

excercise1

2024-01-15

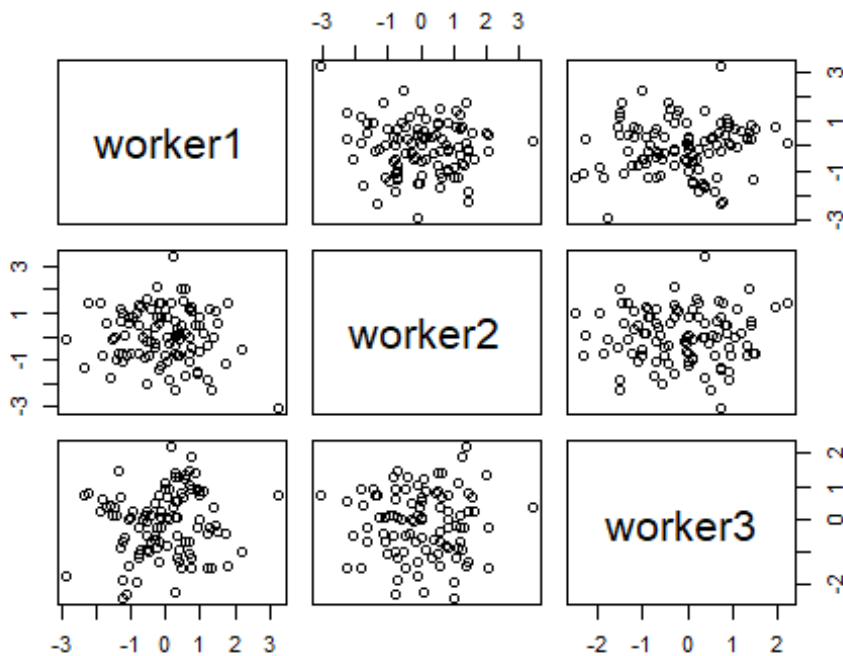
Import data:

```
df = read.csv('C:/Users/tsong/Downloads/performance_data (1).csv')  
  
head(df, 20)
```

Including Plots

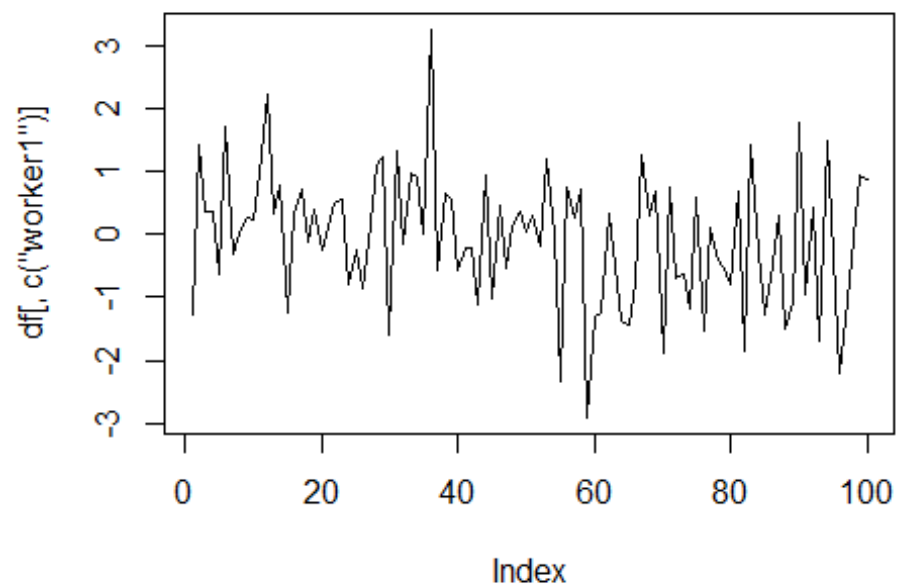
You can also embed plots, for example:

```
df = read.csv('C:/Users/tsong/Downloads/performance_data (1).csv')  
plot(df[,c("worker1", "worker2", "worker3")])
```

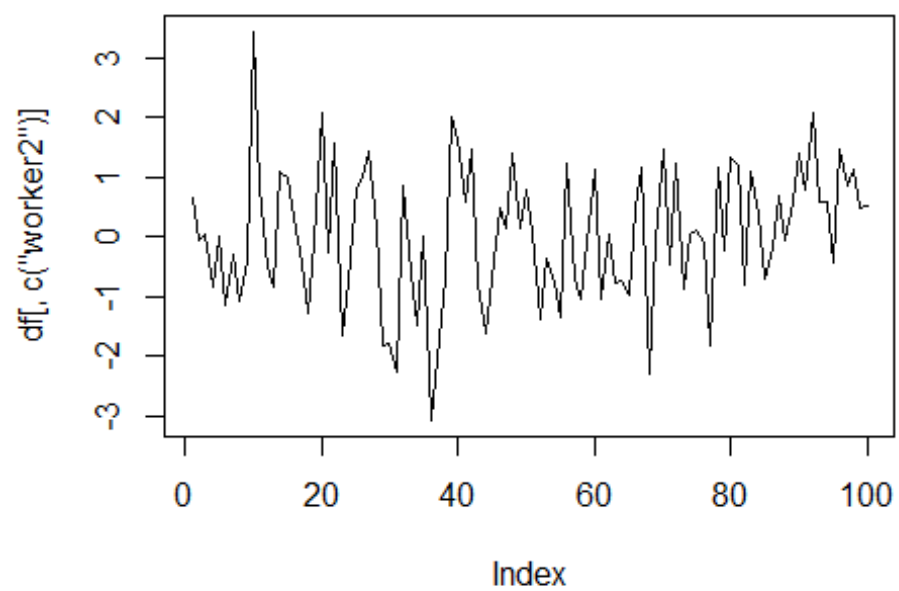


Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

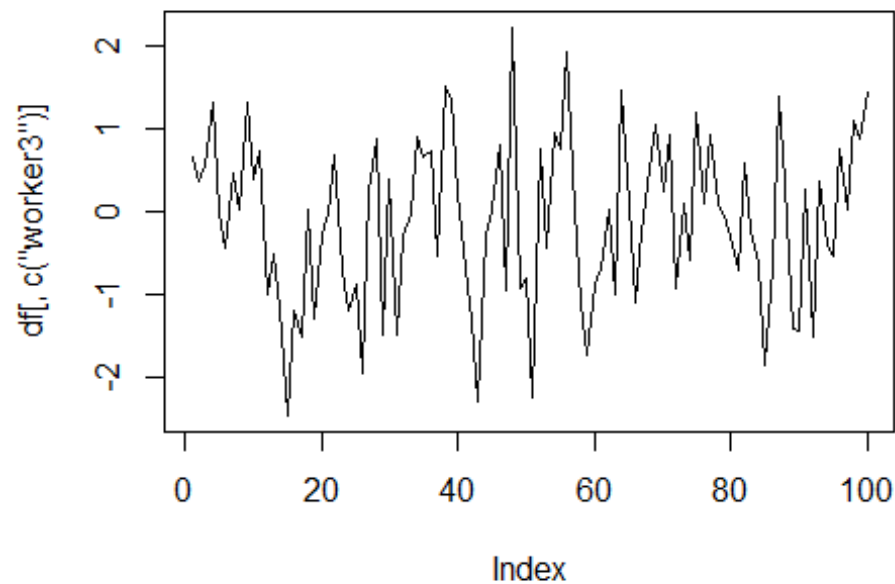
```
df = read.csv('C:/Users/tsong/Downloads/performance_data (1).csv')  
plot(x=df[,c("worker1")], type = 'l')
```



```
plot(x=df[,c("worker2")], type = 'l')
```



```
plot(x=df[,c("worker3")], type = 'l')
```

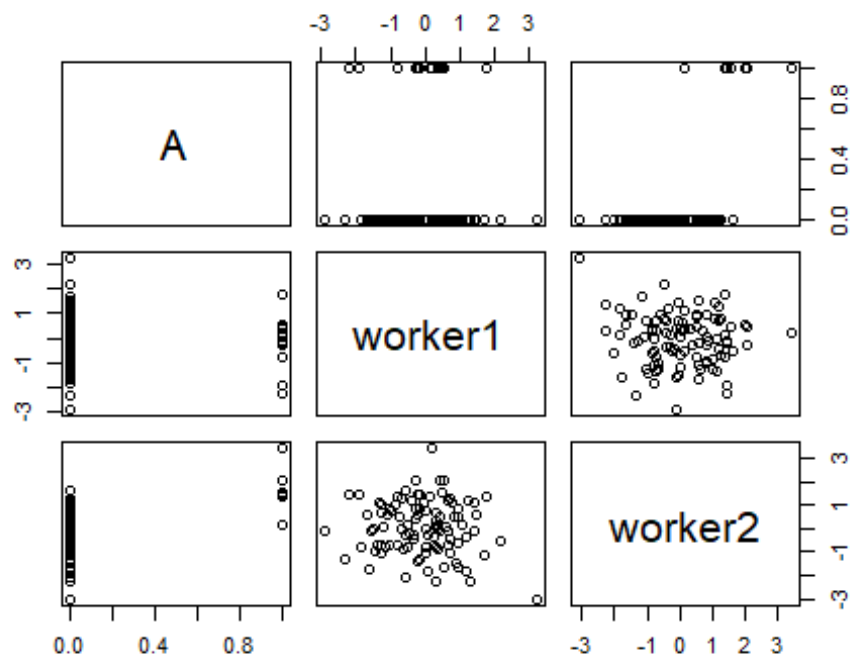


Create two columns to denote interventions on w2 and w3 numerically

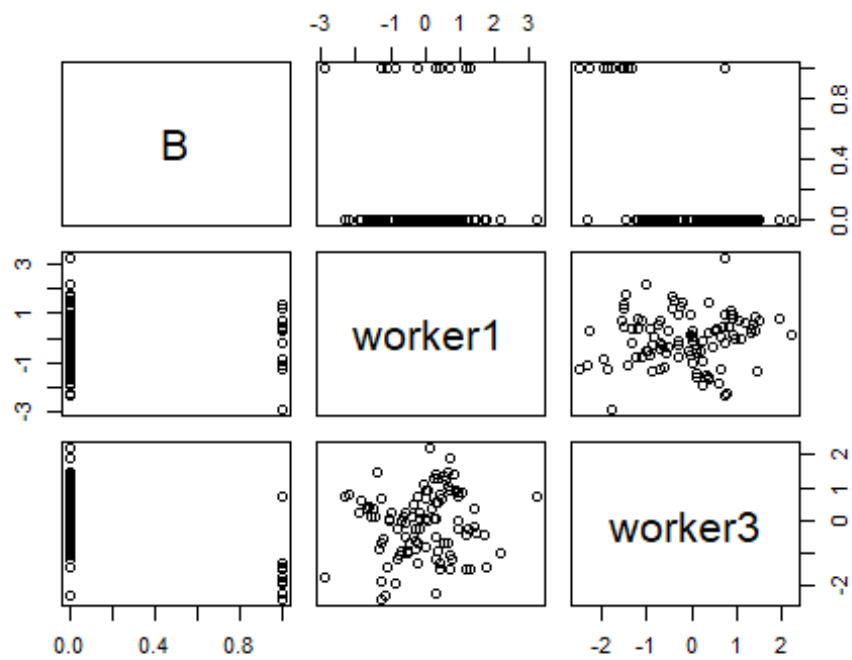
```
df = read.csv('C:/Users/tsong/Downloads/performance_data (1).csv')
df['A'] = ifelse(df["w2_intervention"]=="A",1,0)
df['B'] = ifelse(df["w3_intervention"]=="B",1,0)
```

By plotting the following three variables on scatter plots, we can see that intervention A did have a positive effect on performance as 0 and 1 have different distributions for worker 2

```
df = read.csv('C:/Users/tsong/Downloads/performance_data (1).csv')
df['A'] = ifelse(df["w2_intervention"]=="A",1,0)
df['B'] = ifelse(df["w3_intervention"]=="B",1,0)
plot(df[,c('A', 'worker1', 'worker2')])
```



```
df = read.csv('C:/Users/tsong/Downloads/performance_data (1).csv')
df['A'] = ifelse(df["w2_intervention"]=="A",1,0)
df['B'] = ifelse(df["w3_intervention"]=="B",1,0)
plot(df[,c('B','worker1','worker3')])
```



The intervention on worker has negative effect as the the scatter plot shows the distribution moves downwards