

Using Power BI to present the QC status for studies

Spencer Renyard, PHASTAR, Macclesfield, United Kingdom

ABSTRACT

PHASTAR hosts the QC trackers for studies on a SharePoint® site, which has multiple benefits including multi-user edit access and version history. Power BI® can access the QC tracker files and use them as data sources, providing the ability to analyse and summarise this data to provide information on the QC status of the studies in real time. With a centrally located, easy to use dashboard created with Power BI, the study or studies of interest can be selected, and a standard, validated set of summary tables and plots present the overall QC status for the study. This report can be used by the lead programmer and project manager to track the QC status and provide feedback to interested parties such as clients/customers and portfolio leads.

QC TRACKERS HOSTED ON SHAREPOINT

Within PHASTAR the trackers that record the QC status of studies are stored as Excel files hosted on a SharePoint site. This brings multiple benefits to the users, e.g., native multi-user access; administrators, the files are within the company security framework, two-factor authentication, centralised storage, and easily managed access control including named users; and the company, e.g., easy review by non-study team members such as project management.

OBJECTIVE

Provide a summary of the QC status for a study considering the technical environment used to store and access the QC trackers. The summary needs to include for each type of program (SDTM and ADaM dataset programs, and output programs), graphical displays of the QC status, supporting tabular displays, and an overall percentage complete. The technical solution developed needs to, in addition to presenting the information detailed above, be easy to use and require minimal training and support for users to be able to utilise fully.

Considering these requirements and the QC tracker Excel files being hosted on SharePoint the technical solution chosen was Power BI.

WHAT IS POWER BI?

Power BI is a collection of software services, apps, and connectors that work together to turn your unrelated sources of data into coherent, visually immersive, and interactive insights [1] [2], and it is the leading BI tool according to Gartner [3].

Power BI comprises two components, with some shared functionality but also distinct usage: Power BI Desktop, and Power BI service:

- Power BI Desktop is a data analysis and report creation tool that includes Power Query Editor (starting to sound like there may be some programming involved!) and is a free to download and use desktop application.
- Power BI service is how reports are shared and distributed. It is cloud based, and requires a paid-for license, and whilst it allows some report editing, the functionality is limited compared to Power BI Desktop

NOW WE KNOW WHAT POWER BI IS, WHAT'S NEXT?

This is the journey from looking at individual Excel files hosted in a SharePoint library to having a report which is available to all within PHASTAR and requires minimal learning to use and share. Alongside details of the technical implementation and development there will also be some personal reflections on the journey from a long-standing SAS programmer.

ACCESSING THE DATA

Working with Power BI Desktop, the first task is accessing the data stored in the QC trackers. Power BI can connect to the SharePoint library containing the Excel QC tracker files. This is achieved by opening the "Get Data" screen (Figure 1) and selecting "SharePoint folder" from the sources available. Note that this source may not be visible on opening the screen but using the search facility shows the SharePoint data sources available in Power BI.

Get Data

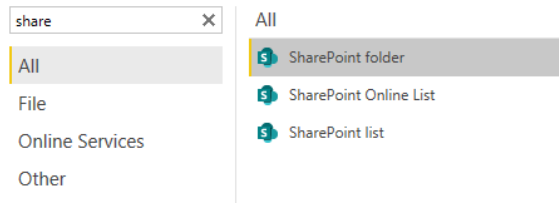


Figure 1 Getting access to the data on SharePoint

In the next screen (Figure 2) enter the root site URL for the SharePoint folder is specified; note that this is not the full URL for the folder/library:

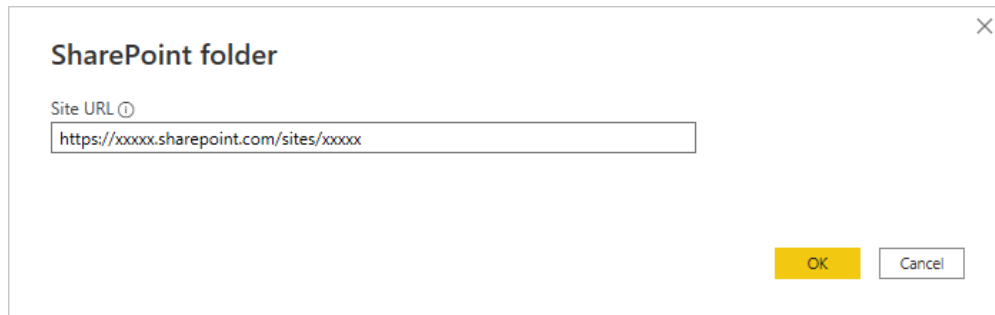


Figure 2 Selecting the SharePoint site containing the folder/library containing the QC trackers

This allows the loading of not only a single Excel workbook but also combining the contents of the same worksheet from multiple workbooks into a single data table. This is key to creating an output where the user can select a study (or multiple studies if reviewing a project) to present the QC status for. In the “Combine Files” window (Figure 3) selecting a workbook containing all the worksheets required allows the import of, in turn, the worksheets required containing the data from all the QC tracker files in the library, in this case: RAW2SDTM Datasets, ADAM Datasets, and Outputs.

Combine Files

Select the object to be extracted from each file. [Learn more](#)

Sample File: Beta QC plan template SP-002-T01...

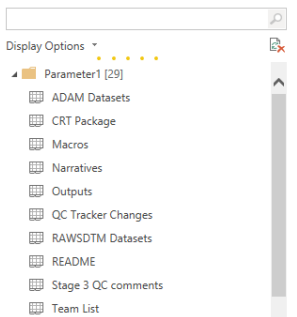


Figure 3 Combine Files

TRANSFORMING THE DATA

The data in the QC trackers requires some transformation to enable it to be used as the basis for the report. This includes:

- Some of the data columns need to be updated before they can be used as desired
- Filtering, renaming columns, data cleaning, removing blank rows, etc.

This is achieved using Power Query Editor, which is a point-and-click interface. The transformation steps applied to a data table can be seen in the “Applied Steps” window of the Power Query Editor window (Figure 4).

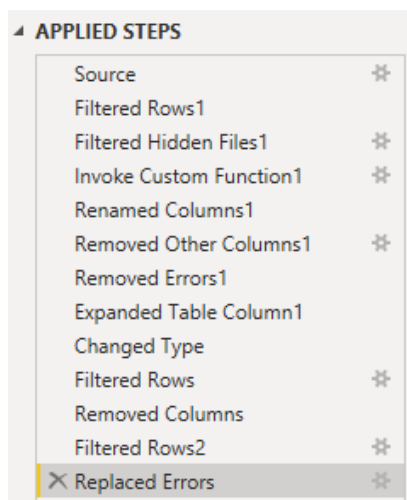


Figure 4 Applied Steps

These steps translate into queries written in Data Analysis Expressions (DAX) [4], which is a library of functions and operators that can be combined to build formulae and expressions in Power BI. The DAX queries exhibit similarities to both Excel functions and SQL, introducing another area of this project that feels somewhat familiar to a SAS programmer. By clicking on an individual step in the “Applied Steps” window it is possible to see both the DAX syntax and the data output for the step. This aids review and debugging of the data transformation by allowing a stepwise process akin to the steps in a SAS program or R script.

The full DAX query can be viewed and edited using the Advanced Editor window (Figure 5), where some examples of the data transformation steps can be seen. Whilst the DAX syntax was unfamiliar to me, once viewed like this the similarities to Excel functions and SQL allowed the modification of and addition to the overall query without recourse to the point-and-click interface.

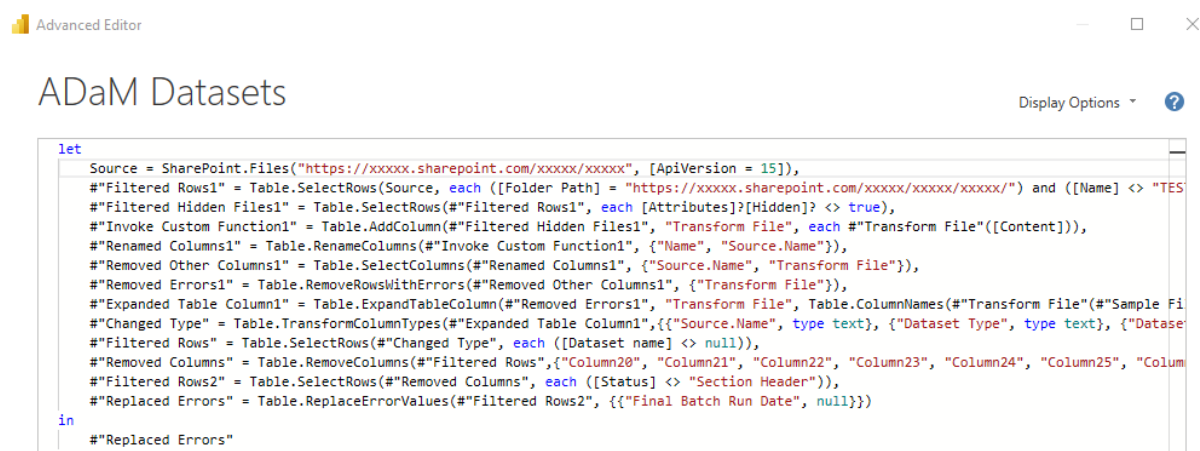


Figure 5 Advanced Editor showing a complete DAX data transformation query

BUILDING THE QC STATUS REPORT

With access to the data in the QC trackers and the required transformation of the data completed, building the QC status report can start.

KEY COMPONENTS

The QC status report needs to include for each type of program (SDTM and ADaM dataset programs, and output programs):

- For each program type
 - Graphical displays of the QC status
 - Supporting tabular displays
 - Overall percentage complete
- QC tracker name (equivalent to the study name)
- Date and time of data transfer

KEY REQUIREMENTS

- Ability to easily filter by study (or studies)
- Good performance
- Easy to present to other parties

DATA TRANSFORMATION STEPS

All the data tables created from importing the worksheets from the Excel files include the QC tracker file name – this is a key ID column allowing the data to be linked to a specific QC tracker. As with any automated data import process it can be necessary to check that the data types of the columns are correct and reflects the data present. For example, character to numeric and vice versa, or converting a date/time column to one containing just the date. The data may require some data cleaning activities to be performed too, for example, remove blank rows or filtering out records with values not required.

The date and time that the data was refreshed is required to ensure that users know exactly when the data they are querying and reporting was loaded into the Power BI report. This was added by creating a blank query and adding the code to get the date and time that the report was refreshed [5], allowing the value to be displayed in the report.

BUILDING THE REPORT

Once the required data has been imported and transformed as necessary for use in the report, work can start on building the report in the “Report” window of Power BI Desktop (Figure 6). The blank screen can look intimidating, but the panes on the right for visualizations and fields, amongst other elements, make it easy to add the first visual and specify the data to be presented.

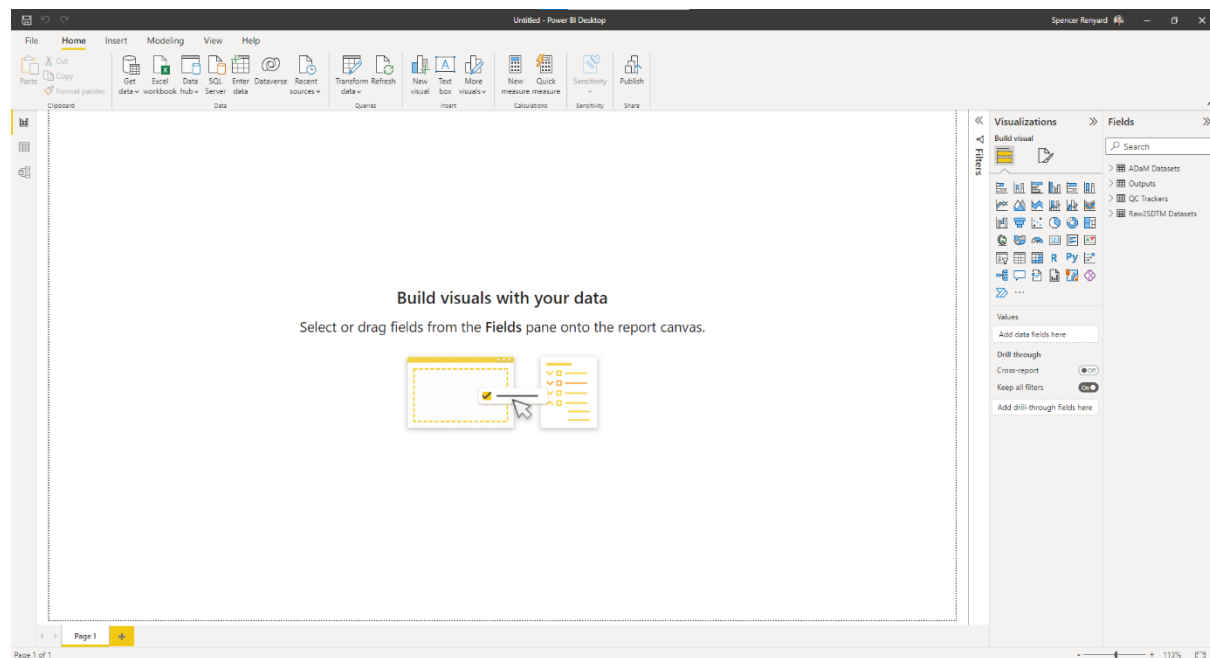


Figure 6 Power BI Desktop report editor

The first item to add to the report is a slicer [6], which can be used to display the filter to be used and makes it easier to see the current filtered state without having to open a drop-down list. This will display the QC tracker file names and allows the user to select which QC tracker they would like to view the status of. The slicer filters the data for all visualization objects on the report page.

The key options for the slicer behaving as desired include choosing the type of slicer to be “dropdown” and selecting the visual settings to allow “Multi-select with CTRL” and “Show ‘Select All’ option” (Figure 7).

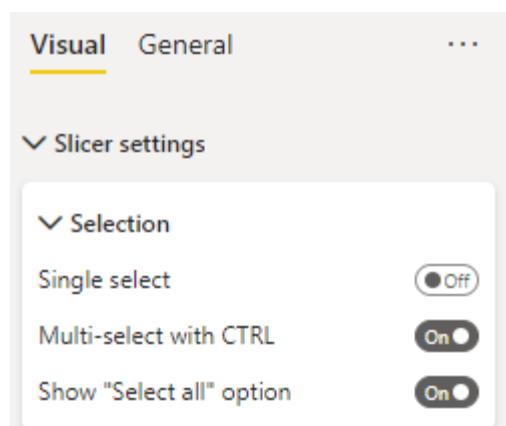


Figure 7 Slicer settings

The second item to add is a pie chart displaying the counts per QC status for a program type (Figure 8). The QC tracker contains conditional formatting which based on the value of the QC status colours a row. To ensure consistency between the QC tracker and the status report the colours of wedges match those used to the QC tracker.

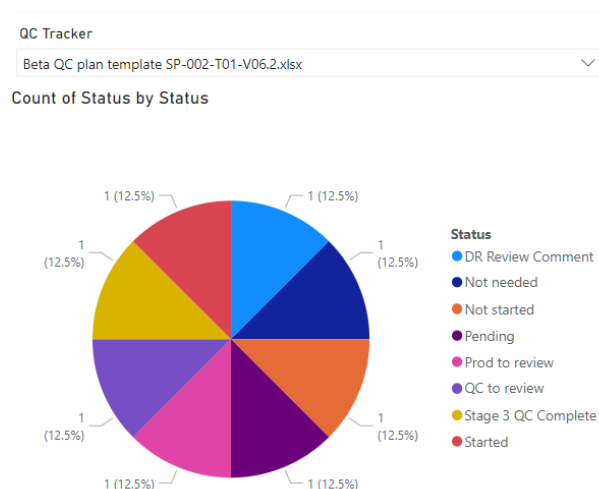


Figure 8 Slicer and Pie Chart

Next a table can be added to present the counts of programs by QC status, which is fine for the SDTM and ADaM programs (Figure 9). However, a table is not the correct visualization to use to summarise the output programs as these grouped into output type, e.g., table, listing, figure. For the outputs, a matrix (Figure 10) allows the reporting of the counts by QC status and output type.

Status	Count of Status
DR Review Comment	1
Not needed	1
Not started	1
Pending	1
Prod to review	1
QC to review	1
Stage 3 QC Complete	1
Started	1
Total	8

Figure 9 Summary Table

The default presentation for a matrix shows the grouping variable (Output Type) collapsed. A group can be expanded by clicking on the + to the left of the group (Figure 10). The default alphabetical sort order of the QC status does not match the order required, where, for example, "Started" occurs before "QC to review". It is possible to sort a column

by the value in another column [7], so by creating a numeric order variable the values can be presented in the desired order.

Output Type	Count of Status
Figure	4
Not started	1
Pending	1
Prod to review	1
Started	1
Table	4
DR Review Comment	1
Not needed	1
QC to review	1
Stage 3 QC Complete	1
Total	8

Figure 10 Summary Matrix

CALCULATING THE WEIGHTED OVERALL PERCENTAGE

The weighted overall percentage is calculated by determining the number of programs with each value of the QC status and assigning a weighting to the count depending on the QC status, such that the further through the programming and QC process a program is the greater its contribution to the overall percentage completed. This is achieved by creating a new “measure” for each status value (Figure 11) which is calculated at the time the data is queried. These measures can be combined in an equation based on the weighting requirements to calculate the overall weighted percentage. This can be added to the report using the “card” visualization which along with the pie chart and table/matrix described above the value reported here is based on the data selected by the slicer.

```

1 Outputs_Completed =
2 CALCULATE(
3     COUNT(Outputs[Status]),
4     FILTER(Outputs,
5         Outputs[Status] = "Completed"
6     )
7 )

```

Figure 11 Example of a calculated measure

The calculated measure above counts the number of records in the “outputs” table where Status = “Completed”.

FINALISING THE REPORT

The visualization objects described above can be repeated for each program type as required and the full report is ready, taking advantage of the interactive elements provided by Power BI.

PUBLISHING THE REPORT

The QC status report has been created in Power BI Desktop; however, this does not allow the report to be easily shared across the company. To achieve this the report is published to the second component of Power BI, Power BI service, which allows others to see and interact with the report. Once published to a central location, the Power BI workspace, the report can be embedded into a SharePoint page, thereby enabling users to access and interact with the report without requiring access to the data library or the underlying Power BI report. Further functionality enables the user to export the report to an external file in either PDF or PowerPoint format.

TAKING THE REPORT TO THE NEXT LEVEL

The report (Figure 12) fulfils all the requirements described earlier in this paper. Utilising the interactivity included as standard in a Power BI report the user can select a specific study to report the QC status of, select a category to highlight which is reflected in the outputs for a program type, e.g., SDTM programs, and export for distribution or presentation.

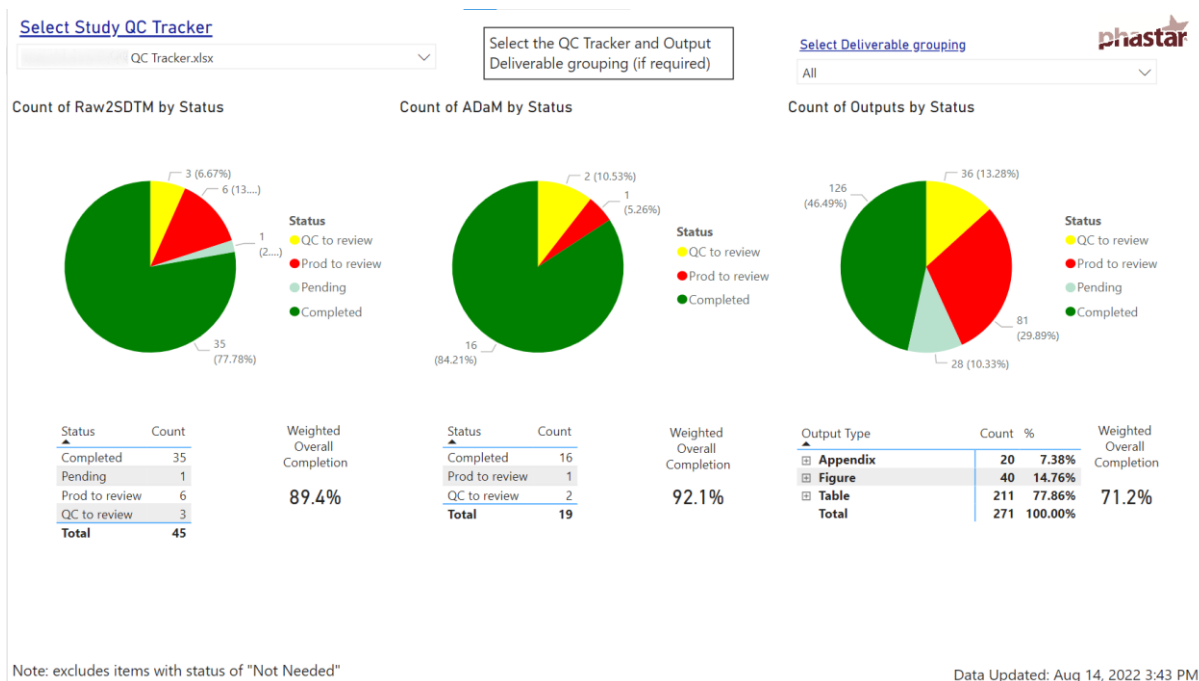


Figure 12 QC Status report

It is possible to add another level of interactivity to the report by presenting drillthrough [8] tables which are displayed when hovering over summary items in the pie chart and the summary table/matrix – in Power BI these are called “tooltips”. Each tooltip is created on a new page in the report, and if needed the tooltips can include visuals, images, and any other collection of items [9]. These pages are not intended to be viewed by the user and the option to hide the page is selected and can be seen by the symbol displayed to the left of the page name (Figure 13).



Figure 13 Tooltips pages are hidden

For this report a tooltip is created for each program type, and this contains a table of the data presenting key information such as output type, name, and number (Figure 14).

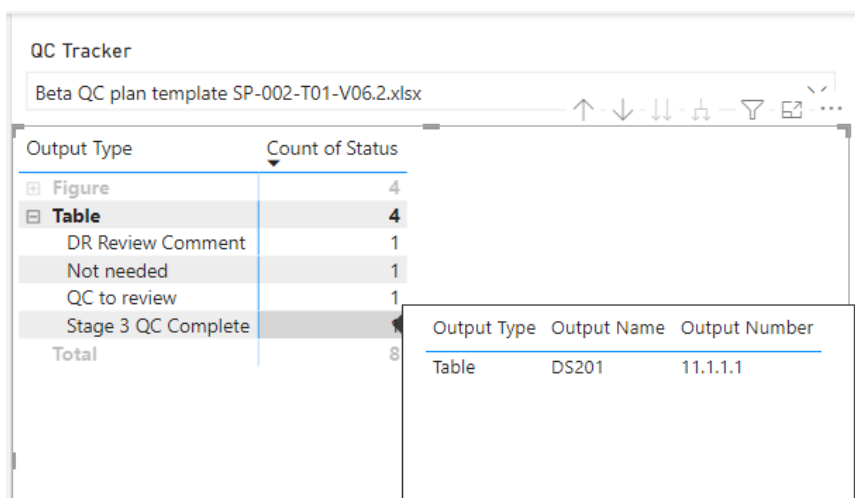


Figure 14 Example of a tooltip

The tooltip is not designed to present many rows of data, so it most useful when investigating a small number of programs, as above.

CONCLUSION

As an introduction to Power BI this has covered several areas of Power BI. It allows data to be accessed from many sources and used to quickly create usable reports. Looking further into the options available to manipulate and present the data, in addition to the point-and-click interface, there is access to the DAX queries and expressions which underly this and provide the ability for some programming to take place. The Power BI report objects provide many useful interaction options as standard, including filtering of the data, and the automatic updating of counts and percentages based on the data and calculated measures.

When the report is published and embedded in a SharePoint page, the end users – lead programmers and statisticians, project managers – do not require access to the underlying data or report to interact with it. It is easy to select the study of interest and can focus on specific items and drilldown to see more detail.

From the personal perspective, whilst there is much new to learn with Power BI, there is also plenty that feels familiar to an experienced programmer, with DAX being another programming/query syntax to investigate and learn. The transformation of source data into report ready data and the design and creation of the report elements and overall report are also familiar activities.

REFERENCES

- [1] Power BI overview: <https://docs.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview>
- [2] What is Power BI?: <https://powerbi.microsoft.com/en-us/what-is-power-bi/>
- [3] 2022 Gartner® Magic Quadrant™ for Analytics and Business Intelligence Platforms: <https://info.microsoft.com/ww-landing-2022-gartner-mq-report-on-bi-and-analytics-platforms.html?lcid=en-us>
- [4] Data Analysis Expressions (DAX) Reference: <https://docs.microsoft.com/en-us/dax/>
- [5] Add a Refresh Time Stamp To Your Power BI Reports: <https://www.ehansalytics.com/blog/2020/7/29/add-a-refresh-time-stamp-to-your-power-bi-reports>
- [6] Slicers in Power BI: <https://docs.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-slicers?tabs=powerbi-desktop>
- [7] Sort one column by another column in Power BI: <https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-sort-by-column?tabs=powerbi-desktop>
- [8] Set up drillthrough in Power BI reports: <https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-drillthrough>
- [9] Create tooltips based on report pages: <https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-tooltips?tabs=powerbi-desktop>

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Spencer Renyard

PHASTAR

2D Bollo Ln,

Chiswick,

London, W4 5LE

Email: spencer.renyard@phastar.com

Web: phastar.com

Brand and product names are trademarks of their respective companies.