonly used to extract insights for decision making and to gain competitive advantage and risk mitigation. Some of the applications include risk profiling, sales forecasting, fraud detection, stock market prediction and the list goes on. Similarly, data Science can find ample application in government. It can help in extracting useful information and knowledge from large volumes of data in order to improve government decision-making or provide the insights to make data-driven decisions.

In recent years, there is a growing awareness that data science can provide new sources of evidence for policy-making. However, there is still a dearth of digitised and disaggregated data. The importance of data analytics in the field of governance or public administration can be categorised broadly into 4 areas.

**Evidence-based decision making:** Data Analytics acts as the supporting basis for decisions as it enables policy decisions based on or informed by objective facts rather than ideologies, perceptions or common sense.

**Resource optimisation:** Resource optimization is the set of processes and methods to match the available resources (human, machinery, financial) with the needs. It can be used while allocating resources in the various sectors of the government such as health, education, agriculture etc.

**Measuring governance:** The availability of statistical data helps in measuring the performance of the government in various aspects of governance. Key Performance Indicators (KPI) are a set of quantifiable measures that can be defined to measure its performance over time.

**Effective Policy implementation:** Data Analytics allows for better prediction of social, economic, environmental, cultural and other phenomena and improves causal inference. This helps the government to design better policies.

India has been one of the fastest-growing economies in the world for more than two decades. However, his growth has been unequal with the top 1% of earners receiving 22% of income and the top 10% earners accounting for 56%. It continues to fare poorly in human development indicators such as malnutrition, learning outcomes, maternal mortality, infant mortality, etc. The application of data science can help determine the factors impacting human development and recommend specific interventions for improvement. Data science can also be applied to monitor India's progress in achieving sustainable goals.

## Objectives of the Training

The specific objectives of the data training workshop explored the concepts of data and data science and their importance in policymaking and governance.

- Enhanced the participants' understanding of data and different types of data
- Provided a comprehensive understanding of the data science process
- Explored case studies to elucidate the potential application of data science for governance and policymaking.

## Structure of the Programme and Methodology

The structure of the programme was designed to introduce the participants to the application of data science in the field of for governance and policy making.

- An overview of data science applications and its potential in the field of governance with emphasis on the Indian context
- Data and its types
  - Qualitative and Quantitative data
  - Structured, Semi-Structured and Unstructured data
  - Meta data
  - Big data
- Data Science and Data Science Process
  - Framing Problem
  - Identifying Research Questions
  - Developing Conceptual and Data framework
  - Data Collection
  - Data Preparation
  - Exploratory Data Analysis
  - Models and Algorithms
  - Reporting
- Case Studies
  - Public Affairs Index
  - Multi-Dimensional Poverty Index
- The methodology included
  - Theoretical presentations of the concepts

Instructor-led exercises on the data science methods

## Participants

The participants included Senior Manager, Deputy Manager, AGM, Accounts Manager, Systems Engineer, Supervisors from KSRTC, KSIIDC, KSBCL, KSFC, etc. The detailed list is given in the annexure.

### Activities

The activities included the following instructor-led group activities.

- Identifying qualitative and quantitative data  
The participants were given an exercise to identify quantitative and qualitative data from a data set consisting of 10 variables.
- Identifying structured, semi-structured and unstructured data  
The participants were presented different documents with numbers, text, audio, images, html tags, etc, and made to classify the documents as containing structured, semi-structured and unstructured data.
- Calculation of mean, median, mode and standard deviation  
The participants were given a dataset containing details of Sachin Tendulkar's performance in one-day matches. The participants were made to identify and apply univariate analysis to answer the following questions.
  - What is the average batting score of Sachin Tendulkar?
  - Which batting position has Tendulkar batted the most?
  - Graphically illustrate Tendulkar's different modes of dismissals

Another dataset containing employee details from a job portal was given and the participants were again asked to apply univariate analysis for the following question.

- What is the median Annual Salary?

### Correlation analysis

The participants were given a dataset containing fuel consumption and 10 aspects of automobile design and performance for 32 automobiles. They were asked to establish association using both correlation technique and scatter plot for the following variables

- Fuel Consumption (defined as miles per gallon) vs weight of the car
- Fuel Consumption (defined as miles per gallon) vs horse power
- Fuel Consumption (defined as miles per gallon) vs number of cylinders

### Linear regression

Using the same car dataset mentioned above, the instructor showed the application of linear regression to determine the impact of various factors on Fuel Consumption.

## Case Studies

### Case Study 1

#### Public Affairs Index

The participants of the training were introduced to the methodology followed to create the Public Affairs Index 2021 (PAI 2021). The participants were walked through the ideation of indicators, preparation of data and the basics of Principal Component Analysis (PCA) technique of dimensionality reduction."

### Case Study 2

#### Multidimensional Poverty Index

The participants of the training were introduced to the concept of multiple, competing and simultaneously existing poverties and briefed about the methodology to capture the same in the form of the Multidimensional Poverty Index (MPI) developed by Alkire-Foster and endorsed by the United Nations Development Programme (UNDP) in 2010. The calculations within the methodology were elaborately discussed. Further, a demonstrative example of an MPI analysis conducted by PAC-CODR was shared."

## Questions and Answers

1. How to use data science in the field of accounts, especially in budgeting and daily transactions.

The budgets are usually prepared by evaluating the results of the current year. Data science enables a more scientific preparation by analysing the trends of multiple years that help account for shocks, surprises, etc.

Data science helps uncover valuable insights within the transactions.

2. When to use mean, median, and mode?

Mean is the most commonly used measure of central tendency. It is used when the variables are numeric. If the numeric variable contains extreme values, median is preferred while mode is mainly used for categorical variables.





## Annexures

### Annexure 1: List of Participants

**DEPARTMENT OF PUBLIC ENTERPRISES**  
**INFORMATION OF PARTICIPANTS FOR ONLINE TRAINING**

<b>SUBJECT</b>		<b>Data Analytics for Governance</b>						
<b>NAME OF INSTITUTE</b>		<b>Public Affairs Foundation, Bangalore(PAF)</b>						
<b>DATE OF TRAINING</b>		<b>10/01/2022 - 11/01/2022</b>						
<b>Sl No</b>	<b>Name Mr/Mrs</b>	<b>Designation</b>	<b>e-mail ID</b>	<b>Mobile Number</b>	<b>Day 1</b>	<b>Day 2</b>	<b>Comments from Trainer</b>	<b>Name of PSE</b>
1	P.S.Nandisha	DGM- Non Technical	<a href="mailto:nandi9creative@gmail.com">nandi9creative@gmail.com</a>	9448290616	P	P	No Comments	MCA
2	Nagappa H.Kittur	Manager-CSD	<a href="mailto:nagarajkittur69@gmail.com">nagarajkittur69@gmail.com</a>	9448290613	P	P	No Comments	
3	R.Rangawamy	A/c Officer- Finance	<a href="mailto:md@kavika.co.in">md@kavika.co.in</a>	7019880169	P	P	No Comments	KAVIKA
4	Surjith	Manager	<a href="mailto:surjith046@gmail.com">surjith046@gmail.com</a>	9945298240	P	P	No Comments	KFDC
5	C.Vidyananda	GM(M&C)	<a href="mailto:mktg.kssc@gmail.com">mktg.kssc@gmail.com</a>	9448990385	P	P	No Comments	KSSC
6	Nagabhushana N	Manager (A/c) I/c	<a href="mailto:abhushan07@gmail.com">abhushan07@gmail.com</a>	9535321007	P	P	No Comments	MPVL
7	Mahaboobi A Badekan	Manager	<a href="mailto:mahaboobi@rediffmail.com">mahaboobi@rediffmail.com</a>	9448287546	P	P	No Comments	RGHCL
8	Shruthi.S	Development Officer	<a href="mailto:md.dbcdc@gmail.com">md.dbcdc@gmail.com</a>	7795119577	P	P	No Comments	DBCDC
9	Shwetha.S	Development Officer	<a href="mailto:dmchamarajinagar@gmail.com">dmchamarajinagar@gmail.com</a>	9986562851	P	P	No Comments	
10	M. Kiran kumar	Computer Programmer	<a href="mailto:md.dbcdc@gmail.com">md.dbcdc@gmail.com</a>	9964111915	P	P	No Comments	
11	Ganesh Kundar	AEE(Ele)	<a href="mailto:sepmescom@rediffmail.com">sepmescom@rediffmail.com</a>	9448289539	P	P	No Comments	MESCOM
12	Sanjeev Ranagatti	AE(Ele)	<a href="mailto:aeet1@mesco.in">aeet1@mesco.in</a>	9480841361	P	P	No Comments	



## DPE Report on Data Analytics

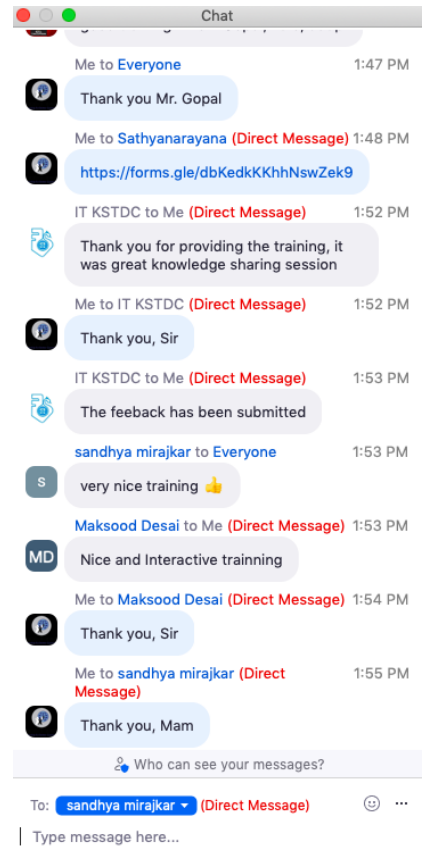
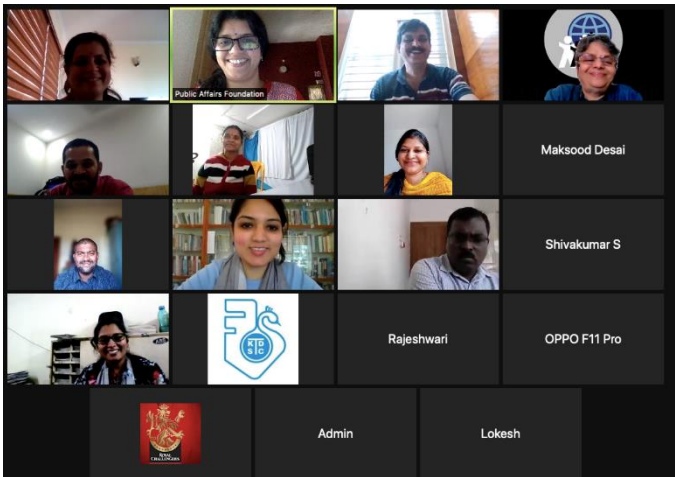
13	Rajendra	System Engineer	<a href="mailto:info.kmdc@gmail.com">info.kmdc@gmail.com</a>	8123277463	P	P	No Comments	KMDC
14	B K Praksah	EE(M)	<a href="mailto:nodalofficerkerp@gmail.com">nodalofficerkerp@gmail.com</a>	7892554837	P	P	No Comments	KPCL
15	Ganesh pondve	AEE(Sys)	<a href="mailto:pandave.ganesh@gmail.com">pandave.ganesh@gmail.com</a>	9972292420	P	P	No Comments	
16	M.R.Munji	CSM	<a href="mailto:csm.nwkrctc@karnataka.gov.in">csm.nwkrctc@karnataka.gov.in</a>	7760991519	P	P	No Comments	NWKRCTC
17	Ismail S Navalagunda	Statistical Officer	<a href="mailto:cpso.nwkrctc@karnataka.gov.in">cpso.nwkrctc@karnataka.gov.in</a>	7760982550	P	P	No Comments	
18	Sathisha Kadannavara	Asst Statistical Officer	<a href="mailto:incharaksrc@gmail.com">incharaksrc@gmail.com</a>	7760991905	P	P	No Comments	
19	Siddappa G	DC	<a href="mailto:dcklb1.nekrctc@karnataka.gov.in">dcklb1.nekrctc@karnataka.gov.in</a>	7760992014	P	P	No Comments	KKRCTC
20	R V Puranik	Dy.CME	<a href="mailto:cme.nekrctc@karnataka.gov.in">cme.nekrctc@karnataka.gov.in</a>	7760992010	P	P	No Comments	
21	Sathyanarayana	Sr.Manager	<a href="mailto:sathyanarayanadg@rediffmail.com">sathyanarayanadg@rediffmail.com</a>	9449728415	P	P	Active	KSFC
22	Srinivasa P C	Sr.Manager	<a href="mailto:srinivas1871965@gmail.com">srinivas1871965@gmail.com</a>	9448155768	P	P	Active	
23	Shivakumar S	Manager	<a href="mailto:sugurskr@gmail.com">sugurskr@gmail.com</a>	9449138641	P	P	No Comments	
24	Sathyanarayana	Dy.Manager	<a href="mailto:sathyakksfc@yahoo.com">sathyakksfc@yahoo.com</a>	9880434320	P	P	No Comments	
25	Gopala	Asst.Manager	<a href="mailto:gopalkunder@yahoo.co.in">gopalkunder@yahoo.co.in</a>	9731877517	P	P	No Comments	
26	Dakshina Murthy	AGM	<a href="mailto:murthy.ksiidc@gmail.com">murthy.ksiidc@gmail.com</a>	9845833566	P	P	Active	KSIIDC
27	Chandrashekar Chorgi	DGM	<a href="mailto:cschoragi1962@gmail.com">cschoragi1962@gmail.com</a>	9845478279	P	P	No Comments	KSMCL
28	Yogarajaiah S	Asst.Manager	<a href="mailto:yogarajaiah1962@gmail.com">yogarajaiah1962@gmail.com</a>	9980106233	P	P	No Comments	
29	Daksha Uthappa	IT consultant	<a href="mailto:it@kstdc.com">it@kstdc.com</a>	9980970906	P	P	Active	KSTDC
30	Maksood Desai	Supervisor	<a href="mailto:syscomdesai@ksbcl.com">syscomdesai@ksbcl.com</a>	9448994005	P	P	Active	KSBCCL
31	M Anand	Dy Eng-CAMS	<a href="mailto:camsngcf@yahoo.com">camsngcf@yahoo.com</a>	9483674897	P	P	No Comments	NGEF
32	Girish Purantar	Dy.Manager(Mktg)	<a href="mailto:mktngcf@yahoo.co.in">mktngcf@yahoo.co.in</a>	9483674900	P	P	No Comments	
33	Manjunath Hansbhavi	Asst Eng-MPS	<a href="mailto:mpsnger@yahoo.co.in">mpsnger@yahoo.co.in</a>	9738652175	P	P	No Comments	
34	Anthony Maria Emmanuel M	M.D.	<a href="mailto:tame1972@gmail.com">tame1972@gmail.com</a>	9448412935	P	P	No Comments	KSPAML
35	Mallappa B	A.O.	<a href="mailto:mdktamlqlb@gmail.com">mdktamlqlb@gmail.com</a>	9902951960	P	P	No Comments	



## DPE Report on Data Analytics

36	N. Chandrashekar	Dy. Manager	<a href="mailto:ncs.n36@gmail.com">ncs.n36@gmail.com</a>	9481094407	P	P	No Comments	Lidkar
37	H.M.Shankaraiah	Dy. Manager	<a href="mailto:shankaraiahhm@gmail.com">shankaraiahhm@gmail.com</a>	9845397282	P	P	No Comments	MSIL
38	B. Krishnamurthy	Dy. Manager	<a href="mailto:kmurthy1246@yahoo.com">kmurthy1246@yahoo.com</a>	9480812960	P	P	No Comments	
39	Lokesh	Dy. Manager (Met)	<a href="mailto:lokeshhgml@gmail.com">lokeshhgml@gmail.com</a>	9902814070	P	P	Active	HGMCL
40	G.Suresh	Jr. Engineer (Mining)	<a href="mailto:suresh.gopala@ka.gov.in">suresh.gopala@ka.gov.in</a>	9986633328	P	P	No Comments	
41	T.S. Manjunath	Dy. Manager (Mining)	<a href="mailto:manjunath.ts@ka.gov.in">manjunath.ts@ka.gov.in</a>	9986253936	P	P	No Comments	
42	Sandhya C. Mirajkar	AEE(Ele)	<a href="mailto:sandhyamrajkar83@gmail.com">sandhyamrajkar83@gmail.com</a>	9535990013	P	p	Active	KPTCL
43	Pushpavath G.K.	AE(Ele)	<a href="mailto:pushpa.suraj482@gmail.com">pushpa.suraj482@gmail.com</a>	9743020866	P	P	Active	
44	Azeemuddin	Assistant	<a href="mailto:azeem.mrb@gmail.com">azeem.mrb@gmail.com</a>	9739497786	P	P	No Comments	
45	Vajrakumar T	Jr.Eng(EI)	<a href="mailto:vajrakumart1111@gmail.com">vajrakumart1111@gmail.com</a>	9448275982	P	P	No Comments	
46	H.R.Puneeth Kumar	Jr.Eng(EI)	<a href="mailto:puneethmandya@gmail.com">puneethmandya@gmail.com</a>	9743802424	P	p	No Comments	
47	Sadhana V	Asst.System Manager	<a href="mailto:asmit2@ksrtc.org">asmit2@ksrtc.org</a>	7760282624	P	P	No Comments	KSRTC
48	Neetha	Supervisors	<a href="mailto:neethajs.11@gmail.com">neethajs.11@gmail.com</a>	7760557990	P	P	Active	
49	Sadanand	Supervisors	<a href="mailto:sadha.ks56@gmail.com">sadha.ks56@gmail.com</a>	8095268274	P	P	No Comments	
50	Rajeshwari	Supervisors	<a href="mailto:itmusr@ksrtc.org">itmusr@ksrtc.org</a>	8105874597	P	P	Active	
51	Karunakar	Supervisors	<a href="mailto:itrmn@ksrtc.org">itrmn@ksrtc.org</a>	9738385938	P	P	Active	

## Annexure 2: Snapshots







Conceptualised & Designed by

**Public Affairs Foundation**

#15, KIADB Industrial Area,  
Jigani-Bommasandra Link Road,  
Jigani Post, Anekal Taluk,  
Bangalore - 560105, Karnataka, India

**Telefax:** +918027839918/19/20

**Website:** [www.pafglobal.org](http://www.pafglobal.org)