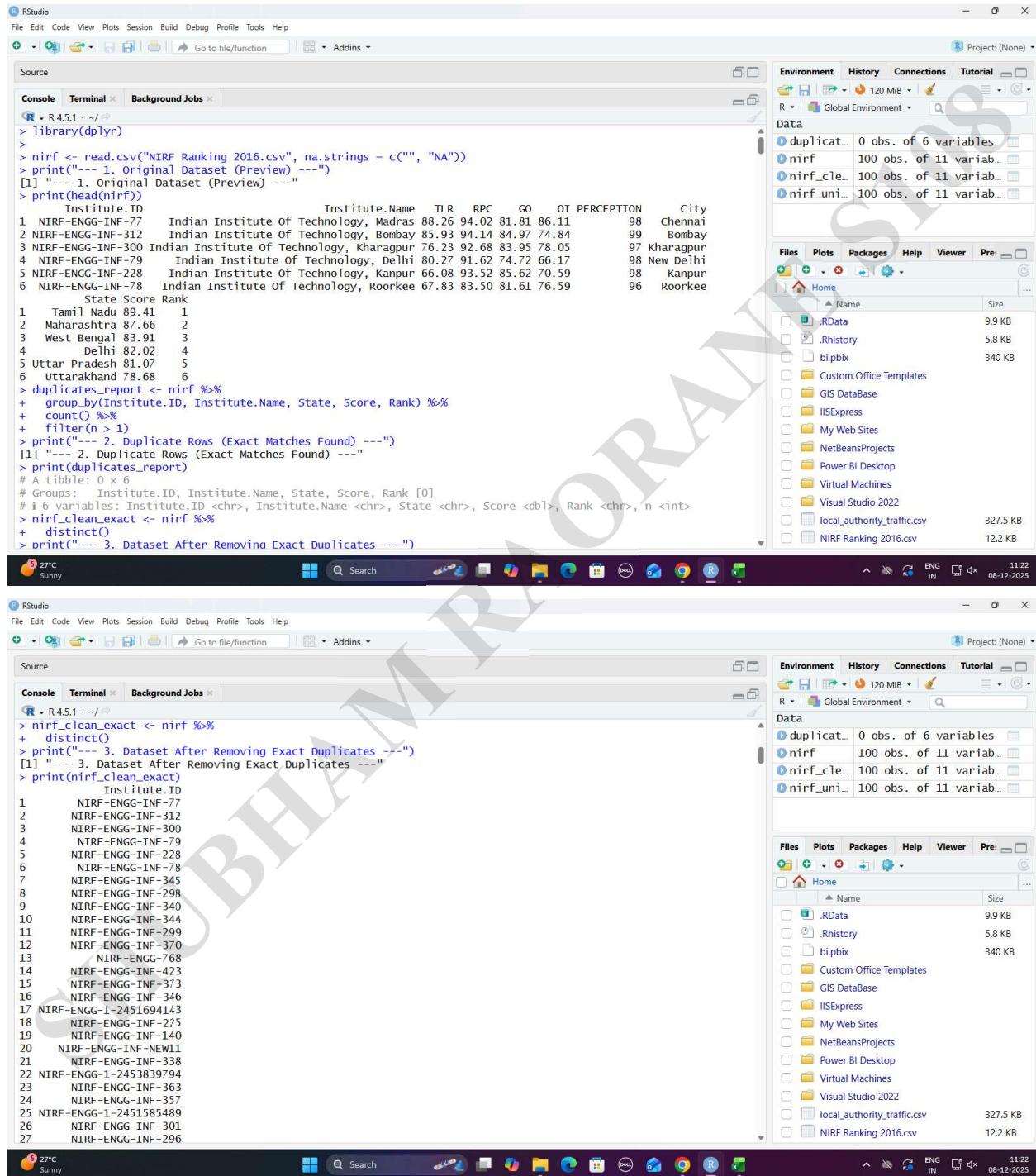


SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS /R

AIM : Identifying and handling duplicates using distinct() (R).

OUTPUT



The screenshot shows two instances of the RStudio interface. Both instances have the following setup:

- Console Tab:** Displays R code and its output. The code reads a CSV file ('NIRF Ranking 2016.csv'), prints the original dataset, prints the head of the dataset, identifies duplicates, and prints the dataset after removing exact duplicates.
- Environment Tab:** Shows the global environment with objects like 'nirf' (100 obs. of 11 variables), 'nirf_clean_exact' (100 obs. of 11 variables), and 'nirf_unique' (100 obs. of 11 variables).
- File Explorer:** Shows a directory structure with files like 'RData', '.Rhistory', 'bipbix', 'Custom Office Templates', 'GIS DataBase', 'IISExpress', 'My Web Sites', 'NetBeansProjects', 'Power BI Desktop', 'Virtual Machines', 'Visual Studio 2022', 'local_authority_traffic.csv', and 'NIRF Ranking 2016.csv'.
- Bottom Status Bar:** Shows system information including temperature (27°C), battery level (120 MiB), language (ENG IN), date (08-12-2025), and time (11:22).

Output from the R Console (Top Window):

```

> library(dplyr)
> nirf <- read.csv("NIRF Ranking 2016.csv", na.strings = c("", "NA"))
> print("--- 1. Original Dataset (Preview) ---")
[1] "--- 1. Original Dataset (Preview) ---"
> print(head(nirf))
#> #> #> #> #> #>
Institute.ID      Institute.Name    TLR   RPC   GO   OI PERCEPTION City
1 NIRF-ENGG-INF-77 Indian Institute of Technology, Madras 88.26 94.02 81.81 86.11 98 Chennai
2 NIRF-ENGG-INF-312 Indian Institute Of Technology, Bombay 85.93 94.14 84.97 74.84 99 Bombay
3 NIRF-ENGG-INF-300 Indian Institute Of Technology, Kharagpur 76.23 92.68 83.95 78.05 97 Kharagpur
4 NIRF-ENGG-INF-79 Indian Institute Of Technology, Delhi 80.27 91.62 74.72 66.17 98 New Delhi
5 NIRF-ENGG-INF-228 Indian Institute Of Technology, Kanpur 66.08 93.52 85.62 70.59 98 Kanpur
6 NIRF-ENGG-INF-78 Indian Institute Of Technology, Roorkee 67.83 83.50 81.61 76.59 96 Roorkee
#> #> #> #> #> #>
State Score Rank
1 Tamil Nadu 89.41 1
2 Maharashtra 87.66 2
3 West Bengal 83.91 3
4 Delhi 82.02 4
5 Uttar Pradesh 81.07 5
6 Uttarakhand 78.68 6
> duplicates_report <- nirf %>%
+ group_by(Institute.ID, Institute.Name, State, Score, Rank) %>%
+ count() %>%
+ filter(n > 1)
> print("--- 2. Duplicate Rows (Exact Matches Found) ---")
[1] "--- 2. Duplicate Rows (Exact Matches Found) ---"
> print(duplicates_report)
#> #> #> #> #> #>
# A tibble: 0 × 6
# Groups: Institute.ID, Institute.Name, State, Score, Rank [0]
#> #> #> #> #> #>
> nirf_clean_exact <- nirf %>%
+ distinct()
> print("--- 3. Dataset After Removing Exact Duplicates ---")

```

Output from the R Console (Bottom Window):

```

> nirf_clean_exact <- nirf %>%
+ distinct()
> print("--- 3. Dataset After Removing Exact Duplicates ---")
[1] "--- 3. Dataset After Removing Exact Duplicates ---"
> print(nirf_clean_exact)
#> #> #> #> #> #>
Institute.ID
1 NIRF-ENGG-INF-77
2 NIRF-ENGG-INF-312
3 NIRF-ENGG-INF-300
4 NIRF-ENGG-INF-79
5 NIRF-ENGG-INF-228
6 NIRF-ENGG-INF-78
7 NIRF-ENGG-INF-345
8 NIRF-ENGG-INF-298
9 NIRF-ENGG-INF-340
10 NIRF-ENGG-INF-344
11 NIRF-ENGG-INF-299
12 NIRF-ENGG-INF-370
13 NIRF-ENGG-INF-768
14 NIRF-ENGG-INF-423
15 NIRF-ENGG-INF-373
16 NIRF-ENGG-INF-346
17 NIRF-ENGG-1-2451694143
18 NIRF-ENGG-INF-225
19 NIRF-ENGG-INF-140
20 NIRF-ENGG-INF-NEW11
21 NIRF-ENGG-INF-338
22 NIRF-ENGG-1-2453839794
23 NIRF-ENGG-INF-363
24 NIRF-ENGG-INF-357
25 NIRF-ENGG-1-2451585489
26 NIRF-ENGG-INF-301
27 NIRF-ENGG-INF-296

```

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS /R

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

Console Terminal Background Jobs

```
R 4.5.1 - ~/ 
80 NIRF-ENGG-1-2451417043
81 NIRF-ENGG-INF-410
82 NIRF-ENGG-1-2452796944
83 NIRF-ENGG-1-2455130573
84 NIRF-ENGG-1-2453523111
85 NIRF-ENGG-1-2450743734
86 NIRF-ENGG-1-2454009090
87 NIRF-ENGG-INF-367
88 NIRF-ENGG-773
89 NIRF-ENGG-1-2453802770
90 NIRF-ENGG-1-2454313873
```

Institut

e.Name	Indian Institute Of Technology, M
1	Indian Institute Of Technology, B
adras	Indian Institute Of Technology, Khar
2	Indian Institute Of Technology, K
ombay	Indian Institute Of Technology, Ro
3	Indian Institute Of Technology, Hyde
apgur	Indian Institute Of Technology, Gandh
4	Indian Institute of Technology, Ropar-Rup
Delhi	Indian Institute Of Technology, Patna
5	
anpur	
6	
orkee	
7	
rabad	
8	
nagar	
9	
nagar	
10	
Patna	

27°C Sunny

Search

Files Plots Packages Help Viewer Preferences

Environment History Connections Tutorial

R Project: (None)

Data

- duplicat... 0 obs. of 6 variables
- nirf 100 obs. of 11 variab...
- nirf_cle... 100 obs. of 11 variab...
- nirf_uni... 100 obs. of 11 variab...

Files Plots Packages Help Viewer Preferences

Home Name Size

- RData 9.9 KB
- .Rhistory 5.8 KB
- bipbx 340 KB
- Custom Office Templates
- GIS DataBase
- IISExpress
- My Web Sites
- NetBeansProjects
- Power BI Desktop
- Virtual Machines
- Visual Studio 2022
- local_authority_traffic.csv 327.5 KB
- NIRF Ranking 2016.csv 12.2 KB

11:22 ENG IN 08-12-2025

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Source

Console Terminal Background Jobs

```
R - R 4.5.1 - ~/ 
arded
90
atore
    TLR RPC GO OI PERCEPTION
1 88.26 94.02 81.81 86.11 98
2 85.93 94.14 84.97 74.84 99
3 76.23 92.68 83.95 78.05 97
4 80.27 91.62 74.72 66.17 98
5 66.08 93.52 85.62 70.59 98
6 67.83 83.50 81.61 66.59 96
7 76.76 76.83 80.89 63.76 95
8 80.36 72.09 64.91 71.94 96
9 89.90 73.56 86.42 60.67 38
10 79.80 68.29 74.13 68.78 93
11 71.92 78.66 71.82 65.62 94
12 56.99 80.37 81.50 79.50 92
13 66.95 74.72 77.38 77.11 91
14 59.08 80.63 96.80 56.73 96
15 59.02 85.24 78.43 68.81 83
16 86.75 75.78 89.85 39.23 46
17 55.38 81.06 82.47 65.50 90
18 59.98 73.64 69.05 82.70 85
19 60.78 75.69 81.49 73.21 74
20 57.76 83.07 76.07 59.50 87
21 85.83 70.35 81.05 47.29 43
22 58.61 66.40 82.92 74.67 95
23 60.26 67.86 77.29 70.22 88
24 54.63 72.07 74.08 75.04 82
25 56.72 64.67 77.30 72.10 94
26 73.02 66.41 60.71 60.11 86
27 67.40 75.73 61.06 51.57 85
28 49.43 78.70 77.69 64.58 83
29 58.20 71.82 75.41 61.20 82
```

Sri Ramakrishna Engineering College-Coimb

City	State	Score	Rank
Chennai	Tamil Nadu	89.41	1
Bombay	Maharashtra	87.66	2
Kharagpur	West Bengal	83.91	3
New Delhi	Delhi	82.02	4
Kanpur	Uttar Pradesh	81.07	5
Roorkee	Uttarakhand	78.68	6
Hyderabad	Telangana	77.22	7
Ahmedabad	Gujarat	75.20	8
Rupnagar	Punjab	74.88	9
Patna	Bihar	74.68	10
Guwahati	Assam	74.62	11
Tiruchirappalli	Tamil Nadu	74.45	12
Vellore	Tamil Nadu	74.40	13
Varanasi	Uttar Pradesh	74.39	14
Surat	Gujarat	73.13	15
Indore	Madhya Pradesh	72.00	16
RANCHI	Jharkhand	71.80	17
Nagpur	Maharashtra	71.29	18
Rourkela	Odisha	70.80	19
Howrah	West Bengal	70.35	20A
Mandi	Himachal Pradesh	70.32	20
PUNE	Maharashtra	69.71	21
Mangalore	Karnataka	68.95	22
Allahabad	Uttar Pradesh	67.94	23
COIMBATORE	Tamil Nadu	67.80	24
Jodhpur	Rajasthan	67.68	25
Bhubaneswar	Odisha	67.58	26
PATIALA	Punjab	67.51	27
Waranao	Telangana	67.12	28

27°C Sunny

Search

Files Plots Packages Help Viewer Preferences

Environment History Connections Tutorial

R Project: (None)

Data

- duplicat... 0 obs. of 6 variables
- nirf 100 obs. of 11 variab...
- nirf_cle... 100 obs. of 11 variab...
- nirf_uni... 100 obs. of 11 variab...

Files Plots Packages Help Viewer Preferences

Home Name Size

- RData 9.9 KB
- .Rhistory 5.8 KB
- bipbx 340 KB
- Custom Office Templates
- GIS DataBase
- IISExpress
- My Web Sites
- NetBeansProjects
- Power BI Desktop
- Virtual Machines
- Visual Studio 2022
- local_authority_traffic.csv 327.5 KB
- NIRF Ranking 2016.csv 12.2 KB

11:23 ENG IN 08-12-2025

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help, and a Project dropdown set to (None). The main window has tabs for Source, Console, Terminal, and Background Jobs. The Source tab displays R code for data manipulation, including reading CSV files, filtering unique institutes, and printing results. The Environment tab shows the Global Environment with objects like 'nirf' (100 obs. of 11 variables) and 'nirf_uniq...' (100 obs. of 11 variables). The Data tab lists data frames: 'duplicat...' (0 obs. of 6 variables), 'nirf' (100 obs. of 11 variables), 'nirf_cle...' (100 obs. of 11 variables), and 'nirf_uni...' (100 obs. of 11 variables). The Files tab shows the file structure: Home, RData, .Rhistory, bipbip, Custom Office Templates, GIS DataBase, IISExpress, My Web Sites, NetBeansProjects, Power BI Desktop, Virtual Machines, Visual Studio 2022, local_authority_traffic.csv (327.5 KB), and NIRF Ranking 2016.csv (12.2 KB). The bottom status bar shows the date (08-12-2025), time (11:23), and system information (ENG IN).

```
R - R 4.5.1 - ~/r/nirf/nirf.R
87 46.73 52.37 66.90 55.69      37          Patna        Bihar 51.08 87
88 49.21 50.65 72.09 67.19      0           Guntur    Andhra Pradesh 50.65 88
89 48.74 47.61 67.42 51.75     41          NANDED   Maharashtra 50.64 89
90 46.62 56.15 67.80 65.62      6       COIMBATORE   Tamil Nadu 50.55 90
[ reached 'max' / getoption("max.print") -- omitted 10 rows ]
> nirf_unique_institutes <- nirf %>
+  distinct(Institute.Name, .keep_all = TRUE)
> print("--- 4. Unique Institutes Only (1 entry per college) ---")
[1] "--- 4. Unique Institutes Only (1 entry per college) ---"
> print(nirf_unique_institutes)
Institute.ID
1      NIRF-ENGG-INF-77
2      NIRF-ENGG-INF-312
3      NIRF-ENGG-INF-300
4      NIRF-ENGG-INF-79
5      NIRF-ENGG-INF-228
6      NIRF-ENGG-INF-78
7      NIRF-ENGG-INF-345
8      NIRF-ENGG-INF-298
9      NIRF-ENGG-INF-340
10     NIRF-ENGG-INF-344
11     NIRF-ENGG-INF-299
12     NIRF-ENGG-INF-370
13      NIRF-ENGG-768
14     NIRF-ENGG-INF-423
15     NIRF-ENGG-INF-373
16     NIRF-ENGG-INF-346
17 NIRF-ENGG-1-2451694143
18     NIRF-ENGG-INF-225
19     NIRF-ENGG-INF-140
20     NIRF-ENGG-INF-NEW11
21     NIRF-ENGG-INF-338
22 NIRF-ENGG-1-2453839794
```

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE
SUBJECT : Data Analysis with SAS / SPSS /R

The screenshot shows the RStudio interface with two data frames displayed:

- NIRF Ranking Data:**

	TLR	RPC	GO	OI	PERCEPTION
1	88.26	94.02	81.81	86.11	98
2	85.93	94.14	84.97	74.84	99
3	76.23	92.68	83.95	78.05	97
4	80.27	91.62	74.72	66.17	98
5	66.08	93.52	85.62	70.59	98
6	67.83	83.50	81.61	76.59	96
7	76.76	76.83	80.89	63.76	95
8	80.36	72.09	64.91	71.94	96
9	89.96	73.56	86.42	60.67	38
10	79.80	68.29	74.13	68.78	93
11	71.92	78.66	71.82	65.62	94
12	56.99	80.37	81.50	79.50	92
13	66.95	74.72	77.38	77.11	91
14	59.08	80.63	96.80	56.73	96
15	59.02	85.24	78.43	68.81	83
16	86.75	75.78	89.85	39.23	46
17	55.39	81.08	82.47	65.50	90
18	59.98	73.64	69.05	82.70	85
19	60.78	75.69	81.49	73.21	74
20	57.76	83.07	76.07	59.50	87
21	85.83	70.38	81.05	47.29	43
22	58.61	66.48	82.92	74.67	95
23	60.26	67.89	77.29	70.22	88
24	54.63	72.07	74.08	75.04	82
25	56.72	64.67	77.30	72.10	94
26	73.02	66.41	60.71	60.11	86
27	67.40	75.73	61.06	51.57	85
28	40.10	79.70	77.77	60.64	59
- Sri Ramakrishna Engineering College- Coimbatore Data:**

	City	State	Score	Rank
1	Chennai	Tamil Nadu	89.41	1
2	Bombay	Maharashtra	87.66	2
3	Kharagpur	West Bengal	83.91	3
4	New Delhi	Delhi	82.02	4
5	Kanpur	Uttar Pradesh	81.07	5
6	Roorkee	Uttarakhand	78.68	6
7	Hyderabad	Telangana	77.22	7
8	Ahmedabad	Gujarat	75.20	8
9	Rupnagar	Punjab	74.88	9
10	Patna	Bihar	74.68	10
11	Guwahati	Assam	74.62	11
12	Tiruchirappalli	Tamil Nadu	74.45	12
13	Vellore	Tamil Nadu	74.40	13
14	Varanasi	Uttar Pradesh	74.39	14
15	Surat	Gujarat	73.13	15
16	Indore	Madhya Pradesh	72.00	16
17	RANCHI	Jharkhand	71.80	17
18	Nagpur	Maharashtra	71.29	18
19	Rourkela	Odisha	70.80	19
20	Howrah	West Bengal	70.35	20A
21	Mandi	Himachal Pradesh	70.32	20
22	PUNE	Maharashtra	69.71	21
23	Mangalore	Karnataka	68.95	22
24	Allahabad	Uttar Pradesh	67.94	23
25	COIMBATOR	Tamil Nadu	67.80	24
26	Jodhpur	Rajasthan	67.68	25
27	Bhubaneswar	Odisha	67.58	26
28	DATTAI	MP	67.51	27

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function | Addins | Source
Console Terminal Background Jobs
R - R 4.5.1 - /- ~
61 45.21 59.86 79.38 69.05 40 Noida Uttar Pradesh 56.79 60
62 55.60 34.04 90.52 75.50 50 PUNE Maharashtra 56.56 61
63 51.61 58.11 77.22 62.78 27 SATHYAMANGALAM Tamil Nadu 56.21 62
64 46.48 62.13 75.15 63.84 31 Raipur Chhattisgarh 55.93 63
65 43.67 64.68 77.11 48.56 48 PUNE Maharashtra 55.70 64
66 55.75 58.12 66.08 71.32 12 Silchar Assam 55.49 65
67 51.73 53.40 50.05 70.40 59 Kanyakumari Tamil Nadu 55.45 66
68 50.37 59.18 72.01 60.42 33 Kashm... Jammu and Kashmir 55.30 67
69 37.05 64.47 67.19 46.69 77 BERHAMPUR Odisha 54.81 69
70 41.66 50.29 74.67 65.68 63 TUMKUR Karnataka 54.67 70
71 52.98 51.91 72.17 66.37 30 HYDERABAD Telangana 54.66 71
72 44.10 47.74 69.08 64.91 73 Chennai Tamil Nadu 54.32 72
73 52.71 41.67 69.01 79.65 42 BHIMAVARAM Andhra Pradesh 54.02 73
74 50.02 48.05 73.75 73.45 23 Bhubaneswar Odisha 53.91 74
75 46.77 56.81 81.69 55.36 23 ISLAMPUR Maharashtra 53.69 75
76 43.98 41.32 73.78 66.09 75 Ponda Goa 53.65 76
77 57.70 38.61 66.13 46.27 85 Jharkhand Madhya Pradesh 53.37 77
78 47.15 62.91 79.00 54.79 12 Jhansi Bihar 53.26 78
79 36.73 67.70 79.14 48.80 31 Jamshedpur Jharkhand 53.35 78
80 43.45 56.73 82.68 58.11 21 KOLKATA West Bengal 53.26 79
81 48.23 41.22 86.82 61.99 38 Chennai Tamil Nadu 52.64 80
82 50.24 53.01 67.05 75.26 12 JATNI Odisha 52.58 81
83 53.51 61.80 50.46 49.20 34 UDAIPUR Rajasthan 52.39 82
84 32.29 57.85 79.62 59.06 47 MUMBAI Maharashtra 52.17 84
85 37.52 45.97 70.43 60.73 80 NASHIK Maharashtra 51.92 85
86 46.81 47.19 72.78 61.03 34 NAGPUR Maharashtra 51.50 86
87 46.73 52.37 66.90 55.69 37 Patna Bihar 51.08 87
88 49.21 50.65 72.09 67.19 0 Guntur Andhra Pradesh 50.65 88
89 48.74 47.61 67.42 51.75 41 Nanded Maharashtra 50.64 89
90 46.62 56.15 67.80 65.62 6 COIMBATORE Tamil Nadu 50.55 90
[reached 'max' / getOption("max.print")] -- omitted 10 rows]
|
Environment History Connections Tutorial Project: (None)
Source
Console Terminal Background Jobs
R - R 4.5.1 - /- ~
114 MB 114 MB
Global Environment
Data
duplicat... 0 obs. of 6 variables
nirf_cle... 100 obs. of 11 variab...
nirf_uni... 100 obs. of 11 variab...
Files Plots Packages Help Viewer Pre
Home Name Size
RData 9.9 kB
.History 5.8 kB
bi.pbix 340 kB
Custom Office Templates
GIS DataBase
IISExpress
My Web Sites
NetBeansProjects
Power BI Desktop
Virtual Machines
Visual Studio 2022
local_authority_traffic.csv 327.5 kB
NIRF Ranking 2016.csv 12.2 kB
27°C ENG 08-12-2024 11:24

SHETH L.U.J & SIR M.V COLLEGE OF SCIENCE

SUBJECT : Data Analysis with SAS / SPSS /R

Data Frame: nirf_unique_institutes

Institute.ID	Institute.Name	TLR	RPC	GO	OI	PERCEPTION	City
1	Indian Institute Of Technology, Madras	88.26	94.02	81.81	86.11	98	Chennai
2	Indian Institute Of Technology, Bombay	85.93	94.14	84.97	74.84	99	Bombay
3	Indian Institute Of Technology, Kharagpur	76.23	92.68	83.95	78.05	97	Kharagpur
4	Indian Institute Of Technology, Delhi	80.27	91.62	74.72	66.17	98	New Delhi
5	Indian Institute Of Technology, Kanpur	66.08	93.52	85.62	70.59	98	Kanpur
6	Indian Institute Of Technology, Roorkee	67.83	83.50	81.61	76.59	96	Roorkee
7	Indian Institute Of Technology, Hyderabad	76.76	76.83	80.89	63.76	95	Hyderabad
8	Indian Institute Of Technology, Gandhinagar	80.36	72.09	64.91	71.94	96	Ahmedabad
9	Indian Institute Of Technology, Ropar-Rupnagar	89.96	73.56	86.42	60.67	38	Rupnagar
10	Indian Institute Of Technology, Patna	79.80	68.29	74.13	68.78	93	Patna
11	Indian Institute Of Technology, North Guwahati	71.92	78.66	71.82	65.62	94	Guwahati
12	National Institute Of Technology, Tiruchirappalli	56.99	80.37	81.50	79.50	92	Tiruchirappalli
13	Vellore Institute Of Technology	66.95	74.72	77.38	77.11	91	Vellore
14	Indian Institute Of Technology (Banaras Hindu University), Varanasi	59.08	80.63	96.80	56.73	96	Varanasi
15	Sardar Vallabhbhai National Institute Of Technology	59.02	85.24	78.43	68.81	83	Surat
16	Indian Institute Of Technology, Indore	86.75	75.78	89.85	39.23	46	Indore
17	Birla Institute Of Technology	55.39	81.06	82.47	65.50	90	RANCHI

Showing 1 to 17 of 100 entries, 11 total columns

Console Output:

```
R - R 4.5.1 - ~/r
90 46.62 56.15 67.80 65.62      6 COIMBATORE      Tamil Nadu 50.55  90
[ reached 'max' / getOption("max.print") -- omitted 10 rows ]
```