

R version 4.3.3 (2024-02-29) -- "Angel Food Cake"
 Copyright (C) 2024 The R Foundation for Statistical Computing
 Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
 You are welcome to redistribute it under certain conditions.
 Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
 Type 'contributors()' for more information and
 'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
 'help.start()' for an HTML browser interface to help.
 Type 'q()' to quit R.

```
> read.csv("hotel_bookings.csv")
  hotel is_canceled lead_time
1  Resort Hotel      0      342
2  Resort Hotel      0      737
3  Resort Hotel      0         7
4  Resort Hotel      0        13
5  Resort Hotel      0        14
6  Resort Hotel      0        14
7  Resort Hotel      0         0
8  Resort Hotel      0         9
9  Resort Hotel      1        85
10 Resort Hotel      1        75
11 Resort Hotel      1        23
12 Resort Hotel      0        35
13 Resort Hotel      0        68
14 Resort Hotel      0        18
15 Resort Hotel      0        37
16 Resort Hotel      0        68
17 Resort Hotel      0        37
18 Resort Hotel      0        12
19 Resort Hotel      0         0
20 Resort Hotel      0         7
21 Resort Hotel      0        37
22 Resort Hotel      0        72
23 Resort Hotel      0        72
24 Resort Hotel      0        72
25 Resort Hotel      0       127
26 Resort Hotel      0        78
27 Resort Hotel      0        48
28 Resort Hotel      1        60
29 Resort Hotel      0        77
30 Resort Hotel      0        99
31 Resort Hotel      0       118
  arrival_date_year arrival_date_month
1          2015          July
2          2015          July
3          2015          July
4          2015          July
5          2015          July
6          2015          July
7          2015          July
8          2015          July
9          2015          July
10         2015          July
11         2015          July
12         2015          July
13         2015          July
14         2015          July
```

15	2015	July
16	2015	July
17	2015	July
18	2015	July
19	2015	July
20	2015	July
21	2015	July
22	2015	July
23	2015	July
24	2015	July
25	2015	July
26	2015	July
27	2015	July
28	2015	July
29	2015	July
30	2015	July
31	2015	July

arrival_date_week_number

1	27
2	27
3	27
4	27
5	27
6	27
7	27
8	27
9	27
10	27
11	27
12	27
13	27
14	27
15	27
16	27
17	27
18	27
19	27
20	27
21	27
22	27
23	27
24	27
25	27
26	27
27	27
28	27
29	27
30	27
31	27

arrival_date_day_of_month

1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1

16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1

	stays_in_weekend_nights	stays_in_week_nights
1	0	0
2	0	0
3	0	1
4	0	1
5	0	2
6	0	2
7	0	2
8	0	2
9	0	3
10	0	3
11	0	4
12	0	4
13	0	4
14	0	4
15	0	4
16	0	4
17	0	4
18	0	1
19	0	1
20	0	4
21	1	4
22	2	4
23	2	4
24	2	4
25	2	5
26	2	5
27	2	5
28	2	5
29	2	5
30	2	5
31	4	10

	adults	children	babies	meal	country
1	2	0	0	BB	PRT
2	2	0	0	BB	PRT
3	1	0	0	BB	GBR
4	1	0	0	BB	GBR
5	2	0	0	BB	GBR
6	2	0	0	BB	GBR
7	2	0	0	BB	PRT
8	2	0	0	FB	PRT
9	2	0	0	BB	PRT
10	2	0	0	HB	PRT
11	2	0	0	BB	PRT
12	2	0	0	HB	PRT
13	2	0	0	BB	USA
14	2	1	0	HB	ESP
15	2	0	0	BB	PRT
16	2	0	0	BB	IRL

17	2	0	0	BB	PRT
18	2	0	0	BB	IRL
19	2	0	0	BB	FRA
20	2	0	0	BB	GBR
21	1	0	0	BB	GBR
22	2	0	0	BB	PRT
23	2	0	0	BB	PRT
24	2	0	0	BB	PRT
25	2	0	0	HB	GBR
26	2	0	0	BB	PRT
27	2	0	0	BB	IRL
28	2	0	0	BB	PRT
29	2	0	0	BB	PRT
30	2	0	0	BB	PRT
31	1	0	0	BB	NULL

	market_segment	distribution_channel
1	Direct	Direct
2	Direct	Direct
3	Direct	Direct
4	Corporate	Corporate
5	Online TA	TA/TO
6	Online TA	TA/TO
7	Direct	Direct
8	Direct	Direct
9	Online TA	TA/TO
10	Offline TA/TO	TA/TO
11	Online TA	TA/TO
12	Online TA	TA/TO
13	Online TA	TA/TO
14	Online TA	TA/TO
15	Online TA	TA/TO
16	Online TA	TA/TO
17	Offline TA/TO	TA/TO
18	Online TA	TA/TO
19	Corporate	Corporate
20	Direct	Direct
21	Online TA	TA/TO
22	Direct	Direct
23	Direct	Direct
24	Direct	Direct
25	Offline TA/TO	TA/TO
26	Offline TA/TO	TA/TO
27	Offline TA/TO	TA/TO
28	Online TA	TA/TO
29	Online TA	TA/TO
30	Online TA	TA/TO
31	Direct	Direct

	is_repeated_guest	previous_cancellations
1	0	0
2	0	0
3	0	0
4	0	0
5	0	0
6	0	0
7	0	0
8	0	0
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0
15	0	0
16	0	0
17	0	0

18	0	0
19	0	0
20	0	0
21	0	0
22	0	0
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
30	0	0
31	0	0

previous_bookings_not_canceled

1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	0
19	0
20	0
21	0
22	0
23	0
24	0
25	0
26	0
27	0
28	0
29	0
30	0
31	0

reserved_room_type assigned_room_type

1	C	C
2	C	C
3	A	C
4	A	A
5	A	A
6	A	A
7	C	C
8	C	C
9	A	A
10	D	D
11	E	E
12	D	D
13	D	E
14	G	G
15	E	E
16	D	E
17	E	E
18	A	E

19	A	G
20	G	G
21	F	F
22	A	A
23	A	A
24	D	D
25	D	I
26	D	D
27	D	D
28	E	E
29	A	A
30	D	D
31	A	A

	booking_changes	deposit_type	agent	company
1	3	No Deposit	NULL	NULL
2	4	No Deposit	NULL	NULL
3	0	No Deposit	NULL	NULL
4	0	No Deposit	304	NULL
5	0	No Deposit	240	NULL
6	0	No Deposit	240	NULL
7	0	No Deposit	NULL	NULL
8	0	No Deposit	303	NULL
9	0	No Deposit	240	NULL
10	0	No Deposit	15	NULL
11	0	No Deposit	240	NULL
12	0	No Deposit	240	NULL
13	0	No Deposit	240	NULL
14	1	No Deposit	241	NULL
15	0	No Deposit	241	NULL
16	0	No Deposit	240	NULL
17	0	No Deposit	8	NULL
18	0	No Deposit	240	NULL
19	0	No Deposit	NULL	110
20	0	No Deposit	250	NULL
21	0	No Deposit	241	NULL
22	1	No Deposit	250	NULL
23	1	No Deposit	250	NULL
24	1	No Deposit	250	NULL
25	0	No Deposit	115	NULL
26	0	No Deposit	5	NULL
27	0	No Deposit	8	NULL
28	0	No Deposit	240	NULL
29	0	No Deposit	240	NULL
30	0	No Deposit	240	NULL
31	2	No Deposit	NULL	NULL

	days_in_waiting_list	customer_type	adr
1	0	Transient	0.00
2	0	Transient	0.00
3	0	Transient	75.00
4	0	Transient	75.00
5	0	Transient	98.00
6	0	Transient	98.00
7	0	Transient	107.00
8	0	Transient	103.00
9	0	Transient	82.00
10	0	Transient	105.50
11	0	Transient	123.00
12	0	Transient	145.00
13	0	Transient	97.00
14	0	Transient	154.77
15	0	Transient	94.71
16	0	Transient	97.00
17	0	Contract	97.50
18	0	Transient	88.20
19	0	Transient	107.42

20	0	Transient	153.00
21	0	Transient	97.29
22	0	Transient	84.67
23	0	Transient	84.67
24	0	Transient	99.67
25	0	Contract	94.95
26	0	Transient	63.60
27	0	Contract	79.50
28	0	Transient	107.00
29	0	Transient	94.00
30	0	Transient	87.30
31	0	Transient	62.00

required_car_parking_spaces

1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	0
19	0
20	0
21	0
22	0
23	0
24	0
25	0
26	1
27	0
28	0
29	0
30	1
31	0

total_of_special_requests reservation_status

1	0	Check-Out
2	0	Check-Out
3	0	Check-Out
4	0	Check-Out
5	1	Check-Out
6	1	Check-Out
7	0	Check-Out
8	1	Check-Out
9	1	Canceled
10	0	Canceled
11	0	Canceled
12	0	Check-Out
13	3	Check-Out
14	1	Check-Out
15	0	Check-Out
16	3	Check-Out
17	0	Check-Out
18	0	Check-Out
19	0	Check-Out
20	1	Check-Out

21	1	Check-Out
22	1	Check-Out
23	1	Check-Out
24	1	Check-Out
25	1	Check-Out
26	0	Check-Out
27	0	Check-Out
28	2	Canceled
29	0	Check-Out
30	1	Check-Out
31	2	Check-Out

reservation_status_date

1	2015-07-01
2	2015-07-01
3	2015-07-02
4	2015-07-02
5	2015-07-03
6	2015-07-03
7	2015-07-03
8	2015-07-03
9	2015-05-06
10	2015-04-22
11	2015-06-23
12	2015-07-05
13	2015-07-05
14	2015-07-05
15	2015-07-05
16	2015-07-05
17	2015-07-05
18	2015-07-02
19	2015-07-02
20	2015-07-05
21	2015-07-06
22	2015-07-07
23	2015-07-07
24	2015-07-07
25	2015-07-01
26	2015-07-08
27	2015-07-08
28	2015-05-11
29	2015-07-08
30	2015-07-08
31	2015-07-15

[reached 'max' / getOption("max.print") -- omitted 119359 rows]

```
> hotel_bookings <- read.csv("hotel_bookings.csv")
```

```
> head(hotel_bookings)
```

	hotel	is_canceled	lead_time
1	Resort Hotel	0	342
2	Resort Hotel	0	737
3	Resort Hotel	0	7
4	Resort Hotel	0	13
5	Resort Hotel	0	14
6	Resort Hotel	0	14

	arrival_date_year	arrival_date_month
1	2015	July
2	2015	July
3	2015	July
4	2015	July
5	2015	July
6	2015	July

	arrival_date_week_number
1	27
2	27
3	27
4	27

5	27				
6	27				
	arrival_date_day_of_month				
1	1				
2	1				
3	1				
4	1				
5	1				
6	1				
	stays_in_weekend_nights	stays_in_week_nights			
1	0	0			
2	0	0			
3	0	1			
4	0	1			
5	0	2			
6	0	2			
	adults	children	babies	meal	country
1	2	0	0	BB	PRT
2	2	0	0	BB	PRT
3	1	0	0	BB	GBR
4	1	0	0	BB	GBR
5	2	0	0	BB	GBR
6	2	0	0	BB	GBR
	market_segment	distribution_channel			
1	Direct	Direct			
2	Direct	Direct			
3	Direct	Direct			
4	Corporate	Corporate			
5	Online TA	TA/TO			
6	Online TA	TA/TO			
	is_repeated_guest	previous_cancellations			
1	0	0			
2	0	0			
3	0	0			
4	0	0			
5	0	0			
6	0	0			
	previous_bookings_not_canceled				
1	0				
2	0				
3	0				
4	0				
5	0				
6	0				
	reserved_room_type	assigned_room_type			
1	C	C			
2	C	C			
3	A	C			
4	A	A			
5	A	A			
6	A	A			
	booking_changes	deposit_type	agent	company	
1	3	No Deposit	NULL	NULL	
2	4	No Deposit	NULL	NULL	
3	0	No Deposit	NULL	NULL	
4	0	No Deposit	304	NULL	
5	0	No Deposit	240	NULL	
6	0	No Deposit	240	NULL	
	days_in_waiting_list	customer_type	adr		
1	0	Transient	0		
2	0	Transient	0		
3	0	Transient	75		
4	0	Transient	75		
5	0	Transient	98		
6	0	Transient	98		

```

  required_car_parking_spaces
1          0
2          0
3          0
4          0
5          0
6          0
  total_of_special_requests reservation_status
1          0          Check-Out
2          0          Check-Out
3          0          Check-Out
4          0          Check-Out
5          1          Check-Out
6          1          Check-Out
  reservation_status_date
1      2015-07-01
2      2015-07-01
3      2015-07-02
4      2015-07-02
5      2015-07-03
6      2015-07-03

```

```

> install.packages('ggplot2')
Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.3'
(as 'lib' is unspecified)
also installing the dependency 'scales'

```

```

trying URL 'http://rspm/default/__linux__/focal/latest/src/contrib/scales_1.3.0.tar.gz'
Content type 'application/x-gzip' length 697062 bytes (680 KB)
=====
downloaded 680 KB

```

```

trying URL 'http://rspm/default/__linux__/focal/latest/src/contrib/ggplot2_3.5.1.tar.gz'
Content type 'application/x-gzip' length 4941778 bytes (4.7 MB)
=====
downloaded 4.7 MB

```

```

* installing *binary* package 'scales' ...
* DONE (scales)
* installing *binary* package 'ggplot2' ...
* DONE (ggplot2)

```

```

The downloaded source packages are in
  '/tmp/RtmpzBUTp0/downloaded_packages'

```

```

> library(ggplot2)
>
> ggplot(data = hotel_bookings) +
+ geom_point(mapping = aes(x = lead_time, y = children))
Warning message:
Removed 4 rows containing missing values or
values outside the scale range ('geom_point()').
>
> ggplot(data = hotel_bookings) +
+   geom_point(mapping = aes(x = 'stays_in_weekend_nights', y = 'children'))
Warning message:
In geom_point(mapping = aes(x = "stays_in_weekend_nights", y = "children")) :
  All aesthetics have length 1, but the data has
  119390 rows.
i Please consider using `annotate()` or provide
  this layer with data containing a single row.
> ggplot(data = hotel_bookings) +
+   geom_point(mapping = aes(x = stays_in_weekend_nights, y = children))
Warning message:
Removed 4 rows containing missing values or
values outside the scale range ('geom_point()').
> save.image("/cloud/project/hotel_bookings Project.RData")

```