

Dataframe-merging

Khusaal Giri

March 25, 2018

Initializing the dataframes:

```
total_values=3
a=c()
b=c()
c=c()
for(i in c(1:total_values)){
  a[i]=i-1
  b[i]=i-1
  c[i]=i-1
}
a=paste('A',a,sep='')
b=paste('B',b,sep='')
c=paste('C',c,sep='')
df1=data.frame(A=a,B=b)
df1
```

```
##      A  B
## 1 A0 B0
## 2 A1 B1
## 3 A2 B2
```

```
df2=data.frame(A=a,C=c)
df2
```

```
##      A  C
## 1 A0 C0
## 2 A1 C1
## 3 A2 C2
```

Paste command will paste the character before or after every element of the vector separated by quotes.

To concat two dataframes in R, we need to have equal number of columns in both the dataframes. So, appending extra columns in both the dataframes & assigning it 'NA'

```
df1_temp=df1
df1_temp$C <- NA
df1_temp
```

```
##      A  B  C
## 1 A0 B0 NA
## 2 A1 B1 NA
## 3 A2 B2 NA
```

```
df2_temp=df2
df2_temp$B <- NA
df2_temp
```

```
##      A  C  B
## 1 A0 C0 NA
## 2 A1 C1 NA
```

```
## 3 A2 C2 NA
```

Concatenating the dataframes along the rows(Vertically):

```
not_so_good=rbind(df1_temp,df2_temp)
not_so_good
```

```
##      A      B      C
## 1 A0     B0 <NA>
## 2 A1     B1 <NA>
## 3 A2     B2 <NA>
## 4 A0 <NA>    C0
## 5 A1 <NA>    C1
## 6 A2 <NA>    C2
```

Concatenating the dataframes along the columns(Horizontally) and with repetition of the common column:

```
much_better=cbind(df1,df2)
much_better
```

```
##      A  B  A  C
## 1 A0 B0 A0 C0
## 2 A1 B1 A1 C1
## 3 A2 B2 A2 C2
```

Concatenating the dataframes along the columns(Horizontally) without the repetition of the common column(Inner join):

```
best_result=merge(df1,df2,by='A')
best_result
```

```
##      A  B  C
## 1 A0 B0 C0
## 2 A1 B1 C1
## 3 A2 B2 C2
```