

# facebook-metrics

Ryan Nie

9/29/2023

```
sapply(df, function(x) sum(is.na(x)))
```

```
##                               Page.total.likes
##                               0
##                               Type
##                               0
##                               Category
##                               0
##                               Post.Month
##                               0
##                               Post.Weekday
##                               0
##                               Post.Hour
##                               0
##                               Paid
##                               1
##                               Lifetime.Post.Total.Reach
##                               0
##                               Lifetime.Post.Total.Impressions
##                               0
##                               Lifetime.Engaged.Users
##                               0
##                               Lifetime.Post.Consumers
##                               0
##                               Lifetime.Post.Consumptions
##                               0
##                               Lifetime.Post.Impressions.by.people.who.have.liked.your.Page
##                               0
##                               Lifetime.Post.reach.by.people.who.like.your.Page
##                               0
##                               Lifetime.People.who.have.liked.your.Page.and.engaged.with.your.post
##                               0
##                               comment
##                               0
##                               like
##                               1
##                               share
##                               4
##                               Total.Interactions
##                               0
```

```
data.frame(do.call(cbind, lapply(na.omit(df), summary)))
```

##	Page.total.likes	Type	Category	Post.Month
## Min.	81370	495	1	1
## 1st Qu.	112324	character	1	4
## Median	129600	character	2	7
## Mean	123173.268686869	495	1.88686868686869	7.02828282828283
## 3rd Qu.	136393	character	3	10
## Max.	139441	character	3	12
##	Post.Weekday	Post.Hour	Paid	
## Min.	1	1	0	
## 1st Qu.	2	3	0	
## Median	4	9	0	
## Mean	4.13333333333333	7.84444444444444	0.280808080808081	
## 3rd Qu.	6	11	1	
## Max.	7	23	1	
##	Lifetime.Post.Total.Reach	Lifetime.Post.Total.Impressions		
## Min.	238	570		
## 1st Qu.	3331	5798		
## Median	5290	9084		
## Mean	14028.101010101	29856.9818181818		
## 3rd Qu.	13248	22503		
## Max.	180480	1110282		
##	Lifetime.Engaged.Users	Lifetime.Post.Consumers		
## Min.	9	9		
## 1st Qu.	399	335		
## Median	630	555		
## Mean	926.830303030303	804.155555555556		
## 3rd Qu.	1062	969		
## Max.	11452	11328		
##	Lifetime.Post.Consumptions			
## Min.	9			
## 1st Qu.	512.5			
## Median	861			
## Mean	1425.92121212121			
## 3rd Qu.	1479			
## Max.	19779			
##	Lifetime.Post.Impressions.by.people.who.have.liked.your.Page			
## Min.	567			
## 1st Qu.	4073.5			
## Median	6282			
## Mean	16916.2787878788			
## 3rd Qu.	15143			
## Max.	1107833			
##	Lifetime.Post.reach.by.people.who.like.your.Page			
## Min.	236			
## 1st Qu.	2213			
## Median	3478			
## Mean	6641.35555555556			
## 3rd Qu.	8018			
## Max.	51456			
##	Lifetime.People.who.have.liked.your.Page.and.engaged.with.your.post			
## Min.	9			
## 1st Qu.	297.5			
## Median	416			
## Mean	614.135353535353			

```
## 3rd Qu. 658.5
## Max. 4376
## comment like share Total.Interactions
## Min. 0 0 0 0
## 1st Qu. 1 57 10 72
## Median 3 101 19 125
## Mean 7.55757575757576 179.145454545455 27.2646464646465 213.967676767677
## 3rd Qu. 7 188 32.5 231
## Max. 372 5172 790 6334
```

```
colnames(df) <- sub("Lifetime", "LT", colnames(df))
colnames(df) <- sub("Post", "P", colnames(df))
colnames(df) <- sub("post", "P", colnames(df))
colnames(df) <- sub("Total", "Tot", colnames(df))
colnames(df) <- sub("total", "tot", colnames(df))
colnames(df) <- sub("Page", "Pg", colnames(df))
colnames(df) <- sub("people", "ppl", colnames(df))
colnames(df) <- sub("People", "ppl", colnames(df))
colnames(df) <- sub("engaged", "eng", colnames(df))
colnames(df) <- sub("Consumptions", "Consump", colnames(df))
colnames(df) <- sub("have.", "", colnames(df))
colnames(df) <- sub("Impressions", "Imp", colnames(df))
colnames(df) <- sub("your.", "", colnames(df))
```

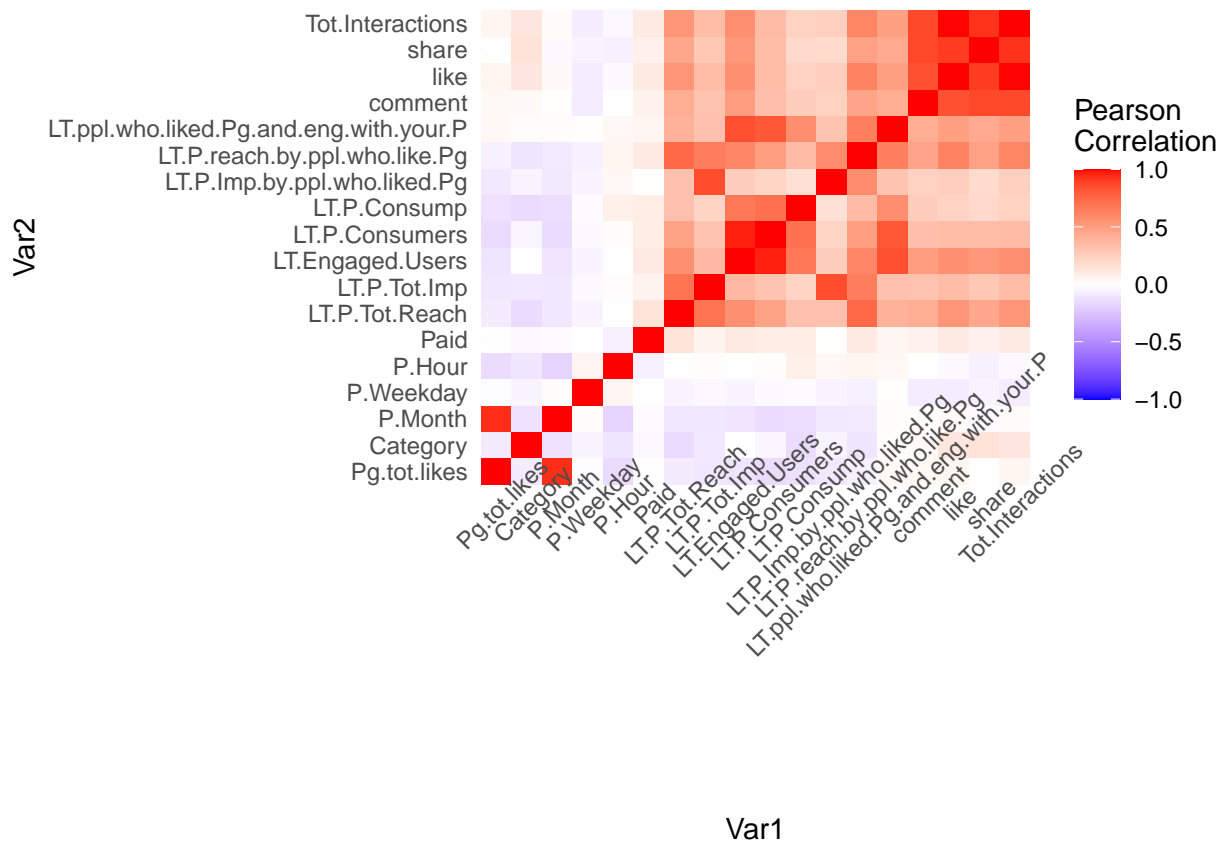
*# can prob map this*

```
library(reshape2)
```

```
##
## Attaching package: 'reshape2'
## The following object is masked from 'package:tidyr':
##
## smiths
```

```
numerical_df <- df[sapply(df, is.numeric)]
corr_mat <- cor(numerical_df, use = "complete.obs")

melted_corr <- melt(corr_mat)
ggplot(data = melted_corr, aes(x = Var1, y = Var2, fill = value)) +
  geom_tile() +
  scale_fill_gradient2(
    low = "blue",
    high = "red",
    mid = "white",
    midpoint = 0,
    limit = c(-1, 1),
    space = "Lab",
    name = "Pearson\nCorrelation"
  ) +
  scale_x_discrete(expand = c(0, 0)) +
  scale_y_discrete(expand = c(0, 0)) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45))
```



```
# ggsave(file="cor.png", width=20, height=25, dpi=300)
```

```
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
```

```
##   method from
```

```
##   +.gg    ggplot2
```

```
num_df <- numerical_df %>% select(!c('Post.Month', 'Post.Weekday', 'Post.Hour', 'Paid', 'Category'))
ggpairs(num_df)
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removing 1 row that contained a missing value
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removing 1 row that contained a missing value
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removed 4 rows containing missing values
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removing 1 row that contained a missing value
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removing 1 row that contained a missing value
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :  
## Removed 4 rows containing missing values
```

```
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
```

```

## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).

## Warning: Removed 1 rows containing non-finite values (`stat_density()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

```

```

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 2 rows containing missing values

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 5 rows containing missing values

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning: Removed 1 rows containing missing values (`geom_point()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

```

```

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning: Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Warning: Removed 2 rows containing missing values (`geom_point()`).

## Warning: Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).

## Warning: Removed 1 rows containing non-finite values (`stat_density()`).

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values

## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removing 1 row that contained a missing value

## Warning: Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).

## Warning: Removed 5 rows containing missing values (`geom_point()`).

## Warning: Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).

```

```
## Removed 4 rows containing missing values (`geom_point()`).
## Removed 4 rows containing missing values (`geom_point()`).
## Warning: Removed 4 rows containing non-finite values (`stat_density()`).
## Warning in ggally_statistic(data = data, mapping = mapping, na.rm = na.rm, :
## Removed 4 rows containing missing values
## Warning: Removed 1 rows containing missing values (`geom_point()`).
## Removed 1 rows containing missing values (`geom_point()`).
## Warning: Removed 4 rows containing missing values (`geom_point()`).
```

