

dmw

July 19, 2020

```
[1]: # importing libraries
from utilities import *

%matplotlib inline
%load_ext autoreload
%autoreload 2
```

Importing plotly failed. Interactive plots will not work.

```
[2]: # file paths
catalogs_file_path = r"data\DMW_Clicks_Catalogs_2.xlsx"
clicks_file_path = r"data\DMW_Clicks_Catalogs_2.xlsx"
```

```
[3]: # parameters
freq = "M"
```

```
[4]: # read catalogs data set
dmw_catalogs = pd.read_excel(catalogs_file_path, sheet_name="Catalogs_03_04",
    ↳ parse_dates=["Date"], index_col=0)
dmw_catalogs.columns = dmw_catalogs.columns.str.lower().str.replace(" ", "_")
dmw_catalogs.index.name = dmw_catalogs.index.name.lower()
dmw_catalogs_ts = dmw_catalogs.resample(freq).sum()
# dmw_catalogs_ts.head(2)
```

```
[5]: # read clicks data
dmw_clicks = pd.read_excel(clicks_file_path, sheet_name="Clicks_03_04",
    ↳ parse_dates=["Date"], index_col=0)
dmw_clicks.columns = dmw_clicks.columns.str.lower().str.replace(" ", "_")
dmw_clicks.index.name = dmw_clicks.index.name.lower()
dmw_clicks_ts = dmw_clicks.resample(freq).sum()
# dmw_clicks_ts.head(2)
```

```
[6]: # merge catalogs data and clicks data
dmw_catalogs_clicks_ts = dmw_catalogs_ts.merge(dmw_clicks_ts, left_index=True,
    ↳ right_index=True, how="right")
dmw_catalogs_clicks_ts
```

```
[6]:      cats_mailed  google  microsoft_bing  yahoo  aol_search  \
date
2012-12-31      66401      560              85      59          10
2013-01-31     13045     9624             1819     1070         248
2013-02-28         0     8310             1545     941         234
2013-03-31     94170     9394             1623     977         247
2013-04-30      6050    11444             1899    1085         294
2013-05-31     22375    12095             1929     980         261
2013-06-30    190871    12994             2465    1415         294
2013-07-31     76990    14302             2302    1376         303
2013-08-31    109940    15519             2805    1620         370
2013-09-30     25965    11639             1803    1040         196
2013-10-31     12475    11039             1674    1131         201
2013-11-30     32195    10869             1602    1313         237
2013-12-31     97135    11489             1520    1238         223
2014-01-31     13960    13768             2246     982         278
2014-02-28         0    13298             2271     826         197
2014-03-31         0    13534             2285    1401         140
2014-04-30         0    12949             2670    1681         181
2014-05-31     29590    13267             2372    1426         154
2014-06-30    225835     5166             1136     693          52
```

```
      ask_jeeves  total
date
2012-12-31      15     729
2013-01-31     270    13031
2013-02-28     293    11323
2013-03-31     337    12578
2013-04-30     360    15082
2013-05-31     312    15577
2013-06-30     309    17477
2013-07-31     432    18715
2013-08-31     422    20736
2013-09-30     149    14827
2013-10-31     134    14179
2013-11-30      92    14113
2013-12-31      44    14514
2014-01-31       7    17281
2014-02-28      12    16604
2014-03-31       5    17365
2014-04-30      10    17491
2014-05-31       8    17227
2014-06-30       6     7053
```

1 In Feb, 2013, and Feb, Mar and April, 2014, there were no catalogs mailed but we had clicks

```
[7]: # correlation of the catalogs mailed to the number of clicks
corr, p = stats.pearsonr(dmw_catalogs_clicks_ts['cats_mailed'],
    ↪dmw_catalogs_clicks_ts['total'])
print(corr, p)
```

-0.2067703215530065 0.39569342590733964

2 There is weak negative correlation between the catalogs mailed to the numbers of clicks

```
[8]: # correlation of the variables
dmw_catalogs_clicks_ts.corr()
```

```
[8]:
```

	cats_mailed	google	microsoft_bing	yahoo	aol_search	\
cats_mailed	1.000000	-0.236497	-0.152836	-0.067028	-0.084291	
google	-0.236497	1.000000	0.939863	0.833226	0.720003	
microsoft_bing	-0.152836	0.939863	1.000000	0.855283	0.653341	
yahoo	-0.067028	0.833226	0.855283	1.000000	0.589586	
aol_search	-0.084291	0.720003	0.653341	0.589586	1.000000	
ask_jeeves	0.138611	0.235300	0.238370	0.230599	0.732762	
total	-0.206770	0.995875	0.958090	0.865723	0.742301	

	ask_jeeves	total
cats_mailed	0.138611	-0.206770
google	0.235300	0.995875
microsoft_bing	0.238370	0.958090
yahoo	0.230599	0.865723
aol_search	0.732762	0.742301
ask_jeeves	1.000000	0.284025
total	0.284025	1.000000

```
[ ]:
```