## A/B Testing the Udacity Website

# Exercise 1

In [12]:

```python
# Import libraries
import warnings
warnings.simplefilter('ignore')
import pandas as pd
import numpy as np
from scipy.stats import ttest_ind #t-test
import statsmodels.formula.api as smf #regression

# Load data
control = pd.read_csv("https://media.githubusercontent.com/media/nickeub
ank/MIDS_Data/master/udacity_AB_testing/control_data.csv")
experimental = pd.read_csv("https://media.githubusercontent.com/media/ni
ckeubank/MIDS_Data/master/udacity_AB_testing/experiment_data.csv")

control.head()
```

Out[12]:

| | Date | Pageviews | Clicks | Enrollments | Payments |
|---|---|---|---|---|---|
| 0 | Sat, Oct 11 | 7723 | 687 | 134.0 | 70.0 |
| 1 | Sun, Oct 12 | 9102 | 779 | 147.0 | 70.0 |
| 2 | Mon, Oct 13 | 10511 | 909 | 167.0 | 95.0 |
| 3 | Tue, Oct 14 | 9871 | 836 | 156.0 | 105.0 |
| 4 | Wed, Oct 15 | 10014 | 837 | 163.0 | 64.0 |

# Exercise 2

The unit of observation is the Date. We are observing the number of pageviews, clicks, enrollments, and payments across all Udacity visitors for that specific date.

# Exercise 3

In [2]:

```
# Add column for treatment
control['treatment'] = 0
experimental['treatment'] = 1

# Combine dataframes
combined = pd.concat([control,experimental])
combined.sample(10)
```

Out[2]:

| | Date | Pageviews | Clicks | Enrollments | Payments | treatment |
|---|---|---|---|---|---|---|
| 1 | Sun, Oct 12 | 9288 | 785 | 116.0 | 91.0 | 1 |
| 4 | Wed, Oct 15 | 10014 | 837 | 163.0 | 64.0 | 0 |
| 21 | Sat, Nov 1 | 8460 | 681 | 156.0 | 93.0 | 0 |
| 6 | Fri, Oct 17 | 9008 | 748 | 146.0 | 76.0 | 0 |
| 13 | Fri, Oct 24 | 9434 | 673 | 220.0 | 122.0 | 0 |
| 14 | Sat, Oct 25 | 8669 | 669 | 127.0 | 81.0 | 1 |
| 32 | Wed, Nov 12 | 10042 | 802 | NaN | NaN | 1 |
| 29 | Sun, Nov 9 | 9656 | 825 | NaN | NaN | 0 |
| 13 | Fri, Oct 24 | 9402 | 697 | 194.0 | 94.0 | 1 |
| 10 | Tue, Oct 21 | 10551 | 864 | 143.0 | 71.0 | 1 |

# Exercise 4

The experiment evaluated the effect on the number of payments after selecting the "Start Free Trial" option. Therefore, our predictor variable is Payments.

We are interested in seeing if the treatment led to a higher number of payments per clicks. We will likely need to create a hybrid variable of payments per click or enrollments per click to standardize by clicks.

# Exercise 5

In [3]:

```
views_control=np.mean(control['Pageviews'])
views_experiment=np.mean(experimental['Pageviews'])
print (f'The mean number of pageviews in the control group is {views_control:.2f}')
print (f'The mean number of pageviews in the control group in experimental group is {views_experiment:.2f}')
```

```
The mean number of pageviews in the control group is 9339.00
The mean number of pageviews in the control group in experimental group
is 9315.14
```

The average number of pageviews for the control and experimental groups are fairly similar (around 9300).

## Exercise 6

```
In [4]: ttest_ind(control.Pageviews.values, experimental.Pageviews.values)
```

```
Out[4]: Ttest_indResult(statistic=0.1417118298287496, pvalue=0.887703406865090
        2)
```

The t-test shows a p-value of 0.88 which is greater than 0.05. There is no significant difference in pageviews between the experimental and control groups.

## Exercise 7

Clicks is another pre-treatment variable because users clicked to review the page before that were forwarded to two versions of the website.

## Exercise 8

```
In [5]: clicks_control=np.mean(control['Clicks'])
        clicks_experiment=np.mean(experimental['Clicks'])
        print (f'The mean number of clicks in the control group is:{clicks_contr
        ol:.2f}')
        print (f'The mean number of clicks in the experimental group is {clicks_
        experiment:.2f}')
        ttest_ind(control.Clicks.values, experimental.Clicks.values)
```

```
        The mean number of clicks in the control group is:766.97
        The mean number of clicks in the experimental group is 765.54

Out[5]: Ttest_indResult(statistic=0.09270642968639531, pvalue=0.926394264248270
        3)
```

The difference between treatment and control group means is also **not** statistically significant.

## Exercise 9

```
In [6]: # Determine number of nulls in payments
        combined['Payments'].isnull().sum()
```

```
Out[6]: 28
```

The data has 28 missing values for Payment column. We will drop the missing values to run the t-test to estimate whether the effect of new website design is statistically significant.

In [7]:

```
# Drop NA values
combined_cleaned = combined.dropna(subset = ['Payments'])

# Create column for payments per click
combined_new = combined_cleaned.copy()
combined_new['Payments_per_Click'] = combined_new['Payments']/combined_n
ew['Clicks']
combined_new

# T test
ttest_ind(combined_new[combined_new['treatment'] == 1].Payments_per_Clic
k.values,
          combined_new[combined_new['treatment'] == 0].Payments_per_Clic
k.values)
```

Out[7]:

```
Ttest_indResult(statistic=-0.5387777625331603, pvalue=0.592755861426802
4)
```

The p-value is 0.592, which indicates that there is no significant difference in payments per click between the experimental and control groups.

# Exercise 10

For the bivariate regression, the x variable will be "Clicks" and the y-variable will be "Payments per click".

In [8]:

```
# Create regression model between payments per click and treatment
model=smf.ols(formula='Payments_per_Click~treatment', data=combined_new)
res=model.fit()
print (res.summary())
print(res.pvalues)
```

```
                            OLS Regression Results
==========================================================================
=======
Dep. Variable:      Payments_per_Click   R-squared:
0.007
Model:                             OLS   Adj. R-squared:
-0.016
Method:                  Least Squares   F-statistic:
0.2903
Date:                 Tue, 04 Feb 2020   Prob (F-statistic):
0.593
Time:                         13:59:14   Log-Likelihood:
95.810
No. Observations:                   46   AIC:
-187.6
Df Residuals:                       44   BIC:
-184.0
Df Model:                            1
Covariance Type:             nonrobust
==========================================================================
=======
                 coef    std err          t      P>|t|      [0.025
0.975]
--------------------------------------------------------------------------
-------
Intercept      0.1183      0.006     18.403      0.000       0.105
0.131
treatment     -0.0049      0.009     -0.539      0.593      -0.023
0.013
==========================================================================
=======
Omnibus:                         0.968   Durbin-Watson:
1.092
Prob(Omnibus):                   0.616   Jarque-Bera (JB):
0.985
Skew:                            0.316   Prob(JB):
0.611
Kurtosis:                        2.662   Cond. No.
2.62
==========================================================================
=======

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
Intercept    2.736586e-22
treatment    5.927559e-01
dtype: float64
```

Indeed, the p-values between the t-test from the previous exercise and the regression model are equivalent and equal to 0.59276.

In [9]:

```
# Create regression model accounting for date
model=smf.ols(formula='Payments_per_Click~treatment+C(Date)', data=combined_new)
res=model.fit()
print (res.summary())
print(res.pvalues)
```

```
                            OLS Regression Results                            
=======================================================================
=======
Dep. Variable:     Payments_per_Click   R-squared:                       
0.743
Model:                            OLS   Adj. R-squared:                  
0.475
Method:                 Least Squares   F-statistic:                     
2.770
Date:                Tue, 04 Feb 2020   Prob (F-statistic):             
0.00991
Time:                        13:59:15   Log-Likelihood:                 
126.94
No. Observations:                  46   AIC:                            
-205.9
Df Residuals:                      22   BIC:                            
-162.0
Df Model:                          23                                         
Covariance Type:            nonrobust                                         
=======================================================================
===================
                             coef    std err          t      P>|t|      
[0.025      0.975]
-----------------------------------------------------------------------
-------------------
Intercept                  0.0815      0.016      5.090      0.000       
0.048       0.115
C(Date)[T.Fri, Oct 24]     0.0791      0.022      3.569      0.002       
0.033       0.125
C(Date)[T.Fri, Oct 31]     0.0777      0.022      3.507      0.002       
0.032       0.124
C(Date)[T.Mon, Oct 13]     0.0179      0.022      0.809      0.427      
-0.028       0.064
C(Date)[T.Mon, Oct 20]     0.0343      0.022      1.548      0.136      
-0.012       0.080
C(Date)[T.Mon, Oct 27]     0.0799      0.022      3.604      0.002       
0.034       0.126
C(Date)[T.Sat, Nov 1]      0.0612      0.022      2.763      0.011       
0.015       0.107
C(Date)[T.Sat, Oct 11]    -0.0033      0.022     -0.148      0.884      
-0.049       0.043
C(Date)[T.Sat, Oct 18]     0.0239      0.022      1.080      0.292      
-0.022       0.070
C(Date)[T.Sat, Oct 25]     0.0742      0.022      3.347      0.003       
0.028       0.120
C(Date)[T.Sun, Nov 2]      0.0405      0.022      1.826      0.081      
-0.005       0.086
C(Date)[T.Sun, Oct 12]     0.0239      0.022      1.078      0.293      
-0.022       0.070
C(Date)[T.Sun, Oct 19]     0.0196      0.022      0.887      0.385      
-0.026       0.066
C(Date)[T.Sun, Oct 26]     0.0673      0.022      3.038      0.006       
0.021       0.113
C(Date)[T.Thu, Oct 16]     0.0095      0.022      0.430      0.672      
-0.036       0.055
C(Date)[T.Thu, Oct 23]     0.0161      0.022      0.725      0.476      
-0.030       0.062
```

```
C(Date)[T.Thu, Oct 30]     0.0181      0.022      0.817      0.423
-0.028       0.064
C(Date)[T.Tue, Oct 14]     0.0394      0.022      1.779      0.089
-0.007       0.085
C(Date)[T.Tue, Oct 21]     0.0226      0.022      1.022      0.318
-0.023       0.069
C(Date)[T.Tue, Oct 28]     0.0643      0.022      2.904      0.008
0.018       0.110
C(Date)[T.Wed, Oct 15]     0.0157      0.022      0.709      0.486
-0.030       0.062
C(Date)[T.Wed, Oct 22]     0.0196      0.022      0.884      0.386
-0.026       0.066
C(Date)[T.Wed, Oct 29]     0.0452      0.022      2.040      0.054
-0.001       0.091
treatment                 -0.0049      0.007     -0.750      0.461
-0.018       0.009
=======================================================================
=======
Omnibus:                        4.114   Durbin-Watson:
1.713
Prob(Omnibus):                  0.128   Jarque-Bera (JB):
1.796
Skew:                           0.000   Prob(JB):
0.407
Kurtosis:                       2.032   Cond. No.
27.3
=======================================================================
=======

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
Intercept                 0.000042
C(Date)[T.Fri, Oct 24]    0.001717
C(Date)[T.Fri, Oct 31]    0.001990
C(Date)[T.Mon, Oct 13]    0.426989
C(Date)[T.Mon, Oct 20]    0.135859
C(Date)[T.Mon, Oct 27]    0.001576
C(Date)[T.Sat, Nov 1]     0.011346
C(Date)[T.Sat, Oct 11]    0.883664
C(Date)[T.Sat, Oct 18]    0.292025
C(Date)[T.Sat, Oct 25]    0.002919
C(Date)[T.Sun, Nov 2]     0.081384
C(Date)[T.Sun, Oct 12]    0.292706
C(Date)[T.Sun, Oct 19]    0.384842
C(Date)[T.Sun, Oct 26]    0.006036
C(Date)[T.Thu, Oct 16]    0.671727
C(Date)[T.Thu, Oct 23]    0.476235
C(Date)[T.Thu, Oct 30]    0.422735
C(Date)[T.Tue, Oct 14]    0.089062
C(Date)[T.Tue, Oct 21]    0.318067
C(Date)[T.Tue, Oct 28]    0.008233
C(Date)[T.Wed, Oct 15]    0.485587
C(Date)[T.Wed, Oct 22]    0.386368
C(Date)[T.Wed, Oct 29]    0.053501
treatment                 0.461482
dtype: float64
```

After including the Date as a predictor variable to our model, the standard error dropped from 0.009 to 0.007, which is about a 22% decrease.

## Exercise 12

Udacity's trial was not statistically effective/significant. Even after including the date as a predictor, the p-value for treatment was still above 0.05 which is a standard significance level. There is no evidence to suggest that the treatment increased payments after clicking "Start Free Trial."

## Exercise 13

```
In [10]: # Look at datatypes of objects
         combined_new.dtypes
```

```
Out[10]: Date                      object
         Pageviews                  int64
         Clicks                     int64
         Enrollments              float64
         Payments                 float64
         treatment                  int64
         Payments_per_Click       float64
         dtype: object
```

In [11]:

```
# Create day column and dictionary
combined_new['Day'] = combined_new['Date'].str[:3]
word_dict = {'Mon':'Monday','Tue':'Tuesday','Wed':'Wednesday','Thu':'Thu
rsday',
             'Fri':'Friday','Sat':'Saturday','Sun':'Sunday'}

combined_new['Day'] = combined_new['Day'].replace(word_dict)
combined_new.head()

# Create model with day of the week
# Create regression model accounting for date
model2=smf.ols(formula='Payments_per_Click~treatment+C(Day)', data=combi
ned_new)
res2=model2.fit()
print(res2.summary())
print(res2.pvalues)
```

```
                            OLS Regression Results
=======================================================================
=======
Dep. Variable:     Payments_per_Click   R-squared:
0.138
Model:                            OLS   Adj. R-squared:
-0.021
Method:                 Least Squares   F-statistic:
0.8687
Date:                Tue, 04 Feb 2020   Prob (F-statistic):
0.540
Time:                        13:59:59   Log-Likelihood:
99.073
No. Observations:                  46   AIC:
-182.1
Df Residuals:                      38   BIC:
-167.5
Df Model:                           7
Covariance Type:            nonrobust
=======================================================================
================
                          coef    std err          t      P>|t|
[0.025      0.975]
-----------------------------------------------------------------------
----------------
Intercept               0.1337      0.013      9.971      0.000
0.107       0.161
C(Day)[T.Monday]       -0.0082      0.018     -0.461      0.647      -
0.044       0.028
C(Day)[T.Saturday]     -0.0133      0.017     -0.794      0.432      -
0.047       0.021
C(Day)[T.Sunday]       -0.0144      0.017     -0.865      0.393      -
0.048       0.019
C(Day)[T.Thursday]     -0.0377      0.018     -2.113      0.041      -
0.074      -0.002
C(Day)[T.Tuesday]      -0.0101      0.018     -0.568      0.574      -
0.046       0.026
C(Day)[T.Wednesday]    -0.0254      0.018     -1.425      0.162      -
0.062       0.011
treatment              -0.0049      0.009     -0.538      0.594      -
0.023       0.014
=======================================================================
=======
Omnibus:                        0.703   Durbin-Watson:
0.986
Prob(Omnibus):                  0.704   Jarque-Bera (JB):
0.237
Skew:                          -0.155   Prob(JB):
0.888
Kurtosis:                       3.168   Cond. No.
9.27
=======================================================================
=======

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.
```

```
Intercept                 3.701754e-12
C(Day)[T.Monday]          6.472255e-01
C(Day)[T.Saturday]        4.319236e-01
C(Day)[T.Sunday]          3.925824e-01
C(Day)[T.Thursday]        4.118523e-02
C(Day)[T.Tuesday]         5.735312e-01
C(Day)[T.Wednesday]       1.622519e-01
treatment                 5.940538e-01
dtype: float64
```

When we use day of the week, the p-value for treatment is 0.594, which is still not statistically significant. It does not appear that Udacity's treatment was effective.