**Python Script for scrapping Openinsider.com**

[**Append A with Email Structure**](#Appendix_A)

1. **Control.csv**

Control.csv will provide Controlling parameters.

|  |  |  |
| --- | --- | --- |
| 1 | Officer Min. in Thousands | 100 |
| 2 | Director, 10% & Others Min. in Thousands | 250 |
| 3 | Min. Volume in Thousands | 500 |
| 4 | Max. % Discount received by Insiders | 10% |
| 5 | Email SMTP Server Name / IP Address | 192.168.68.204 |
| 6 | Email SMTP Server Port | 25 |
| 7 | Email SMTP ID | [send@vj.loc](mailto:send@vj.loc) |
| 8 | Email SMTP Password | 123456 |
| 9 | Require logon using Secure Password Authentication (SPA) | No |
| 10 | Use following type of Encrypted Connection | None |
| 11 | Email TO Email ID | task18327@gmail.com |
| 12 | From | OpenInsider |
| 13 | News Excel File Path | I:\investment\1\_ShortCuts\Python\_Programs\_New\Latest\_Used\_Executables\FinViz\Data |
| 14 | News Excel File Name | News\_Excel\_File.xls |
| 15 | Output.csv File Path | I:\investment\1\_ShortCuts\Python\_Programs\_New\Latest\_Used\_Executables\OpenInsider\1\_Daily\Data |

1. **URL;**

Officer 100K ;

[http://openinsider.com/screener?s=&o=&pl=&ph=&ll=&lh=&fd=1&fdr=&td=0&tdr=&fdlyl=&fdlyh=&daysago=&xp=1&excludeDerivRelated=1&vl=**100**&vh=&ocl=&och=&sic1=-1&sicl=100&sich=9999&isofficer=1&iscob=1&isceo=1&ispres=1&iscoo=1&iscfo=1&isgc=1&isvp=1&grp=0&nfl=&nfh=&nil=&nih=&nol=&noh=&v2l=&v2h=&oc2l=&oc2h=&sortcol=0&cnt=100&page=1](http://openinsider.com/screener?s=&o=&pl=&ph=&ll=&lh=&fd=1&fdr=&td=0&tdr=&fdlyl=&fdlyh=&daysago=&xp=1&excludeDerivRelated=1&vl=100&vh=&ocl=&och=&sic1=-1&sicl=100&sich=9999&isofficer=1&iscob=1&isceo=1&ispres=1&iscoo=1&iscfo=1&isgc=1&isvp=1&grp=0&nfl=&nfh=&nil=&nih=&nol=&noh=&v2l=&v2h=&oc2l=&oc2h=&sortcol=0&cnt=100&page=1)

Director, 10% & Others; 250K

[http://openinsider.com/screener?s=&o=&pl=&ph=&ll=&lh=&fd=1&fdr=&td=0&tdr=&fdlyl=&fdlyh=&daysago=&xp=1&excludeDerivRelated=1&vl=**250**&vh=&ocl=&och=&sic1=-1&sicl=100&sich=9999&isdirector=1&istenpercent=1&isother=1&grp=0&nfl=&nfh=&nil=&nih=&nol=&noh=&v2l=&v2h=&oc2l=&oc2h=&sortcol=0&cnt=100&page=1](http://openinsider.com/screener?s=&o=&pl=&ph=&ll=&lh=&fd=1&fdr=&td=0&tdr=&fdlyl=&fdlyh=&daysago=&xp=1&excludeDerivRelated=1&vl=250&vh=&ocl=&och=&sic1=-1&sicl=100&sich=9999&isdirector=1&istenpercent=1&isother=1&grp=0&nfl=&nfh=&nil=&nih=&nol=&noh=&v2l=&v2h=&oc2l=&oc2h=&sortcol=0&cnt=100&page=1)

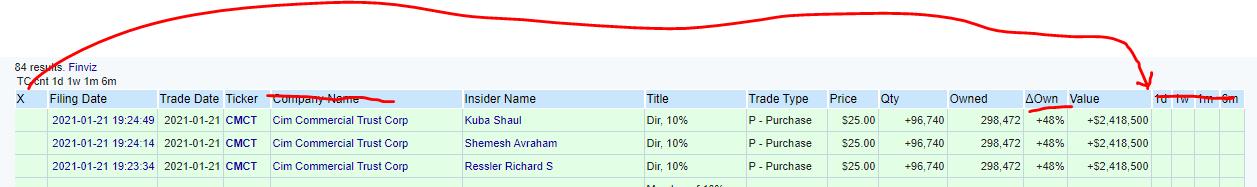
**Control.csv Note**; 100 & 250 values will be provided by Control.csv labels; “Officer Min. in Thousands” & “Director, 10% & Others Min. in Thousands”. These values will be used to substitute the values in URL’s.

**Website Note;** Above URL’s are basically 2 different search schemes on the same web page. Some of the rows may appear in both searches. This is because of the fact that some of the officers (buyers) also fall under the Others category too. Output.csv should only contain unique rows. There should be only one email sent in case of duplicate results (under the category of Officer). That means, if we get a new row in the first search and get the same row in second search, we will disregard the row, obtained in the second search.

1. **Columns;**

### 1; Filing Date, 2; Trade Date, 3; Ticker, 4; Insider Name, 5; Title, 6; Trade Type, 7; Price, 8; Qty, 9; Owned, 10; Own Inc, 11; Value, 12; X

### Filing Date, Trade Date, Ticker, Insider Name, Title, Trade Type, Price, Qty, Owned, Own Inc, Value, X



1. **Steps.**
   1. Insert new rows (not already there in Output.csv) into Output.csv located in “Output.csv File Path” value in Control.csv. New rows should be added at the top. If new row found. (Format of Output.csv in Appendix B)
   2. Open News\_Excel\_File.xls.
   3. If Ticker found in first column (Column A) in News\_Excel\_File.xlsx
   4. If “Volume T” in J Column (10th) Greater than **500** in News\_Excel\_File.xlsx
   5. If Price in Column 7 (webpage) is not less than **90%** (1-10%)\* “Price” in News\_Excel\_File.xlsx in price column AD (Column 30th)
   6. Send an email using **credentials** in Control.csv (Email Format Example in [Appendix A](#Appendix_A))

**Email Subject**; Officer / Other based on URL used, Ticker from website, Value from website, Column X value from website, PerfWeek; Column G value, Days; Column Q value, Short; Column R Value, OPT; Column AE Value

**Email Body**;

Reformatted relevant row (as shown in [2. Columns](#Col_2)) from website & New\_Excel\_File.xlsx with column headings in Table with border format.

**Control.csv Note**;

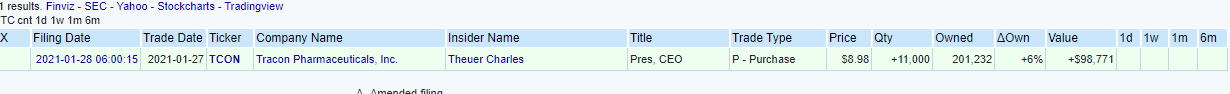
* File Path & File name for New Excel File will be provided by values in the rows 13 & 14 in Control.csv
* Value for Min. Volume (500) in 4.3 will be provided by “Min. Volume in Thousands” label in Control.csv
* Value for Max. discount (10%) in 4.4 will be provided by “Max. % Discount received by Insiders” label in Control.csv
* Email credentials will be provided by labels in rows 5 to 12 in Control.csv

**News\_Excel\_File.xls Note**; News\_Excel\_File.xls format can change in future. However, first column Ticker, will always be there as first column. Please develop in order to pick everything for the email in table format. We are currently using “10 Volume T” (Column J) & “30 Price” (Column AD) in News\_Excel\_File.xls. Position and numeric portion of the Column headings can change in future. However, non-numeric label would remain the same.

**Appendix A**

**Email Structure**

Web results for modified Office query 25 instead of 100

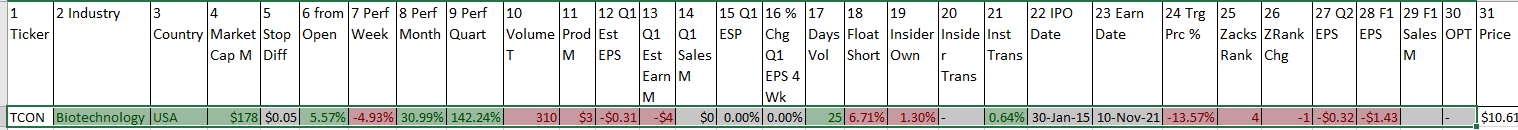


Subject;

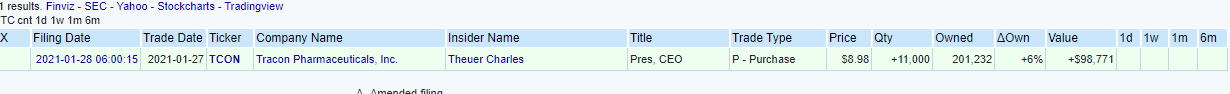
**Officer:TCON:$98,881:X;:PerfWeek;-4.93%:Days;25: Short;6.71%:OPT;**

Body;





**Appendix B**



|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Filing Date** | **Trade Date** | **Ticker** | **Company Name** | **Insider Name** | **Title** | **Trade Type** | **Price** | **Qty** | **Owned** | **Own Inc** | **Value** | **X** |
| 28-Jan-21 6:00:15 AM | 27-Jan-21 | TCON | Tracon Pharmaceuticals Inc. | Theuer Charles | Pres, CEO | P - Purchase | $8.98 | 11,000 | 201,232 | 6% | 98,771 |  |

**Sample Output.csv attached**