

DATA HACKATHON



DATA HACKATHON

Data. Data Everywhere

Bengaluru



DATA HACKATHON

Data. Data Everywhere

Bengaluru



DATA HACKATHON

Data. Data Everywhere



DATA HACKATHON

Data. Data Everywhere



DATA HACKATHON

Data. Data Everywhere





DATA HACKATHON

- Back Row (standing): 18 individuals, including men and women, wearing t-shirts in red, blue, black, and grey, all featuring the 'DATA' logo and a Nike swoosh.
- Front Row (seated): 8 individuals, including men and women, wearing t-shirts in black, grey, and red, all featuring the 'DATA' logo and a Nike swoosh.

RTC BENGALURU



RTC BENGALURU

DATA HACKATHON







Not everything that can be counted counts, and not everything that can be counted can be counted.

DATA
MARATHON
FOUNTEEN

DATA
MARATHON
FOUNTEEN

We who would
search for pearls
must dive below.



TEAM STRATEGY

- Our task started with assessment of the work in hand.
- It was noticed that we had to start with unzipping 24 RAR files with the total file size of 3.74 GB.
- Since it would involve a lot of time to extract these on individual team members' machines, we decided to use the common folder to extract the RAR files therein, instead of in the individual machines.
- Further, the team was divided into two groups to analyse Service Tax and Central Excise files separately.
- As the files supplied were without header, our team decided to utilise the Column Headers_amended file supplied with the data set. The headers were inserted using 'insert column header' node.

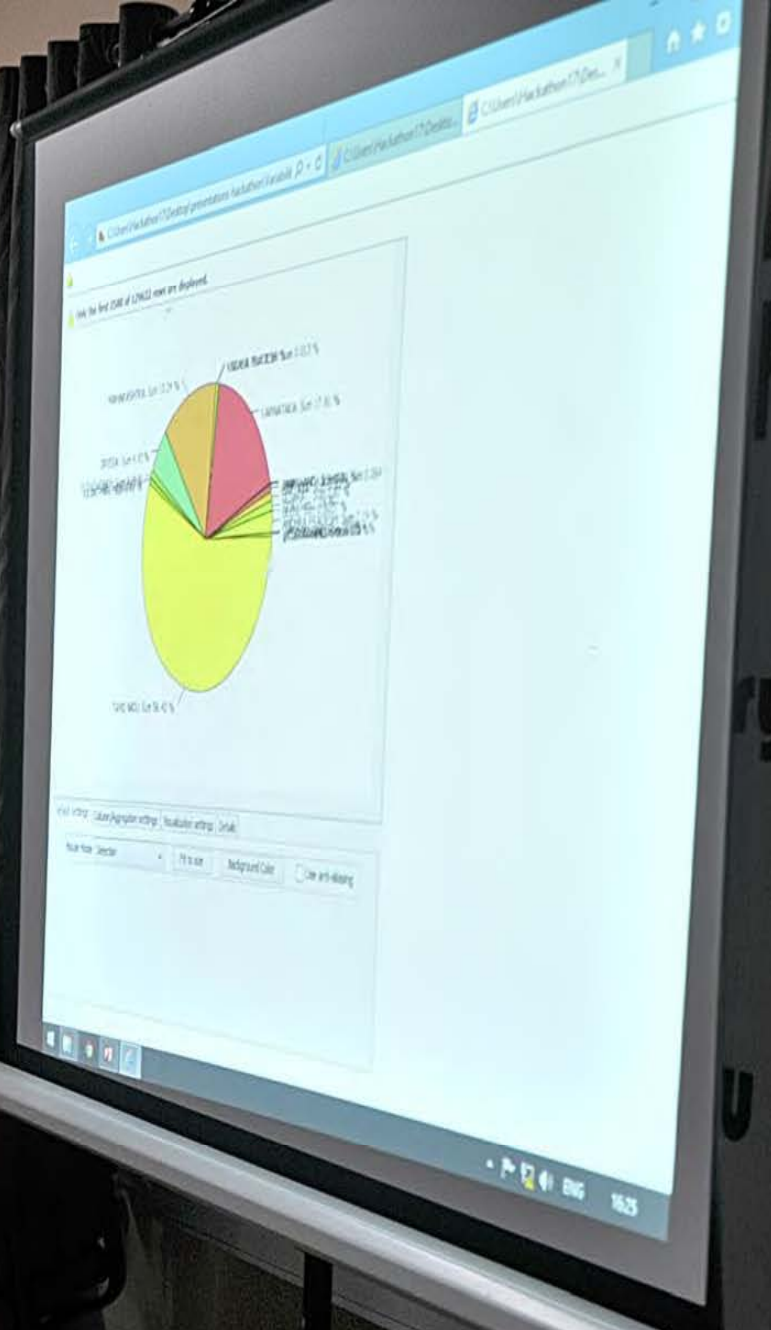


Data Matters



Not everything that can be counted counts, and not everything that counts can be counted.

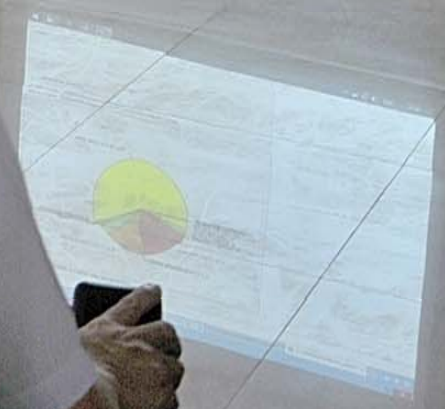
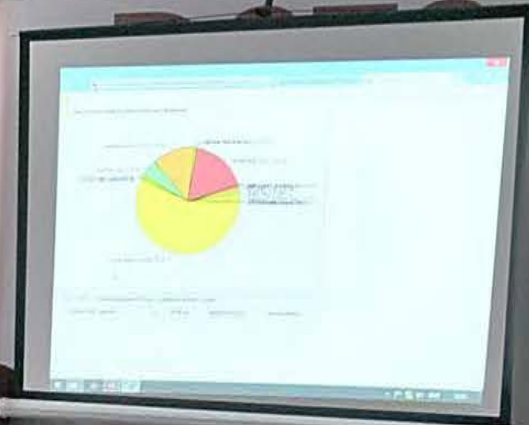
We who search for must dive



LOGIC

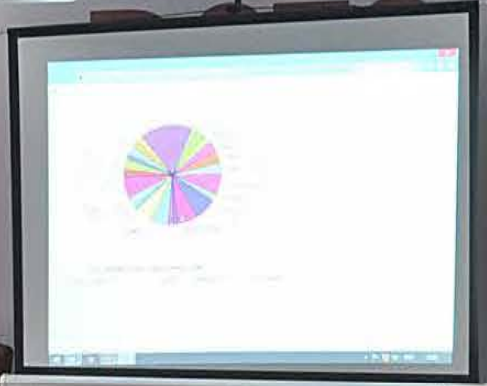


We who would search for pearls, must dive below.





We who would
search for pearls
must dive below.





We who would search for pearls must dive below.



Analysis of Service Tax data

- In the 'REGSTRN_ISSUE_DT' table, we compared the fields 'REGSTRN_APCTN_DT' and 'REGSTRN_ISSUE_DT' to compare delay in issue of registration certificate beyond 15 days. The 'Time to State' Node was used to convert these two fields to date format and the 'Time Difference' Node was used to calculate the delay between the filing and issue of certificate.
- In the reporting that there was no data validation in this field.
- The following issues were noticed:
 - While a threshold of 15 days fetched 1.73 lakh rows out of a total of 29.77 lakh rows in the table, 50% of rows that there was delay of more than 7 days in 25% of records.
 - A comparison of the data also revealed that there were 1,070 records where the delay was exactly 300 days, indicating incorrect data entry/absence of data validation in these columns.
 - Using the 'Extract Time Window' Node, it was noticed that there was no delay, i.e. 0 days delay in 11.23 lakh records, or precisely 20% of the records pertaining to registrations done upto 11-Jan-2012, which appears doubtful.
 - The lower bound and upper bound in the 'REGSTRN_CERTFCT_ISSUE_DT' were '01-Jan-0001' and '14-Apr-2007', respectively. Similarly, the lower bound and upper bound were found to be '01-Jan-0001' and '11-Jul-1951', respectively in the 'REGSTRN_APCTN_DT' field.
 - It was also observed that there was no data validation in these date fields and they were not validated.

Not everything that can be counted counts,
and
not everything that counts
can be counted.



Albert Einstein
Physicist





We who would
search for pearls
must dive below.



Data Hackathon
Bangalore
20-22 Nov 2016

Presentation by Team Variety - Tableau

We who would
search for pearls
must dive below.



Team Variety

- Sasiidhar Karanam, SrAO, Office of DGA (Central), Hyderabad
- D. Vaidyanadham, AAO, Office of DGA (Central), Hyderabad
- N.S. Negi, AO, Office of PDA (Central), Ahmedabad
- A.K.Shukla, AAO, Office of PDA (Central), Ahmedabad

We who would
search for pearls
must dive below.



Spade work

- Downloading trial version of Tableau
- Emails received from Hqs. And PDA (Central), Bengaluru
- Collection of guidelines issued by Hqs on various parameters for Audit Planning
- Analysis of Data pertaining to Duty paid in Hyderabad Zone through Tableau
- Analysis of MCA 21 open source Data through Tableau
- Interaction in Data Analytics Whatsapp sub- group consisting of members of Team and exchange of information, strengths and weaknesses of each individual Team member

We who would
search for pearls
must dive below.



Scabe work

- Downloading a version of Tableau
- Data is loaded from Hqs. And PDA (Central), Bengaluru
- Collection of guidelines issued by Hqs on various parameters for Audit Planning
- Analysis of Data pertaining to Duty call in Hyderabad Zone through Tableau
- Analysis of NCA 21 open source Data through Tableau
- Formation of Data Analysis Whatsapp sub-group consisting of members of Team for exchange of information, strengths and weaknesses of each individual team member.

We who would
search for pearls
must dive below.



We who would
search for pearls
must dive



ACID Test of Tableau:

- ❖ Problem: It took more time to extract joined table.
- ❖ Solution #1: this limitation can be alleviated by appending the files (with similar structure) using knime and the resultant file/table can be converted into the tableau readable format
- ❖ Knime has two special nodes: viz "tableau writer (TDE)" and "send to tableau server" which convert knime output to tableau readable format




Not everything that can be counted counts, and not everything that counts can be counted.



AI & Analytics
All on
Smart Energy data

Not everything that can be counted counts, and not everything that counts can be counted.



Albert Einstein
Physicist





Not everything that can be counted counts, and not everything that counts can be counted.

Data Matters

We who would search for pearls must dive below.



We who would
search for pearls
must dive below.



Introduction

- 1. Introduction
- 2. Data Science & Analytics
- 3. Data Science & Analytics
- 4. Data Science & Analytics



Not everything that can be counted counts, and not everything that counts can be counted.

Data Matter

Data Matter

We who would
search for pearls
must dive

ACID Test of Tableau:

- Problem: It took more time to extract joined table.
- Solution #1: this limitation can be alleviated by appending the files (with similar structure) using knime and the resultant file/table can be converted into the tableau readable format
- Knime has two special nodes, viz "tableau writer (TDE)" and "send to tableau server" which convert knime output to tableau readable format

Not everything that can be counted counts,
and
not everything that counts
can be counted.




Albert Einstein
Physicist



Data M



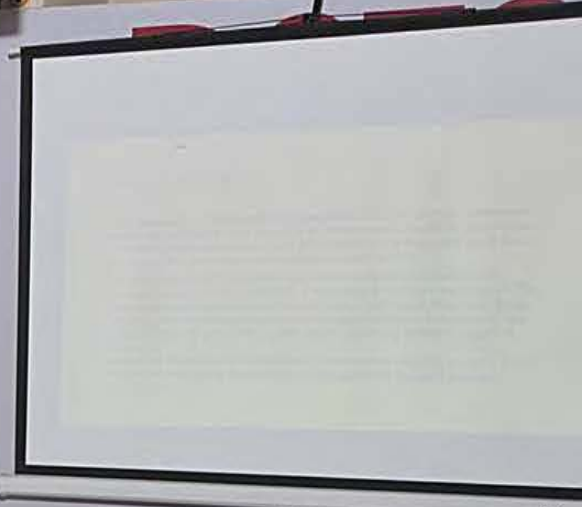
Not everything that can be counted counts,
and
not everything that counts
can be counted.



Albert Einstein
Physicist



We who would search for pearls must dive below.





Not everything that can be counted counts, and not everything that counts can be counted.

Data Matter

We who would search for pearls must dive below.

Matter

We who would search for pearls must dive below

Continued.....

- THE OCCURRING ARE THE BREAK UP FOR 2015-16:
 - 101 Gal - 89 Assessment
 - 101 Gal - 572 Assessment
 - 101 Gal - 444 Assessment
 - 101 Gal - 586 Assessment
 - 101 Gal - 187 Assessment
 - 101 Gal - 743 Assessment
 - 101 Gal - 696 Assessment
 - 101 Gal - 785 Assessment
 - 101 Gal - 170 Assessment
 - 101 Gal - 802 Assessment
 - 101 Gal - 1583 Assessment



We who would
search for pearls
must dive below...

WAT State Master
plastic data of 2014-15
Out of 30786 IT Assessments in
Chemical and Composites
3980 IT Assessments doing plastic
business do not figure in WAT
plastic data of 2014-15.



Not everything that can be counted counts, and not everything that counts can be counted.

Data Matter

We who search must c



Data Matter



We who would search for pearls must dive below.



Not everything that counts can be counted, and not everything that can be counted isn't counted.





Not everything that can be counted counts, and not everything that counts can be counted.

Data Matter

We who would search for pearls must dive below.



Tru

DATA
HACKATHON
VARIETY



DRY-FIT

DATA
HACKATHON
VARIETY



DRY-FIT

DATA
HACKATHON
TOURNAMENT



DRY-FIT

11/18/14

Data Matter



In Data we Trust



Not everything that can be
counted counts,
and
not everything that
counts can be counted



Albert Einstein,
Physicist





Data!
Data!
Data!
I can't make bricks without
Clay!
Sir Arthur Conan Doyle

Sixth
Password

DATA



DATA
HACKATHON
Data. Data Everywhere

DATA
HACKATHON
Data. Data Everywhere

Data!
Data!
Data!
can't make bricks without Clay!

Without big data analytics, companies are blind and deaf, wandering out onto the Web like deer on a freeway.



Geoffrey Moore
Author of Crossing the Chasm & The Tornado



In God We Trust,
All others
must bring data

W.Edv







DATA HACKATHON

Data. Data Everywhere



Ben

DATA HACKATHON

Data. Data everywhere



DATA HACKATHON

Data. Everywhere





In God We Trust,
All others
must bring data

Not everything that can be
counted counts,
and
not everything that counts
can be counted.

Data Matters



data analytics, companies
and deaf, wandering out
like deer on a freeway.

In Data we Trust

DATA
HACK
Data. Do



In Data we Trust

Without I are blind onto th

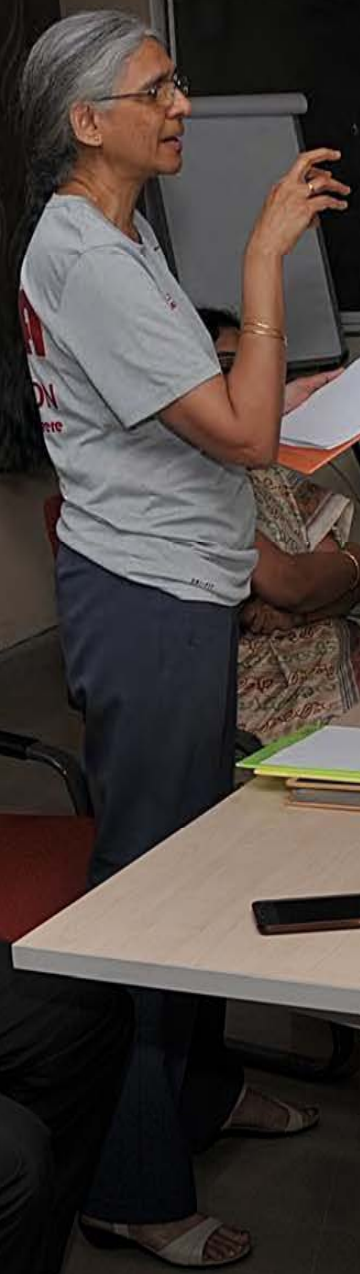
DATA CKATHON

Data Everywhere

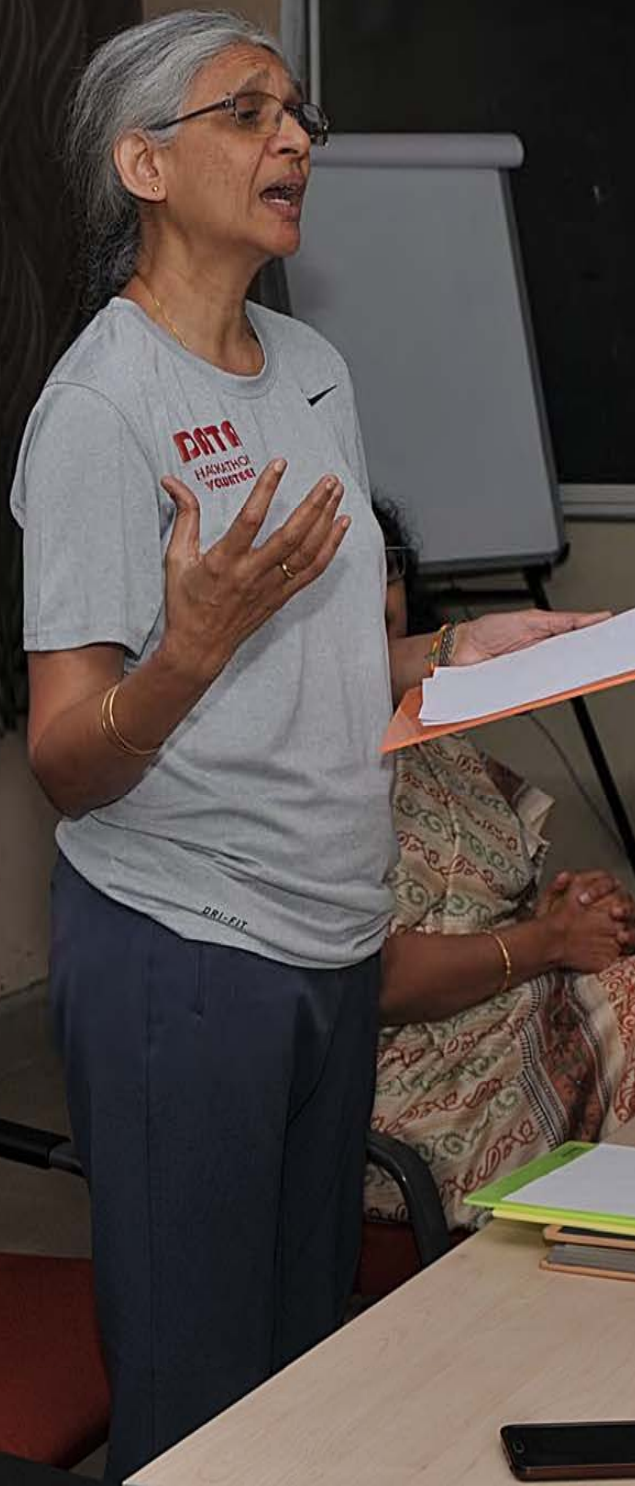
Bengaluru

Data!
Data!
Data!
I can't make bricks without
Clay!

Without big d
are blind and
only the Web

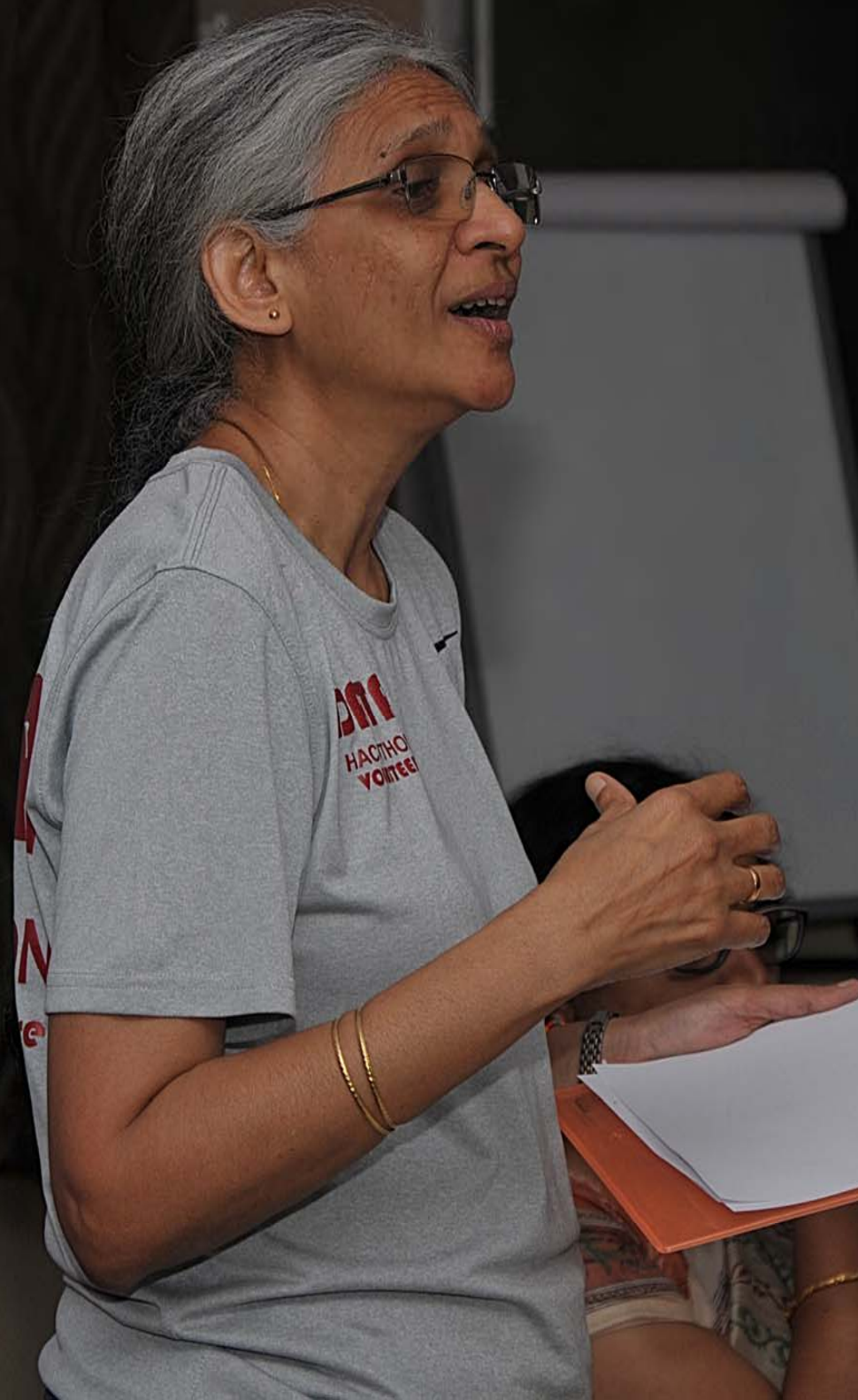


ON
here



Data!
Data!
Data!
I can't make bricks without
Clay!





Data!
Data!
Data!
I can't make

In Data we Trust



In God We Trust,
All others
must bring data

W. Edwards Deming





Clay!

are blind and deaf, wandering on
onto the Web like deer on a freeway

Geoffrey M...
Author of Cr...
the Chasm &
the Tornado

DATA



DATA
NIKE

DATA HACKATHON

Data. Data. Everywhere

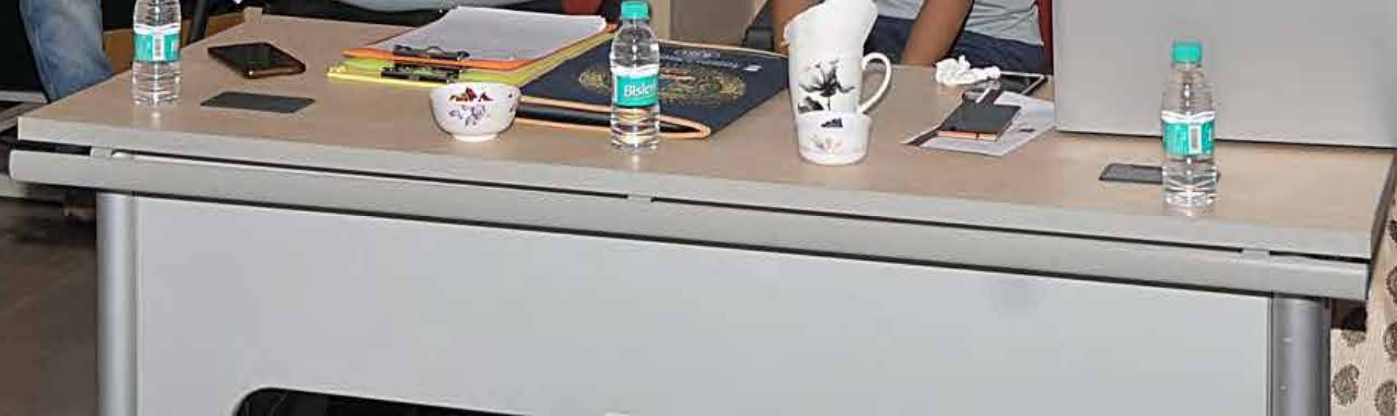




DATA HACKATHON

Data. Data Everywhere

Bengaluru



DATA HACKATHON

Data. Data Everywhere

Bengaluru



DATA HACKATHON

Data. Data Everywhere



Without big data analytics, companies are blind and deaf, wandering out onto the Web like deer on a freeway.




In Data we Trust





Not everything that can be counted counts,
and
not everything that counts
can be counted.



Albert Einstein



Data Matters





Not everything that can be counted counts, and not everything that counts can be counted.

Data Matter

We who would search for pearls must dive below.



Data Matter

We who would search for pearls must dive below.



Data Matter

DATA
INSPIRATION
INNOVATION

DATA
INSPIRATION
INNOVATION

DATA
INSPIRATION
INNOVATION

Bisleri



We who would
search for pearls
must dive below.



DATA HACKATHON

Data. Data Everywhere

Beng





We who would
search for pearls
must dive below.
John Dryden





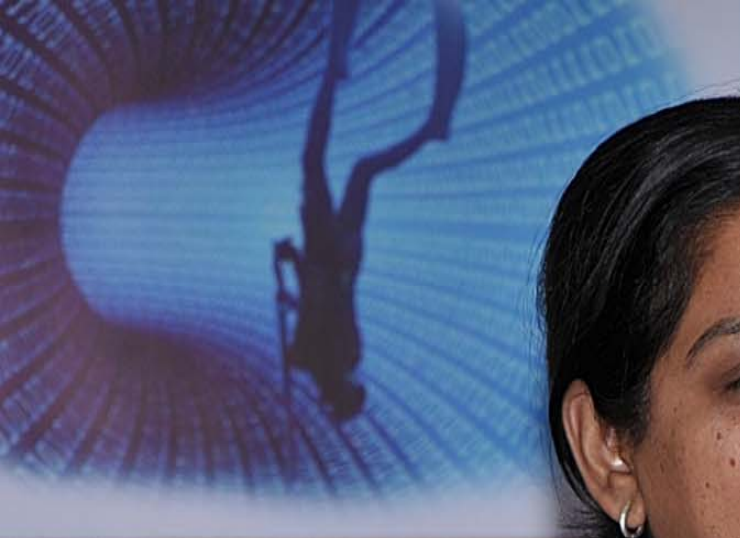
Data.





We who would
search for pearls
must dive below.







We who would
search for pearls
must dive below.



DATA HACKATHON

Data. Data Everywhere

Bengal



DATA HACKATHON

Data. Data Everywhere

Bangalore





everywhere

uru



erywhere

uru



DATA HACKATHON

Data. Data Everywhere

engaluru



DATA HACKATHON

Data. Data Everywhere

Bengaluru



DATA HACKATHON

Data. Data Everywhere

Bengaluru





Not everything that can be counted counts, and not everything that counts can be counted.

Data Matte

We who would search for pearls must dive below.





Not everything that can be counted counts, and not everything that counts can be counted.



Data Matters





Data Matters



Not everything that can be counted counts, and not everything that counts can be counted.



We who would search for pearls must dive below.

ve who would
search for pearls
must dive b





earch for pearls
must dive below.

John Dwyer

DITA
HAKAITHON
WUITEER



DATA HACKATHON

Data. Data Everywhere

Bengaluru



DATA HACKATHON

Data. Data. Data. Where





DATA HACKATHON

DATA
HACKATHON
YOUWITER

DATA
HACKATHON
YOUWITER

DATA
HACKATHON
YOUWITER

DATA
HACKATHON
YOUWITER



DATA HACKATHON

data. Data everywhere





DATA
HACKATHON

DATA
HACKATHON

Data.

DATA
HACKATHON

DRI-FIT

DRI-FIT

Bisleri

DATA HACKATHON







DATA HACKATHON

DATA
HACKATHON
VOLUNTEER

Be

A
N

DATA
HACKATHON
MIT





DATA HACKATHON

Data. Data. Ever.

Bengaluru

DATA HACKATHON

... Data Everywhere

Bengaluru



DATA HACKATHON

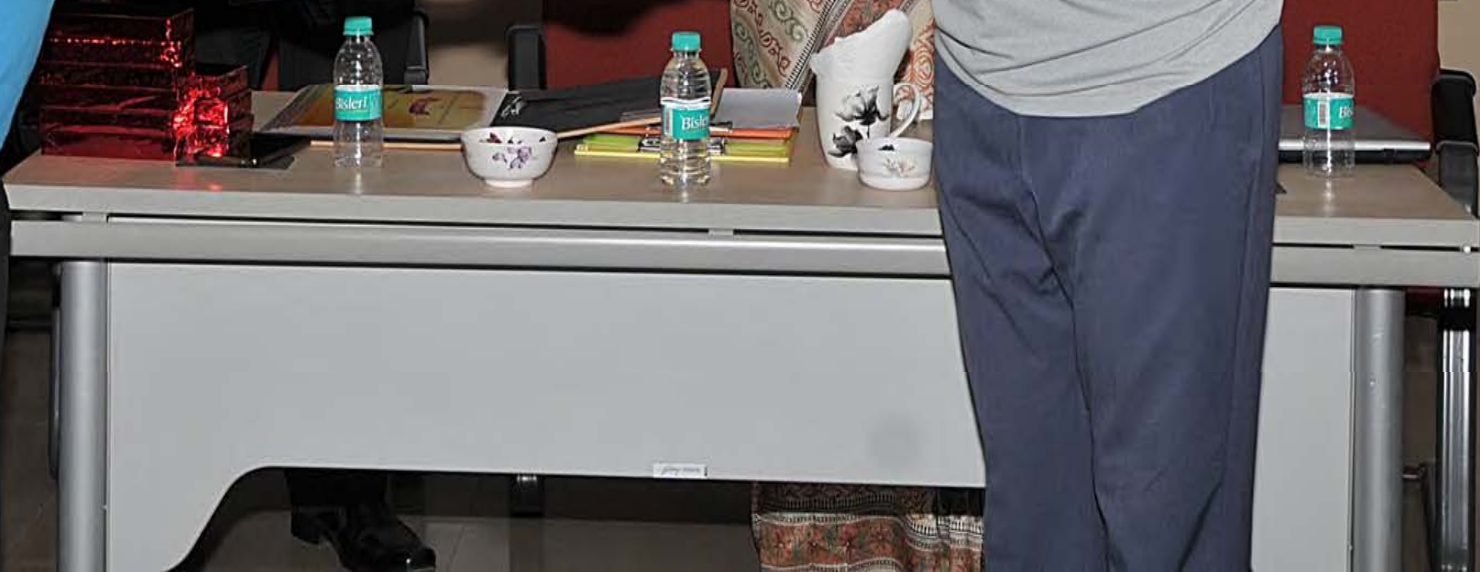
Data. Data. Everywhere.

Bengaluru



DATA HACKATHON

Data. Data.





DATA HACKATHON

Date. Data.

DATA HACKATHON

DATA HACKATHON

Be

DATA HACKATHON

Data. Hack. Here



DATA HACKATHON

ta Everywhere

URU



DATA HACKATHON

Everywhere



DATA HACKATHON

anywhere



DATA HACKATHON

everywhere



DATA MARATHON

DATA
where



DATA HACKATHON



DATA HACKATHON



DATA HACKATHON





DATA HACKATHON

Everywhere

DATA

DATA

DATA



DATA HACKATHON

DATA
HACKATHON
YOUWEEK

DATA
HACKATHON
YOUWEEK

DATA
HACKATHON
YOUWEEK



