Product Case Study

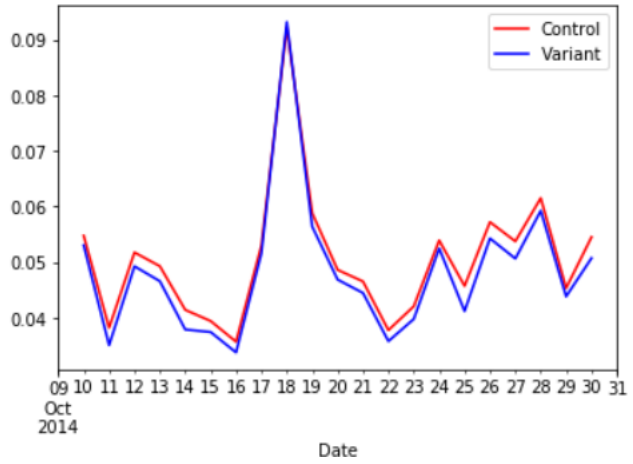
I-Lin Liu

Report of case study

- We could not see the significant differences in conversion rate between control(0.05056) and variant (0.04827), the control has 3771 more times purchase than variant group. The sanity check shows not significance between control and variant, even considering the type of user and different channels.
- The new version of homepage increases bounce rate from sanity check. The bounce rate between control(0.396) and variant (0.4119) is not significant under alpha = 0.05. The variant group has 61453 more times of bounce than control group.
- The homepage seems not changes the behaviors of returning users. There is no significant change in relative trend between new and returning users. The change of the homepage might not affect any search. It will need further studies such as staying time and what cookies users clicked in homepage. The new version of homepage might help users search info they like but the info provided could not increase the purchase rate.
- For users who tend to purchase directly through link, email, and other media, they might not even notice the change of homepage. These data might dilute the significance of the testing.
- The homepage might not be efficient for users purchasing through email, directly purchase. We saw clear the bounce rate has negatively correlation with conversion rate.

Conversion rate:
purchase rate in home page land

```
Probility of purchase in control group:
count    21.000000
mean      0.050560
std       0.012084
min       0.035699
25%       0.042077
50%       0.049284
75%       0.054522
max       0.092231
Name: Visitors_Control, dtype: float64
Conversion rate in Control group: 0.026875360000000004 - 0.07424464
Probility of purchase in variant group:
count    21.000000
mean      0.048270
std       0.012715
min       0.033765
25%       0.039781
50%       0.046874
75%       0.052477
max       0.093210
Name: Visitors_Variant, dtype: float64
Conversion rate in Variant group: 0.0233486 - 0.0731914
```

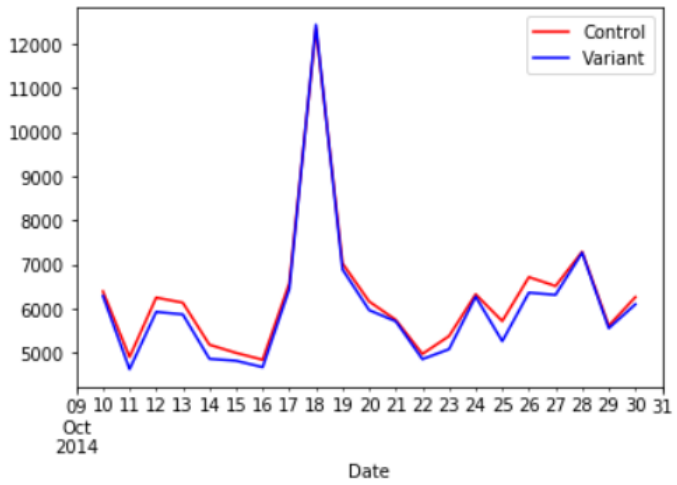


We did not see the obvious enhancement of Variant group.
The conversion rate of the control and variant groups are 0.05056 and 0.04827
The Confidence interval case:
Control: 0.02688-0.07424
Variant: 0.02335-0.07319
Both lower and higher margin of Variant group are lower than control group. We need to do further studies before launch the new version of homepage.

aggregate relative difference of purchase

```
Total amounty of purchase in control group:
count       21.000000
mean      6258.380952
std       1567.031604
min       4847.000000
25%       5382.000000
50%       6166.000000
75%       6516.000000
max      12351.000000
Name: Visitors_Control, dtype: float64
Total amounty of purchase in variant group:
count       21.000000
mean      6078.809524
std       1635.266847
min       4633.000000
25%       5086.000000
50%       5930.000000
75%       6316.000000
max      12431.000000
Name: Visitors_Variant, dtype: float64
```

The aggregate amount of purchase in control group has 179.5714* 21 =3771 more times than variant group over these data.




```
count      21.000000
mean      179.571429
std       129.318820
min       -80.000000
25%       101.000000
50%       168.000000
75%       276.000000
max       455.000000
dtype: float64
```

Sanity check: conversion rate

d = P(variant)-P(Control)

```
Date
2014-10-10   -0.001720
2014-10-11   -0.003140
2014-10-12   -0.002491
2014-10-13   -0.002702
2014-10-14   -0.003540
2014-10-15   -0.001982
2014-10-16   -0.001934
2014-10-17   -0.001434
2014-10-18    0.000979
2014-10-19   -0.002427
2014-10-20   -0.001753
2014-10-21   -0.002073
2014-10-22   -0.001980
2014-10-23   -0.002296
2014-10-24   -0.001447
2014-10-25   -0.004528
2014-10-26   -0.002934
2014-10-27   -0.003112
2014-10-28   -0.002275
2014-10-29   -0.001505
2014-10-30   -0.003785
dtype: float64
```

d- Z*SE(control+variant)

```
Date
2014-10-10    0.000105
2014-10-11   -0.001697
2014-10-12   -0.000743
2014-10-13   -0.001030
2014-10-14   -0.002021
2014-10-15   -0.000491
2014-10-16   -0.000564
2014-10-17    0.000316
2014-10-18    0.003178
2014-10-19   -0.000569
2014-10-20   -0.000095
2014-10-21   -0.000446
2014-10-22   -0.000554
2014-10-23   -0.000760
2014-10-24    0.000360
2014-10-25   -0.002939
2014-10-26   -0.001078
2014-10-27   -0.001354
2014-10-28   -0.000374
2014-10-29    0.000110
2014-10-30   -0.001980
dtype: float64
count    21.000000
mean     -0.000601
std       0.001204
min      -0.002939
25%      -0.001078
50%      -0.000564
75%      -0.000095
max       0.003178
dtype: float64
```

Here, we mix control and variant data into a pool and calculate the SE(control+variant). Here we calculate the difference of conversion rate between control and variant, the difference between two distributions (d). Here, I took 95% of confidence interval in normal distribution and the Z is 1.96. The expectation value of d- Z*SE(control+variant) is 0.0006 +- 0.0012, which covers 0. This means that variant data is not significant from the data set including control and variant.

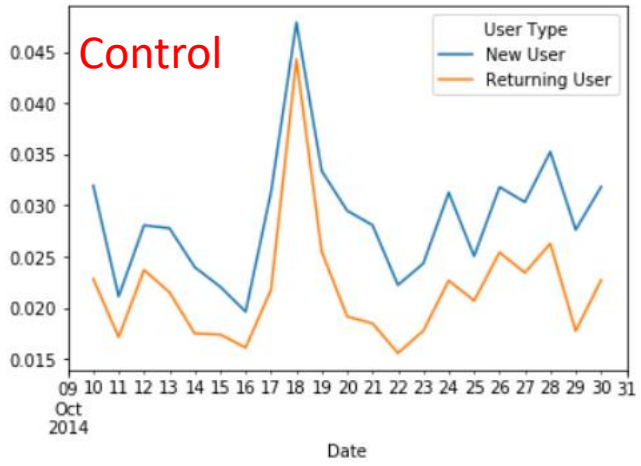
User Type: conversion rate

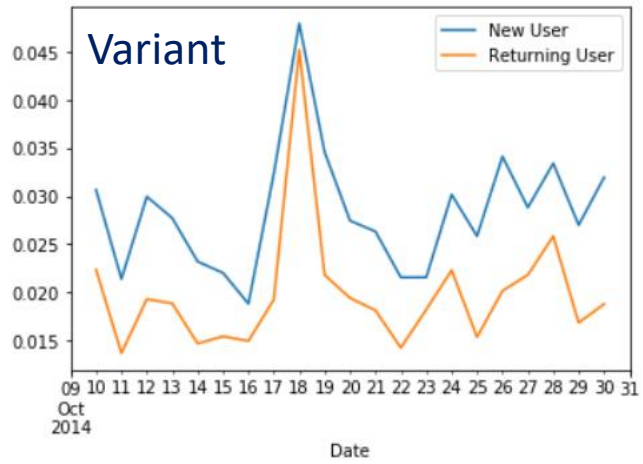
conversion rate from returning user and New user in control group:

User Type	New User	Returning User
count	21.000000	21.000000
mean	0.028784	0.021776
std	0.006185	0.006110
min	0.019606	0.015556
25%	0.024338	0.017714
50%	0.028077	0.021499
75%	0.031813	0.023421
max	0.047919	0.044312

conversion rate from returning user and New user in variant group:

User Type	New User	Returning User
count	21.000000	21.000000
mean	0.028430	0.019840
std	0.006462	0.006626
min	0.018820	0.013681
25%	0.023209	0.015422
50%	0.027732	0.018850
75%	0.031954	0.021808
max	0.047981	0.045229





New User: Conversion rate:
0.0288+-0.0121(0.0167-0.0409)Control
0.0284+-0.0127(0.0157-0.0411)Variant
Returning User
0.02178+-0.0121(0.0097-0.0339)Control
0.01984+-0.0130(0.0069-0.0328)Variant
For returning users, the confidence interval is within the control range. Therefore, the new version of homepage does not increase the purchase rate for returning users. For new users, the higher margin is slight above the control group. It might be worth further studies.

Also, the relative ratio between new and returning users is not significant between control and variant groups.

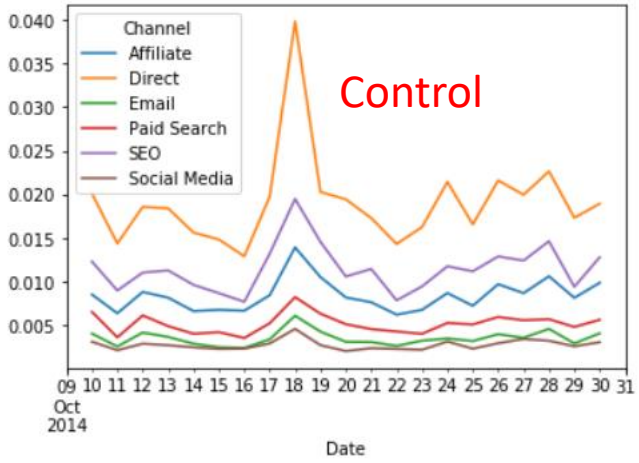
Channel: conversion rate

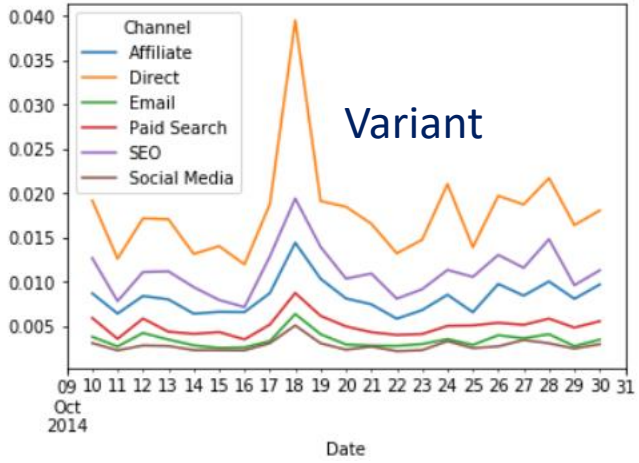
Probility of purchase in control group:

Channel	Affiliate	Direct	Email	Paid Search	SEO	Social Media
count	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000
mean	0.008434	0.019078	0.003546	0.005207	0.011512	0.002782
std	0.001823	0.005439	0.000863	0.001111	0.002700	0.000579
min	0.006245	0.012926	0.002431	0.003572	0.007719	0.002074
25%	0.006794	0.016269	0.002954	0.004318	0.009475	0.002350
50%	0.008210	0.018581	0.003419	0.005150	0.011311	0.002746
75%	0.008848	0.020109	0.004078	0.005721	0.012819	0.003089
max	0.013942	0.039809	0.006123	0.008252	0.019490	0.004615

Probility of purchase in variant group:

Channel	Affiliate	Direct	Email	Paid Search	SEO	Social Media
count	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000
mean	0.008234	0.017812	0.003366	0.004996	0.011107	0.002755
std	0.001927	0.005713	0.000868	0.001147	0.002788	0.000646
min	0.005781	0.011914	0.002491	0.003449	0.007101	0.002128
25%	0.006558	0.013978	0.002753	0.004262	0.009348	0.002242
50%	0.008071	0.017116	0.003259	0.004976	0.011051	0.002661
75%	0.008679	0.019065	0.003744	0.005504	0.012623	0.003012
max	0.014366	0.039448	0.006321	0.008690	0.019353	0.005031

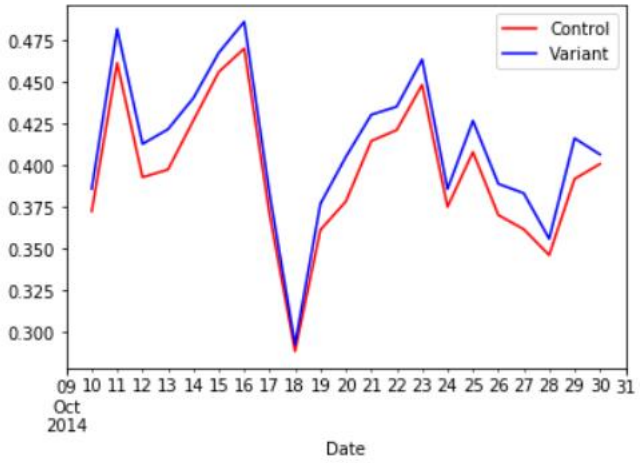




No significant changes for control and variant groups.

Bounce rate:
Bounce rate from home page land

```
Probility of purchase in control group:
count    21.000000
mean      0.396011
std       0.043229
min       0.288506
25%       0.370583
50%       0.392988
75%       0.421335
max       0.470314
Name: Visitors_Control, dtype: float64
Bounce rate in Control group: 0.31127116000000005 - 0.48072884
Probility of purchase in variant group:
count    21.000000
mean      0.411870
std       0.044992
min       0.292601
25%       0.385897
50%       0.412970
75%       0.435362
max       0.486368
Name: Visitors_Variant, dtype: float64
Bounce rate in Variant group: 0.32368568000000003 - 0.50005432
```

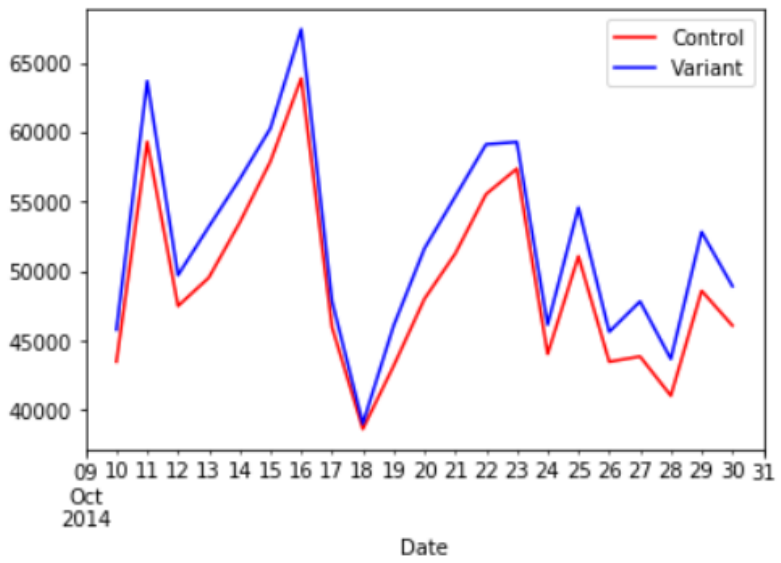


We did not see the obvious enhancement of Variant group.
The bounce rate of the control and variant groups are 0.396 and 0.4119.
The Confidence interval case:
Control: 0.3113-0.4807
Variant: 0.3237-0.5
Both lower and higher margin(95% normal distribution) of Variant group are lower than control group. We need to do further studies before launch the new version of homepage.

aggregate relative difference of bounce

```
Total amount of Bounce in control group:
count       21.000000
mean     49188.714286
std       6653.947829
min      38635.000000
25%      43849.000000
50%      47989.000000
75%      53472.000000
max      63856.000000
Name: Visitors_Control, dtype: float64
Bounce rate in variant group:
count       21.000000
mean     52115.047619
std       7150.290816
min      39023.000000
25%      46144.000000
50%      51611.000000
75%      56601.000000
max      67398.000000
Name: Visitors_Variant, dtype: float64
```

The variant group has 2926.333*21 = 61453 more times of bounce than control group.




```
count       21.000000
mean     -2926.333333
std        989.940924
min      -4374.000000
25%      -3622.000000
50%      -2893.000000
75%      -2221.000000
max       -388.000000
dtype: float64
```

Sanity check: conversion rate

d = P(variant)-P(Control)

```
Date
2014-10-10    0.013536
2014-10-11    0.020430
2014-10-12    0.019982
2014-10-13    0.024218
2014-10-14    0.013457
2014-10-15    0.011544
2014-10-16    0.016054
2014-10-17    0.012369
2014-10-18    0.004095
2014-10-19    0.015713
2014-10-20    0.026768
2014-10-21    0.015760
2014-10-22    0.014027
2014-10-23    0.015217
2014-10-24    0.010735
2014-10-25    0.018795
2014-10-26    0.018778
2014-10-27    0.021620
2014-10-28    0.009832
2014-10-29    0.024402
2014-10-30    0.005706
dtype: float64
```

d- Z*SE(control+variant)

```
2014-10-10    0.009614
2014-10-11    0.016595
2014-10-12    0.016067
2014-10-13    0.020368
2014-10-14    0.009600
2014-10-15    0.007679
2014-10-16    0.012315
2014-10-17    0.008562
2014-10-18    0.000652
2014-10-19    0.011864
2014-10-20    0.022972
2014-10-21    0.011903
2014-10-22    0.010277
2014-10-23    0.011356
2014-10-24    0.006825
2014-10-25    0.014951
2014-10-26    0.014851
2014-10-27    0.017797
2014-10-28    0.006021
2014-10-29    0.020559
2014-10-30    0.001739
dtype: float64
count    21.000000
mean      0.012027
std       0.005890
min       0.000652
25%       0.008562
50%       0.011864
75%       0.016067
max       0.022972
dtype: float64
```

Here, we mix control and variant data into a pool and calculate the SE(control+variant). Here we calculate the difference of bounce rate between control and variant, the difference between two distributions (d). Here, I took 95% of confidence interval in normal distribution and the Z is 1.96. The expectation value of d- Z*SE(control+variant) is 0.012 +- 0.00589. This means that the true difference of variant and control is always larger than margin. Therefore, the variant data is significant from control. The bounce rate from version of home page is higher than original version(control).

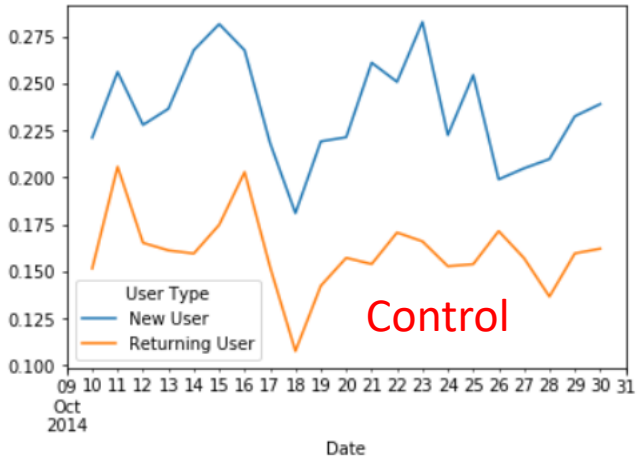
User Type: bounce rate

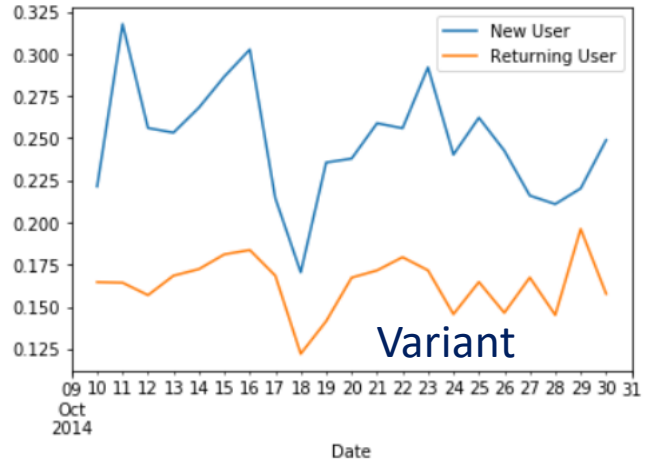
bounce rate from returning user and New user in control group:

User Type	New User	Returning User
count	21.000000	21.000000
mean	0.235877	0.160134
std	0.027570	0.020405
min	0.180930	0.107577
25%	0.219077	0.152691
50%	0.232466	0.159452
75%	0.256020	0.165891
max	0.282576	0.205588

bounce rate from returning user and New user in variant group:

User Type	New User	Returning User
count	21.000000	21.000000
mean	0.248232	0.163638
std	0.034296	0.016671
min	0.170508	0.122093
25%	0.221417	0.156894
50%	0.248948	0.167284
75%	0.262221	0.171623
max	0.317760	0.196146





New User: Conversion rate:

0.2359+-0.0540(0.1819-0.2900)Control

0.2482+-0.0672(0.1810-0.3154)Variant

Returning User

0.1601+-0.0400(0.1201-0.2001)Control

0.1636+-0.0327(0.1309-0.1963)Variant

For returning users, the confidence interval of the new version is within the control one. However, for new users, the confidence interval of bounce rate in new version is larger than control group. Compare with sanity check, for new users, the bounce rate increase in variant group.

Channel: conversion rate

Probility of purchase in control group:

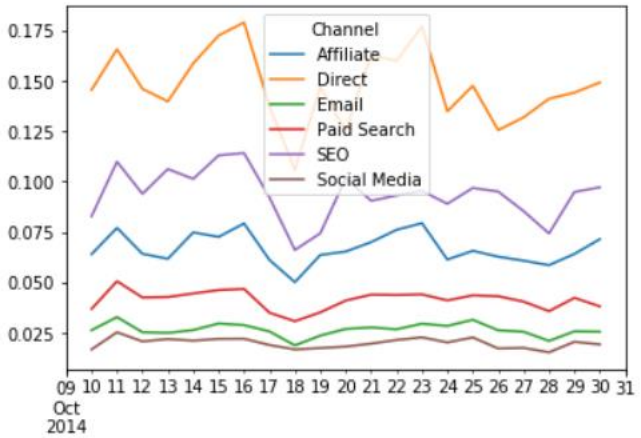
Channel	Affiliate	Direct	Email	Paid Search	SEO	Social Media
count	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000
mean	0.066897	0.147548	0.026541	0.041377	0.093739	0.019909
std	0.007610	0.018272	0.003161	0.004674	0.012450	0.002524
min	0.050144	0.105904	0.018773	0.030773	0.066139	0.015332
25%	0.061743	0.137561	0.025542	0.038135	0.089045	0.017514
50%	0.064243	0.146115	0.026321	0.042558	0.094912	0.020253
75%	0.072619	0.159964	0.028440	0.043974	0.101410	0.021916
max	0.079455	0.179056	0.032830	0.050634	0.114220	0.025294

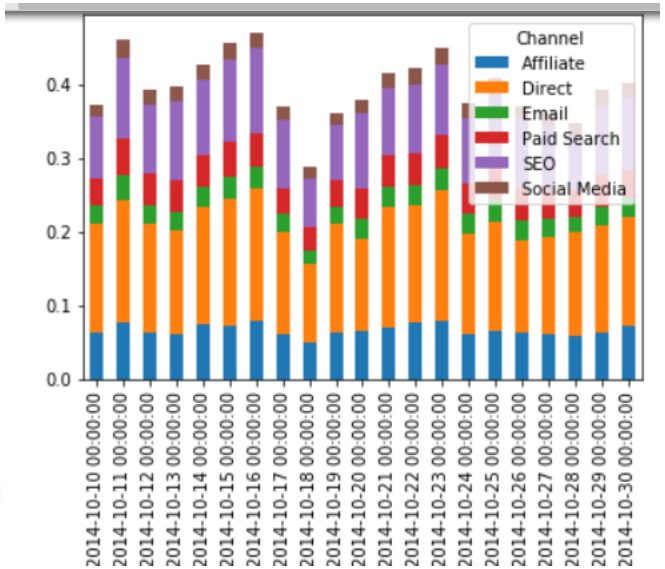
Probility of purchase in variant group:

Channel	Affiliate	Direct	Email	Paid Search	SEO	Social Media
count	21.000000	21.000000	21.000000	21.000000	21.000000	21.000000
mean	0.067634	0.154437	0.027836	0.043392	0.098615	0.019957
std	0.007133	0.016660	0.003526	0.005493	0.014810	0.002455
min	0.050860	0.107419	0.019038	0.031215	0.067064	0.015254
25%	0.063675	0.145248	0.026317	0.039968	0.091507	0.017893
50%	0.066457	0.157739	0.028333	0.043714	0.097509	0.020221
75%	0.071455	0.168397	0.029565	0.046923	0.108422	0.021181
max	0.080572	0.180301	0.034963	0.052196	0.124605	0.024659

No significant changes for control and variant groups.

Control





Variant