

COMPSCI 2034b / DIGIHUM 2144b
Data Analytics: Principles and Tools

Assignment #1

Market Analysis



Western
UNIVERSITY • CANADA

Posted:	January 22nd 2024
Due:	February 7th 2023 11:55PM
Total:	100 Points (7% of Final Grade)

Learning Outcomes

By completing this assignment, you will gain and demonstrate skills relating to:

- Retrieving web data.
- Formatting spreadsheets.
- Computing basic statistics using Excel.
- Using conditional statements in Excel formulas.
- Performing an RSI Analysis.

Instructions

In this assignment, you will download historical stock market data from [Yahoo Finance](#) and import it into Excel. You will then format the data, perform a series of calculations on the data and a Relative Strength Index (RSI) Analysis.

You are required to follow each step in this assignment and submit an Excel .xlsx copy of your assignment. You must assume that the data in your sheet can change (i.e. you may not hardcode your answers). Each step must be followed precisely including the file naming convention given in Part 9. You will be assessed on the following:

- Retrieving the correct data.
- Your formatting of the spreadsheet.
- Your Excel formulas.
- Your final analysis.
- Completion of each task correctly.
- Assignment submission via OWL.

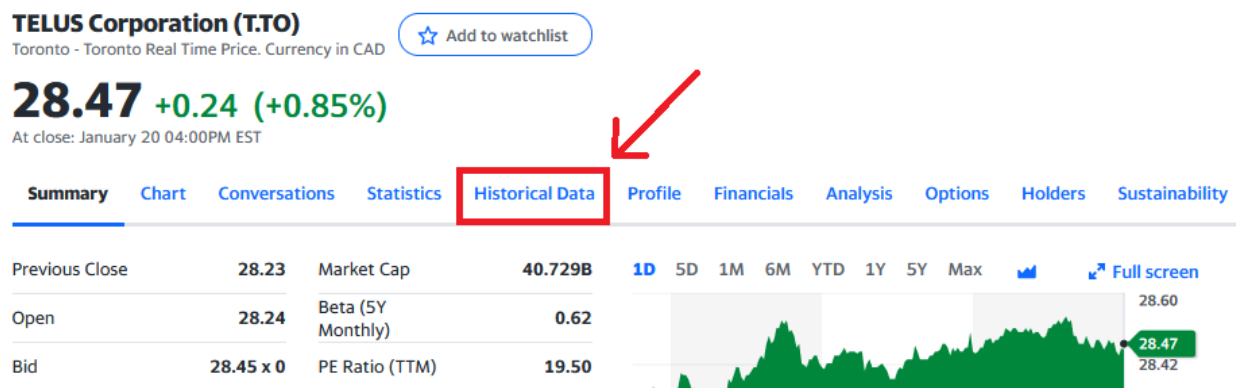
Tasks

1. Get The Historical Data For Your Stock (5 Marks)

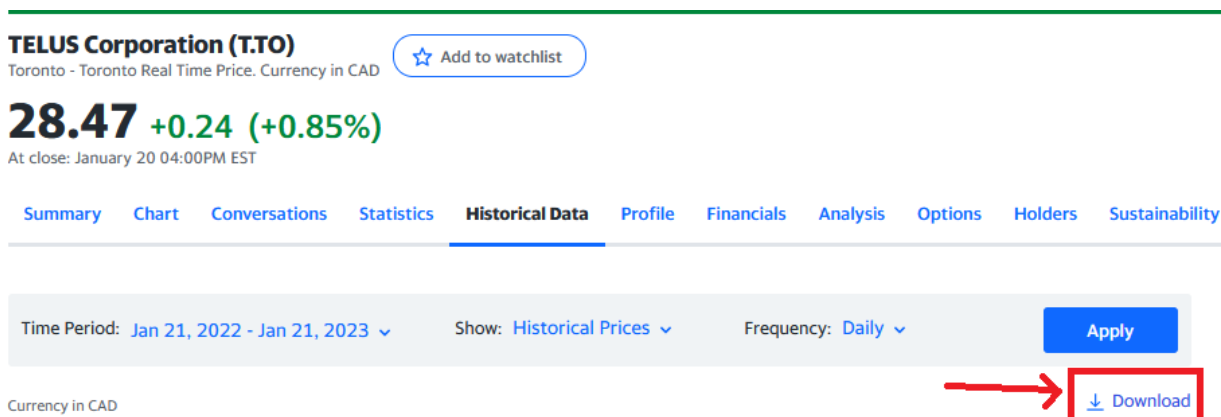
[Yahoo Finance](#) provides historical stock data for most stocks trading on the New York Stock Exchange (NYSE), Toronto Stock Exchange, as well as many other markets including some cryptocurrencies. Visit the following link and select a **stock** of your choosing that has **at least one year of historical data** (*you can not pick Telus as it is used in the examples in this assignment*). For example you could pick IBM (for International Business Machines Corporation (IBM)), SHOP.TO (for Shopify Inc.), TSLA (for Tesla), or AMZN (for Amazon). The choice is up to you and it can be any **stock** (*except for Telus*) on Yahoo Finance with at least 1 year of historical data. It can **not** be a cryptocurrency, bond, Exchange-Traded Fund (ETF), or any security other than a stock.

<https://ca.finance.yahoo.com/>

Enter the company name or stock symbol into the search bar and then select the historical data tab as shown below:



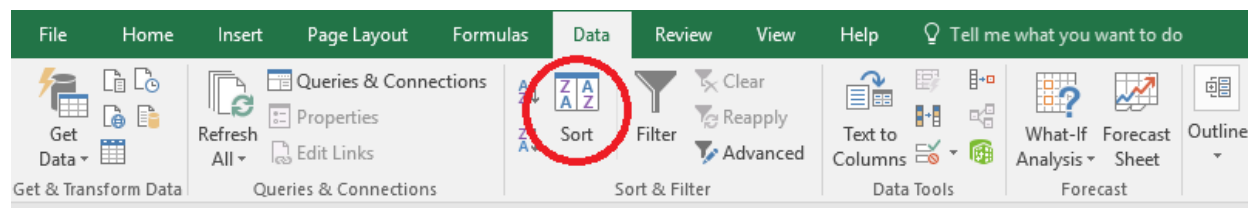
Make sure you have a years worth of data selected (for example, Jan 22 2023 to Jan 22 2024) and click the Apply button. Then select "Download" to get a Comma-Separated Value (CSV) export of the data.



Open the downloaded CSV file in Excel and **immediately save it as an Excel Workbook (*.xlsx) file**. Failing to save as a .xlsx file and continuing to work on the file as a .csv file can lead to data loss.

2. Sort the Data by Date

Highlight all of the data (including headers) in your table, right click on the "Sort" button on the Data tab. It should look like this:



Sort the data by date from oldest to newest. Make sure you have “My data has headers” checked.

3. Format the Data Nicely

Adjust the column widths so that all data is visible, i.e. there are no cells showing “#####”. Adjust the precision of all the cells in the “Open”, “High”, “Low”, “Close”, and “Adj Close” columns to show two decimal places and format them as currency. Format the dates to be in “Month Day, Year” style (e.g. January 1st, 2023 for 2023-01-01). Format the column headers by making them centered, in bold, with a grey background. Put black grid lines around the entries in the table. **Insert 6 blank lines above the table and 1 new column before column A.**

Add a title in cell A1 with the text “[Stock] RSI Analysis by [Your Name]” in a large font and bold. Of course “[Your Name]” should be replaced with your name, and [Stock] should be replaced with the name of the stock you are exploring. For example, if your name is Daniel Servos and the stock is Telus your title would be “Telus RSI Analysis by Daniel Servos”. Merge and center A1 with cells A1 to R1 to make the title centred across the first row.

Make the title **bold**, *italic*, *blue*, any font other than the default, and font size 18.

It should look something like the following screen shot:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	<i>Telus RSI Analysis by Daniel Servos</i>																		
2																			
3																			
4																			
5																			
6																			
7		Date	Open	High	Low	Close	Adj Close	Volume											
8		January 20, 2023	\$28.24	\$28.54	\$28.13	\$28.47	\$26.88	3539400											
9		January 23, 2023	\$28.50	\$28.60	\$28.24	\$28.48	\$26.89	1396200											
10		January 24, 2023	\$28.35	\$28.75	\$28.35	\$28.52	\$26.92	2537700											
11		January 25, 2023	\$28.38	\$28.58	\$28.20	\$28.48	\$26.89	1610000											
12		January 26, 2023	\$28.46	\$28.62	\$28.37	\$28.55	\$26.95	3453700											
13		January 27, 2023	\$28.50	\$28.67	\$28.33	\$28.40	\$26.81	2484700											
14		January 30, 2023	\$28.35	\$28.65	\$28.35	\$28.57	\$26.97	3145400											
15		January 31, 2023	\$28.63	\$28.81	\$28.35	\$28.67	\$27.06	2522000											
16		February 1, 2023	\$28.65	\$28.96	\$28.52	\$28.94	\$27.32	3113500											
17		February 2, 2023	\$28.95	\$29.00	\$28.66	\$28.72	\$27.11	2833500											
18		February 3, 2023	\$28.70	\$28.86	\$28.47	\$28.62	\$27.02	2866300											
19		February 6, 2023	\$28.53	\$28.58	\$28.35	\$28.47	\$26.88	2107800											
20		February 7, 2023	\$28.43	\$28.48	\$28.04	\$28.07	\$26.50	3190300											
21		February 8, 2023	\$28.06	\$28.15	\$27.87	\$28.09	\$26.52	2051700											
22		February 9, 2023	\$27.48	\$27.75	\$26.81	\$27.03	\$25.52	8956700											
23		February 10, 2023	\$27.00	\$27.53	\$26.66	\$27.25	\$25.72	5176800											
24		February 13, 2023	\$27.20	\$27.42	\$27.04	\$27.27	\$25.74	3935900											
25		February 14, 2023	\$27.32	\$27.48	\$27.11	\$27.38	\$25.85	3325300											
26		February 15, 2023	\$27.22	\$27.45	\$27.07	\$27.24	\$25.84	2721400											

4. Compute the Range

Put the text “Range:” in cell A3. In cell B2 put the text “Low” and in cell C2 put the text “High”. In cell B3 put a formula that calculates the minimum value from the “Low” column. In cell C3 put a formula that calculates the maximum value from the “High” column.

Put the text “Range Date:” in cell A4. In cell B4 use the **INDEX** and **MATCH** functions to find the date with the smallest low. In cell C4 use the same functions to find the date with the largest high. Format the resulting values as dates.

Hint: If you fail to format the cells as dates **INDEX** and **MATCH** will return a number rather than a date.

Center the values in the cells B2 to C4.

5. Measure the Volatility

In column J, compute the Daily Logarithmic Return for each date by using equation 1:

$$DailyLogReturn = \ln\left(\frac{C_i}{C_{i-1}}\right) \quad (1)$$

where C_i is the closing price for the current day, C_{i-1} is the closing price for the previous day and \ln is the natural logarithm. There should be no value for the first row as there is no preceding day. Give this column the header “Daily Return” and format the values as a percentage with two decimal places. Also add borders as in Step 3 and make the width of column I about half it’s default size.

In cell E3 put the text “STDev:” and in cell F3 calculate the standard deviation of the daily returns. In cell E4 put the text “Volatility:” and compute $\sqrt{days} \times STDev$ in cell F4, where $days$ is the number of days in the year that your stock is traded (make this dynamic based on the number of dates we have in our data set and not a hardcoded value) and $STDev$ is a reference to the cell in which we calculated the standard deviation. This gives us the annualized historical volatility. Format the volatility value as a percentage with two decimal places.

Volatility is a technical indicator of the amount of uncertainty or risk about the size of changes in a currency, stock or other security’s value. A higher volatility means the price is more likely to change dramatically over a short time period. A lower volatility means the price is more stable and less likely to fluctuate dramatically in a short time period. For comparison, gold has an average annualized historical volatility of approximately 15% whereas Bitcoin has a volatility closer to 81%.

6. Relative Strength Index Analysis

Relative Strength Index (RSI) is a technical indicator that compares the magnitude of recent gains and losses over a set time to give an indication of the momentum of stock, currency or other security. It is used by traders to identify securities that are potentially overbought or oversold. In theory, an overbought security is overvalued and prime for a downwards reversal (the price will drop), likewise an oversold security is undervalued and prime for an upwards reversal (the price will rise).

The relative strength index is calculated using the following formula:

$$RSI = 100 - \frac{100}{1 + RS} \quad (2)$$

where RS is the relative strength and is calculated as follows:

$$RS = \frac{\text{Average gain of up periods}}{\text{Average loss of down periods}} \quad (3)$$

6.1. Find the Gains and Losses

To make this calculation in Excel we will use columns K and L to determine the Gain or Loss for the day. Give column K the header “Loss” and the second (column L) the header “Gain”. In the gains column enter an equation that will find the gain (today’s closing price minus yesterday’s closing price) if and only if the daily change is positive, otherwise the result will be zero. Similarly, in the loss column enter an equation that will find the loss (also based on today’s and yesterday’s closing price) if and only if the daily change is negative, otherwise the result will be zero. Make sure both gain and loss are positive numbers.

Hint: You will need to use the Excel *IF* function.

6.2. Compute the 14-Day Average Gains and Losses

Relative strength is traditionally calculated over a 14 day period, so in the next two columns (columns M and N) we will find the average loss and gain over a 14 day period. Add the headers “Avg Loss” and “Avg Gain” to these columns and format them appropriately. As we are calculating the average for the first 14 days, the first 14 rows in the table will be blank. In cell M22 find the average of the first 14 losses using the AVERAGE function. In cell N22 find the average of the first 14 losses using the AVERAGE function.

To calculate the averages for the remaining rows following this formula:

$$14 \text{ Day Average} = \frac{(\text{Last Average} \times 13) + \text{Today's Gain or Loss}}{14} \quad (4)$$

6.3. Compute the Relative Strength (RS) and Relative Strength Index (RSI)

Now that we have the average losses and gains we can calculate the relative strength (RS) by dividing the average gain by the average loss for that day (as shown in Equation 3). In column O add the header “RS”. In this column find the RS for each day by dividing the value you calculated for Avg Gain by the value you calculated for Avg Loss on the same row.

We can now finally compute the 14-day RSI for each date by following Equation 2. In column P add the header “RSI”. In this column calculate the RSI using Equation 2 and the RS value you calculated.

6.4. Find the Oversold and Overbought Days

Generally an RSI above 70 is considered to be overbought and a RSI below 30 to be oversold. In cell J3 enter the text “Oversold:” in cell J4 enter the text “Overbought:”. In cell K3 put the number 30, this will be the value over which we consider an RSI to indicate oversold. In cell K4 put the number 70, this will be the value under which we consider an RSI to indicate overbought.

In column Q add the header “Over Sold or Bought?”. In this column you will enter a formula that will return the text “Overbought” if the RSI for that row is over or equal to the value in cell K4 and put the “Oversold” if the RSI for that row is under or equal to the value in cell K3. Otherwise no value should be printed (**Hint:** an empty string, “”, will output no value). Ensure that your values are not hard coded and that the result will change if the values in K3, K4 or the original data changes. You should be able to copy and paste (or drag) your formula into any cell in the column and have it print the correct value.

Hint: You will need to either use nested *IF functions* or the *IFS function*.

Make the text in this column (the “Oversold” and “Overbought” outputs) bold, red, and centered.

7. On-Balance Volume

On-balance volume (OBV) is a technical trading momentum indicator that uses volume flow to predict changes in stock price. [You can read more about it here.](#)

In column R add the header “OBV” and enter the value 0 into cell R8. In this column in the remaining rows, starting in cell R9 write a formula for OBV using this equation:

$$OBV = OBV_{prev} + \begin{cases} volume, & \text{if } close > close_{prev} \\ 0, & \text{if } close = close_{prev} \\ -volume, & \text{if } close < close_{prev} \end{cases} \quad (5)$$

where:

OBV = Current on-balance volume level (for the current day).

OBV_{prev} = Previous on-balance volume level (for the previous day with stock data).

$volume$ = Latest trading volume amount (from column H for the current day).

$close$ = The closing price for the current day (from column F).

$close_{prev}$ = The closing price for the previous day with stock data.

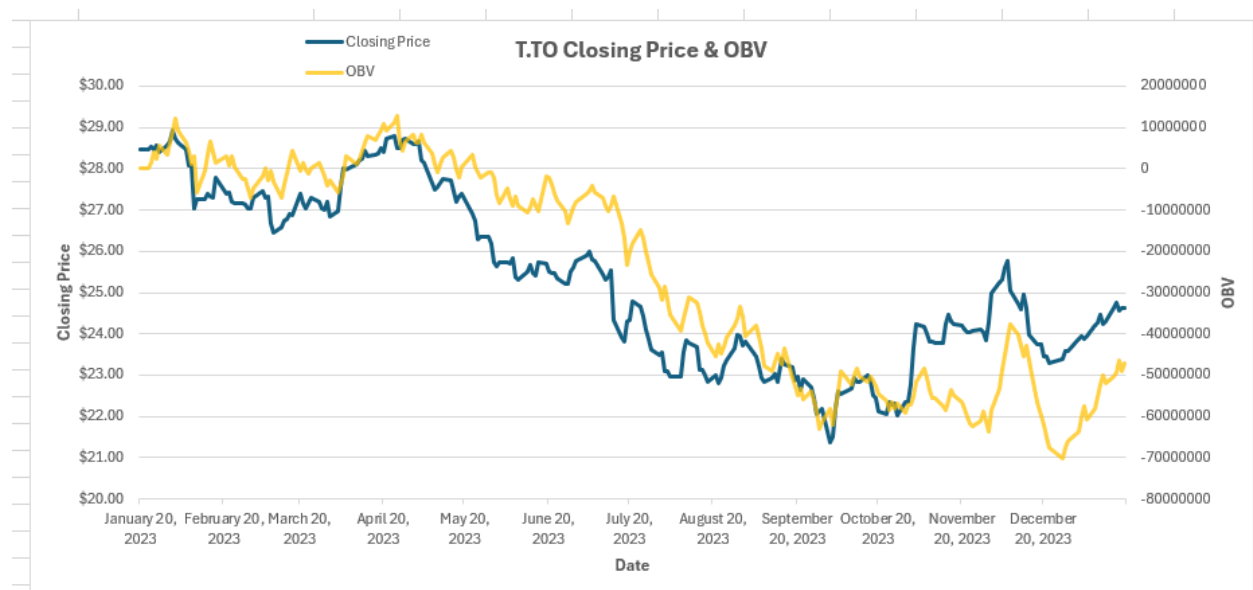
In simple terms this means you add the volume for the current day if the previous closing price was lower than today's and subtract it if it was less than today's. If they are the same, there is no change (add zero).

8. Analysis

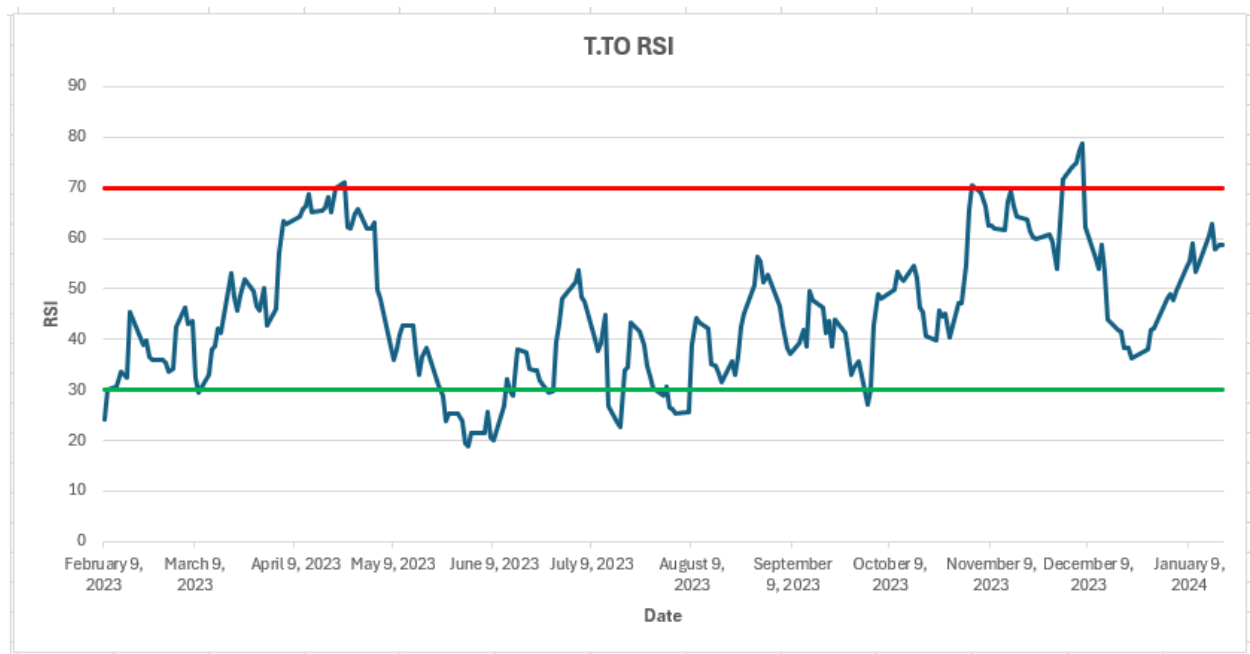
Name your current worksheet "RSI" and add a new worksheet named "Analysis". On this sheet add a line graph the shows a series for both Closing Price and OBV for each day on the same graph. The Closing Price and OBV should use different axes and your graph should include a legend, axis titles, and an overall title. Play around with the different formatting settings to create a good looking and clear visual.

You may have to research how to add axis titles, legends, series on different axes, etc. This is intentional and to encourage you to learn how to use the Excel documentation and online resources. This guide is a good starting point: [Present your data in a scatter chart or a line chart](#).

It should look something like this but for your own data:



Added a second line graph of the RSI for each day. Make sure the graph is properly titled and the axes are labelled. Add horizontal lines at 70 and 30 on the RSI graph to denote overbought and oversold levels. Note that this can be difficult in Excel. You may want to follow the [How to add a line to an existing Excel graph tutorial](#) but any way you can get the lines on the graph is acceptable. In our case we need to add two new columns to our data that simply contain the Overbought and Oversold values (one for each line) rather than a target value. The result should look something like this:



Finally, add a text box with your analysis of your stock as an investment. Base your analysis on the technical indicators we have calculated (namely Volatility, RSI, Overbought/sold, and OBV) and write a short (one or two paragraph) analysis of your stock as an investment. At a minimum you should discuss what the volatility means for potential investors, if the market is currently overbought, oversold or neither and what this could mean. If you are not familiar with these indicators you may first wish to read about them on Investopedia:

- [Historical Volatility](#)
- [Relative Strength Index - RSI](#)
- [Understanding Momentum Indicators and RSI](#)
- [Overbought Or Oversold? Using The RSI To Find Out](#)
- [Overbought](#)
- [Oversold](#)
- [On-Balance Volume \(OBV\): Definition, Formula, and Uses As Indicator](#)

9. Submission

Failing to follow the submission instructions can lead to a mark penalty. Marks will be deducted for incorrectly named files, files not submitted to OWL correctly, etc.

Before submitting take a look at the *example_soultion.pdf* file attached to the assignment on OWL. Make sure your formatting matches this document as closely as possible (fonts and colours do not have to match exactly but should be close). Note that this solution will have different data than you. Make sure all headings and columns are formatted as specified in Step 3 and that all values have a sensible format (currency, percentage, etc.). Also do one final check to make sure your numbers make sense and that you have not hardcoded any values that should be dynamic.

Save your Excel file as an .xlsx file and name it “*userid_assign1.xlsx*” where *userid* is your user id. For example, if your UWO e-mail was “*dservos5@uwo.ca*”, the file should be named “*dservos5_assign1.xlsx*”.

This must be an .xlsx workbook file created by a current version of Excel included in Office 365. No other file formats will be accepted. Submitting a .csv, pages, or Google Docs file will result in a zero grade. Also make sure that what you are submitting is a stand alone Excel workbook and not a link to online copy in OneDrive or Sharepoint. Submitting a link to a file or a file that is not self contained locally will result in a zero grade.

Make sure you leave enough time before the deadline to account for any technical issues. If you are having computer issues, the GenLabs computers on campus have Excel installed and can access OWL. You are required to backup your work (a remote backup on something like OneDrive is highly recommended). Lost work or technical issues will not be grounds for an extension.

All assignments are to be your own work. You may not work on this assignment with another person (including tutors), copy from an on-line source, or use any kind of tool to generate solutions for you. At no time should you ever share your work with another student or receive another student’s work. All submissions will be subject to similarity and metadata analysis.

Submit your work on the Assignments tab on our OWL course site. Submissions will not be accepted via e-mail or by any other means.