

Business Analytics for Strategic Decision Making

Capstone

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 8:42 PM

lms.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio This Lab will get reset on 15th August 2024, 9:30 AM

Learning Track Notes Help

rstudio

R File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R marketing_campaign df df

Source on Save Run Source

1
2 # data science process.
3
4 # set the objective - segment the customer , using clustering
5
6 # 2.import the data set
7
8 # 3.data cleaning missing value treatment
9 #outlier treatment
10
11 # 4.data manipulation - encoding categorical vars
12 # create new variables sales_jan,feb,mar = sales_quarter
13 (Top Level) :

R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
1212\t1973\tGraduation\tMarried\t52845\t1\t0\t13-08-2013\t7\t384\t25\t292\t130\t41\t64\t3\t8\t8\n\t6\t6\t1\t0\t0\t0\t0\t0\t3\t11\t0\n1000\n9097\t1956\tGraduation\tDivorced\t46086\t0\t1\t03-11-2013\t34\t244\t8\t32\t7\t2\t50\t4\t3\t1\t8\n\t4\t0\t0\t0\t0\t0\t0\t0\t3\t11\t0\n[ reached 'max' / getOption("max.print") -- omitted 1240 rows ]\n> View(df)\n> df <- read.csv('/home/labsuser/marketing_campaign.csv',sep = '\t')\n> View(df)\n> 
```

Environment History Connections Tutorial

Import Dataset 173 MiB

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home Name Size Modified

- CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM
- marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM
- R 1.4 KB Feb 29, 2024, 6:39 AM
- R_1.R 10.4 KB Mar 8, 2024, 8:00 AM
- R_2 (1).R 10.4 KB Mar 8, 2024, 9:54 AM
- R_2.R 4 KB Mar 8, 2024, 9:53 AM
- R_3.R 4 KB Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 8:42 PM

lms.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio This Lab will get reset on 15th August 2024, 9:30 AM

Learning Track Notes Help

rstudio

R File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R marketing_campaign df df

Source on Save Run Source

13
14 # 5.Exploratory data variables - finding out the relationship between the variables
15
16 # plots
17
18 # building the model
19
20 # accuracy of the model
21
22 # telecom customer churn problem
23
24 # import the dataset
24.19 (Top Level) : Spellcheck

R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
1212\t1973\tGraduation\tMarried\t52845\t1\t0\t13-08-2013\t7\t384\t25\t292\t130\t41\t64\t3\t8\t8\n\t6\t6\t1\t0\t0\t0\t0\t0\t3\t11\t0\n1000\n9097\t1956\tGraduation\tDivorced\t46086\t0\t1\t03-11-2013\t34\t244\t8\t32\t7\t2\t50\t4\t3\t1\t8\n\t4\t0\t0\t0\t0\t0\t0\t0\t3\t11\t0\n[ reached 'max' / getOption("max.print") -- omitted 1240 rows ]\n> View(df)\n> df <- read.csv('/home/labsuser/marketing_campaign.csv',sep = '\t')\n> View(df)\n> 
```

Environment History Connections Tutorial

Import Dataset 173 MiB

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home Name Size Modified

- CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM
- marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM
- R 1.4 KB Feb 29, 2024, 6:39 AM
- R_1.R 10.4 KB Mar 8, 2024, 8:00 AM
- R_2 (1).R 10.4 KB Mar 8, 2024, 9:54 AM
- R_2.R 4 KB Mar 8, 2024, 9:53 AM
- R_3.R 4 KB Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 8:37 PM

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 ► Start Lab ■ End Lab X

Learning Track Certificate

R studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R marketing_campaign df df

```
22 # telecom customer churn problem
23
24 # import the dataset
25
26 df <- read.csv('/home/labsuser/marketing_campaign.csv')
27
28 df
29
30 View(df)
31
32 # pass the delimiter
33
34 (Top Level) > R Script S
```

Console Terminal Jobs

R 4.0.2 - ~

```
11188:t1957:tGraduation:tTogether:t26091:t1:t1:t25-02-2014:t84:t15:t10:t19:t8:t17:t20:t3:t2:t1:t3:t5:t0:t0:t0:t0:t0:t3:t11:t0
999
1212:t1973:tGraduation:tMarried:t52845:t1:t0:t13-08-2013:t7:t384:t25:t292:t130:t41:t64:t3:t8:t8:t6:t6:t1:t0:t0:t0:t0:t3:t11:t0
1000
9097:t1956:tGraduation:tDivorced:t46086:t0:t1:t03-11-2013:t34:t244:t8:t32:t7:t2:t50:t4:t3:t1:t8
:t4:t0:t0:t0:t0:t0:t0:t3:t11:t0
[ reached 'max' /getOption("max.print") -- omitted 1240 rows ]
```

Environment History Connections Tutorial

Import Dataset - 158 MB R Global Environment

Data df 2240 obs. of 1 variable

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home Name Size Modified

- CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM
- marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM
- R R_1.R 1.4 KB Feb 29, 2024, 6:39 AM
- R_2 (1).R 10.4 KB Mar 8, 2024, 8:00 AM
- R_2.R 10.4 KB Mar 8, 2024, 9:54 AM
- R_3.R 4 KB Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 8:38 PM

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 ► Start Lab ■ End Lab X

Learning Track Certificate

R studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R marketing_campaign df df

```
1 5524 1957 Graduation Single 58138 0 0 04-09-2012...
2 2174 1954 Graduation Single 46344 1 1 08-03-2014...
3 4141 1965 Graduation Together 71613 0 21-08-2...
4 6182 1984 Graduation Together 26646 1 0 10-02-2...
5 6173 1981 RHD Married 58293 1 0 19-01-2014 94 ...
6 7446 1967 Master Together 62513 0 1 09-09-2013 ...
7 965 1971 Graduation Divorced 55635 0 1 13-11-20...
8 6173 1981 RHD Married 33454 1 1 08-05-2013 33 ...
Showing 1 to 8 of 2,240 entries. 1 total columns
```

Console Terminal Jobs

R 4.0.2 - ~

```
3:t5:t0:t0:t0:t0:t0:t3:t11:t0
999
1212:t1973:tGraduation:tMarried:t52845:t1:t0:t13-08-2013:t7:t384:t25:t292:t130:t41:t64:t3:t8:t8:t6:t6:t1:t0:t0:t0:t0:t3:t11:t0
1000
9097:t1956:tGraduation:tDivorced:t46086:t0:t1:t03-11-2013:t34:t244:t8:t32:t7:t2:t50:t4:t3:t1:t8
:t4:t0:t0:t0:t0:t0:t3:t11:t0
[ reached 'max' /getOption("max.print") -- omitted 1240 rows ]
> View(df)
> |
```

Environment History Connections Tutorial

Import Dataset - 170 MB R Global Environment

Data df 2240 obs. of 1 variable

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home Name Size Modified

- CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM
- marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM
- R R_1.R 1.4 KB Feb 29, 2024, 6:39 AM
- R_2 (1).R 10.4 KB Mar 8, 2024, 8:00 AM
- R_2.R 10.4 KB Mar 8, 2024, 9:54 AM
- R_3.R 4 KB Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

Data View:

ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer
1	1957	Graduation	Single	58138	0	0	04-09-2012
2	1954	Graduation	Single	46344	1	1	08-03-2014
3	1965	Graduation	Together	71613	0	0	21-08-2013
4	1984	Graduation	Together	26646	1	0	10-02-2014
5	1981	PhD	Married	58293	1	0	19-01-2014
6	1967	Master	Together	62513	0	1	09-09-2013
7	1971	Graduation	Divorced	55635	0	1	13-11-2012
8	1984	PhD	Married	33454	1	0	08-06-2013

Showing 1 to 8 of 2,240 entries, 29 total columns

Console:

```
R 4.0.2 - ~/  
1212\1973\Graduation\Married\52845\t1\t0\t13-08-2013\t7\t384\t25\t292\t130\t41\t64\t3\t8\t8\t6\t6\t1\t0\t0\t0\t0\t0\t0\t3\t11\t0  
1000  
9097\1956\Graduation\Divorced\46086\t0\t1\t03-11-2013\t34\t244\t8\t32\t7\t2\t50\t4\t3\t1\t8\t4\t8\t0\t0\t0\t0\t0\t3\t11\t0  
[ reached 'max' / getOption("max.print") -- omitted 1240 rows ]  
> View(df)  
> df <- read.csv('/home/labsuser/marketing_campaign.csv',sep = '\t')  
> View(df)  
>
```

Environment:

Project: (None)

Global Environment - Data df 2240 obs. of 29 variables

Files:

- CAPSTONE PROJECT.R (2.5 KB, Aug 4, 2024, 10:04 AM)
- marketing_campaign.csv (215 KB, Aug 4, 2024, 5:55 AM)
- R (1.4 KB, Feb 29, 2024, 6:39 AM)
- R_1.R (10.4 KB, Mar 8, 2024, 8:00 AM)
- R_2.R (10.4 KB, Mar 8, 2024, 9:54 AM)
- R_3.R (4 KB, Mar 8, 2024, 9:53 AM)

Chrome File Edit View History Bookmarks Profiles Tab Window Help

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

R Script:

```
30 View(df)  
31  
32 # pass the delimiter  
33  
34 df <- read.csv('/home/labsuser/marketing_campaign.csv',sep = '\t')  
35  
36 View(df)  
37  
38 # initial analysis  
39  
40 # no of rows cols  
41  
38:1 (Top Level) : R Script :  
Console Terminal Jobs
```

Console:

```
R 4.0.2 - ~/  
1212\1973\Graduation\Married\52845\t1\t0\t13-08-2013\t7\t384\t25\t292\t130\t41\t64\t3\t8\t8\t6\t6\t1\t0\t0\t0\t0\t0\t0\t3\t11\t0  
1000  
9097\1956\Graduation\Divorced\46086\t0\t1\t03-11-2013\t34\t244\t8\t32\t7\t2\t50\t4\t3\t1\t8\t4\t8\t0\t0\t0\t0\t0\t3\t11\t0  
[ reached 'max' / getOption("max.print") -- omitted 1240 rows ]  
> View(df)  
> df <- read.csv('/home/labsuser/marketing_campaign.csv',sep = '\t')  
> View(df)  
>
```

Environment:

Project: (None)

Global Environment - Data df 2240 obs. of 29 variables

Files:

- CAPSTONE PROJECT.R (2.5 KB, Aug 4, 2024, 10:04 AM)
- marketing_campaign.csv (215 KB, Aug 4, 2024, 5:55 AM)
- R (1.4 KB, Feb 29, 2024, 6:39 AM)
- R_1.R (10.4 KB, Mar 8, 2024, 8:00 AM)
- R_2.R (10.4 KB, Mar 8, 2024, 9:54 AM)
- R_3.R (4 KB, Mar 8, 2024, 9:53 AM)

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

R Script

```
40 # no of rows cols
41
42 nrow(df)
43
44 ncol(df)
45
46 dim(df)
47
48 # first rows of the data
49
50 head(df,5)
51
48:1 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/>
> df <- read.csv('/home/labsuser/marketing_campaign.csv',sep = '\t')
> View(df)
> View(df)
> nrow(df)
[1] 2240
> ncol(df)
[1] 29
> dim(df)
[1] 2240 29
>
```

Environment History Connections Tutorial

Import Dataset - 174 MiB

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Name Size Modified

- CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM
- marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM
- R
- R_1.R 1.4 KB Feb 29, 2024, 6:39 AM
- R_2_(1).R 10.4 KB Mar 8, 2024, 8:00 AM
- R_2.R 10.4 KB Mar 8, 2024, 9:54 AM
- R_3.R 4 KB Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

R Script

```
44 ncol(df)
45
46 dim(df)
47
48 # first rows of the data
49
50 head(df,5)
51
52 # last 5 rows
53
54 tail(df,5)
55
48:25 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/>
> nrow(df)
[1] 2240
> ncol(df)
[1] 29
> dim(df)
[1] 2240 29
> head(df,5)
  ID Year_Birth Education Marital_Status Income Kidhome Teenhome Dt_Customer Recency
1 5524 1957 Graduation Single 58138 0 0 04-09-2012 58
2 2174 1954 Graduation Single 46344 1 1 08-03-2014 38
3 4141 1965 Graduation Together 71613 0 0 21-08-2013 26
```

Environment History Connections Tutorial

Import Dataset - 174 MiB

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Name Size Modified

- CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM
- marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM
- R
- R_1.R 1.4 KB Feb 29, 2024, 6:39 AM
- R_2_(1).R 10.4 KB Mar 8, 2024, 8:00 AM
- R_2.R 10.4 KB Mar 8, 2024, 9:54 AM
- R_3.R 4 KB Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

R Script

```
48 # first 5 rows of the data
49
50 head(df,5)
51
52 # last 5 rows
53
54 tail(df,5)
55
56 # names of the columns
57
58 colnames(df)
59
60 # row names
61
61.1 (Top Level) ↓
```

Console Terminal Jobs

```
R 4.0.2 - /-
> tail(df,5)
   ID Year_Birth Education Marital_Status Income Kidhome Teenhome Dt_Customer
2236 10870    1967 Graduation Married 61223     0      1 13-06-2013
2237 4001     1946 PhD Together 64014     2      1 10-06-2014
2238 7270     1981 Graduation Divorced 56981     0      0 25-01-2014
2239 8235     1956 Master Together 69245     0      1 24-01-2014
2240 9405     1954 PhD Married 52869     1      1 15-10-2012
  Recency MntWines MntFruits MntMeatProducts MntFishProducts MntSweetProducts
```

Environment History Connections Tutorial

R - Global Environment

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

R Script

```
50 head(df,5)
51
52 # last 5 rows
53
54 tail(df,5)
55
56 # names of the columns
57
58 colnames(df)
59
60 # row names
61
62 rownames(df)
63
63.1 (Top Level) ↓
```

Console Terminal Jobs

```
R 4.0.2 - /-
> tail(df,5)
   ID Year_Birth Education Marital_Status Income Kidhome Teenhome Dt_Customer
2238 0     3     11     0
2239 0     3     11     0
2240 0     3     11     1
> colnames(df)
[1] "ID"          "Year_Birth"    "Education"
[4] "Marital_Status" "Income"       "Kidhome"
[7] "Teenhome"     "Dt_Customer"  "Recency"
[10] "MntWines"     "MntFruits"   "MntMeatProducts"
```

Environment History Connections Tutorial

R - Global Environment

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

The screenshot shows an RStudio interface with the following details:

- Header:** "IITR BA: Foundations of Business Analytics" and "Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2".
- Top Bar:** "Community", "Notes", and "Help".
- Right Panel:** Shows the status "Used 4.7 of 50 hours in Aug, 2024" and buttons for "Start Lab" and "End Lab". It also displays the message "This Lab will get reset on 15th August 2024, 9:30 AM".
- Environment Tab:** Shows the global environment with a data frame named "df" containing 2240 observations and 29 variables.
- Files Tab:** Shows the project structure with files like "CAPSTONE PROJECT.R", "marketing_campaign.csv", and various R script files ("R_1.R", "R_2.R", "R_2.R", "R_3.R").
- Code Editor:** The main workspace shows R code for reading a CSV file and inspecting its structure:

```
tail(df,5)
# names of the columns
colnames(df)
# row names
rownames(df)
# data types of the variables
typeof(df$ID)
# (Top Level) t
```

The code output shows the first few rows of the "marketing_campaign" dataset:

1	2	3	4	5	6	7	8	9	10	11	12
"1"	"2"	"3"	"4"	"5"	"6"	"7"	"8"	"9"	"10"	"11"	"12"
"13"	"14"	"15"	"16"	"17"	"18"	"19"	"20"	"21"	"22"	"23"	"24"
"25"	"26"	"27"	"28"	"29"	"30"	"31"	"32"	"33"	"34"	"35"	"36"
"37"	"38"	"39"	"40"	"41"	"42"	"43"	"44"	"45"	"46"	"47"	"48"

Below the rows, there is a footer note: "FACT: "Age" = "E41" "E42" "E43" "E44" "E45" "E46" "E47" "E48" "E49" "E50" "E51"

The screenshot shows the PG R Studio interface. The top navigation bar includes 'File', 'Edit', 'View', 'History', 'Bookmarks', 'Profiles', 'Tab', 'Window', and 'Help'. The address bar points to 'lms.simplilearn.com/courses/6395/IITR-BA--Foundations-of-Business-Analytics/practice-labs'. A sidebar on the left features 'Learning Track' and 'Certificate' sections. The main area is titled 'IITR BA: Foundations of Business Analytics' with a progress message 'Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2'. A 'Community' button is in the top right. The central workspace is titled 'PG R Studio' and contains an 'rstudio' icon. The status bar at the bottom right says 'Used 4.7 of 50 hours in Aug, 2024' and includes 'Start Lab' and 'End Lab' buttons. The RStudio interface has tabs for 'File', 'Edit', 'Code', 'View', 'Plots', 'Session', 'Build', 'Debug', 'Profile', 'Tools', and 'Help'. The code editor shows R code for reading a CSV file and printing its structure. The 'Environment' tab in the sidebar lists variables like 'df' (2240 obs. of 29 variables) and 'marketing_campaign.csv'. The 'Files' tab shows files in the project directory, including 'CAPSTONE PROJECT.R' (2.5 KB), 'marketing_campaign.csv' (215 KB), and several R script files ('R_1.R', 'R_2(1).R', 'R_2.R', 'R_3.R') with their respective sizes and modification dates.

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 9:02 PM

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

CAPSTONE PROJECT.R* marketing_campaign df df

62 rownames(df)

63

64 # data types of the variables

65

66 typeof(df\$ID)

67

68 str(df)

69

70 df\$Dt_Customer[1:10]

71

72 as.Date(df\$Dt_Customer,format = "%d-%m-%Y")

73

74 # lubridate

75

72:1 (Top Level) t R Script

Console Terminal Jobs

R 4.0.2 - ~/

\$ Complain : int 0 0 0 0 0 0 0 0 0 ...

\$ Z_CostContact : int 3 3 3 3 3 3 3 3 3 ...

\$ Z_Revenue : int 11 11 11 11 11 11 11 11 11 ...

\$ Response : int 1 0 0 0 0 0 0 1 0 ...

> df\$Dt_Customer[1:10]

[1] "04-09-2012" "08-03-2014" "21-08-2013" "10-02-2014" "19-01-2014" "09-09-2013"

[7] "13-11-2012" "08-05-2013" "06-06-2013" "13-03-2014"

[7] >

Environment History Connections Tutorial

Import Dataset 174 MiB

Data Global Environment

df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name Size Modified

CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM

marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM

R

R_1.R 1.4 KB Feb 29, 2024, 6:39 AM

R_2_(1).R 10.4 KB Mar 8, 2024, 8:00 AM

R_2.R 10.4 KB Mar 8, 2024, 9:54 AM

R_3.R 4 KB Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 9:02 PM

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

CAPSTONE PROJECT.R* marketing_campaign df df

64 # data types of the variables

65

66 typeof(df\$ID)

67

68 str(df)

69

70 df\$Dt_Customer[1:10]

71

72 as.Date(df\$Dt_Customer,format = "%d-%m-%Y")

73

74 # lubridate

75

76 install.packages('lubridate')

77

74:1 (Top Level) t R Script

Console Terminal Jobs

R 4.0.2 - ~/

> as.Date(df\$Dt_Customer,format = "%d-%m-%Y")

[1] "2012-09-04" "2014-03-08" "2013-08-21" "2014-02-10" "2014-01-19" "2013-09-09"

[7] "2012-11-13" "2013-05-08" "2013-06-06" "2014-03-13" "2013-11-15" "2012-11-13"

[13] "2013-11-15" "2012-11-10" "2012-11-24" "2012-12-24" "2012-08-31"

[19] "2013-03-28" "2012-11-03" "2012-08-08" "2013-01-06" "2012-12-23" "2014-01-11"

[25] "2013-03-18" "2013-01-02" "2013-05-27" "2013-02-20" "2013-05-31" "2013-11-22"

[31] "2014-05-22" "2013-05-11" "2012-10-10" "2012-10-29" "2013-02-20" "2013-08-29"

[37] "2013-12-31" "2013-09-02" "2014-02-11" "2013-02-01" "2013-04-29" "2013-03-12"

[43] "2012-12-24" "2013-11-05" "2013-10-02" "2014-06-28" "2012-11-09" "2013-05-24"

Environment History Connections Tutorial

Import Dataset 174 MiB

Data Global Environment

df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name Size Modified

CAPSTONE PROJECT.R 2.5 KB Aug 4, 2024, 10:04 AM

marketing_campaign.csv 215 KB Aug 4, 2024, 5:55 AM

R

R_1.R 1.4 KB Feb 29, 2024, 6:39 AM

R_2_(1).R 10.4 KB Mar 8, 2024, 8:00 AM

R_2.R 10.4 KB Mar 8, 2024, 9:54 AM

R_3.R 4 KB Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 9:05 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

This Lab will get reset on 15th August 2024, 9:30 AM

PG R Studio

rstudio

Used 4.7 of 50 hours in Aug, 2024 Start Lab End Lab

R Script

```
70 dfSDt_Customer[1:10]
71 
72 as.Date(dfSDt_Customer,format = "%d-%m-%Y")
73 
74 # lubridate
75 
76 install.packages('lubridate')
77 
78 library('lubridate')
79 
80 # to check all packages have been loaded in the RAM
81 
82 search()
83 
84 dfSDt_Customer[1:10]
85 
86 dmy(dfSDt_Customer)
87 
```

Console Terminal Jobs

R 4.0.2 - ~/ ~

* installing *binary* package 'lubridate' ...

* DONE (lubridate)

The downloaded source packages are in
'/tmp/RtmpXImKo3/downloaded_packages'
library('lubridate')

Environment History Connections Tutorial

Import Dataset - 174 MiB

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Name Size Modified

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 9:05 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

This Lab will get reset on 15th August 2024, 9:30 AM

PG R Studio

rstudio

Used 4.7 of 50 hours in Aug, 2024 Start Lab End Lab

R Script

```
74 # lubridate
75 
76 install.packages('lubridate')
77 
78 library('lubridate')
79 
80 # to check all packages have been loaded in the RAM
81 
82 search()
83 
84 dfSDt_Customer[1:10]
85 
86 dmy(dfSDt_Customer)
87 
```

Console Terminal Jobs

R 4.0.2 - ~/ ~

1: package 'lubridate' was built under R version 4.0.5

2: In system("timedatectl", intern = TRUE) :

running command 'timedatectl' had status 1

> search()

[1] ".GlobalEnv" "package:lubridate" "tools:rstudio" "package:stats"

[5] "package:graphics" "package:grDevices" "package:utils" "package:datasets"

[9] "package:methods" "Autoloads" "package:base"

Environment History Connections Tutorial

Import Dataset - 213 MiB

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Name Size Modified

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help New tab Mon 5 Aug 9:06 PM

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 ► Start Lab ■ End Lab X

Learning Track Certificate

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
76 install.packages('lubridate')
77
78 library('lubridate')
79
80 # to check all packages have been loaded in the RAM
81
82 search()
83
84 df$Dt_Customer[1:10]
85
86 dmy(df$Dt_Customer)
87
88 # for example
89
90 y <- '25July2024'
91
```

R Script

Console Terminal Jobs

R 4.0.2 - / -

```
> search()
[1] ".GlobalEnv" "package:lubridate" "tools:rstudio" "package:stats"
[5] "package:graphics" "package:grDevices" "package:utils" "package:datasets"
[9] "package:methods" "Autoloads" "package:base"
> df$Dt_Customer[1:10]
[1] "04-09-2012" "08-03-2014" "21-08-2013" "10-02-2014" "19-01-2014" "09-09-2013"
[7] "13-11-2012" "08-05-2013" "06-06-2013" "13-03-2014"
>
```

Environment History Connections Tutorial

Import Dataset - 213 MiB

R - Global Environment

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2_(1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help New tab Mon 5 Aug 9:06 PM

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 ► Start Lab ■ End Lab X

Learning Track Certificate

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
78 library('lubridate')
79
80 # to check all packages have been loaded in the RAM
81
82 search()
83
84 df$Dt_Customer[1:10]
85
86 dmy(df$Dt_Customer)
87
88 # for example
89
90 y <- '25July2024'
91
```

R Script

Console Terminal Jobs

R 4.0.2 - / -

```
> search()
[1] ".GlobalEnv" "package:lubridate" "tools:rstudio" "package:stats"
[5] "package:graphics" "package:grDevices" "package:utils" "package:datasets"
[9] "package:methods" "Autoloads" "package:base"
> df$Dt_Customer[1:10]
[1] "2012-09-04" "2014-03-08" "2013-08-21" "2014-02-10" "2014-01-19" "2013-09-09"
[7] "2012-11-13" "2013-05-08" "2013-06-06" "2014-03-13" "2013-11-15" "2012-11-13"
[13] "2013-11-15" "2013-11-15" "2012-10-10" "2012-11-24" "2012-12-24" "2012-08-31"
[19] "2013-03-28" "2012-11-03" "2012-08-08" "2013-01-06" "2012-12-23" "2014-01-11"
[25] "2013-03-18" "2013-01-02" "2013-05-27" "2013-02-20" "2013-05-31" "2013-11-22"
[31] "2014-05-22" "2013-05-11" "2012-10-18" "2012-10-29" "2013-02-20" "2013-08-29"
[37] "2013-12-31" "2013-09-02" "2014-02-11" "2013-02-01" "2013-04-29" "2013-03-12"
```

Environment History Connections Tutorial

Import Dataset - 213 MiB

R - Global Environment

Data df 2240 obs. of 29 variables

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2_(1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 9:07 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

This Lab will get reset on 15th August 2024, 9:30 AM

PG R Studio

rstudio

Used 4.7 of 50 hours in Aug, 2024 Start Lab End Lab

R Script

```
88 # for example
89
90 y <- '25July2024'
91
92 dmy(y)
93
94 # YYYY-MM-DD - date format in R
95
96 a <- '25/7/24'
97
98 dmy(a)
99
100 str(df)
101
```

(Top Level) :

Console Terminal Jobs

```
R 4.0.2 - ~/ ~
[1] "2014-03-21" "2012-11-29" "2012-08-31" "2014-02-18" "2014-06-22" "2014-03-26"
[991] "2013-11-09" "2014-02-25" "2013-08-13" "2013-11-03"
[997] "2013-11-09" "2014-02-25" "2013-08-13" "2013-11-03"
[ reached 'max' / getOption("max.print") -- omitted 1240 entries ]
> y <- '25July2024'
> dmy(y)
[1] "2024-07-25"
> a <- '25/7/24'
>
```

Environment History Connections Tutorial

Import Dataset 213 MiB

Data df 2240 obs. of 29 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

Project: (None)

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

Chrome File Edit View History Bookmarks Profiles Tab Window Help Mon 5 Aug 9:08 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2

This Lab will get reset on 15th August 2024, 9:30 AM

PG R Studio

rstudio

Used 4.7 of 50 hours in Aug, 2024 Start Lab End Lab

R Script

```
92 dmy(y)
93
94 # YYYY-MM-DD - date format in R
95
96 a <- '25/7/24'
97
98 dmy(a)
99
100 str(df)
101
102 # count the number of missing value
103 sapply(df,function(x)(sum(is.na(x))))
104
105
```

(Top Level) :

Console Terminal Jobs

```
R 4.0.2 - ~/ ~
[1] "2024-07-25"
> a <- '25/7/24'
> dmy(a)
[1] "2024-07-25"
> str(df)
'data.frame': 2240 obs. of 29 variables:
 $ ID : int 5524 2174 4141 6182 5324 7446 965 6177 4855 5899 ...
 $ Year_Birth : int 1957 1954 1965 1984 1981 1967 1971 1985 1974 1950 ...
 $ Education : chr "Graduation" "Graduation" "Graduation" "Graduation" ...
```

Environment History Connections Tutorial

Import Dataset 213 MiB

Data df 2240 obs. of 29 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

Project: (None)

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

Community Notes Help

R Studio

CAPSTONE PROJECT.R*

marketing_campaign

df

df

102 # count the number of missing value
103
104 sapply(df,function(x)(sum(is.na(x))))
105
106 24/(dim(df)[1])
107
108 na.omit(df)
109
110 df <- na.omit(df)
111
112 dim(df)
113
114 # oldest
115

(Top Level) R Script

Console Terminal Jobs

R 4.0.2 - / ~

> sapply(df,function(x)(sum(is.na(x))))

ID	Year_Birth	Education	Marital_Status
0	0	0	0
Income	Kidhome	Teenhome	Dt_Customer
24	0	0	0
Recency	MntFwines	MntFruits	MntMeatProducts
0	0	0	0
MntFishProducts	MntSweetProducts	MntGoldProds	NumDealsPurchases
0	0	0	0

Environment History Connections Tutorial

Import Dataset - 213 MB

R - Global Environment

Data df 2216 obs. of 29 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R	1.4 KB	Feb 29, 2024, 6:39 AM
R_1.R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2_(1).R	10.4 KB	Mar 8, 2024, 9:54 AM
R_2.R	4 KB	Mar 8, 2024, 9:53 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

Community Notes Help

R Studio

CAPSTONE PROJECT.R*

marketing_campaign

df

df

102 # count the number of missing value
103
104 sapply(df,function(x)(sum(is.na(x))))
105
106 24/(dim(df)[1])
107
108 na.omit(df)
109
110 df <- na.omit(df)
111
112 dim(df)
113
114 # oldest
115

(Top Level) R Script

Console Terminal Jobs

R 4.0.2 - / ~

> response

> 0

> 24/(dim(df)[1])

[1] 0.01071429

> na.omit(df)

ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	
1	5524	1957	Graduation	Single	58138	0	0	04-09-2012	58
2	2174	1954	Graduation	Single	46344	1	1	08-03-2014	38
3	4141	1965	Graduation	Together	71613	0	0	21-08-2013	26

Environment History Connections Tutorial

Import Dataset - 213 MB

R - Global Environment

Data df 2216 obs. of 29 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R	1.4 KB	Feb 29, 2024, 6:39 AM
R_1.R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2_(1).R	10.4 KB	Mar 8, 2024, 9:54 AM
R_2.R	4 KB	Mar 8, 2024, 9:53 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

```
182 # count the number of missing value
183
184 sapply(df,function(x)(sum(is.na(x))))
185
186 24/(dim(df)[1])
187
188 na.omit(df)
189
190 df <- na.omit(df)
191
192 dim(df)
193
194 # oldest
195
196 min(df$Dt_Customer)
197
198 # newest
199
200 max(df$Dt_Customer)
```

R Script

Console Terminal Jobs

R 4.0.2 - /~

34	0	3	11	1
35	0	3	11	0
36	0	3	11	0

[reached 'max' / getOption("max.print") -- omitted 2182 rows]

> df <- na.omit(df)

> dim(df)

dim(df)

Environment History Connections Tutorial

Import Dataset 213 MiB

Data df 2216 obs. of 29 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2_(1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

```
185 24/(dim(df)[1])
186
187 na.omit(df)
188
189 df <- na.omit(df)
190
191 dim(df)
192
193 # oldest
194
195 min(df$Dt_Customer)
196
197 # newest
198
199 max(df$Dt_Customer)
```

R Script

Console Terminal Jobs

R 4.0.2 - /~

> dim(df)

[1] 2216 29

> dim(df)

[1] 2216 29

> df <- na.omit(df)

> dim(df)

[1] 2216 29

> str(df)

Environment History Connections Tutorial

Import Dataset 213 MiB

Data df 2216 obs. of 29 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2_(1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

The screenshot shows the PG R Studio interface. The top navigation bar includes 'File', 'Edit', 'Code', 'View', 'Plots', 'Session', 'Build', 'Debug', 'Profile', 'Tools', and 'Help'. A 'Community' button is also present. The main area displays an R script titled 'CAPSTONE PROJECT.R' with the following code:

```
112 dim(df)
113
114 # oldest
115 min(df$Dt_Customer)
116
117 # newest
118 max(df$Dt_Customer)
119
120 #create a feature "Customer_for" of the number of days the customers
121 #started to shop in the store relative to the last recorded date
122
123 df$Customer_for <- max(df$Dt_Customer) - df$Dt_Customer
124
125 (Top Level)
```

The 'Environment' pane on the right shows the global environment with variables 'df' (2216 obs. of 29 variables), 'a' ("25/7/24"), and a file list containing 'CAPSTONE PROJECT.R', 'marketing_campaign.csv', 'R', 'R_1.R', 'R_2.R', 'R_2.R', and 'R_3.R'. The 'Plots' tab is selected.

The screenshot shows a web-based RStudio environment within a browser window. The title bar indicates the URL is lms.simplilearn.com/courses/6395/IITR-BA-:-Foundations-of-Business-Analytics/practice-labs. The main header says "IITR BA: Foundations of Business Analytics". Below it, a progress bar shows "Class completed: 1 | 51% of Self-Learning Completed | Projects completed: 1/2". On the right, there are "Community", "Notes", and "Help" buttons.

The main workspace is titled "PG R Studio". It features a toolbar with "rstudio" and "File", "Edit", "Code", "View", "Plots", "Session", "Build", "Debug", "Profile", "Tools", "Help" menus. A status bar at the bottom left shows "R 4.0.2 - ~/".

The code editor contains R script code:

```
123 #Started to shop in the store relative to the last recorded date
124
125 df$Customer_for <- max(df$DT_Customer) - df$DT_Customer
126
127 # age
128 df$Year_Birth
129 Sys.Date()
130
131 year(Sys.Date())
132
133 year(Sys.Date()) - df$Year_Birth
134
135 year(Sys.Date()) - df$Year_Birth
136
137 (Top Level) :
```

The R Script pane shows the current file structure:

- Home
 - Name
 - Size
 - Modified
- CAPSTONE PROJECT.R (2.5 KB, Aug 4, 2024, 10:04 AM)
- marketing_campaign.csv (215 KB, Aug 4, 2024, 5:55 AM)
- R
 - R_1.R (1.4 KB, Feb 29, 2024, 6:39 AM)
 - R_2 (1).R (10.4 KB, Mar 8, 2024, 8:00 AM)
 - R_2.R (10.4 KB, Mar 8, 2024, 9:54 AM)
 - R_3.R (4 KB, Mar 8, 2024, 9:53 AM)

The Environment pane shows the global environment:

- Data: df (2216 obs. of 29 variables)
- Values:
 - a "25/7/24"

The History, Connections, and Tutorial panes are also visible on the right side of the interface.

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

Community Notes Help

File Edit Code View Plots Session Build Debug Profile Tools Help

File Edit Code View Plots Session Build Debug Profile Tools Help

R studio

CAPSTONE PROJECT.R marketing_campaign df df

```
129 df$Year_Birth
130 Sys.Date()
131 year(Sys.Date())
132
133 year(Sys.Date()) - df$Year_Birth
134
135 df$age <- year(Sys.Date()) - df$Year_Birth
136
137 # create a feature "spent" indicating total amount spent by the customer in
138 # various categories over two years
139
140 install.packages("dplyr")
141
142 (Top Level) t
```

R Script

Console Terminal Jobs

R 4.0.2 - /~ [reached getOption("max.print") -- omitted 1216 entries]
> Sys.Date()
[1] "2024-08-05"
> year(Sys.Date())
[1] 2024
> year(Sys.Date()) - df\$Year_Birth
[1] 67 70 59 40 43 57 53 39 50 74 48 65 72 37 78 44
[17] 78 75 39 42 45 75 70 73 55 48 35 59 35 61 54 72
[33] 78 73 54 48 51 81 44 39 67 49 48 28 56 70 67 68

Environment History Connections Tutorial

Import Dataset - 213 MiB +

R - Global Environment

Data df 2216 obs. of 30 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

Community Notes Help

File Edit Code View Plots Session Build Debug Profile Tools Help

File Edit Code View Plots Session Build Debug Profile Tools Help

R studio

CAPSTONE PROJECT.R marketing_campaign df df

```
129 df$Year_Birth
130 Sys.Date()
131 year(Sys.Date())
132
133 year(Sys.Date()) - df$Year_Birth
134
135 df$age <- year(Sys.Date()) - df$Year_Birth
136
137 # create a feature "spent" indicating total amount spent by the customer in
138 # various categories over two years
139
140 install.packages("dplyr")
141
142 (Top Level) t
```

R Script

Console Terminal Jobs

R 4.0.2 - /~ [reached getOption("max.print") -- omitted 1216 entries]
> df\$age <- year(Sys.Date()) - df\$Year_Birth
> str(df)

Environment History Connections Tutorial

Import Dataset - 213 MiB +

R - Global Environment

Data df 2216 obs. of 30 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

New Folder New Blank File Upload Delete Rename More

Home

Name	Size	Modified
CAPSTONE PROJECT.R	2.5 KB	Aug 4, 2024, 10:04 AM
marketing_campaign.csv	215 KB	Aug 4, 2024, 5:55 AM
R		
R_1.R	1.4 KB	Feb 29, 2024, 6:39 AM
R_2 (1).R	10.4 KB	Mar 8, 2024, 8:00 AM
R_2.R	10.4 KB	Mar 8, 2024, 9:54 AM
R_3.R	4 KB	Mar 8, 2024, 9:53 AM

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
145 library('dplyr')
146
147 select(df,starts_with("Mnt"))
148
149 # select all column begin with Mnt
150
151 df1 <- select(df,starts_with("Mnt"))
152
153 head(df1)
154
155 # sum of the rows
156
152:1 (Top Level) :
```

R Script

Console Terminal Jobs

R 4.0.2 - ~/

> library('dplyr')

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

Environment History Connections Tutorial

Import Dataset - 224 MiB

Data Global Environment -

df 2216 obs. of 30 variables

df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
145 library('dplyr')
146
147 select(df,starts_with("Mnt"))
148
149 # select all column begin with Mnt
150
151 df1 <- select(df,starts_with("Mnt"))
152
153 head(df1)
154
155 # sum of the rows
156
150:1 (Top Level) :
```

R Script

Console Terminal Jobs

R 4.0.2 - ~/

package 'dplyr' was built under R version 4.0.5

> select(df,starts_with("Mnt"))

	MntWines	MntFruitProducts	MntMeatProducts	MntFishProducts	MntSweetProducts	MntGoldProducts
1	635	88	546	172	88	88
2	11	1	6	2	1	6
3	426	49	127	111	21	42
4	11	4	20	10	3	5
5	173	43	118	46	27	15
6	520	42	98	0	42	14
7	235	65	164	50	49	27
e	-e	-e	-e	-e	-e	-e

Environment History Connections Tutorial

Import Dataset - 224 MiB

Data Global Environment -

df 2216 obs. of 30 variables

df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

The screenshot shows an RStudio interface with the following details:

- Title Bar:** PG R Studio
- Header:** This Lab will get reset on 15th August 2024, 9:30 AM
- Code Editor:** The code is part of a file named "CAPSTONE PROJECT.R". It includes the following R code:

```
147 select(df,starts_with("Mnt"))
148
149 # select all column begin with Mnt
150
151 df1 <- select(df,starts.with("Mnt"))
152
153 head(df1)
154
155 # sum of the rows
156
157 # apply(df1,MARGIN =1,apply the function at rows,2=column),function)
158
158:1 (Top Level) :
```
- Environment View:** Shows two data frames:
 - df: 2216 obs. of 30 variables
 - df1: 2216 obs. of 6 variables
- Console View:** Displays the output of the executed code, including the creation of df1 and its head.
- Data View:** Shows the structure of df1:

MntWines	MntFruits	MntMeatProducts	MntFishProducts	MntSweetProducts	MntGoldProds	
1	635	88	546	172	88	88
2	11	1	6	2	1	6
3	426	49	127	111	21	42
4	11	4	28	10	3	5
5	173	43	118	46	27	15
6	520	42	98	0	42	14

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

This Lab will get reset on 15th August 2024, 9:30 AM

PG R Studio

rstudio

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

File Edit Code View Plots Session Build Debug Profile Tools Help

R CAPSTONE PROJECT.R marketing_campaign df df

Source on Save Run Source

```
153 head(df1)
154 
155 # sum of the rows
156 
157 # apply(df1,MARGIN = 1{apply the function at rows,2=column},function)
158 | 
159 apply(df1,MARGIN = 1,FUN = sum)
160 
161 df$amt_spent <- apply(df1,MARGIN = 1,FUN = sum)
162 
163 View(df)
164 
```

158:1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
> apply(df1,MARGIN = 1,FUN = sum)
```

1	2	3	4	5	6	7	8	9	10	12	13	14	15	16	17	18	19
1617	27	776	53	422	718	598	169	46	49	61	1102	310	46	1315	96	317	1782
20	21	22	23	24	25	26	27	29	30	31	32	33	34	35	36	37	38
133	316	1730	972	544	444	75	257	131	1672	30	318	128	302	1196	65	913	81
39	40	41	42	43	45	46	47	48	50	51	52	53	54	55	56	57	58
67	982	1395	53	22	31	984	122	55	1319	587	1693	72	1617	606	1957	1093	29
60	61	62	63	64	65	66	67	68	69	70	71	73	74	75	76	77	78
518	1438	612	884	606	1076	34	11	1274	653	1562	1253	863	661	65	13	1890	2209
79	80	81	82	83	84	85	86	87	88	89	90	94	95	96	97	98	99

Environment History Connections Tutorial

Import Dataset 224 MiB

Global Environment

df 2216 obs. of 38 variables

df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMAP(d,q)	1.5-3
ifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

User Library

Learning Track

Certificate

The screenshot shows a RStudio interface with the following details:

- Header:** PG R Studio, Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2.
- Toolbar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Environment:** Shows variables df, df1, and Global Environment.
- Data View:** A data frame named "marketing_campaign" with 2,216 observations and 31 columns. The columns include ID, Year_Birth, Education, Marital_Status, Income, Kidhome, Teenhome, and Dt_Customer.
- Console:** Displays R code and its output. The output includes:
 - Showing 1 to 8 of 2,216 entries, 31 total columns
 - R 4.0.2 - ~
 - 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985
 - 93 23 662 1198 81 92 491 1564 926 1910 411 94 1085 23 6 54 1150 1415
 - 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003
 - 405 1908 2486 1315 45 1789 19 573 805 108 23 1493 89 936 343 528 1901 724
 - 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013
 - 58 297 488 859 1574 69 17 2194 116 279
- Notes:** This Lab will get reset on 15th August 2024, 9:30 AM.
- Help:** Used 4.7 of 50 hours in Aug, 2024.

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug. 2024

Start Lab End Lab

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R

```

157 # apply(df1,MARGIN = 1,apply the function at rows,2=column),function)
158
159 apply(df1,MARGIN = 1,FUN = sum)
160
161 df$amt_spent <- apply(df1,MARGIN = 1,FUN = sum)
162
163 View(df)
164
165 sum(df$MtnlNines)
166
167 df$Marital_Status
168
169:1 (Top Level) R Script
```

Console Terminal Jobs

	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985
93	23	662	1198	81	92	491	1564	926	1910	411	94	1085	23	6	54	1150	1415	
986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	
405	1908	2486	1315	45	1789	19	573	805	108	23	1493	89	936	343	528	1901	724	
1004	1805	1806	1067	1068	1069	1010	1011	1012	1013	58	297	484	859	1574	69	17	2194	116
279	[reached getOption("max.print") -- omitted 1216 entries]	df\$amt_spent <- apply(df1,MARGIN = 1,FUN = sum)	View(df)	>														

Community Notes Help

Learning Track Certificate

Project: (None)

Chrome File Edit View History Bookmarks Profiles Tab Window Help

Tue 6 Aug 12:02 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
161 df$amt_spent <- apply(df1,MARGIN = 1,FUN = sum)
162 
163 View(df)
164 
165 sum(df$MntWines)
166 
167 df$Marital_Status
168 
169 unique(df$Marital_Status)
170 
171 # partner - married or together, rest all alone
172 
169:1 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/ ...
465 1908 2486 1315 45 1789 19 573 805 108 23 1493 89 936 343 528 1901 724
1004 1005 1006 1007 1008 1009 1010 1011 1012 1013
58 297 488 859 1574 69 17 2194 116 279
[ reached getOption("max.print") -- omitted 1216 entries ]
> df$amt_spent <- apply(df1,MARGIN = 1,FUN = sum)
> View(df)
> sum(df$MntWines)
[1] 676083
> df$Marital_Status
[1] "Single" "Together" "Together" "Married" "Together" "Divorced"
[8] "Married" "Together" "Together" "Married" "Divorced" "Divorced" "Married"
```

Environment History Connections Tutorial

R - Global Environment

Data

- df 2216 obs. of 31 variables
- df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Chrome File Edit View History Bookmarks Profiles Tab Window Help

Tue 6 Aug 12:03 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
163 View(df)
164 
165 sum(df$MntWines)
166 
167 df$Marital_Status
168 
169 unique(df$Marital_Status)
170 
171 # partner - married or together, rest all alone
172 
173 # ifelse(condition, the statement to execute if the condition is met, the statement to execute if the condition is not met)
174 
171:1 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/ ...
[960] "Married" "Divorced" "Together" "Married" "Together" "Married" "Married"
[967] "Together" "Widow" "Divorced" "Divorced" "Divorced" "Married" "Married"
[974] "Married" "Married" "Single" "Widow" "Divorced" "Married" "Together"
[981] "Single" "Divorced" "Single" "Together" "Together" "Married" "Divorced"
[988] "Widow" "Single" "Married" "Together" "Married" "Single" "Married"
[995] "Together" "Married" "Single" "Together" "Together" "Single" "Married"
[ reached getOption("max.print") -- omitted 1216 entries ]
> unique(df$Marital_Status)
[1] "Single" "Together" "Married" "Divorced" "Widow" "Alone" "Absurd" "YOLO"
>
```

Environment History Connections Tutorial

R - Global Environment

Data

- df 2216 obs. of 31 variables
- df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

Community Notes Help

Learning Track Certificate

R Studio

R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
170 # partner - married or together, rest all alone
171 # ifelse(condition, the statement to execute if the condition is met, the statement to execute if the condition is not met)
172 df$Living_With<-ifelse(df$Marital_Status %in% c("Married", "Together"), "Partner", "Alone")
173 df$Living_With
174 df$Living_With<-unique(df$Living_With)
175 table(df$Living_With)
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

Import Dataset 224 MiB Global Environment

Data df 2216 obs. of 32 variables df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

Community Notes Help

Learning Track Certificate

R Studio

R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
170 # partner - married or together, rest all alone
171 # ifelse(condition, the statement to execute if the condition is met, the statement to execute if the condition is not met)
172 df$Living_With<-ifelse(df$Marital_Status %in% c("Married", "Together"), "Partner", "Alone")
173 df$Living_With
174 df$Living_With<-unique(df$Living_With)
175 table(df$Living_With)
```

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

Import Dataset 224 MiB Global Environment

Data df 2216 obs. of 32 variables df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 1:19 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio This Lab will get reset on 15th August 2024, 9:30 AM

rstudio

Used 4.7 of 50 hours in Aug, 2024 ► Start Lab ■ End Lab

Learning Track Certificate

R Studio Environment History Connections Tutorial

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

df1\$Parent <- ifelse(df\$Children>0,1,0)

keep only two categories in the Field - 'Education' - Undergraduate, Graduate

table(df\$Education)

#2n,basic- UG #Graduation,Master,PhD- Graduate

df\$Education <- ifelse(df\$Education %in% c("2n cycle","Basic"),"Undergraduate","Graduate")

table(df\$Education)

2n Cycle Basic Graduation Master PhD

200 54 1116 365 481

> df\$Education <- ifelse(df\$Education %in% c("2n cycle","Basic"),"Undergraduate","Graduate")

> table(df\$Education)

Graduate Undergraduate

2162 54

R Script

Console Terminal Jobs

R 4.0.2 - ~/

> table(df\$Education)

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMAP(p,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 2:17 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio This Lab will get reset on 15th August 2024, 9:30 AM

rstudio

Used 4.7 of 50 hours in Aug, 2024 ► Start Lab ■ End Lab

Learning Track Certificate

R Studio Environment History Connections Tutorial

File Edit Code View Plots Session Build Debug Profile Tools Help

Project: (None)

df

table(df\$Education)

#2n,basic- UG #Graduation,Master,PhD- Graduate

df\$Education <- ifelse(df\$Education %in% c("2n cycle","Basic"),"Undergraduate","Graduate")

table(df\$Education)

d <- "data science"

d

change data - analytics

names(df)

[1] "ID" "Year_Birth" "Education"

[4] "Marital_Status" "Income" "Kidhome"

[7] "Teenhome" "Dt_Customer" "Recency"

[10] "MntWines" "MntFruits" "MntMeatProducts"

[13] "MntFishProducts" "MntSweetProducts" "MntGoldProds"

[16] "NumDealsPurchases" "NumWebPurchases" "NumCatalogPurchases"

[19] "NumStorePurchases" "NumWebVisitsMonth" "AcceptedCmp3"

R Script

Console Terminal Jobs

R 4.0.2 - ~/

[1] "analytics"

> sub(d,"data","analytics")

[1] "analytics"

> names(df)

[1] "ID" "Year_Birth" "Education"

[4] "Marital_Status" "Income" "Kidhome"

[7] "Teenhome" "Dt_Customer" "Recency"

[10] "MntWines" "MntFruits" "MntMeatProducts"

[13] "MntFishProducts" "MntSweetProducts" "MntGoldProds"

[16] "NumDealsPurchases" "NumWebPurchases" "NumCatalogPurchases"

[19] "NumStorePurchases" "NumWebVisitsMonth" "AcceptedCmp3"

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMAP(p,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

CAPSTONE PROJECT.R* marketing_campaign df df

```
206 table(df$Education)
207 
208 d <- "data science"
209 
210 # change data - analytics
211 
212 sub("data","analytics",d)
213 
214 # Mnt- remove it from the column names
215 
216 names(df)
217
```

(Top Level) :

R Script :

Console Terminal Jobs

R 4.0.2 - ~

```
[22] "AcceptedCmp4"      "AcceptedCmp5"      "AcceptedCmp1"
[23] "AcceptedCmp2"      "Complain"        "Z_CostContact"
[28] "Z_Revenue"          "Response"        "age"
[31] "mkt_spent"          "Living_With"     "Children"
[34] "no_adults"          "Family_Size"    "Is_Parent"
> sub(d,"data","analytics",d)
[1] "analytics"
> sub("data","analytics",d)
[1] "analytics science"
> sub("Mnt","",names(df))
[1] "ID"                 "Year_Birth"      "Education"
```

Environment History Connections Tutorial

Import Dataset 224 MiB

Data df 2216 obs. of 30 variables df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

Name Description Version

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(P,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024

Used 4.7 of 50 hours in Aug, 2024

Start Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 52% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

CAPSTONE PROJECT.R* marketing_campaign df df

```
206 table(df$Education)
207 
208 d <- "data science"
209 
210 # change data - analytics
211 
212 sub("data","analytics",d)
213 
214 # Mnt- remove it from the column names
215 
216 names(df)
217
```

(Top Level) :

R Script :

Console Terminal Jobs

R 4.0.2 - ~

```
[34] "no_adults"          "Family_Size"    "Is_Parent"
> names(df) <- sub("Mnt","",names(df))
> names(df)
[1] "ID"                 "Year_Birth"      "Education"
[4] "Marital_Status"     "Income"         "Kidhome"
```

Environment History Connections Tutorial

Import Dataset 224 MiB

Data df 2216 obs. of 30 variables df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

Name Description Version

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0
cli	Helpers for Developing Command Line Interfaces	3.6
desc	Manipulate DESCRIPTION Files	1.4
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.1
dplyr	A Grammar of Data Manipulation	1.1
forecast	Forecasting Functions for Time Series and Linear Models	8.2
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(P,d,q)	1.5

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 ▶ Start Lab ■ End Lab

```

222 names(df)
223
224 # question 33. Drop the redundant columns: "Marital_Status", "Dt_Customer", "Z_CostContact"
225 df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
226 df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
227
228 View(df)
229
230 names(df)
231
232 names(df)
233
234.12 (Top Level) ▾ R Script ▾
```

R 4.0.2 - ~/

```

> df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
  Marital_Status Dt_Customer Z_CostContact Z_Revenue Year_Birth ID
1      Single 04-09-2012       3     11    1957 5524
2      Single 08-03-2014       3     11    1954 2174
3    Together 21-08-2013       3     11    1965 4141
4    Together 10-02-2014       3     11    1984 6182
5     Married 19-01-2014       3     11    1981 5324
6    Together 09-09-2013       3     11    1967 7446
7   Divorced 13-11-2012       3     11    1971 965
8     Married 08-05-2013       3     11    1985 6177
```

Environment History Connections Tutorial

Data

- df 2216 obs. of 30 variables
- df1 2216 obs. of 6 variables

Values

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024 ▶ Start Lab ■ End Lab

```

222 names(df)
223
224 # question 33. Drop the redundant columns: "Marital_Status", "Dt_Customer", "Z_CostContact"
225 df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
226 df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
227
228 View(df)
229
230 names(df)
231
232 names(df)
233
234.12 (Top Level) ▾ R Script ▾
```

R 4.0.2 - ~/

```

> df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
  Marital_Status Dt_Customer Z_CostContact Z_Revenue Year_Birth ID
1      Single 04-09-2012       3     11    1957 5524
2      Single 08-03-2014       3     11    1954 2174
3    Together 21-08-2013       3     11    1965 4141
4    Together 10-02-2014       3     11    1984 6182
5     Married 19-01-2014       3     11    1981 5324
6    Together 09-09-2013       3     11    1967 7446
7   Divorced 13-11-2012       3     11    1971 965
8     Married 08-05-2013       3     11    1985 6177
```

Environment History Connections Tutorial

Data

- df 2216 obs. of 30 variables
- df1 2216 obs. of 6 variables

Values

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

RStudio

CAPSTONE PROJECT.R*

marketing_campaign df df

```
228 df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
229
230 View(df)
231
232 names(df)
233
234 df[,-1]
235
236
237
238
239
```

233.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
[28] "no_adults" "Family_Size" "Is_Parent"
> df[,-1]
```

	Income	Kidhome	Teenhome	Recency	Wines	MeatProducts	FishProducts	SweetProducts	
1	58138	0	0	58	635	88	546	172	88
2	46344	1	1	38	11	1	6	2	1
3	71613	0	0	26	426	49	127	111	21
4	26646	1	0	26	11	4	20	10	3
5	58293	1	0	94	173	43	118	46	27
6	62513	0	1	16	528	42	98	8	42
7	55635	0	1	34	235	65	164	50	49

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 4.7 of 50 hours in Aug, 2024

Start Lab End Lab

RStudio

CAPSTONE PROJECT.R*

marketing_campaign df df

```
228 df[,c("Marital_Status", "Dt_Customer", "Z_CostContact", "Z_Revenue", "Year_Birth", "ID")]
229
230 View(df)
231
232 names(df)
233
234 df[,-1]
235
236
237
238
239
```

233.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
[29] 0 0 0 35 131 Partner 1 2 3 1
30 0 0 59 1672 Partner 0 2 2 0
31 0 0 35 30 Alone 0 1 1 0
32 0 0 61 318 Partner 0 2 2 0
33 0 0 54 120 Partner 1 2 3 1
34 0 1 72 302 Alone 2 1 3 1
35 0 0 78 1196 Partner 0 2 2 0
36 0 0 73 65 Alone 2 1 3 1
[ reached 'max' / getOption("max.print") -- omitted 2182 rows ]
```

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
234 df[,1]
235
236 # 14. boxplot for age
237
238 dfSage
239
240 boxplot(dfSage)
241
242
243
244
245
```

241.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

```
[ reached 'max' / getOption("max.print") -- omitted 2182 rows ]
> dfSage
[1] 67 70 59 40 43 57 53 39 50 74 48 65 72 37 78 44 78 75 39 42 45
[22] 75 70 73 55 48 35 59 35 61 54 72 78 73 54 48 51 81 44 39 67 49
[43] 48 28 56 70 67 60 47 47 46 69 58 36 56 42 49 72 73 76 53 57 45
[64] 66 54 70 65 49 72 47 52 48 48 59 39 55 51 61 37 64 54 67 51 57
[85] 52 54 41 64 52 73 41 64 65 45 79 65 45 60 55 46 64 45 33 61 33
[106] 44 62 54 46 57 55 71 57 39 48 41 59 40 71 68 72 66 55 60 57 48
[127] 65 51 59 32 48 62 50 36 48 49 43 51 58 71 55 52 36 72 75 66 46
[148] 47 69 66 79 49 51 58 71 65 36 67 59 42 54 65 57 49 76 54 65
```

Environment History Connections Tutorial

Data

- df 2216 obs. of 30 variables
- df1 2216 obs. of 6 variables

Values

Plots

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
235
236 # 14. boxplot for age
237
238 dfSage
239
240 boxplot(dfSage)
241
242 boxplot(dfSage,horizontal = TRUE)
243
244
245
246
```

243.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

```
[883] 52 68 66 50 38 66 54 57 51 55 47 63 53 50 38 58 32 61 52 30 70
[904] 64 42 70 69 31 55 76 38 45 46 42 31 48 50 35 73 35 72 53 53
[925] 54 53 49 72 62 41 71 38 37 42 45 47 49 45 48 51 53 48 48 59 72
[946] 51 67 57 62 55 49 44 65 51 51 50 44 36 56 36 65 53 62 65 59 72
[967] 70 64 59 45 46 74 58 46 55 48 68 70 52 43 35 70 29 46 67 51 68
[988] 68 37 43 37 70 70 71 72 56 61 45 48 46
[ reached getOption("max.print") -- omitted 1216 entries ]
> boxplot(dfSage)
> boxplot(dfSage,horizontal = TRUE)
> |
```

Environment History Connections Tutorial

Data

- df 2216 obs. of 30 variables
- df1 2216 obs. of 6 variables

Values

Plots

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 6:15 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

RStudio

CAPSTONE PROJECT.R* marketing_campaign df df

```

239
240 boxplot(df$age)
241
242 boxplot(df$age, horizontal = TRUE)
243
244 boxplot(df$age, horizontal = TRUE, xlabel="Age")
245
246 boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
247
248
249
250
247:1 (Top Level) R Script

```

Console Terminal Jobs

```

R 4.0.2 - ~/ ...
[946] 51 67 57 62 55 49 44 65 51 51 50 44 36 56 36 65 53 62 65 59 72
[967] 70 64 59 45 46 74 50 46 55 48 68 70 52 43 35 70 29 46 67 51 68
[988] 68 37 43 37 70 70 71 72 56 61 45 48 46
[ reached getOption("max.print") -- omitted 1216 entries ]
> boxplot(df$age)
> boxplot(df$age, horizontal = TRUE)
> boxplot(df$age, horizontal = TRUE, xlabel="Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
>

```

Environment History Connections Tutorial

Data

- df 2216 obs. of 30 variables
- df1 2216 obs. of 6 variables

Values

Plots Packages Help Viewer

Boxplot of Age

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 6:22 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

RStudio

CAPSTONE PROJECT.R* marketing_campaign df df

```

243
244 boxplot(df$age, horizontal = TRUE, xlabel="Age")
245
246 boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
247
248 # Boxplot for income
249
250 boxplot(df$income)
251
252
253
254
252:1 (Top Level) R Script

```

Console Terminal Jobs

```

R 4.0.2 - ~/ ...
[967] 70 64 59 45 46 74 50 46 55 48 68 70 52 43 35 70 29 46 67 51 68
[988] 68 37 43 37 70 70 71 72 56 61 45 48 46
[ reached getOption("max.print") -- omitted 1216 entries ]
> boxplot(df$age)
> boxplot(df$age, horizontal = TRUE)
> boxplot(df$age, horizontal = TRUE, xlabel="Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
> boxplot(df$income)
>

```

Environment History Connections Tutorial

Data

- df 2216 obs. of 30 variables
- df1 2216 obs. of 6 variables

Values

Plots Packages Help Viewer

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 6:23 PM

ims.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Community Notes Help

Learning Track Certificate

R studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

Source on Save Go to file/function Addins

```
243 boxplot(df$age, horizontal = TRUE, xlabel="Age")
244
245 boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
246
247 # Boxplot for income
248
249 boxplot(df$income)
250
251 boxplot(df$income, horizontal = TRUE)
252
253
254
```

253.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

```
[988] 68 37 43 37 70 70 71 72 56 61 45 48 46
[ reached getOption("max.print") -- omitted 1216 entries ]
> boxplot(df$age)
> boxplot(df$age, horizontal = TRUE)
> boxplot(df$age, horizontal = TRUE, xlabel="Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
> boxplot(df$income)
> boxplot(df$income, horizontal = TRUE)
```

Environment History Connections Tutorial

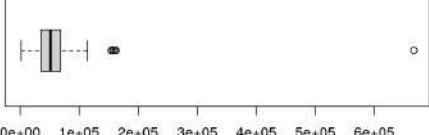
Import Dataset - 241 MiB -

R - Global Environment -

Data df 2216 obs. of 30 variables df1 2216 obs. of 6 variables

Values Files Plots Packages Help Viewer

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05



Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 6:25 PM

ims.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Community Notes Help

Learning Track Certificate

R studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

Source on Save Go to file/function Addins

```
246 boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
247
248 # Boxplot for income
249
250 boxplot(df$income)
251
252 boxplot(df$income, horizontal = TRUE)
253
254 boxplot(df$income, horizontal = TRUE, main="Boxplot of Income")
255
256 boxplot(df$income, horizontal = TRUE, xlabel="Income")
257
```

256.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

```
> boxplot(df$age)
> boxplot(df$age, horizontal = TRUE)
> boxplot(df$age, horizontal = TRUE, xlabel="Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
> boxplot(df$age, horizontal = TRUE, main="Boxplot of Age")
> boxplot(df$income)
> boxplot(df$income, horizontal = TRUE)
> boxplot(df$income, horizontal = TRUE, xlabel="Income")
> boxplot(df$income, horizontal = TRUE, main="Boxplot of Income")
>
```

Environment History Connections Tutorial

Import Dataset - 245 MiB -

R - Global Environment -

Data df 2216 obs. of 30 variables df1 2216 obs. of 6 variables

Values Files Plots Packages Help Viewer

Boxplot of Income

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05



Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 6:32 PM

ims.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

CAPSTONE PROJECT.R* marketing_campaign df df

```
253 boxplot(df$Income, horizontal = TRUE, main="Boxplot of Income")
254
255 # create a function to find out outliers
256 # quartiles
257
258 quantile(df$age,0.25)
259
260 # 3rd quartiles
261 quantile(df$age,0.75)
262
263
264
265:1 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/ ...
> boxplot(df$Income, horizontal = TRUE, main="Boxplot of Income")
> boxplot(df$Income, horizontal = TRUE, main="Boxplot of Income")
> quantile(df$age,0.25)
25%
47
> # 3rd quartiles
> quantile(df$age,0.75)
75%
65
>
```

Environment History Connections Tutorial

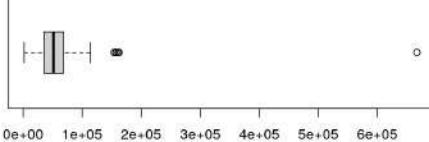
R - Global Environment

Data df 2216 obs. of 30 variables df1 2216 obs. of 6 variables

Values

Files Plots Packages Help Viewer

Boxplot of Income



Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 7:01 PM

ims.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

CAPSTONE PROJECT.R* marketing_campaign df df

```
268
269 def_lim <- function(x){
270   q1 <- quantile(x,0.25)
271   q3 <- quantile(x,0.75)
272   iqr<- IQR(x)
273   ll <- q1-1.5*iqr
274   ul <- q3+1.5*iqr
275   return(c(ll,ul))
276 }
277 def_lim(df$age)
278
279
280:1 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/ ...
+ q1 <- quantile(x,0.25)
+ q3 <- quantile(x,0.75)
+ iqr<- IQR(x)
+ ll <- q1-1.5*iqr
+ ul <- q3+1.5*iqr
+ return(c(ll,ul))
> def_lim(df$age)
25% 75%
20 92
>
```

Environment History Connections Tutorial

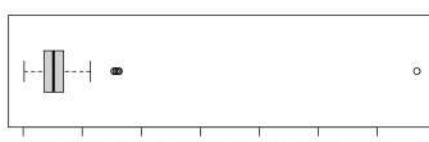
R - Global Environment

a "25/7/24"
d "data science"
y "25july2024"

Functions

Files Plots Packages Help Viewer

Boxplot of Income



Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 7:06 PM

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

RStudio

CAPSTONE PROJECT.R* marketing_campaign df df

```
278 # lower and upper threshold for the age
279 lim_age <- def_lim(df$age)
280
281 # lower and upper threshold for the income
282 lim_income <- def_lim(df$Income)
283
284
285 lim_income <- def_lim(df$Income)
286
287
288
289
286.1 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/ ...
+ iqr<- IQR(x)
+ ll <- q1-1.5*iqr
+ ul <- q3+1.5*iqr
+ return(c(ll,ul))
> def_lim(df$age)
25% 75%
20 92
> lim_age <- def_lim(df$age)
> lim_income <- def_lim(df$Income)
> |
```

Environment History Connections Tutorial

Import Dataset - 334 MiB -

R - Global Environment -

lim_age	Named num [1:2] 20 92
lim_income	Named num [1:2] -14526 118350
y	"25july2024"

Functions

Boxplot of Income

The boxplot displays the distribution of income. The x-axis represents income values from 0e+00 to 6e+05. The y-axis represents the count or frequency. The plot shows a median around 1e+05, with a significant spread and several outliers.

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 7:07 PM

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

RStudio

CAPSTONE PROJECT.R* marketing_campaign df df

```
278 # lower and upper threshold for the age
279 lim_age <- def_lim(df$age)
280
281 # lower and upper threshold for the income
282 lim_income <- def_lim(df$Income)
283
284 lim_income <- def_lim(df$Income)
285
286 lim_income
287
288
289
286.1 (Top Level) :
```

Console Terminal Jobs

```
R 4.0.2 - ~/ ...
+ return(c(ll,ul));
> def_lim(df$age)
25% 75%
20 92
> lim_age <- def_lim(df$age)
> lim_income <- def_lim(df$Income)
> lim_income
25% 75%
-14526.5 118350.5
> |
```

Environment History Connections Tutorial

Import Dataset - 334 MiB -

R - Global Environment -

lim_age	Named num [1:2] 20 92
lim_income	Named num [1:2] -14526 118350
y	"25july2024"

Functions

Boxplot of Income

The boxplot displays the distribution of income. The x-axis represents income values from 0e+00 to 6e+05. The y-axis represents the count or frequency. The plot shows a median around 1e+05, with a significant spread and several outliers.

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

Global Environment

lim_age Named num [1:2] 20 92
lim_income Named num [1:2] -14526 118358
y "25july2024"

Functions

Boxplot of Income

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

Global Environment

lim_age Named num [1:2] 20 92
lim_income Named num [1:2] -14526 118358
y "25july2024"

Functions

Boxplot of Income

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 7:25 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
292 # remove the outliers for the age
294
295 #age >lower limit
296
297 df$age >lim_age[1]
298
299 #age <upper limit
300 df$age > lim_age[2]
301
302
303
```

301.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
[940] TRUE TRUE
[987] TRUE TRUE
[ reached getOption("max.print") -- omitted 1216 entries ]
> #age <upper limit
> df$age > lim_age[2]
[1] FALSE FALSE
[15] FALSE FALSE
[29] FALSE FALSE
[43] FALSE FALSE
[57] FALSE FALSE
[71] FALSE FALSE
```

Environment History Connections Tutorial

Import Dataset 339 MiB

Global Environment

lim_age Named num [1:2] 20 92
lim_income Named num [1:2] -14526 118358
y "25july2024"

Functions

Boxplot of Income

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 7:43 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
301
302 # select those rows which satisfy both the conditions
303
304 TRUE & TRUE
305
306 # Use the &
307
308 df$age >lim_age[1]
309
310
311 df[(df$age>lim_age[1] & df$age<lim_age[2]),]
312
```

311.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
[868] TRUE TRUE
[885] TRUE TRUE
[902] TRUE TRUE
[919] TRUE TRUE
[936] TRUE TRUE
[953] TRUE TRUE
[970] TRUE TRUE
[987] TRUE TRUE
```

[reached getOption("max.print") -- omitted 1216 entries]

Environment History Connections Tutorial

Import Dataset 339 MiB

Global Environment

lim_age Named num [1:2] 20 92
lim_income Named num [1:2] -14526 118358
y "25july2024"

Functions

Boxplot of Income

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05

The screenshot shows a web-based RStudio interface. At the top, there's a navigation bar with tabs like 'File', 'Edit', 'View', 'History', 'Bookmarks', 'Profiles', 'Tab', 'Window', 'Help'. Below it, a URL bar shows 'lms.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs'. The main title 'IITR BA: Foundations of Business Analytics' is displayed, along with 'Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2'. On the right, there are buttons for 'Community', 'Notes', and 'Help'. A message at the top right says 'This Lab will get reset on 15th August 2024, 9:30 AM'. The interface includes a 'PG R Studio' tab, a 'rstudio' icon, and a 'Learning Track' sidebar. The central workspace has an 'Environment' tab open, showing variables like 'lim_age' and 'lim_income'. The 'Console' tab shows R code and its execution results, including a boxplot titled 'Boxplot of Income'.

Chrome File Edit View History Bookmarks Profiles Tab Window Help

lms.simplilearn.com/courses/6395/IITR-BA- Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 ▶ Start Lab ■ End Lab

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R: marketing_campaign df df

```
302 # select those rows which satisfy both the conditions
303
304 TRUE & TRUE
305
306 # Use the &
307
308 df$age > lim_age[1]
309
310 df$age > lim_age[1] & df$age > lim_age[2]
311
312 df[(df$age>lim_age[1] & df$age<lim_age[2]),]
313
```

(Top Level) R Script

Console Terminal Jobs

```
R 4.0.2 : ~
[ reached getOption("max.print") -- omitted 1216 entries ]
> df[(df$age>lim_age[1] & df$age<lim_age[2]),]
```

	Education	Income	Kidhome	Teenhome	Recency	Wines	Fruits	MeatProducts	FishProducts
1	Graduate	58138	0	0	58	635	88	546	172
2	Graduate	46344	1	1	38	11	1	6	2
3	Graduate	71613	0	0	26	426	49	127	111
4	Graduate	26646	1	0	26	11	4	20	10
5	Graduate	58293	1	0	94	173	43	118	46
6	Graduate	62513	0	1	16	520	42	98	0
7	Graduate	55635	0	1	34	235	65	164	50
8	Graduate	33454	1	0	32	76	10	56	3

Environment History Connections Tutorial

Import Dataset - 339 MB

Global Environment

lim_age	Named num [1:2] 20 92
lim_income	Named num [1:2] -14526 118350
y	"25July2024"

Functions

Files Plots Packages Help Viewer

Boxplot of Income

The boxplot displays the distribution of income. The x-axis represents income values from 0e+00 to 6e+05. The y-axis represents the count or frequency. The plot shows a median income around 1e+05, with a significant spread and several outliers at higher income levels.

Chrome File Edit View History Bookmarks Profiles Tab Window Help

ims.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 ▶ Start Lab ■ End Lab

Learning Track Certificate

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

302 # select those rows which satisfy both the conditions
303 TRUE & TRUE
305
306 # Use the &
307
308 df\$age > lim_age[1]
309
310 df\$age > lim_age[1] & df\$age > lim_age[2]
311
312 df[(df\$age>lim_age[1] & df\$age<lim_age[2]),]
313

313.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

	0	0	0	59	776	Partner	0	2	2
3	0	0	0	40	53	Partner	1	2	3
4	0	0	0	43	422	Partner	1	2	3
5	0	0	0	57	716	Partner	1	2	3
6	0	0	0	53	590	Alone	1	1	2
7	0	0	0	39	169	Partner	1	2	3
8	0	0	1	50	46	Partner	1	2	3
9	0	0	0	74	49	Partner	2	2	4
10	0	0	0	48	61	Partner	0	2	2
11	0	0	0	65	1102	Alone	0	1	1
12	0	0	0	72	310	Alone	2	1	3

Boxplot of Income

The boxplot displays the distribution of income. The x-axis represents income values ranging from 0e+00 to 6e+05. The y-axis represents the count of observations. The plot shows a median around 1e+05, with a significant spread and several outliers.

Chrome File Edit View History Bookmarks Profiles Tab Window Help

ims.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 ▶ Start Lab ■ End Lab

Learning Track Certificate

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

302 # select those rows which satisfy both the conditions
303 TRUE & TRUE
305
306 # Use the &
307
308 df\$age > lim_age[1]
309
310 df\$age > lim_age[1] & df\$age > lim_age[2]
311
312 df[(df\$age>lim_age[1] & df\$age<lim_age[2]),]
313

313.1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

	1	1	0	0	1	1	0
27	1	1	0	0	1	1	0
28	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0
33	0	0	0	0	0	0	0
34	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0

[reached 'max' / getOption("max.print") -- omitted 2180 rows]

Boxplot of Income

The boxplot displays the distribution of income. The x-axis represents income values ranging from 0e+00 to 6e+05. The y-axis represents the count of observations. The plot shows a median around 1e+05, with a significant spread and several outliers.

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Boxplot of Income

Unable to connect to Vocareum notification service, please refresh the page to try again

```

311 # df with outliers removal
312
313 df_1 <- df[(df$age>lim_age[1] & df$age<lim_age[2]),]
315
316 # delete the rows with outliers for income
317
318 df$Income > lim_income[1] & df$Income < lim_income[2]
319
320
321
322
319.1 (Top Level) R Script
  
```

Console Terminal Jobs

R 4.0.2 - /

```

[ reached 'max' / getOption("max.print") -- omitted 2180 rows ]
> df_1 <- df[(df$age>lim_age[1] & df$age<lim_age[2]),]
> df_1$Income > lim_income[1] & df_1$Income < lim_income[2]
 [1] TRUE TRUE
 [15] TRUE TRUE
 [29] TRUE TRUE
 [43] TRUE TRUE
 [57] TRUE TRUE
 [71] TRUE TRUE
 [85] TRUE TRUE
  
```

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Boxplot of Income

Low Battery Your Mac will sleep soon unless plugged into a power outlet.

```

315
316 # delete the rows with outliers for income
317
318 df$Income > lim_income[1] & df$Income < lim_income[2]
319
320 df_1$Income > lim_income[1] & df_1$Income < lim_income[2]
321
322 df_1[df_1$Income > lim_income[1] & df_1$Income < lim_income[2],]
323
324
325
326
323.1 (Top Level) R Script
  
```

Console Terminal Jobs

R 4.0.2 - /

```

[ reached getOption("max.print") -- omitted 1216 entries ]
> df_1$Income > lim_income[1] & df_1$Income < lim_income[2]
 [1] TRUE TRUE
 [15] TRUE TRUE
 [29] TRUE TRUE
 [43] TRUE TRUE
 [57] TRUE TRUE
 [71] TRUE TRUE
 [85] TRUE TRUE
 [99] TRUE TRUE
  
```

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

CAPSTONE PROJECT.R* marketing_campaign df df

Source on Save

315
316 # delete the rows with outliers for income
317
318 df\$Income > lim_income[1] & df\$Income < lim_income[2]
319
320 df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2]
321
322 df_1[df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2],]
323
324
325
326

318:54 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

```
[995] TRUE TRUE TRUE TRUE TRUE
[ reached getOption("max.print") -- omitted 1213 entries ]
> df_1[df_1$Income > lim_income[1] & df_1$Income < lim_income[2],]
  Education Income Kidhome Teenhome Recency Wines Fruits MeatProducts FishProducts
1 Graduate 58138 0 0 58 635 88 546 172
2 Graduate 46344 1 1 38 11 1 6 2
3 Graduate 71613 0 0 26 426 49 127 111
4 Graduate 26646 1 0 26 11 4 20 10
5 Graduate 58293 1 0 94 173 43 118 46
6 Graduate 62513 0 1 16 520 42 98 0
7 Graduate 55635 0 1 34 235 65 164 50
```

Environment History Connections Tutorial

Import Dataset 345 MB

R Global Environment

df_1 2213 obs. of 30 variables
df1 2216 obs. of 6 variables

Values a "25/7/24"

Boxplot of Income

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

CAPSTONE PROJECT.R* marketing_campaign df df

Source on Save

315
316 # delete the rows with outliers for income
317
318 df\$Income > lim_income[1] & df\$Income < lim_income[2]
319
320 df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2]
321
322 df_1[df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2],]
323
324
325
326

320:58 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

```
34 22 43 6 4 1
35 68 22 1 3 5
NumStorePurchases NumWebVisitsMonth AcceptedCmp3 AcceptedCmp4 AcceptedCmp5 AcceptedCmp1
1 4 7 0 0 0 0
2 2 5 0 0 0 0
3 10 4 0 0 0 0
4 4 6 0 0 0 0
5 6 5 0 0 0 0
6 10 6 0 0 0 0
7 7 6 0 0 0 0
```

Environment History Connections Tutorial

Import Dataset 345 MB

R Global Environment

df_1 2213 obs. of 30 variables
df1 2216 obs. of 6 variables

Values a "25/7/24"

Boxplot of Income

0e+00 1e+05 2e+05 3e+05 4e+05 5e+05 6e+05

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

Low Battery Your Mac will sleep soon unless plugged into a power outlet.

R Studio

CAPSTONE PROJECT.R*

marketing_campaign

df

df

315
316 # delete the rows with outliers for income
317 |
318 df\$Income > lim_income[1] & df\$Income < lim_income[2]
319
320 df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2]
321
322 df_1[df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2],]
323
324
325
326

317:1 (Top Level) :

Console Terminal Jobs

R 4.0.2 - ~/

AcceptedCmp2 Complain Response age amt_spent Living_With Children no_adults Family_Size

	0	0	1	67	1617	Alone	0	1	1
1	0	0	0	70	27	Alone	2	1	3
2	0	0	0	59	776	Partner	0	2	2
3	0	0	0	40	53	Partner	1	2	3
4	0	0	0	43	422	Partner	1	2	3
5	0	0	0	57	716	Partner	1	2	3
6	0	0	0	53	598	Alone	1	1	2
7	0	0	0	39	169	Partner	1	2	3
8	a	a	1	ea	4c	Partner	1	2	3

Boxplot of Income

The boxplot displays the distribution of income. The x-axis ranges from 0e+00 to 6e+05. The plot shows a median around 1e+05, with a box spanning approximately 1e+05 to 2e+05. There are several outliers represented by individual points above the upper whisker, with one notable outlier at approximately 6e+05.

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

Low Battery Your Mac will sleep soon unless plugged into a power outlet.

R Studio

CAPSTONE PROJECT.R*

marketing_campaign

df

df

315
316 # delete the rows with outliers for income
317 |
318 df\$Income > lim_income[1] & df\$Income < lim_income[2]
319
320 df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2]
321
322 df_1[df_1\$Income > lim_income[1] & df_1\$Income < lim_income[2],]
323
324
325
326

317:1 (Top Level) :

Console Terminal Jobs

R 4.0.2 - ~/

AcceptedCmp2 Complain Response age amt_spent Living_With Children no_adults Family_Size

	0	0	1	67	1617	Alone	0	1	1
1	0	0	0	70	27	Alone	2	1	3
2	0	0	0	59	776	Partner	0	2	2
3	0	0	0	40	53	Partner	1	2	3
4	0	0	0	43	422	Partner	1	2	3
5	0	0	0	57	716	Partner	1	2	3
6	0	0	0	53	598	Alone	1	1	2
7	0	0	0	39	169	Partner	1	2	3
8	a	a	1	ea	4c	Partner	1	2	3

Boxplot of Income

The boxplot displays the distribution of income. The x-axis ranges from 0e+00 to 6e+05. The plot shows a median around 1e+05, with a box spanning approximately 1e+05 to 2e+05. There are several outliers represented by individual points above the upper whisker, with one notable outlier at approximately 6e+05.

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 8:39 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

R Script

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
323
324 #question 15.find the correlation between variables create heatmap to visualize the corre
325
326 # descriptive stats of the data set
327
328 summary(df_1)
329
330
331
332
333
334
```

329:1 (Top Level) :

Console Terminal Jobs

R 4.0.2 - ~/

[reached 'max' /getOption("max.print") -- omitted 2172 rows]

```
> summary(df_1)
```

	Education	Income	Kidhome	Teenhome	Recency
Length:	2213	Min. : 1730	Min. :0.0000	Min. :0.0000	Min. : 0.00
Class :	character	1st Qu.: 35246	1st Qu.:0.0000	1st Qu.:0.0000	1st Qu.:24.00
Mode :	character	Median : 51373	Median :0.0000	Median :0.0000	Median :49.00
	Mean : 52237	Mean :0.4419	Mean :0.5056	Mean :49.01	
	3rd Qu.: 68487	3rd Qu.:1.0000	3rd Qu.:1.0000	3rd Qu.:74.00	
	Max. :666666	Max. :2.0000	Max. :2.0000	Max. :99.00	
Wines	Fruits	MeatProducts	FishProducts	SweetProducts	

R Script

Environment History Connections Tutorial

R - Global Environment

df_1 2213 obs. of 30 variables
df1 2216 obs. of 6 variables
Values a "25/7/24"

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Chrome File Edit View History Bookmarks Profiles Tab Window Help Tue 6 Aug 8:40 PM

ims.simplilearn.com/courses/6395/IITR-BA:-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024 Start Lab End Lab

rstudio

R Script

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R* marketing_campaign df df

```
323
324 #question 15.find the correlation between variables create heatmap to visualize the corre
325
326 # descriptive stats of the data set
327
328 summary(df_1)
329
330
331
332
333
334
```

329:1 (Top Level) :

Console Terminal Jobs

R 4.0.2 - ~/

```
3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:65.00 3rd Qu.:1048
Max. :1.00000 Max. :1.00000 Max. :1.0000 Max. :84.00 Max. :2525
Living_With Children no_adults Family_Size Is_Parent
Length:2213 Min. :0.0000 Min. :1.0000 Min. :1.0000 Min. :0.0000
Class :character 1st Qu.:0.0000 1st Qu.:1.0000 1st Qu.:2.0000 1st Qu.:0.0000
Mode :character Median :1.0000 Median :2.0000 Median :3.0000 Median :1.0000
Mean :0.9476 Mean :1.646 Mean :2.593 Mean :0.7144
3rd Qu.:1.0000 3rd Qu.:2.0000 3rd Qu.:3.0000 3rd Qu.:1.0000
Max. :3.0000 Max. :2.0000 Max. :5.000 Max. :1.0000
```

R Script

Environment History Connections Tutorial

R - Global Environment

df_1 2213 obs. of 30 variables
df1 2216 obs. of 6 variables
Values a "25/7/24"

Files Plots Packages Help Viewer

Install Update

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Chrome File Edit View History Bookmarks Profiles Tab Window Help

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics/practice-labs

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

R studio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

CAPSTONE PROJECT.R* marketing_campaign df df

Source on Save Run Source

323
324 #question 15.find the correlation between variables create heatmap to visualize the correle
325
326 # descriptive stats of the data set
327
328 summary(df_1)
329
330 str(df_1)
331
332 # remove the chr variables
333
334

332:27 (Top Level) <--

R Script

Console Terminal Jobs

R 4.0.2 - ~/

Mean :1.0000 Mean :2.000 Mean :3.000 Mean :1.0000
3rd Qu.:1.0000 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:1.0000
Max. :3.0000 Max. :2.000 Max. :5.000 Max. :1.0000

> str(df_1)
'data.frame': 2213 obs. of 30 variables:
 \$ Education : chr "Graduate" "Graduate" "Graduate" "Graduate" ...
 \$ Income : int 58138 46344 71613 26646 58293 62513 55635 33454 30351 5648 ...
 \$ Kidhome : int 0 1 0 1 1 0 0 1 1 1 ...
 \$ Teenhome : int 0 1 0 0 0 1 1 0 0 1 ...
 \$ Recency : int 58 38 26 26 94 16 34 32 19 68 ...
 \$ WInes : int 635 11 426 11 173 520 235 76 14 28 ...

Environment History Connections Tutorial

Import Dataset 362 MB

Global Environment

df_1 2213 obs. of 30 variables
df1 2216 obs. of 6 variables
Values a "25/7/24"

Files Plots Packages Help Viewer

Install Update

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in RStudio	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(P,d,q)	1.5-3
ifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Chrome File Edit View History Bookmarks Profiles Tab Window Help

lms.simplilearn.com/courses/6395/IITR-BA-Foundations-of-Business-Analytics

To exit full screen, press (Fn) F

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

R studio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

CAPSTONE PROJECT.R* marketing_campaign df df

Source on Save Run Source

331
332 # remove the chr variables
333
334 which()
335
336 which(sapply(df_1,class)=="character")
337
338 drop_vars <- which(sapply(df_1,class)=="character")
339
340 drop_vars
341
342

341:1 (Top Level) <--

R Script

Console Terminal Jobs

R 4.0.2 - ~/

>
>
> which(sapply(df_1,class)=="character")
Education Living_With
1 26
> drop_vars <- which(sapply(df_1,class)=="character")
> drop_vars
Education Living_With
1 26
>

Environment History Connections Tutorial

Import Dataset 362 MB

Global Environment

Values a "25/7/24"
d "data science"
drop_vars Named int [1:2] 1 26

Files Plots Packages Help Viewer

Install Update

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in RStudio	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(P,d,q)	1.5-3
ifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
336 which(sapply(df_1,class)=="character")
337
338 drop_vars <- which(sapply(df_1,class)=="character")
339
340 drop_vars
341
342 df_1[,-drop_vars]
343
344
345
346
347
```

343.1 (Top Level) :

	Income	Kidhome	Teenhome	Recency	Wines	Fruits	MeatProducts	FishProducts	SweetProducts
1	58138	0	0	58	635	88	546	172	88
2	46344	1	1	38	11	1	6	2	1
3	71613	0	0	26	426	49	127	111	21
4	26646	1	0	26	11	4	20	10	3
5	58293	1	0	94	173	43	118	46	27
6	62513	0	1	16	528	42	98	0	42
7	55635	0	1	34	235	65	164	50	49
8	33454	1	0	32	76	18	56	3	1
9	70251	1	0	14	14	8	74	2	4

Environment History Connections Tutorial

R - Global Environment

Values

Name	Description
a	"25/7/24"
d	"data science"
drop_vars	Named int [1:2] 1 26

Files Plots Packages Help Viewer

Install Update

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

R Script

Console Terminal Jobs

R 4.0.2 - ~/

```
336 which(sapply(df_1,class)=="character")
337
338 drop_vars <- which(sapply(df_1,class)=="character")
339
340 drop_vars
341
342 df_1[,-drop_vars]
343
344
345
346
347
```

343.1 (Top Level) :

	GoldProls	NumDealsPurchases	NumWebPurchases	NumCatalogPurchases	NumStorePurchases
1	88	3	8	10	4
2	6	2	1	1	2
3	42	1	8	2	10
4	5	2	2	0	4
5	15	5	5	3	6
6	14	2	6	4	10
7	27	4	7	3	7
8	73	?	4	a	4

Environment History Connections Tutorial

R - Global Environment

Values

Name	Description
a	"25/7/24"
d	"data science"
drop_vars	Named int [1:2] 1 26

Files Plots Packages Help Viewer

Install Update

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q)	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

R - Global Environment

Values

a "25/7/24"
d "data science"
drop_vars Named int [1:2] 1 26

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Console Terminal Jobs

R 4.0.2 - ~/

```
336 which(sapply(df_1,class)=="character")
337 drop_vars <- which(sapply(df_1,class)=="character")
338 drop_vars
339 df_1[drop_vars]
340 df_1[drop_vars]
341 df_1[drop_vars]
342 df_1[drop_vars]
343
344
345
346
347
```

343.1 (Top Level) R Script

	NumWebVisitsMonth	AcceptedCmp3	AcceptedCmp4	AcceptedCmp5	AcceptedCmp1	AcceptedCmp2
1	7	0	0	0	0	0
2	5	0	0	0	0	0
3	4	0	0	0	0	0
4	6	0	0	0	0	0
5	5	0	0	0	0	0

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

rstudio

File Edit Code View Plots Session Build Debug Profile Tools Help

File Edit Code View Plots Session Build Debug Profile Tools Help

Environment History Connections Tutorial

R - Global Environment

Values

a "25/7/24"
d "data science"
drop_vars Named int [1:2] 1 26

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Console Terminal Jobs

R 4.0.2 - ~/

```
336 which(sapply(df_1,class)=="character")
337 drop_vars <- which(sapply(df_1,class)=="character")
338 drop_vars
339 df_1[drop_vars]
340 df_1[drop_vars]
341 df_1[drop_vars]
342 df_1[drop_vars]
343
344
345
346
347
```

343.1 (Top Level) R Script

	Complain	Response	age	amt_spent	Children	no_adults	Family_Size	Is_Parent
1	0	1	67	1617	0	1	1	0
2	0	0	70	27	2	1	3	1
3	0	0	59	776	0	2	2	0
4	0	0	40	53	1	2	3	1
5	0	0	43	422	1	2	3	1
6	0	0	57	716	1	2	3	1
7	0	0	53	590	1	1	2	1

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

Community Notes Help

R Studio

CAPSTONE PROJECT.R* marketing_campaign df df

```

336 which(sapply(df_1,class)=="character")
337 drop_vars <- which(sapply(df_1,class)=="character")
338 drop_vars
339 df_1[-drop_vars]
340 df_1
341 df_1[-drop_vars]
342 df_1<-df_1[-drop_vars]
343 head(df_1,5)
344 df_1<-df_1[-drop_vars]
345 head(df_1,5)
346 head(df_1,5)
347
  
```

343:1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

	0	0	59	1672	0	2	2	0
30	0	0	35	30	0	1	1	0
31	0	0	61	318	0	2	2	0
32	0	0	54	120	1	2	3	1
33	0	0	72	302	2	1	3	1
34	0	0	78	1196	0	2	2	0
35	0	0	73	65	2	1	3	1
36	0	0	54	913	1	2	3	1
37	0	0	54	913	1	2	3	1

[reached 'max' / getOption("max.print") -- omitted 2178 rows]

Environment History Connections Tutorial

R - Global Environment

Values

Name	Description
a	"25/7/24"
d	"data science"
drop_vars	Named int [1:2] 1 26

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Install Update

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

Community Notes Help

R Studio

CAPSTONE PROJECT.R* marketing_campaign df df

```

336 which(sapply(df_1,class)=="character")
337 drop_vars <- which(sapply(df_1,class)=="character")
338 drop_vars
339 df_1[-drop_vars]
340 df_1
341 df_1[-drop_vars]
342 df_1<-df_1[-drop_vars]
343 df_1<-df_1[-drop_vars]
344 head(df_1,5)
345 head(df_1,5)
346 head(df_1,5)
347
  
```

343:1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

	0	0	58	635	88	546	172	88
1	58138	0	1	38	11	1	6	2
2	46344	1	1	26	426	49	127	111
3	71613	0	0	26	26	11	20	10
4	26646	1	0	26	11	4	18	3
5	58293	1	0	94	173	43	118	46

Income Kidhome Teenhome Recency Wines Fruits MeatProducts FishProducts SweetProducts

GoldProds	NumDealsPurchases	NumWebPurchases	NumCatalogPurchases	NumStorePurchases
1	88	3	8	10
2	88	3	8	10
3	88	3	8	10
4	88	3	8	10
5	88	3	8	10

Environment History Connections Tutorial

R - Global Environment

Values

Name	Description
a	"25/7/24"

Files Plots Packages Help Viewer

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
<input checked="" type="checkbox"/> dplyr	A Grammar of Data Manipulation	1.1.4
<input type="checkbox"/> forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
<input type="checkbox"/> fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
<input type="checkbox"/> lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
<input checked="" type="checkbox"/> lubridate	Make Dealing with Dates a Little Easier	1.9.3
<input type="checkbox"/> manipulate	Interactive Plots for RStudio	1.0.1
<input type="checkbox"/> moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Install Update

Chrome File Edit View History Bookmarks Profiles Tab Window Help

Tue 6 Aug 9:02 PM

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

rstudio

R Script

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

CAPSTONE PROJECT.RV marketing_campaign df df

```
336 which(sapply(df_1,class)=="character")
337 drop_vars <- which(sapply(df_1,class)=="character")
338 drop_vars
340 df_1[,-drop_vars]
341 df_1 <- df_1[,-drop_vars]
342 head(df_1,5)
343
344 head(df_1,5)
```

342:17 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

	1	0	94	173	43	118	46	27
GoldProds	NumDealsPurchases	NumWebPurchases	NumCatalogPurchases	NumStorePurchases				
1	88	3	8	10				4
2	6	2	1	1				2
3	42	1	8	2				10
4	5	2	2	0				4
5	15	5	5	3				6

NumWebVisitsMonth AcceptedCmp3 AcceptedCmp4 AcceptedCmp5 AcceptedCmp1 AcceptedCmp2

	1	0	0	0	0	0	0
1	7	0	0	0	0	0	0
2	5	0	0	0	0	0	0
3	4	0	0	0	0	0	0

R 4.0.2 - ~/

3 4 0 0 0 0 0
4 6 0 0 0 0 0
5 5 0 0 0 0 0

Complain Response age amt_spent Children no_adults Family_Size Is_Parent

	0	1	67	1617	0	1	1	0
1	0	0	70	27	2	1	3	1
2	0	0	59	776	0	2	2	0
3	0	0	40	53	1	2	3	1
4	0	0	43	422	1	2	3	1

R 4.0.2 - ~/

3 4 0 0 0 0 0
4 6 0 0 0 0 0
5 5 0 0 0 0 0

Complain Response age amt_spent Children no_adults Family_Size Is_Parent

	0	1	67	1617	0	1	1	0
1	0	0	70	27	2	1	3	1
2	0	0	59	776	0	2	2	0
3	0	0	40	53	1	2	3	1
4	0	0	43	422	1	2	3	1

Environment History Connections Tutorial

Import Dataset 362 MB

Global Environment

df_1 2213 obs. of 28 variables

df 2216 obs. of 6 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

Install Update

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

Chrome File Edit View History Bookmarks Profiles Tab Window Help

Tue 6 Aug 9:02 PM

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

PG R Studio

This Lab will get reset on 15th August 2024, 9:30 AM

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

Learning Track Certificate

rstudio

R Script

File Edit Code View Plots Session Build Debug Profile Tools Help

Addins

CAPSTONE PROJECT.RV marketing_campaign df df

```
336 which(sapply(df_1,class)=="character")
337 drop_vars <- which(sapply(df_1,class)=="character")
338 drop_vars
340 df_1[,-drop_vars]
341 df_1 <- df_1[,-drop_vars]
342 head(df_1,5)
343
344 head(df_1,5)
```

342:17 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~/

	1	0	94	173	43	118	46	27
GoldProds	NumDealsPurchases	NumWebPurchases	NumCatalogPurchases	NumStorePurchases				
1	88	3	8	10				4
2	6	2	1	1				2
3	42	1	8	2				10
4	5	2	2	0				4
5	15	5	5	3				6

NumWebVisitsMonth AcceptedCmp3 AcceptedCmp4 AcceptedCmp5 AcceptedCmp1 AcceptedCmp2

	1	0	0	0	0	0	0
1	7	0	0	0	0	0	0
2	5	0	0	0	0	0	0
3	4	0	0	0	0	0	0

R 4.0.2 - ~/

3 4 0 0 0 0 0
4 6 0 0 0 0 0
5 5 0 0 0 0 0

Complain Response age amt_spent Children no_adults Family_Size Is_Parent

	0	1	67	1617	0	1	1	0
1	0	0	70	27	2	1	3	1
2	0	0	59	776	0	2	2	0
3	0	0	40	53	1	2	3	1
4	0	0	43	422	1	2	3	1

R 4.0.2 - ~/

3 4 0 0 0 0 0
4 6 0 0 0 0 0
5 5 0 0 0 0 0

Complain Response age amt_spent Children no_adults Family_Size Is_Parent

	0	1	67	1617	0	1	1	0
1	0	0	70	27	2	1	3	1
2	0	0	59	776	0	2	2	0
3	0	0	40	53	1	2	3	1
4	0	0	43	422	1	2	3	1

Environment History Connections Tutorial

Import Dataset 362 MB

Global Environment

df_1 2213 obs. of 28 variables

df 2216 obs. of 6 variables

Values a "25/7/24"

Files Plots Packages Help Viewer

Install Update

User Library

Name	Description	Version
caret	Classification and Regression Training	6.0-94
cli	Helpers for Developing Command Line Interfaces	3.6.2
desc	Manipulate DESCRIPTION Files	1.4.3
DescToolsAddins	Interactive Functions to be Used as Shortcuts in 'RStudio'	1.10
dplyr	A Grammar of Data Manipulation	1.1.4
forecast	Forecasting Functions for Time Series and Linear Models	8.22.0
fracdiff	Fractionally Differenced ARIMA aka ARFIMA(p,d,q) Models	1.5-3
lifecycle	Manage the Life Cycle of your Package Functions	1.0.4
lubridate	Make Dealing with Dates a Little Easier	1.9.3
manipulate	Interactive Plots for RStudio	1.0.1
moments	Moments, Cumulants, Skewness, Kurtosis and Related Tests	0.14.1

The screenshot shows the RStudio IDE interface. The top bar displays the title "PG R Studio". The left sidebar includes a "Learning Track" section and a "Certificate" section. The main workspace shows a script editor with the following R code:

```

head(df_1,5)
347
348 cor(df_1)
349
350
351
352
353
354
355
356
357

```

The console output at the bottom shows:

```

R 4.0.2 - /~/
5 0 0 43 422 1 2 3 1
> cor(df_1)
      Income Kidhome Teenhome Recency Wines
Income  1.00000000 -0.42823061  0.019285302 -0.0031106299  0.578481207
Kidhome -0.428230607  1.00000000 -0.039451501  0.0101956438 -0.497407277
Teenhome  0.019285302 -0.039451501  1.000000000  0.0174640376  0.004312308
Recency -0.003110630  0.01019564  0.0147640938  1.0000000000  0.016331704
Wines   0.578481207 -0.497407278  0.004312308  0.0163317040  1.000000000
Fruits  0.430248297 -0.37330487 -0.175736398 -0.0051293913  0.385891705
MeatProducts  0.584361425 -0.43919232 -0.260777836  0.021771277  0.568188849

```

The right sidebar includes sections for Environment, History, Connections, Tutorial, and Packages. The "Environment" section shows variables like "df_1" and "df1". The "Packages" section lists several packages, with "dplyr" checked as the selected package.

IITR BA: Foundations of Business Analytics

Class completed: 1 | 54% of Self-Learning Completed | Projects completed: 1/2

This Lab will get reset on 15th August 2024, 9:30 AM

PG R Studio

rstudio

Used 7.9 of 50 hours in Aug, 2024

Start Lab End Lab

R Studio

File Edit Code View Plots Session Build Debug Profile Tools Help

CAPSTONE PROJECT.R*

marketing_campaign

Source on Save Run Source

349 cor(df_1[1:5,])

350

351

352

353

354

355

356

357

358

359

360

360:1 (Top Level) R Script

Console Terminal Jobs

R 4.0.2 - ~

> cor(df_1[1:5,])

	Income	Kidhome	Teenhome	Recency	Wines	Fruits
Income	1.00000000	-0.68608424	-0.1944313	0.27687349	0.69071600	0.6712272
Kidhome	-0.68608424	1.00000000	0.4082483	0.20392518	-0.93198719	-0.8001524
Teenhome	-0.19443127	0.40824829	1.00000000	-0.20232693	-0.49082600	-0.5599080
Recency	0.27687349	0.20392518	-0.2023269	1.00000000	0.13019113	0.3904453
Wines	0.69071600	-0.93198719	-0.4908260	0.13019113	1.00000000	0.9605908
Fruits	0.67122718	-0.8001524	-0.5599080	0.39044528	0.96059075	1.0000000
MeatProducts	0.39451413	-0.71548378	-0.3984034	0.28712313	0.89491659	0.9198113
FishProducts	0.65437968	-0.92639169	-0.5123471	0.12293617	0.99868933	0.9609871

```
R 4.0.2 . ~/ ~
library("corrplot")

search()
# correlation matrix
cor_mat <- cor(df_2)

corrplot()

names(df_1)
# df_1$x <- NULL

corrplot(cor_mat,main="correlation plot")

df$Education

which(sapply(df_1,class)=="character")

as.factor(df_1$Education)

|
```

< Top Level: R Script >

Files Plots Packages Help Viewer

```
R 4.0.2 . ~/ ~
search()
# correlation matrix
cor_mat <- cor(df_2)

corrplot()

names(df_1)
# df_1$x <- NULL

corrplot(cor_mat,main="correlation plot")

df$Education

which(sapply(df_1,class)=="character")

df_1$Education <- as.factor(df_1$Education)

df$Living_With

|
```

< Top Level: R Script >

Files Plots Packages Help Viewer

7
8 search()
9 # correlation matrix
0 cor_mat <- cor(df_2)
1
2 corplot()
3
4 names(df_1)
5
6 df\$Education
7
8 which(sapply(df_1,class)=="character")
9
0 df_1\$Education <- as.factor(df_1\$Education)
1
2 df_1\$Living_With <- as.factor(df_1\$Living_With)
3
I

R 4.0.2 - /

```
[737] Partner Partner Partner Partner Alone Partner Partner Partner Partner Partner
[745] Partner Partner Partner Partner Partner Alone Partner Partner Partner Partner Partner
[753] Partner Partner Partner Partner Partner Partner Partner Partner Partner Partner
[761] Partner Partner Partner Alone Partner Alone Alone Alone
[769] Partner Alone Partner Partner Partner Partner Partner Alone Partner
[777] Alone Partner Partner Alone Partner Alone Alone Partner
[785] Alone Partner Partner Partner Alone Partner Alone Partner
[793] Partner Partner Partner Partner Alone Partner Alone Partner
[801] Partner Partner Partner Partner Alone Partner Partner Alone
[809] Partner Alone Partner Partner Alone Partner Alone Alone
[817] Partner Alone Alone Partner Partner Partner Partner Alone
[825] Alone Alone Alone Partner Partner Alone Partner Partner
[833] Partner Alone Alone Alone Partner Alone Partner Partner
[841] Partner Partner Alone Partner Partner Partner Partner Partner
[849] Partner Partner Partner Alone Partner Alone Alone Partner
[857] Partner Partner Alone Partner Alone Partner Partner Partner
[865] Partner Alone Partner Alone Partner Alone Partner Alone
[873] Partner Alone Partner Partner Alone Partner Alone Partner
[881] Alone Alone Partner Alone Partner Alone Partner
[889] Alone Partner Partner Partner Alone Alone Alone Partner
[897] Partner Alone Partner Alone Partner Alone Partner Partner
[905] Partner Alone Alone Partner Alone Alone Partner Partner
[913] Partner Partner Alone Partner Partner Alone Alone
[921] Partner Alone Partner Partner Partner Alone Alone
[929] Partner Alone Partner Alone Partner Partner Alone Partner
[937] Alone Partner Partner Alone Alone Partner Partner Partner
[945] Alone Partner Alone Alone Alone Partner Partner Partner
[953] Partner Partner Partner Partner Partner Partner Partner
[961] Alone Partner Partner Partner Partner Partner Partner
[969] Alone Alone Alone Partner Partner Partner Partner
[977] Alone Alone Partner Partner Alone Alone Alone
[985] Partner Partner Alone Alone Alone Partner Partner Partner
[993] Alone Partner Partner Partner Alone Partner Partner Partner
[ reached getOption("max.print") -- omitted 1216 entries ]
```

Levels: Alone Partner

R Script :

File History Connections Tutorial

File Plots Packages Help Viewer

cor_mat <- cor(df_2)
corplot()
names(df_1)
df_1\$x <- NULL
corplot(cor_mat,main="correlation plot")
df\$Education

which(sapply(df_1,class)=="character")
df_1\$Education <- as.factor(df_1\$Education)

df_1\$Living_With <- as.factor(df_1\$Living_With)

str(df_1)
I

R 4.0.2 - /

```
[745] Partner Partner Partner Partner Alone Partner Alone Alone
[753] Partner Partner Partner Partner Partner Alone Partner Partner Partner
[761] Partner Partner Partner Alone Partner Alone Alone
[769] Partner Alone Partner Partner Partner Partner Partner
[777] Alone Partner Partner Alone Partner Alone Alone
[785] Alone Partner Partner Partner Alone Partner Alone
[793] Partner Partner Partner Partner Alone Partner
[801] Partner Partner Partner Partner Alone Partner
[809] Partner Alone Partner Partner Alone Partner
[817] Partner Alone Alone Partner Partner Partner
[825] Alone Alone Alone Partner Partner Alone
[833] Partner Alone Alone Alone Partner Partner
[841] Partner Partner Alone Partner Partner Partner
[849] Partner Partner Partner Alone Partner Alone
[857] Partner Partner Alone Partner Alone
[865] Partner Alone Partner Alone Partner
[873] Partner Alone Partner Partner Alone
[881] Alone Alone Partner Alone Partner
[889] Alone Partner Partner Partner Alone
[897] Partner Alone Partner Alone Partner
[905] Partner Alone Alone Partner Alone
[913] Partner Partner Alone Partner Partner
[921] Partner Alone Partner Partner Partner
[929] Partner Alone Partner Alone Partner
[937] Alone Partner Partner Alone
[945] Alone Partner Alone Alone Partner
[953] Partner Partner Partner Partner Partner
[961] Alone Partner Partner Partner Partner
[969] Alone Alone Alone Partner Partner
[977] Alone Alone Partner Partner
[985] Partner Partner Alone Alone
[993] Alone Partner Partner Partner
[ reached getOption("max.print") -- omitted 1216 entries ]
```

Levels: Alone Partner

> str(df_1)

'data.frame': 2205 obs. of 30 variables

R Script :

File History Connections Tutorial

File Plots Packages Help Viewer

```

cor_mat <- cor(df_2)

corrplot()

names(df_1)

# df_1$x <- NULL

corrplot(cor_mat, main="correlation plot")

df$Education

which(sapply(df_1, class)=="character")

df_1$Education <- as.numeric(as.factor(df_1$Education))

df_1$Living_With <- as.numeric(as.factor(df_1$Living_With))

str(df_1)

# all the variables are numeric

# we have the data ready for clustering

# find the distance b/w the vars

summary(df_1)

(R Script) R 4.0.2 - ~/ ~/
0351 564B ...
$ Kidhome : int 0 1 0 1 1 0 0 1 1 1 ...
$ Teenhome : int 0 1 0 0 0 1 1 0 0 1 ...
$ Recency : int 58 38 26 26 94 16 34 32 19 68 ...
$ Wines : int 635 11 426 11 173 520 235 76 14 28 ...
$ Fruits : int 88 1 49 4 43 42 65 10 0 0 ...
$ MeatProducts : int 546 6 127 20 118 98 164 56 24 6 ...
$ FishProducts : int 172 2 111 10 46 0 58 3 3 1 ...
$ SweetProducts : int 88 1 21 3 27 42 49 1 3 1 ...
$ GoldProducts : int 88 6 42 5 15 14 27 23 2 13 ...
$ NumDealsPurchases : int 3 2 1 2 5 2 4 2 1 1 ...
$ NumWebPurchases : int 8 1 8 2 5 6 7 4 3 1 ...
$ NumCatalogPurchases: int 10 1 2 0 3 4 3 0 0 0 ...
$ NumStorePurchases : int 4 2 10 4 6 10 7 4 2 0 ...
$ NumWebVisitsMonth : int 7 5 4 6 5 6 6 8 9 20 ...
$ AcceptedCmp3 : int 0 0 0 0 0 0 0 0 0 1 ...
$ AcceptedCmp4 : int 0 0 0 0 0 0 0 0 0 0 ...
$ AcceptedCmp5 : int 0 0 0 0 0 0 0 0 0 0 ...
$ AcceptedCmp1 : int 0 0 0 0 0 0 0 0 0 0 ...
$ AcceptedCmp2 : int 0 0 0 0 0 0 0 0 0 0 ...
$ Complain : int 0 0 0 0 0 0 0 0 0 0 ...
$ Response : int 1 0 0 0 0 0 0 0 1 0 ...
$ age : num 67 70 59 49 43 57 53 39 50 74 ...
$ amt_spent : int 1617 27 776 53 422 716 590 169 46 49 ...
$ Living_With : Factor w/ 2 levels "Alone","Partner": 1 1 2 2 2 2 1 2
2 ...
$ Children : int 0 2 0 1 1 1 1 1 1 2 ...
$ no_adults : num 1 1 2 2 2 2 1 2 2 2 ...
$ Family_Size : num 1 3 2 3 3 3 2 3 3 4 ...
$ Is_Parent : num 0 1 0 1 1 1 1 1 1 1 ...
- attr(*, "na.action")= 'omit' Named int [1:24] 11 28 44 49 59 72 91 92 93 12
9 ...
..- attr(*, "names")= chr [1:24] "11" "28" "44" "49" ...
> df_1$Education <- as.numeric(as.factor(df_1$Education))
> df_1$Living_With <- as.numeric(as.factor(df_1$Living_With))

(Top Level) R Script Files Plots Packages Help Viewer
```

```

df$Education

which(sapply(df_1, class)=="character")

df_1$Education <- as.factor(df_1$Education)

df_1$Living_With <- as.factor(df_1$Living_With)

str(df_1)

# all the variables are numeric

# we have the data ready for clustering

# find the distance b/w the vars

summary(df_1)

(R Script) R 4.0.2 - ~/ ~/
$ Living_With : Factor w/ 2 levels "Alone","Partner": 1 1 2 2 2 2 1 2
2 ...
$ Children : int 0 2 0 1 1 1 1 1 1 2 ...
$ no_adults : num 1 1 2 2 2 2 1 2 2 2 ...
$ Family_Size : num 1 3 2 3 3 3 2 3 3 4 ...
$ Is_Parent : num 0 1 0 1 1 1 1 1 1 1 ...
- attr(*, "na.action")= 'omit' Named int [1:24] 11 28 44 49 59 72 91 92 93 12
9 ...
..- attr(*, "names")= chr [1:24] "11" "28" "44" "49" ...
> summary(df_1)
   Education      Income      Kidhome      Teenhome
Graduate :1953  Min.   :1730  Min.   :0.0000  Min.   :0.0000
Undergraduate:252  1st Qu.:35196  1st Qu.:0.0000  1st Qu.:0.0000
                           Median :51287  Median :0.0000  Median :0.0000
                           Mean   :51622  Mean   :0.4422  Mean   :0.5066
                           3rd Qu.:68281  3rd Qu.:1.0000  3rd Qu.:1.0000
                           Max.  :113734  Max.  :2.0000  Max.  :2.0000
   Recency      Wines      Fruits      MeatProducts
Min.   : 0.00  Min.   : 0.0  Min.   : 0.0  Min.   : 0.0
1st Qu.:24.00  1st Qu.: 2.0  1st Qu.: 2.0  1st Qu.: 16.0
Median :49.00  Median : 8.0  Median : 8.0  Median : 68.0
Mean   :149.01  Mean   :306.2  Mean   :26.4  Mean   :165.3
3rd Qu.:174.00 3rd Qu.:507.0 3rd Qu.:33.0 3rd Qu.:232.0
Max.  :199.00  Max.  :1493.0  Max.  :199.0  Max.  :1725.0
   FishProducts  SweetProducts  GoldProducts  NumDealsPurchases
Min.   : 0.00  Min.   : 0.00  Min.   : 0.00  Min.   : 0.000
1st Qu.: 3.00  1st Qu.: 1.00  1st Qu.: 9.00  1st Qu.: 1.000
Median :12.00  Median : 8.00  Median : 25.00  Median : 2.000
Mean   :37.76  Mean   :27.13  Mean   :44.06  Mean   :2.318
3rd Qu.:50.00  3rd Qu.:34.00  3rd Qu.:56.00  3rd Qu.: 3.000
Max.  :1259.00  Max.  :262.00  Max.  :321.00  Max.  :15.000
   NumWebPurchases  NumCatalogPurchases  NumStorePurchases  NumWebVisitsMonth
Min.   : 0.000  Min.   : 0.000  Min.   : 0.000  Min.   : 0.000
1st Qu.: 2.000  1st Qu.: 0.000  1st Qu.: 3.000  1st Qu.: 3.000
Median : 4.000  Median : 2.000  Median : 5.000  Median : 6.000
Mean   : 4.101  Mean   : 2.645  Mean   : 5.824  Mean   : 5.337
3rd Qu.: 6.000  3rd Qu.: 4.000  3rd Qu.: 8.000  3rd Qu.: 7.000
```

which(sapply(df_1, class) == "character")

```
df_1$Education <- as.numeric(as.factor(df_1$Education))

df_1$Living_With <- as.numeric(as.factor(df_1$Living_With))

str(df_1)

# all the variables are numeric

# we have the data ready for clustering

# find the distance b/w the vars

summary(df_1)

str(df_1)

scale(df_1)
```

I

(Top Level) R Script

R 4.0.2 . ~/

	SweetProducts	GoldProds	NumDealsPurchases
2	7.12834e+01	4.405714e+01	2.318367e+00
NumWebPurchases	2.645351e+00	5.823583e+00	NumStorePurchases
NumWebVisitsMonth	AcceptedCmp3	AcceptedCmp4	
5.336961e+00	7.392290e-02	7.437642e-02	
AcceptedCmp5	AcceptedCmp1	AcceptedCmp2	
7.301587e-02	6.439909e-02	1.360544e-02	
Complain	Response	age	
9.070295e-03	1.510204e-01	5.509569e+01	
amt_spent	Living_With	Children	
6.068218e+02	1.644898e+00	9.487528e-01	
no_adults	Family_Size	Is_Parent	
1.644898e+00	2.593651e+00	7.151927e-01	
attr(",scaled:scale")			
Education	Income	Kidhome	
3.182301e-01	2.071306e+04	5.371319e-01	
Teenhome	Recency	Wines	
5.443801e-01	2.893211e+01	3.374938e+02	
Fruits	MeatProducts	FishProducts	
3.978448e+01	2.177845e+02	5.482463e+01	
SweetProducts	GoldProds	NumDealsPurchases	
4.113047e+01	5.173621e+01	1.886107e+00	
NumWebPurchases	NumCatalogPurchases	NumStorePurchases	
2.413535e+00	2.798647e+00	3.241796e+00	
NumWebVisitsMonth	AcceptedCmp3	AcceptedCmp4	
2.413535e+00	2.617047e-01	2.624420e-01	
AcceptedCmp5	AcceptedCmp1	AcceptedCmp2	
2.602216e-01	2.455182e-01	1.158724e-01	
Complain	Response	age	
9.482670e-02	3.581500e-01	1.170580e+01	
amt_spent	Living_With	Children	
6.016753e+02	4.786528e-01	7.492310e-01	
no_adults	Family_Size	Is_Parent	
4.786528e-01	9.061974e-01	4.514250e-01	

File Plots Packages Help Viewer

```
df_1$Living_With <- as.numeric(as.factor(df_1$Living_With))

str(df_1)

# all the variables are numeric

# we have the data ready for clustering

# find the distance b/w the vars

summary(df_1)

str(df_1)

df_sc <- scale(df_1)

kmeans(scaled data, centers= no of clusters, iter.max=max no of iterations)
eg <- kmeans(df_sc,3,iter.max=10)

eg
```

I

(Top Level) R Script

R 4.0.2 . ~/

	Complain	Response	age	
1	9.482670e-02	3.581500e-01	1.170580e+01	
2	6.016753e+02	4.786528e-01	7.492310e-01	
3	4.786528e-01	9.061974e-01	4.514250e-01	
> df_sc <- scale(df_1)				
> eg <- kmeans(df_sc,3,iter.max=10)				
> eg				
K-means clustering with 3 clusters of sizes 1055, 639, 511				
Cluster means:				
1	Education	Income	Kidhome	Teenhome
1	0.12935457	-0.8130457	0.6485268	-0.0930491
2	-0.15750539	-0.4415471	0.84890164	-0.022652777
3	-0.07010398	1.1642805	-0.86944378	0.009900028
1	Fruits	MeatProducts	FishProducts	SweetProducts
1	-0.53533334	-0.64506209	-0.54786213	-0.5323115
2	0.1128684	-0.01984259	0.04877182	0.1068392
3	0.9640975	1.35659476	1.06846449	0.9653988
1	NumDealsPurchases	NumWebPurchases	NumCatalogPurchases	NumStorePurchases
1	-0.09642851	-0.6905229	-0.7376096	-0.7844459
2	0.69079651	0.8687705	0.2995078	0.6781198
3	-0.66474930	0.3392511	1.1483221	0.7715693
1	NumWebVisitsMonth	AcceptedCmp3	AcceptedCmp4	AcceptedCmp5
1	0.48050715	-0.007207277	-0.2400607	0.2805912
2	0.06268873	-0.037294504	0.2174918	-0.2144383
3	-1.07043666	0.061506977	0.2236532	0.8474555
1	AcceptedCmp2	Complain	Response	age
1	-0.101056880	0.07429814	-0.1596581	-0.22008407
2	-0.009371343	-0.01313521	-0.1245404	0.32912444
3	0.220358700	-0.033748001	0.4853632	0.04281444
1	Children	no_adults	Family_Size	Is_Parent
1	0.3973309	0.01313452	0.3354452	0.3579432
2	0.3002493	0.07163495	0.2860793	0.5823737
3	-1.1957797	-0.11669600	-1.0502922	-1.4672542

File Plots Packages Help Viewer

```

Source on Save | Run | Source | R 4.0.2 . ~/ ~

str(df_1)
# all the variables are numeric
# we have the data ready for clustering
# find the distance b/w the vars
summary(df_1)
str(df_1)
df_sc <- scale(df_1)
kmeans(scaled data, centers= no of clusters, iter.max=max no of iterations)
eg <- kmeans(df_sc,3,iter.max=10)
eg
# means of the clusters
eg$centers
I

(Top Level) R Script Files Plots Packages Help Viewer

```

Within cluster sum of squares by cluster:
[1] 15723.88 14003.28 16383.22
(between_SS / total_SS = 30.3 %)

Available components:

```

[1] "cluster"      "centers"       "totss"        "withinss"
[5] "tot.withinss" "betweenss"     "size"         "iter"
[9] "ifault"
> eg$centers
   Education Income Kidhome Teenhome Recency Wines
1 0.12935457 -0.8130457 0.6485268 -0.09304491 0.008925318 -0.7714957
2 -0.15750539 0.4112925 -0.4415471 0.94890164 -0.021652777 0.5211099
3 -0.07010398 1.1642805 -0.7867851 -0.86944378 0.009900028 0.9411718
  Fruits MeatProducts FishProducts SweetProducts GoldProds
1 -0.5353334 -0.64506209 -0.54706213 -0.5323115 -0.5421833
2 0.1128684 -0.01984259 -0.04877182 0.1068392 0.4174749
3 0.9640975 1.35659476 1.06846449 0.9653988 0.5973325
  NumDealsPurchases NumWebPurchases NumCatalogPurchases NumStorePurchases
1 -0.09642851 -0.6905229 -0.7376096 -0.7844459
2 0.69079651 0.8687705 0.2995078 0.6781198
3 -0.66474930 0.3392511 1.1483221 0.7715693
  NumWebVisitsMonth AcceptedCmp3 AcceptedCmp4 AcceptedCmp5 AcceptedCmp1
1 0.48050715 -0.007202727 -0.2400607 -0.2805912 -0.2584380
2 0.06268873 -0.037294504 0.2174918 -0.2144383 -0.1156955
3 -1.07043666 0.061506977 0.2236532 0.8474555 0.6782417
  AcceptedCmp2 Complain Response age amt_spent LivingWith
1 -0.101056880 0.02429814 -0.1596581 -0.22008407 -0.8344946 0.0131452
2 -0.009371343 -0.01313521 -0.1245404 0.32912444 0.3402287 0.07163495
3 0.220358700 -0.03374001 0.4853632 0.04281444 1.2974279 -0.11669600
  Children no_adults Family_Size Is_Parent
1 0.3973309 0.01312452 0.3354452 0.3579432
2 0.3062493 0.07163495 0.2860793 0.5823737
3 -1.1957797 -0.11669600 -1.0502922 -1.4672542

```

```

Source on Save | Run | Source | R 4.0.2 . ~/ ~

summary(df_1)
str(df_1)
df_sc <- scale(df_1)
kmeans(scaled data, centers= no of clusters, iter.max=max no of iterations)
eg <- kmeans(df_sc,3,iter.max=10)
eg
# means of the clusters
eg$centers
I

# which obs goes to which cluster
eg$cluster
table(eg$cluster)

eg$tot.withinss
I

(Top Level) R Script Files Plots Packages Help Viewer

```

1 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830
2 1 2 1 2 2 1 3 1 3 1 3 2 2 1 2 1 2 1 2 1 3
831 832 833 834 835 836 837 838 839 840 841 842 843 844 845
1 3 2 1 1 2 1 3 1 2 1 3 1 2 1 1 3
846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
3 3 2 1 1 2 2 3 1 2 1 2 2 1 2 1 3
861 862 863 864 865 866 867 868 869 870 871 872 873 874 875
1 1 1 1 1 2 2 1 2 1 3 2 1 1 1 2
876 877 878 879 880 881 882 883 884 885 886 887 888 889 890
3 2 1 3 1 1 1 1 1 1 3 1 1 1 2 3
891 892 893 894 895 896 897 898 899 900 901 902 903 904 905
1 2 3 1 3 2 2 3 1 1 1 3 1 1 1 2
906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
3 3 1 1 1 3 3 2 1 3 3 3 3 3 2 3
921 922 923 924 925 926 927 928 929 930 931 932 933 934 935
1 3 1 1 3 1 3 3 2 3 1 2 1 1 3
936 937 938 939 940 941 942 943 944 945 946 947 948 949 950
1 3 2 2 2 2 2 3 3 1 3 2 1 1 2
951 952 953 954 955 956 957 958 959 960 961 962 963 964 965
1 1 1 1 1 2 2 1 2 2 1 2 1 1 1
966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
2 3 1 1 2 3 1 1 2 3 3 3 1 1 2
981 982 983 984 985 986 987 988 989 990 991 992 993 994 995
1 1 1 2 3 2 3 3 1 3 1 2 2 1
996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010
1 3 1 2 2 2 3 2 1 2 2 2 3 1 1
1011 1012 1013 1014 1015 1016 1017 1018 1019 1020
3 1 1 1 1 3 3 1 1 1 1
[reached getOption("max.print") -- omitted 1205 entries]
> table(eg\$cluster)

1 2 3
1055 639 511
> eg\$tot.withinss
[1] 46110.38
>

Source on Save Run Source

```

str(df_1)
df_sc <- scale(df_1)

kmeans(scaled data, centers=3, iter.max=10)
eg <- kmeans(df_sc, 3, iter.max=10)
eg

# means of the clusters
eg$centers

# which obs goes to which cluster
eg$cluster

table(eg$cluster)

eg$tot.withinss

eg$totss

```

R 4.0.2 - ~/

```

831 832 833 834 835 836 837 838 839 840 841 842 843 844 845
1 3 2 1 1 2 1 3 1 2 1 2 1 1 1 3
846 847 848 849 850 851 852 853 854 855 856 857 858 859 860
3 3 2 1 1 2 2 3 1 2 2 1 2 1 1 3
861 862 863 864 865 866 867 868 869 870 871 872 873 874 875
1 1 1 1 2 2 1 2 1 3 2 1 1 2 1 2
876 877 878 879 880 881 882 883 884 885 886 887 888 889 890
3 2 1 3 1 1 1 1 1 3 3 1 1 2 3
891 892 893 894 895 896 897 898 899 900 901 902 903 904 905
1 2 3 1 3 2 2 3 1 1 3 1 3 1 2 2
906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
3 3 1 1 3 3 2 1 3 3 3 3 3 3 2 3
921 922 923 924 925 926 927 928 929 930 931 932 933 934 935
1 3 1 1 3 1 3 3 2 3 1 2 1 3 1 3
936 937 938 939 940 941 942 943 944 945 946 947 948 949 950
1 3 2 2 2 2 3 3 1 3 2 1 1 2 1 2
951 952 953 954 955 956 957 958 959 960 961 962 963 964 965
1 1 1 1 1 2 2 1 2 2 2 1 2 1 1 1
966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
2 3 1 1 2 3 1 1 2 3 3 3 3 1 1 2
981 982 983 984 985 986 987 988 989 990 991 992 993 994 995
1 1 1 2 3 2 3 3 1 3 1 2 2 2 1
996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010
1 3 1 2 2 3 2 1 2 2 2 2 3 1 1 1
1011 1012 1013 1014 1015 1016 1017 1018 1019 1020
3 1 1 1 3 3 1 1 1 1
[ reached getOption("max.print") -- omitted 1205 entries ]
> table(eg$cluster)

1 2 3
1055 639 511
> eg$tot.withinss
[1] 46110.38
> eg$totss
[1] 66120

```

R Script Files Plots Packages Help Viewer

Source on Save Run Source

```

eg <- kmeans(df_sc, 3, iter.max=10)
eg

# means of the clusters
eg$centers

# which obs goes to which cluster
eg$cluster

table(eg$cluster)

eg$tot.withinss

wss <- sapply(1:16, function(x) {kmeans(df_sc, x, iter.max=10)$tot.withinss})
wss

```

R 4.0.2 - ~/

```

1 1 1 1 1 2 2 1 2 1 3 2 1 1 2
876 877 878 879 880 881 882 883 884 885 886 887 888 889 890
3 2 1 3 1 1 1 1 1 3 3 1 1 2 3
891 892 893 894 895 896 897 898 899 900 901 902 903 904 905
1 2 3 1 3 2 2 3 1 1 3 1 3 1 2 2
906 907 908 909 910 911 912 913 914 915 916 917 918 919 920
3 3 1 1 1 3 3 2 1 3 3 3 3 3 2 3
921 922 923 924 925 926 927 928 929 930 931 932 933 934 935
1 3 1 1 3 1 3 3 2 3 1 2 1 3 1 3
936 937 938 939 940 941 942 943 944 945 946 947 948 949 950
1 3 2 2 2 2 3 3 1 3 2 1 1 2 1 2
951 952 953 954 955 956 957 958 959 960 961 962 963 964 965
1 1 1 1 1 2 2 1 2 2 2 1 2 1 1 1
966 967 968 969 970 971 972 973 974 975 976 977 978 979 980
2 3 1 1 2 3 1 1 2 3 3 3 3 1 1 2
981 982 983 984 985 986 987 988 989 990 991 992 993 994 995
1 1 1 2 3 2 3 3 1 3 1 2 2 2 1
996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010
1 3 1 2 2 3 2 1 2 2 2 2 3 1 1 1
1011 1012 1013 1014 1015 1016 1017 1018 1019 1020
3 1 1 1 3 3 1 1 1 1
[ reached getOption("max.print") -- omitted 1205 entries ]
> table(eg$cluster)

1 2 3
1055 639 511
> eg$tot.withinss
[1] 46110.38
> eg$totss
[1] 66120
> wss <- sapply(1:16, function(x) {kmeans(df_sc, x, iter.max=10)$tot.withinss})
> wss
[1] 66120.00 51048.12 46110.38 43685.03 42404.07 40367.74 38843.88
[8] 36657.55 35183.89 34392.28 35153.80 33258.76 33283.85 31485.12
[15] 31456.01 31397.64

```

R Script Files Plots Packages Help Viewer

