How can one extract the year from a date in R? Can you provide some examples?

Authored by

stats writer

April 21, 2024

To extract the year from a date in R, one can use the "format" function with the "as.Date" function. This will convert the date into a date object and then use the "format" function to specify the desired format, such as "%Y" for the year. For example, if we have a date in the format "2019-05-20", we can extract the year by using the code "format(as.Date("2019-05-20"), "%Y")", which will return "2019". Another example would be if we have a date in the format "May 20, 2019", we can extract the year by using the code "format(as.Date("May 20, 2019"), "%Y")", which will also return "2019".

Extract Year from Date in R (With Examples)

There are two ways to quickly extract the year from a date in R:

Method 1: Use format()

```
df$year <- format(as.Date(df$date, format="%d/%m/%Y"),"%Y")
```

Method 2: Use the lubridate package

```
library(lubridate)

df$year <- year(mdy(df$date))
```

This tutorial shows an example of how to use each of these methods in practice.

Method 1: Extract Year from Date Using format()

The following code shows how to extract the year from a date using the format() function combined with the "%Y" argument:

```
#create data frame
df <- data.frame(date=c("01/01/2021", "01/04/2021" , "01/09/2021"),
sales=c(34, 36, 44))

#view data frame
df

date sales
1 01/01/2021 34
2 01/04/2021 36
3 01/09/2021 44

#create new variable that contains year
df$year <- format(as.Date(df$date, format="%d/%m/%Y"),"%Y")

#view new data frame
df

date sales year
1 01/01/2021 34 2021
```

2 01/04/2021 36 2021
3 01/09/2021 44 2021

Note that this format() function works with a variety of date formats. You simply must specify the format:

```
#create data frame
df <- data.frame(date=c("2021-01-01", "2021-01-04" , "2021-01-09"),
sales=c(34, 36, 44))

#view data frame
df

date sales
1 2021-01-01 34
2 2021-01-04 36
3 2021-01-09 44

#create new variable that contains year
df$year <- format(as.Date(df$date, format="%Y-%m-%d"),"%Y")

#view new data frame
df
```

date sales year
1 01/01/2021 34 2021
2 01/04/2021 36 2021
3 01/09/2021 44 2021

Method 2: Extract Year from Date Using Lubridate

We can also use functions from the lubridate package to quickly extract the year from a date:

```
library(lubridate)

#create data frame
df <- data.frame(date=c("01/01/2021", "01/04/2021" , "01/09/2021"),
sales=c(34, 36, 44))

#view data frame
df

date sales
1 01/01/2021 34
2 01/04/2021 36
3 01/09/2021 44

#create new variable that contains year
```

```
df$year <- year(mdy(df$date))

#view new data frame
df

date sales year
1 01/01/2021 34 2021
2 01/04/2021 36 2021
3 01/09/2021 44 2021
```

Lubridate also works with a variety of date formats. You simply must specify the format:

```
#create data frame
df <- data.frame(date=c("2021-01-01", "2021-01-04" ,
"2021-01-09"),
sales=c(34, 36, 44))

#view data frame
df

date sales
1 2021-01-01 34
2 2021-01-04 36
3 2021-01-09 44
```

```
#create new variable that contains year
df$year <- year(ymd(df$date))

#view new data frame
df

date sales year
1 01/01/2021 34 2021
2 01/04/2021 36 2021
3 01/09/2021 44 2021
```

How to Loop Through Column Names in R

How to Remove Outliers from Multiple Columns in R