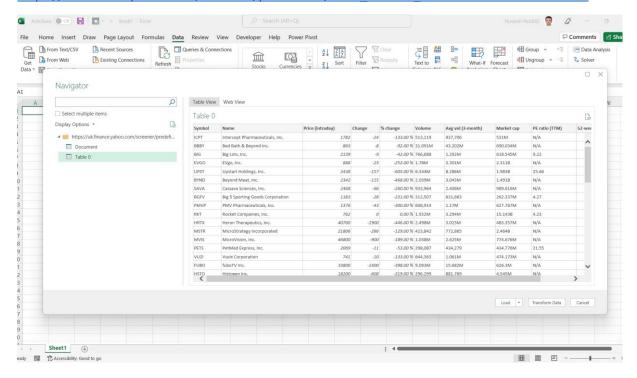
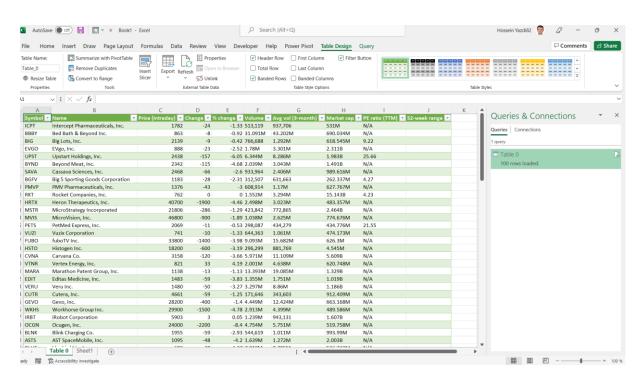
Step 1
The data is downloaded by "Excel/Data Tab/From the Web" from this link of Yahoo Finance: <a href="https://uk.finance.yahoo.com/screener/predefined/most\_shorted">https://uk.finance.yahoo.com/screener/predefined/most\_shorted</a> stocks?offset=0&count=100





## Cleaning, Wrangling

Deleting the "J" column as it is empty and useless.

Substitution of the "N/A" data in the "I" column to be zero as the relevant companies have no P/E ratio.

Symbol	Name	Price (intraday)	Change 💌	% change 💌 Volume 🗈	Avg vol (3-month)	Market cap	PE ratio (TTI ▼
ICPT	Intercept Pharmaceuticals, Inc.	1782	-24	-1.33 513,119	937,706	531M	0
BBBY	Bed Bath & Beyond Inc.	863	-8	-0.92 31.091M	43.202M	690.034M	0
BIG	Big Lots, Inc.	2139	-9	-0.42 766,688	1.292M	618.545M	9.22
EVGO	EVgo, Inc.	888	-23	-2.52 1.78M	3.301M	2.311B	0
UPST	Upstart Holdings, Inc.	2438	-157	-6.05 6.344M	8.286M	1.983B	25.66
BYND	Beyond Meat, Inc.	2342	-115	-4.68 2.039M	3.043M	1.491B	0
SAVA	Cassava Sciences, Inc.	2468	-66	-2.6 933,964	2.406M	989.616M	0
BGFV	Big 5 Sporting Goods Corporation	1183	-28	-2.31 312,507	631,663	262.337M	4.27
PMVP	PMV Pharmaceuticals, Inc.	1376	-43	-3 608,914	1.17M	627.767M	0
RKT	Rocket Companies, Inc.	762	0	0 1.552M	3.294M	15.143B	4.23
HRTX	Heron Therapeutics, Inc.	40700	-1900	-4.46 2.498M	3.023M	483.357M	0
MSTR	MicroStrategy Incorporated	21806	-286	-1.29 423,842	772,865	2.464B	0
MVIS	MicroVision, Inc.	46800	-900	-1.89 1.038M	2.625M	774.676M	0
PETS	PetMed Express, Inc.	2069	-11	-0.53 298,087	434,279	434.776M	21.55
VUZI	Vuzix Corporation	741	-10	-1.33 644,363	1.061M	474.173M	0
FUBO	fuboTV Inc.	33800	-1400	-3.98 9.093M	15.682M	626.3M	0
HSTO	Histogen Inc.	18200	-600	-3.19 296,299	881,769	4.545M	0
CVNA	Carvana Co.	3158	-120	-3.66 5.971M	11.109M	5.609B	0
VTNR	Vertex Energy, Inc.	821	33	4.19 2.001M	4.638M	620.748M	0
MARA	Marathon Patent Group, Inc.	1138	-13	-1.13 13.393M	19.085M	1.329B	0
EDIT	Editas Medicine, Inc.	1483	-59	-3.83 1.355M	1.751M	1.019B	0
VERU	Veru Inc.	1480	-50	-3.27 3.297M	8.86M	1.186B	0
CUTR	Cutera, Inc.	4661	-59	-1.25 171,646	343,603	912.409M	0
GEVO	Gevo, Inc.	28200	-400	-1.4 4.449M	12.424M	663.168M	0
WKHS	Workhorse Group Inc.	29900	-1500	-4.78 2.913M	4.399M	489.586M	0
IRBT	iRobot Corporation	5903	3	0.05 1.239M	943,131	1.607B	0
OCGN	Ocugen, Inc.	24000	-2200	-8.4 4.754M	5.751M	519.758M	0
BLNK	Blink Charging Co.	1955	-59	-2.93 544,619	1.011M	993.99M	0
ASTS	AST SpaceMobile, Inc.	1095	-48	-4.2 1.639M	1.272M	2.003B	0
		500			0.70514		

But there is a problem that some columns contain commas as thousand separators, "M" as millions, and "B" as billions so they must be changed.

To replace commas only in the selected columns from "F" to "I", in the replace box we use Ctrl+A.

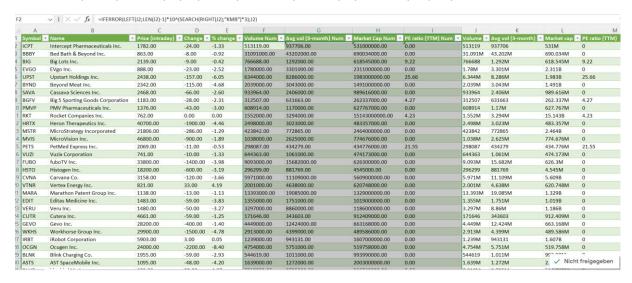
And to convert M and B to their numerical equals, after inserting new columns, the function below is helpful:

=LEFT(J2;LEN(J2)-1)\*10^(SEARCH(RIGHT(J2);"KMB")\*3)

This formula, as does not use absolute reference, can be expanded to all the columns both horizontally and vertically.

But we have another problem relating to those cells without M or B, returning the "Value" error. To solve it is possible to use "iferror" function.

=IFERROR(LEFT(J2;LEN(J2)-1)\*10^(SEARCH(RIGHT(J2);"KMB")\*3);J2)



Query: In what way and how much does the price as the dependent variable depends on the other independent variables, using the multiple linear regression?

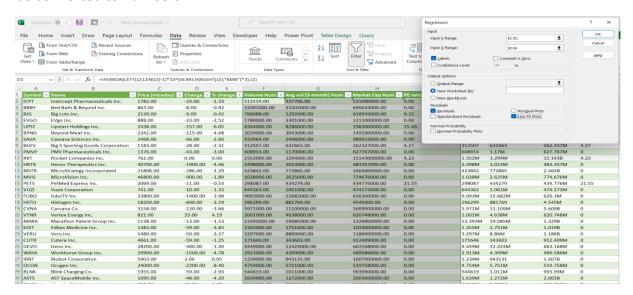
To do so, we use the "Data Tab/Analyze/Data Analysis/Regression".

Input Y range: price in column C

Input X range: other numerical columns

Selecting the options: Labels, Residuals, Line Fit Plots.

There are errors associated with the fact that some data are formatted as text which must be converted to numbers.



#### The result is as follows:

A	В	С	D	Е	F	G	Н	I
SUMMARY OUTPUT								
Regressio	n Statistics							
Multiple R	0.356471191							
R Square	0.12707171							
Adjusted R Square	0.070753756							
Standard Error	13005.08886							
Observations	100							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	6	2289706947	381617824.5	2.256326809	0.044558243			
Residual	93	15729307273	169132336.3					
Total	99	18019014220						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	11947.37381	2040.488754	5.855152983	7.11876E-08	7895.367543	15999.38008	7895.367543	15999.38008
Change	-8.089344025	2.704912848	-2.990611705	0.003563215	-13.46076487	-2.717923183	-13.46076487	-2.717923183
% change	1236.181335	627.3600924	1.970449427	0.051760789	-9.631475294	2481.994144	-9.631475294	2481.994144
Volume Num	1.84724E-05	0.000351175	0.052601806	0.958162003	-0.000678891	0.000715836	-0.000678891	0.000715836
Avg vol (3-month) Num	4.83629E-05	0.000295652	0.163580095	0.870416678	-0.000538744	0.00063547	-0.000538744	0.00063547
Market Cap Num	-6.30463E-07	5.16262E-07	-1.221208222	0.225094077	-1.65566E-06	3.94731E-07	-1.65566E-06	3.94731E-07
PE ratio (TTM) Num	-109.7930088	110.2026053	-0.996283241	0.321697728	-328.6335423	109.0475247	-328.6335423	109.0475247

#### Interpretation

The figures for both "R Square" and "Adjusted R Square" are very low, indicating the fact of not a sufficient correlation between the independent and dependent variables.

The function is 
$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6$$
  
 $Y = 11947.37381 + -8.089344025 X_1 + 1236.181335 X_2 + 1.84724E-05 X_3 + 4.83629E-05 X_4 + -6.30463E-07 X_5 + -109.7930088 X_6$ 

If we check the formula by use of some data from the table, it shows that the prediction is not fine, meaning that the data for "Most Shorted Stocks of screeners" are not so much correlated in total, again annotating not to be good enough predictable.

Assessing the 1 minus P-Value for all elements shows that not all of them are greatly significant.

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%	1-Pvalue
Intercept	11947.37381	2040.488754	5.855152983	7.1188E-08	7895.367543	15999.38008	7895.367543	15999.38008	0.99999999
Change	-8.089344025	2.704912848	-2.99061171	0.00356322	-13.46076487	-2.717923183	-13.4607649	-2.71792318	0.996436785
% change	1236.181335	627.3600924	1.970449427	0.05176079	-9.631475294	2481.994144	-9.63147529	2481.994144	0.948239211
Volume Num	1.84724E-05	0.000351175	0.052601806	0.958162	-0.000678891	0.000715836	-0.00067889	0.000715836	0.041837997
Avg vol (3-month) Num	4.83629E-05	0.000295652	0.163580095	0.87041668	-0.000538744	0.00063547	-0.00053874	0.00063547	0.129583322
Market Cap Num	-6.30463E-07	5.16262E-07	-1.22120822	0.22509408	-1.65566E-06	3.94731E-07	-1.6557E-06	3.94731E-07	0.774905923
PE ratio (TTM) Num	-109.7930088	110.2026053	-0.99628324	0.32169773	-328.6335423	109.0475247	-328.633542	109.0475247	0.678302272