

When I run queries in SQL (or even HiveQL, Spark SQL and so on), it is quite common to use the syntax of `count(case when... else ... end)`. Today, I will provide you an example of how you run this type of commands in `dplyr`.

Let's start:

```
library(sqldf)
library(dplyr)

df<-data.frame(id = 1:10,
               gender = c("m","m","m","f","f","f","m","f","f","f"),
               amt= c(5,20,30,10,20,50,5,20,10,30))
```

df

	id	gender	amt
1	1	m	5
2	2	m	20
3	3	m	30
4	4	f	10
5	5	f	20
6	6	f	50
7	7	m	5
8	8	f	20
9	9	f	10
10	10	f	30

Let's get the `count` and the `sum` per gender in different columns in SQL.

```
sqldf("select count(case when gender='m' then id else null end) as
male_cnt,
               count(case when gender='f' then id else null end) as
female_cnt,
               sum(case when gender='m' then amt else 0 end) as
male_amt,
               sum(case when gender='f' then amt else 0 end) as
female_amt
from df")
```

Output:

	male_cnt	female_cnt	male_amt	female_amt
1	4	6	60	140

Let's get the same output in `dplyr`. We will need to subset the data frame based on one column.

```
df%>%summarise(male_cnt=length(id[gender=="m"]),
               female_cnt=length(id[gender=="f"]),
               male_amt=sum(amt[gender=="m"]),
```

```
female_amt=sum(amt[gender=="f"])
)
```

Output:

	male_cnt	female_cnt	male_amt	female_amt
1	4	6	60	140