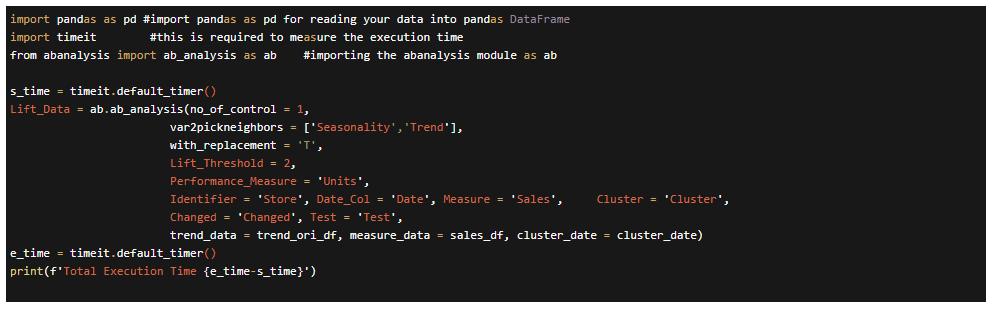
AB Analysis - Python Package

User Guide



Steps:

1. To install the ab\_analysis package follow the steps below –
2. Extract the .rar file and keep the folder in your local system
3. Go to cmd => type   
   **pip install \\...\ab\_analysis\_version\_1\_0\_2** mention the complete path of zip extract  
   *Example*,  
   **pip install C:\Users\Abhishek.Kumar\Desktop\Python\_Modules\ab\_analysis\ab\_analysis\_version\_1\_0\_2**
4. Now you should have abanalysis module available for use.

#Below is the example of test use of the abanalysis module

Calling AB Module with data

**Lift\_Data = ab.ab\_analysis(**

**no\_of\_control = 10,**

**var2pickneighbors = ['Seasonality','Trend'],**

**with\_replacement = 'T',**

**Lift\_Threshold = 2,**

**Performance\_Measure = 'Units',**

**Identifier = 'Store', Date\_Col = 'Date', Measure = 'Sales',**

**Cluster = 'Cluster', Changed = 'Changed', Test = 'Test',**

**trend\_data = trend\_ori\_df, measure\_data = sales\_df, cluster\_date = cluster\_date)**

**User Provided Values / Data**

* *no\_of\_control* = # ( Provide the number of Controls) , # = Control per Treatment Store
* *with\_replacement* = T or F.
* *Performance\_Measure* = Units (Constant Values in the Output)
* *Lift\_Threshold* = # (assign number, so while calculating. the lift at store level, it gives lift no. and threshold is manually assigned.)
* *Identifier* = Store (Map the Store Column from Trend Data table)
* *Date\_Col* = Date (Map Date Column)
* *Measure* = Sales (Map Sales column from Sales data table)
* Cluster = Cluster (Column name should be same as the Cluster Data and Trend Data)
* Lift\_Data (all the results are saves under this name which can be defined as per the user)

**User input data: trend data / Sales data / Cluster data (Sample Data)**

* Avoid using ( \_df ) name like trend\_data = trend\_df or measure\_data = sales\_df , instead use trend\_data = trend\_ori\_df
* Data structure of the data should be same. (Column names)
* Use data format as yyyy-mm-dd
* Cluster Column name of **Cluster Date** table should be same as Cluster Column of **Trend Data** table (user can assign other column names as well but both the table should have same column name)
* Cluster Date
* Use the template for saving dates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Cluster** | **Test\_StartDate** | **Test\_EndDate** | **Prior\_StartDate** | **Prior\_EndDate** | **YoY\_StartDate** | **YoY\_EndDate** |
| A | 9/10/2017 | 10/7/2017 | 6/18/2017 | 9/9/2017 | 9/11/2016 | 10/8/2016 |
| B | 9/10/2017 | 10/7/2017 | 6/18/2017 | 9/9/2017 | 9/11/2016 | 10/8/2016 |
| C | 9/10/2017 | 10/7/2017 | 6/18/2017 | 9/9/2017 | 9/11/2016 | 10/8/2016 |
| All | 9/10/2017 | 10/7/2017 | 6/18/2017 | 9/9/2017 | 9/11/2016 | 10/8/2016 |

* Trend Data
* Avoid using Store values as Integer (1,2, etc ) . Using store no. as (S001, S002. etc.) is recommended

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Store** | **Trend** | **Seasonality** | **Test** | **Cluster** | **Changed** |
| 1 | 0.005192776 | -0.002199118 | FALSE | B | FALSE |
| 2 | 0.066038106 | -0.000428914 | FALSE | B | TRUE |

* Test / Changed = Boolean
* Trend / Seasonality = Integer
* Cluster = String
* Sales Data

|  |  |  |
| --- | --- | --- |
| **Store** | **Date** | **Sales** |
| 1 | 5/24/2015 | 74468.05 |
| 1 | 5/31/2015 | 87829.99 |
| 1 | 6/7/2015 | 88058.13 |