

Data Analysis Project with MS Excel

Title: Automating the update process of pulling data from websites on daily basis using power query in the BI tool advance excel.

By
Mamta M. Natu

Introduction

- ▶ Getting data from a web page allows users to extract data easily from web pages. But it is not easy to extract data from webpages, as data is not often in structured format, getting data from such pages can be challenging.
- ▶ Power Query helps in extracting the data from a web page; it allows the user to transform the data and load it into Excel as an excel table. Power Query automates the process of updating the data from web pages, which saves many resources.
- ▶ This project is about how to automate the process of updating data from a web page in an easy manner using the Power Query editor.

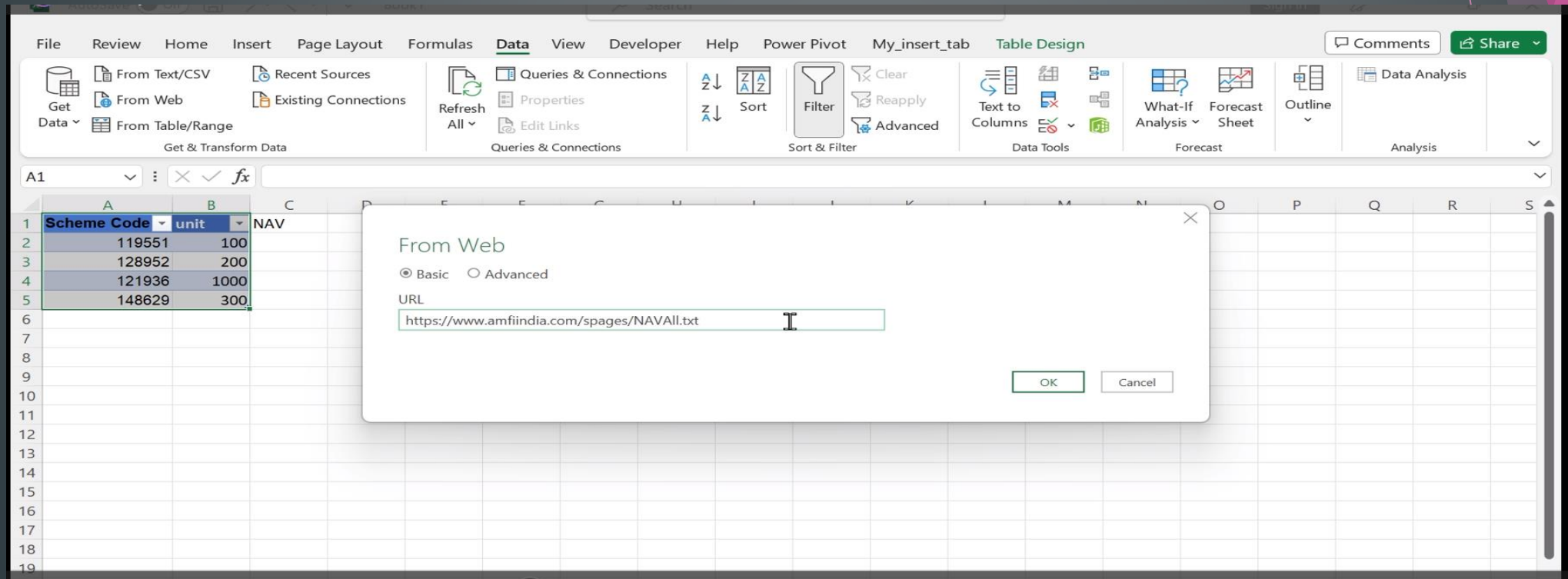
Problem statement

- ▶ A person who invests his money in the stock market is always interested in knowing the latest updates on the net asset value (NAV) of shares in his portfolio.
- ▶ He is using traditional methods of updating the NAV from the web pages of stocks using VLOOKUP. But the traditional method is very lengthy and needs continuous supervision. It includes the repetitive work of extracting data matching his portfolio and getting that NAV on a daily basis, which is quite time-consuming.
- ▶ Help him automate the whole process, which saves time and money. Use any BI tool for automation.
- ▶ Try to simplify the updating process.

Methodology:

Extract data: Using data connector in excel

1. Select **Get Data > Web** in Power BI or **From Web** in the **Data** ribbon in Excel.
2. Choose the **Basic** button and enter a URL address in the text box.
3. Select the authentication method to use for this web site



Extracting data from web page

From the **Navigator** dialog, you can select a table, then either transform the data in the Power Query editor by selecting **Transform Data**, or load the data by selecting **Load**.

https://www.amfiindia.com/spages/NAVAll.txt

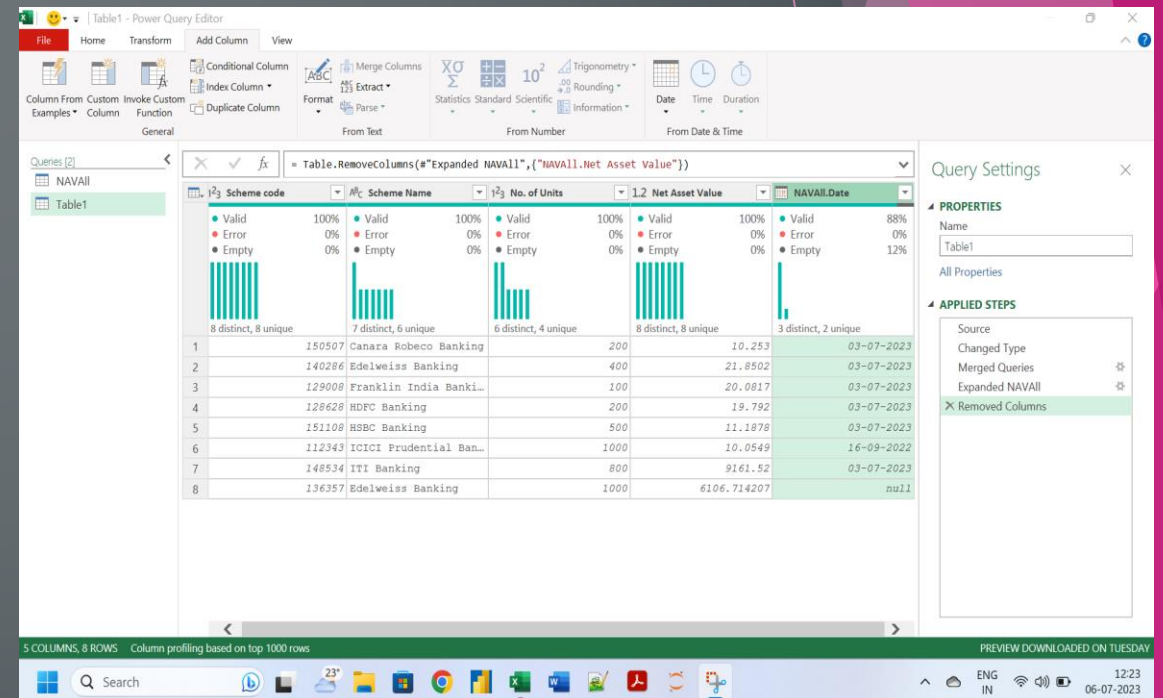
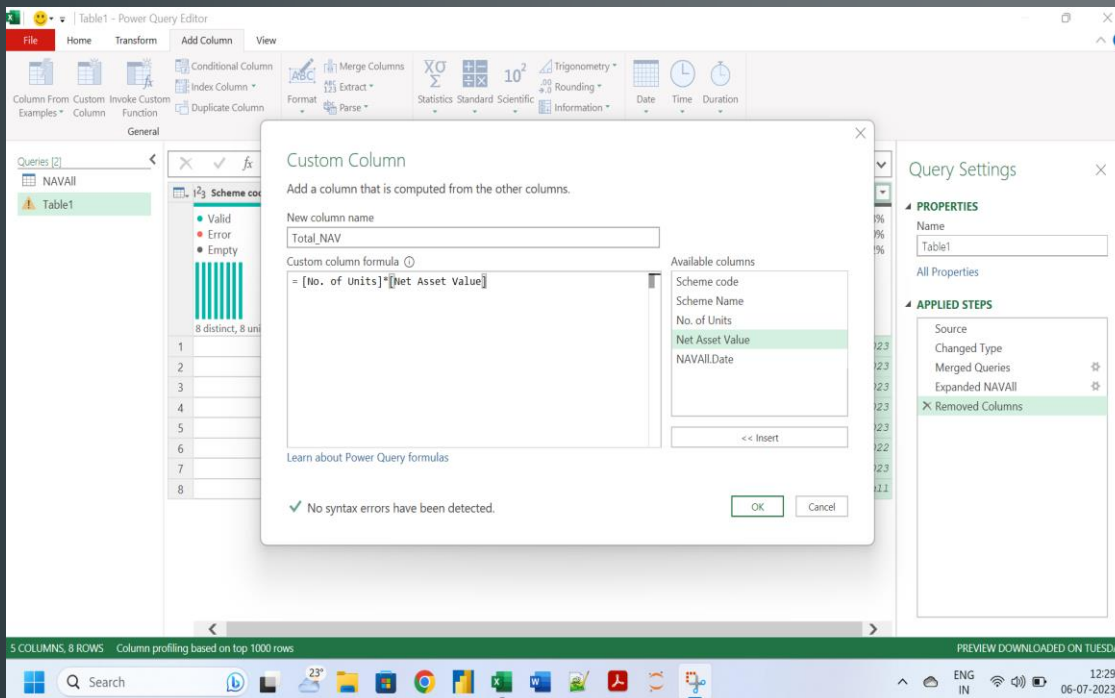
File Origin: 1252: Western European (Windows) | Delimiter: Semicolon | Data Type Detection: Based on first 200 rows

| Scheme Code | ISIN Div Payout/ ISIN Growth | ISIN Div Reinvestment | Scheme Name |
|---|------------------------------|-----------------------|--|
| Open Ended Schemes(Debt Scheme - Banking and PSU... | | | |
| Aditya Birla Sun Life Mutual Fund | | | |
| 119551 | INF209KA12Z1 | INF209KA13Z9 | Aditya Birla Sun Life Banking & PSU Debt |
| 119552 | INF209K01YM2 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 119553 | INF209K01Y08 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 108272 | INF209K01LX6 | INF209KA11Z3 | Aditya Birla Sun Life Banking & PSU Debt |
| 110282 | INF209K01LU2 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 108274 | INF209K01LN7 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 110490 | INF209K01LR8 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 106157 | INF209K01LS6 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 108273 | INF209K01LV0 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 103176 | INF209K01LT4 | - | Aditya Birla Sun Life Banking & PSU Debt |
| 119550 | INF209K01YN0 | - | Aditya Birla Sun Life Banking & PSU Debt |
| Axis Mutual Fund | | | |
| 128952 | INF846K01NF8 | - | Axis Banking & PSU Debt Fund - Direct Pl |

Preview downloaded on Tuesday

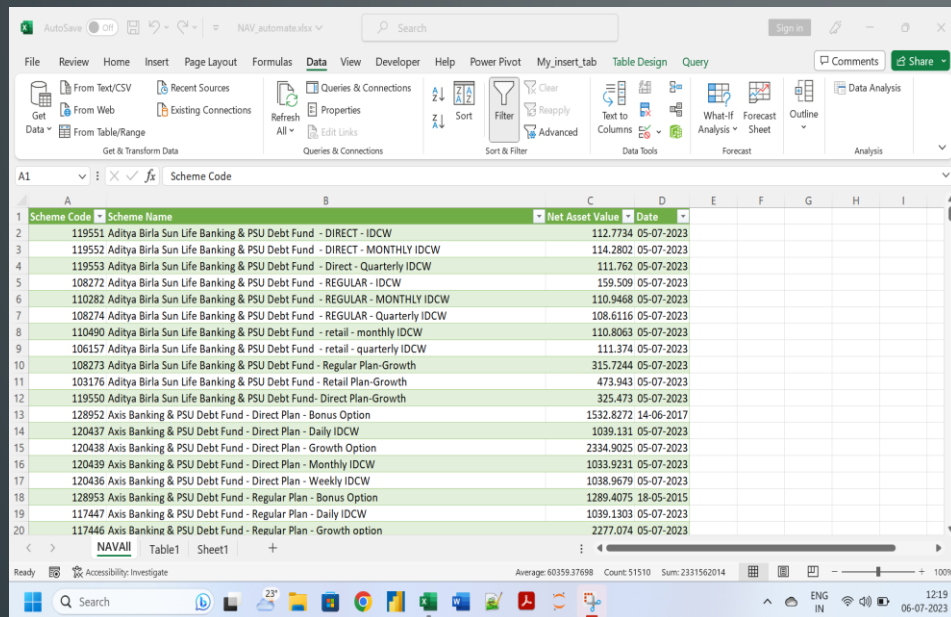
Data transformation

- ▶ Data cleaning: Deleting unwanted columns, changing data types of columns
- ▶ Data integration: Merge query- to join two related tables on common key column(scheme code) and choose columns required rest uncheck.
- ▶ Data enrichment: Extend the source data by creating calculated columns.



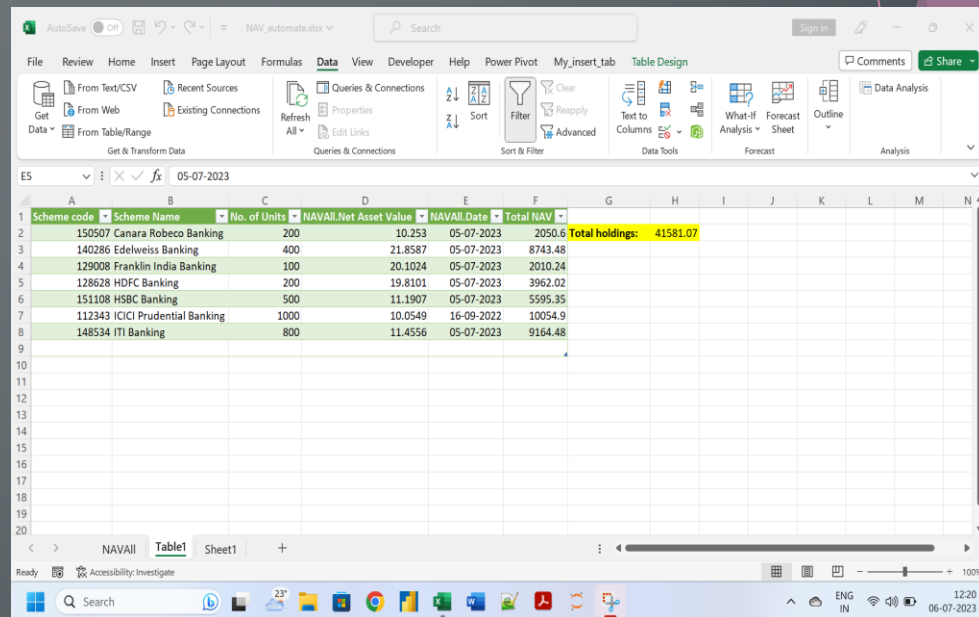
Load data

- ▶ Load transformed data into a worksheet or the Data Model
- ▶ We can refresh all the processes such as re-import the source data, reapply all the data filtering, sorting, and other transformations that we defined in a single click.
- ▶ So, once all of that's set up, we don't need to create it again. Of course, you can also go back and edit each step, and even add steps in between.
- ▶ This is all done within a tool you're already familiar with: Excel.



The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The ribbon includes options for 'Get Data', 'Queries & Connections', 'Sort & Filter', 'Advanced', 'Text to Columns', 'What-If Analysis', 'Forecast Sheet', and 'Outline'. The worksheet displays a table with the following data:

| Scheme Code | Scheme Name | Net Asset Value | Date |
|-------------|--|-----------------|------------|
| 119551 | Aditya Birla Sun Life Banking & PSU Debt Fund - DIRECT - IDCW | 112.7734 | 05-07-2023 |
| 119552 | Aditya Birla Sun Life Banking & PSU Debt Fund - DIRECT - MONTHLY IDCW | 114.2802 | 05-07-2023 |
| 119553 | Aditya Birla Sun Life Banking & PSU Debt Fund - Direct - Quarterly IDCW | 111.762 | 05-07-2023 |
| 108272 | Aditya Birla Sun Life Banking & PSU Debt Fund - REGULAR - IDCW | 159.509 | 05-07-2023 |
| 110282 | Aditya Birla Sun Life Banking & PSU Debt Fund - REGULAR - MONTHLY IDCW | 110.9468 | 05-07-2023 |
| 108274 | Aditya Birla Sun Life Banking & PSU Debt Fund - REGULAR - Quarterly IDCW | 108.6116 | 05-07-2023 |
| 110490 | Aditya Birla Sun Life Banking & PSU Debt Fund - retail - monthly IDCW | 110.8063 | 05-07-2023 |
| 106157 | Aditya Birla Sun Life Banking & PSU Debt Fund - retail - quarterly IDCW | 111.374 | 05-07-2023 |
| 108273 | Aditya Birla Sun Life Banking & PSU Debt Fund - Regular Plan-Growth | 315.7244 | 05-07-2023 |
| 103176 | Aditya Birla Sun Life Banking & PSU Debt Fund - Retail Plan-Growth | 473.943 | 05-07-2023 |
| 119550 | Aditya Birla Sun Life Banking & PSU Debt Fund-Direct Plan-Growth | 325.473 | 05-07-2023 |
| 128952 | Axis Banking & PSU Debt Fund - Direct Plan - Bonus Option | 1532.8272 | 14-06-2017 |
| 120437 | Axis Banking & PSU Debt Fund - Direct Plan - Daily IDCW | 1039.131 | 05-07-2023 |
| 120438 | Axis Banking & PSU Debt Fund - Direct Plan - Growth Option | 2334.9025 | 05-07-2023 |
| 120439 | Axis Banking & PSU Debt Fund - Direct Plan - Monthly IDCW | 1033.9231 | 05-07-2023 |
| 120436 | Axis Banking & PSU Debt Fund - Direct Plan - Weekly IDCW | 1038.9679 | 05-07-2023 |
| 128953 | Axis Banking & PSU Debt Fund - Regular Plan - Bonus Option | 1289.4075 | 18-05-2015 |
| 117447 | Axis Banking & PSU Debt Fund - Regular Plan - Daily IDCW | 1039.1303 | 05-07-2023 |
| 117446 | Axis Banking & PSU Debt Fund - Regular Plan - Growth option | 2277.074 | 05-07-2023 |



The screenshot shows the Microsoft Excel interface with the 'Data' tab selected. The ribbon includes options for 'Get Data', 'Queries & Connections', 'Sort & Filter', 'Advanced', 'Text to Columns', 'What-If Analysis', 'Forecast Sheet', and 'Outline'. The worksheet displays a table with the following data:

| Scheme code | Scheme Name | No. of Units | NAVAIL.Net Asset Value | NAVAIL.Date | Total NAV |
|-----------------|--------------------------|--------------|------------------------|-------------|-----------|
| 150507 | Canara Robeco Banking | 200 | 10.253 | 05-07-2023 | 2050.6 |
| 140286 | Edelweiss Banking | 400 | 21.8587 | 05-07-2023 | 8743.48 |
| 129008 | Franklin India Banking | 100 | 20.1024 | 05-07-2023 | 2010.24 |
| 128628 | HDFC Banking | 200 | 19.8101 | 05-07-2023 | 3962.02 |
| 151108 | HSCB Banking | 500 | 11.1907 | 05-07-2023 | 5595.35 |
| 112343 | ICICI Prudential Banking | 1000 | 10.0549 | 16-09-2022 | 10054.9 |
| 148534 | ITI Banking | 800 | 11.4556 | 05-07-2023 | 9164.48 |
| Total holdings: | | | | | 41581.07 |

Check the video showing the NAV updating .

The screenshot displays the Microsoft Excel interface with the 'Review' tab selected. A 'SECURITY WARNING' banner indicates that external data connections have been disabled. The active worksheet is 'Table1', which contains a table with the following data:

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|----|-------------|--------------------------|--------------|------------------------|-------------|-----------|-----------------|----------|---|---|---|---|---|---|
| 1 | Scheme code | Scheme Name | No. of Units | NAVAII.Net Asset Value | NAVAII.Date | Total NAV | | | | | | | | |
| 2 | 140286 | Edelweiss Banking | 400 | 21.8543 | 04-07-2023 | 8741.72 | Total holdings: | 39523.83 | | | | | | |
| 3 | 129008 | Franklin India Banking | 100 | 20.096 | 04-07-2023 | 2009.6 | | | | | | | | |
| 4 | 128628 | HDFC Banking | 200 | 19.8037 | 04-07-2023 | 3960.74 | | | | | | | | |
| 5 | 151108 | HSBC Banking | 500 | 11.1907 | 04-07-2023 | 5595.35 | | | | | | | | |
| 6 | 112343 | ICICI Prudential Banking | 1000 | 10.0549 | 16-09-2022 | 10054.9 | | | | | | | | |
| 7 | 148534 | ITI Banking | 800 | 11.4519 | 04-07-2023 | 9161.52 | | | | | | | | |
| 8 | | | | | | | | | | | | | | |
| 9 | | | | | | | | | | | | | | |
| 10 | | | | | | | | | | | | | | |
| 11 | | | | | | | | | | | | | | |
| 12 | | | | | | | | | | | | | | |
| 13 | | | | | | | | | | | | | | |
| 14 | | | | | | | | | | | | | | |
| 15 | | | | | | | | | | | | | | |
| 16 | | | | | | | | | | | | | | |
| 17 | | | | | | | | | | | | | | |
| 18 | | | | | | | | | | | | | | |

Overlaid on the right side of the spreadsheet is the text: *Automate Update of NAV using Excel power query get data from web*.

Advantages of using power query and further scope in this project

- ▶ If a person, add new scheme code, or made any change in his portfolio and press refresh button his excel table got automatically update.
- ▶ He can also add more calculated columns in his excel table to get some aggregate values.
- ▶ He can easily analyse the data using different excel tools.
- ▶ It saves your lot of time required previously for fetching data manually from web pages or using VLOOKUP.
- ▶ It other way it saves your money in just refreshing your data in single button click you can get your NAV updated values in front of you.

The background features abstract, overlapping geometric shapes in various shades of pink and magenta, creating a modern, layered effect. A central horizontal pink bar with a thin dark border contains the text.

Thank You