

ITDS 7th Practical

The first screenshot shows the initial R script and environment. The script creates a data frame with columns Name, Maths, and Science. The environment pane shows the data frame and other objects.

```
1 # Transforming Data in R using tidyverse
2
3 library(tidyverse)
4 library(dplyr)
5
6
7 # Example Dataset
8 data <- data.frame(
9   Name = c("Alice", "Bob"),
10  Maths = c(90, 80),
11  Science = c(85, 88)
12 )
13
14
15 print("Original Data:")
16 print(data)
17
```

The second screenshot shows the data being gathered into a long format. The script uses `gather()` to combine Maths and Science into a single column. The environment pane shows the new `long_data` object.

```
19 # a. GATHER() - convert wide to long
20 long_data <- data %>%
21   gather(key = "Subject", value = "Marks", Maths, Science)
22
23
24 print("Data after gather():")
25 print(long_data)
26
27
28
29 # b. SPREAD() - convert long to wide
30 wide_data <- long_data %>%
31   spread(key = "Subject", value = "Marks")
32
33
34 print("Data after spread():")
35 print(wide_data)
36
```

The third screenshot shows the data being spread back into a wide format. The script uses `spread()` to separate Maths and Science into individual columns. The environment pane shows the new `wide_data` object.

```
19 # a. GATHER() - Convert wide to LONG
20 long_data <- data %>%
21   gather(key = "Subject", value = "Marks", Maths, Science)
22
23
24 print("Data after gather():")
25 print(long_data)
26
27
28 # b. SPREAD() - Convert LONG to wide
29 wide_data <- long_data %>%
30   spread(key = "Subject", value = "Marks")
31
32
33 print("Data after spread():")
34 print(wide_data)
35
```

ITDS 7th Practical

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Untested
# b. SPREAD() - Convert long to wide
31 wide_data <- long_data %>%
32   spread(key = "Subject", value = "Marks")
33
34
35 print("Data after spread():")
36 print(wide_data)
37
38
39
40 # c. UNITE() - Combine columns
41 united_data <- data %>%
42   unite(col = "name_maths", name, Maths, sep = "_")
43
44 print("Data after unite():")
45 print(united_data)
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

Environment History Connections Tutorial

- Global Environment
- my_list List of 3
- my_matrix num [1:2, 1:3] 1 2 3 4 5 6
- students 4 obs. of 5 variables
- students_list List of 3
- students_tab 5 obs. of 3 variables
- united_data 2 obs. of 2 variables
- wide_data 2 obs. of 3 variables

Files Plots Packages Help Viewer Presentation

Console Terminal Background Jobs

```

R 4.5.1 - /
> # c. UNITE() - Combine columns
> united_data <- data %>%
+   unite(col = "name_maths", name, Maths, sep = "_")
>
> print("Data after unite():")
[1] "Data after unite():"
> print(united_data)
  name_maths Science
1 Alice_90    85
2 Bob_80      88

```

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Untested
# b. SPREAD() - Convert long to wide
31 wide_data <- long_data %>%
32   spread(key = "Subject", value = "Marks")
33
34
35 print("Data after spread():")
36 print(wide_data)
37
38
39
40 # c. UNITE() - Combine columns
41 united_data <- data %>%
42   unite(col = "name_maths", name, Maths, sep = "_")
43
44 print("Data after unite():")
45 print(united_data)
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

Environment History Connections Tutorial

- Global Environment
- my_matrix num [1:2, 1:3] 1 2 3 4 5 6
- separated_da 2 obs. of 3 variables
- students 4 obs. of 5 variables
- students_list List of 3
- students_tab 5 obs. of 3 variables
- united_data 2 obs. of 2 variables
- wide_data 2 obs. of 3 variables

Files Plots Packages Help Viewer Presentation

Console Terminal Background Jobs

```

R 4.5.1 - /
> # d. SEPARATE() - Split column
> separated_data <- united_data %>%
+   separate(col = "name_maths", into = c("name", "maths"), sep = "_")
> print("Data after separate():")
[1] "Data after separate():"
> print(separated_data)
  name maths Science
1 Alice  90    85
2 Bob    80    88

```

```

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Source Untested
# Includes tidyverse and dplyr
65 library(tidyverse)
66
67
68
69 # Step 1: create employee data set (wide format)
70 employees <- data.frame(
71   empid = 1:10,
72   name = c("Alice", "Bob", "Charlie", "David", "Eva", "Frank", "Grace", "Hank", "Ivy", "Jack"),
73   Department = c("HR", "Finance", "IT", "HR", "Finance", "IT", "HR", "Finance", "IT", "HR"),
74   salary_2023 = c(50000, 55000, 60000, 62000, 58000, 65000, 68000, 70000, 72000, 75000),
75   salary_2024 = c(52000, 57000, 62000, 64000, 60000, 67000, 70000, 73000, 76000, 79000)
76 )
77
78 print("Original Data (Wide Format):")
79 print(employees)
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

Environment History Connections Tutorial

- Global Environment
- employees 10 obs. of 5 variables
- long_data 4 obs. of 3 variables
- my_list List of 3
- my_matrix num [1:2, 1:3] 1 2 3 4 5 6
- separated_da 2 obs. of 3 variables
- students 4 obs. of 5 variables
- students_list List of 3

Files Plots Packages Help Viewer Presentation

Console Terminal Background Jobs

```

R 4.5.1 - /
> # Step 1: create employee data set (wide format)
> employees <- data.frame(
+   empid = 1:10,
+   name = c("Alice", "Bob", "Charlie", "David", "Eva", "Frank", "Grace", "Hank", "Ivy", "Jack"),
+   Department = c("HR", "Finance", "IT", "HR", "Finance", "IT", "HR", "Finance", "IT", "HR"),
+   salary_2023 = c(50000, 55000, 60000, 62000, 58000, 65000, 68000, 70000, 72000, 75000),
+   salary_2024 = c(52000, 57000, 62000, 64000, 60000, 67000, 70000, 73000, 76000, 79000)
+ )
> print("Original Data (Wide Format):")
> print(employees)
  empid name Department salary_2023 salary_2024
1     1 Alice    HR      50000      52000
2     2 Bob     Finance    55000      57000
3     3 Charlie IT      60000      62000
4     4 David   HR      58000      60000
5     5 Eva    Finance    65000      67000
6     6 Frank  IT      68000      70000
7     7 Grace  HR      70000      72000
8     8 Hank   Finance    72000      75000
9     9 Ivy    IT      75000      79000
10    10 Jack   HR      78000      82000

```