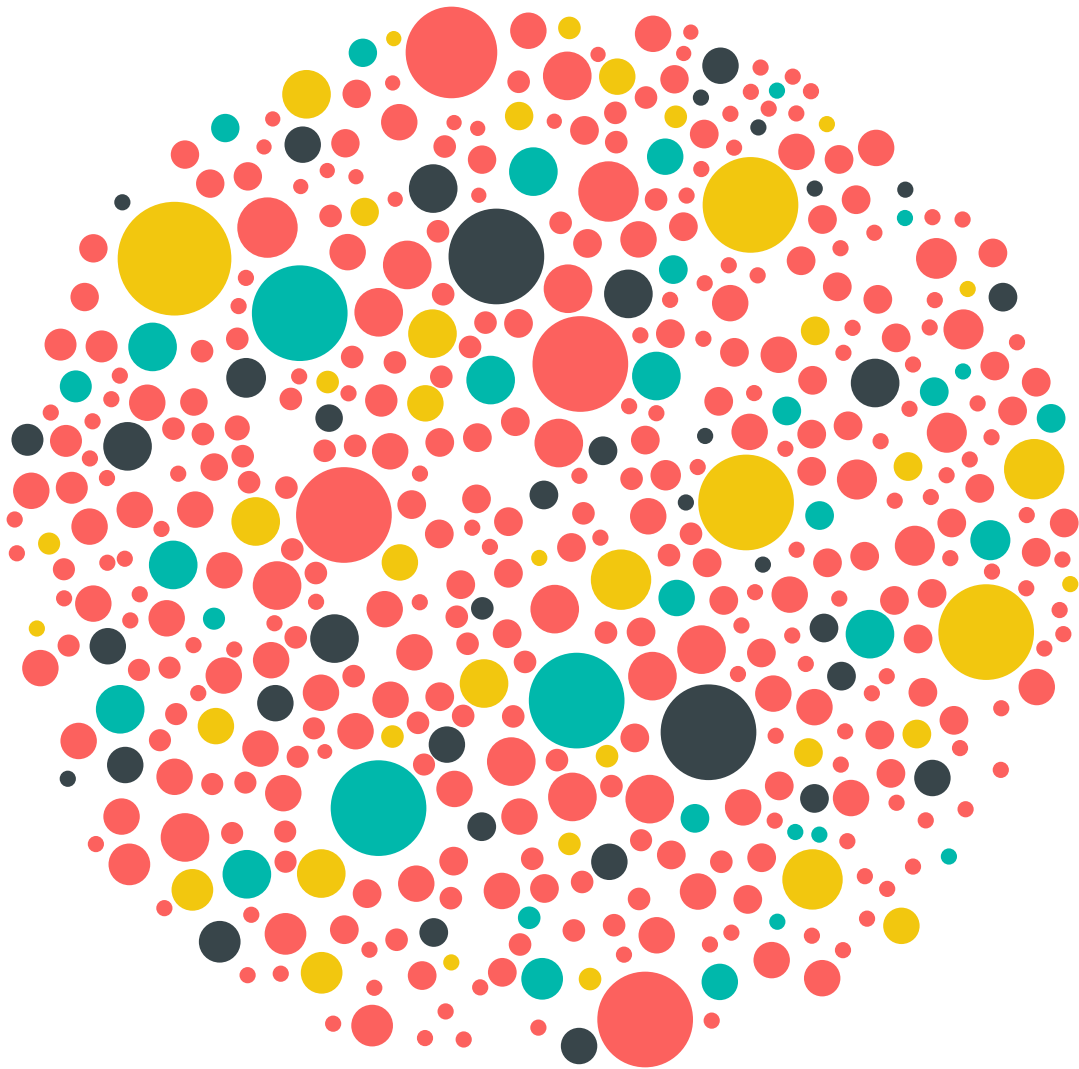




Advanced Analytics with Power BI



Data is everywhere. The world contains an astronomical amount of data, an amount that grows larger and larger each day. This vast collection of information has changed the way the world interacts, uncovered breakthroughs in medicine, and revealed new ways to understand trends in business and in our daily lives. With the increasing availability of data comes new challenges and opportunities as business leaders seek to gain important insights and transform information into actionable and meaningful results.

As data becomes more accessible, manipulating vast amounts of available data to drive insights and make business decisions can be a challenge. Business leaders at every level need to become data literate and be able to understand data and analytical concepts that may have previously seemed out of reach, including statistical methods, machine learning, and data manipulation. With this spread of data literacy comes the powerful ability to make educated business decisions that rely on the smart use of data, rather than on an individual's opinions. In the past, these tasks were extremely complex and would be handed off to engineers. With the tools that exist today, business leaders are able to dive into their own analytics and uncover powerful insights.

Microsoft Power BI brings advanced analytics to the daily business decision process, allowing users to extract useful knowledge from data to solve business problems. This white paper will cover the advanced analytic capabilities of Power BI, including predictive analytics, data visualizations, R integration, and data analysis expressions.

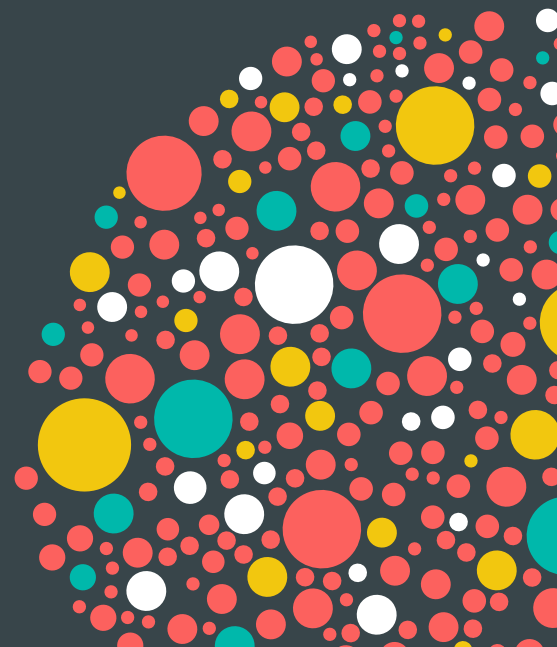


Table of contents

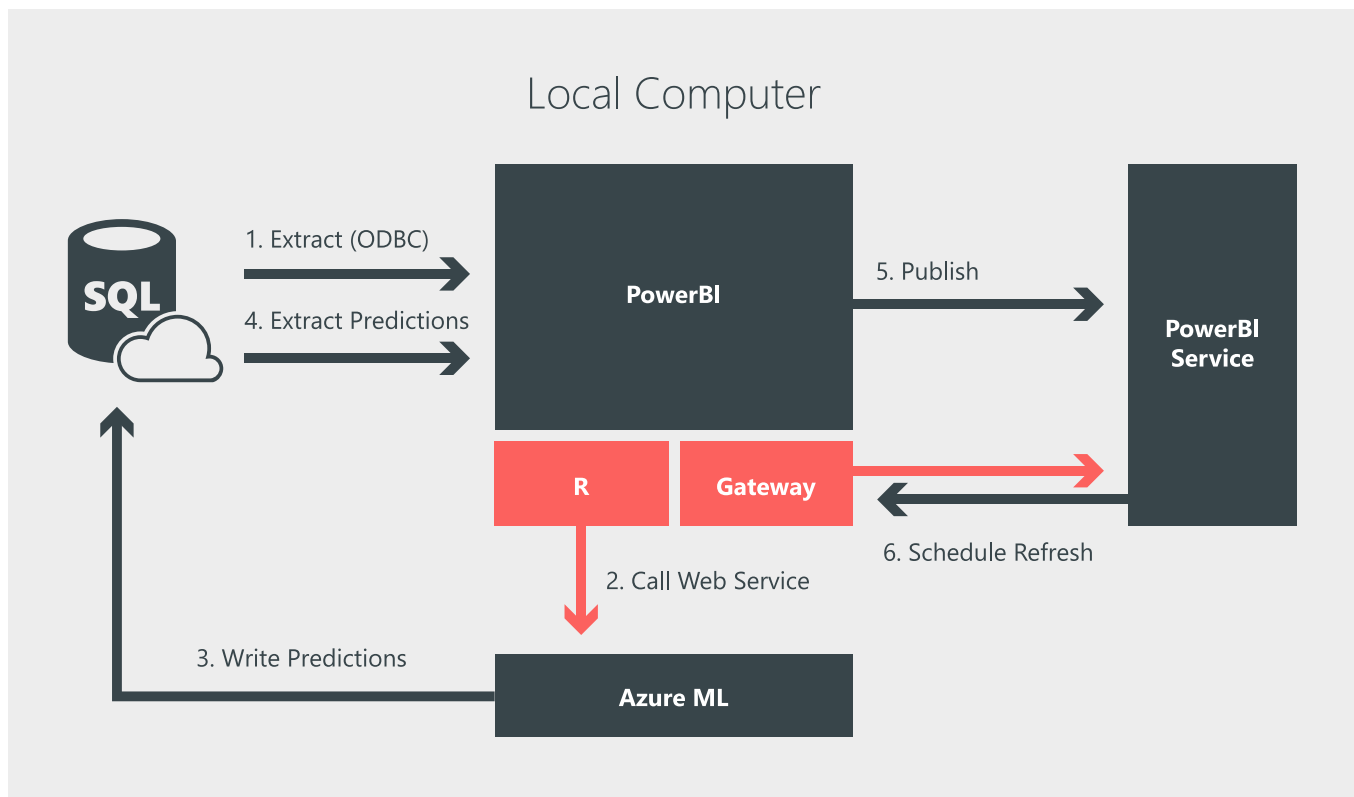
Advanced analytics in Power BI	4
Predictive analytics with Azure	
R integration	
Quick Insights feature	
Segmentation and cohort analysis	9
Data grouping and Binning	
Data streaming in Power BI	11
Real-time dashboards	
Setup of real-time streaming data sets	
Visualizations in Power BI	12
Community-sourced visualizations	
R visualizations	
Custom visualizations	
Data connection and shaping	14
Azure services	
DirectQuery	
Data fetching with the R connector	
Data shaping in Power Query with R	
Data Analysis Expressions	17
Conclusion	18

Advanced analytics in Power BI

Predictive analytics with Azure

Imagine if you could review the latest output of your organization's fraud model on demand, or analyze the sentiment of social media users who tweet or post about your products. Power BI brings the predictive power of advanced analytics to allow users to create predictive models from their data, enabling organizations to make data-based decisions across all aspects of their business.

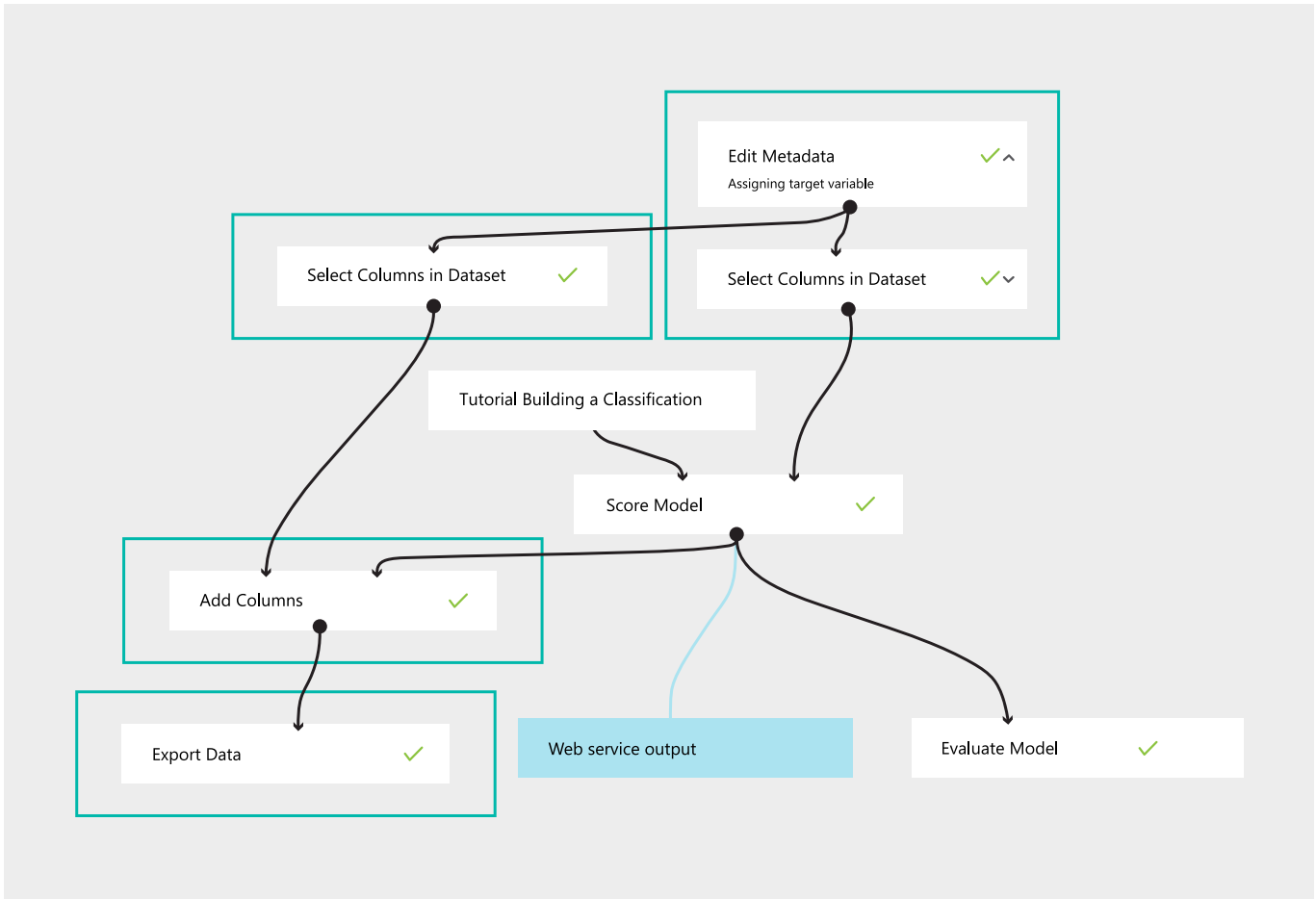
Through machine learning, computers are able to act without being explicitly programmed. Instead, they can teach themselves to grow and change when exposed to new data. Once the work of science fiction, machine learning is rapidly becoming part of our daily lives—through practical speech recognition programs, more effective web searches, and even self-driving cars. Using Azure Machine Learning Studio, users can quickly create predictive models by dragging, dropping, and connecting data modules. Power BI then allows users to visualize the results of their machine learning algorithm.



From <<https://powerbi.microsoft.com/en-us/blog/power-bi-azure-ml/>>

To accomplish this in Power BI, first use R to extract data from Azure SQL that has not yet been scored by the machine learning model. Next, use R to call the Azure Machine Learning web service and send it the unscored data. Write the output of the Azure Machine Learning model back into

SQL and use R to read scored data into Power BI. Then, publish the Power BI file to the Power BI service. Finally, use the Personal Gateway to schedule a refresh of the data, which triggers a scheduled rerun of the R script and brings in the new predictions.



From <<https://powerbi.microsoft.com/en-us/blog/power-bi-azure-ml/>>

R integration

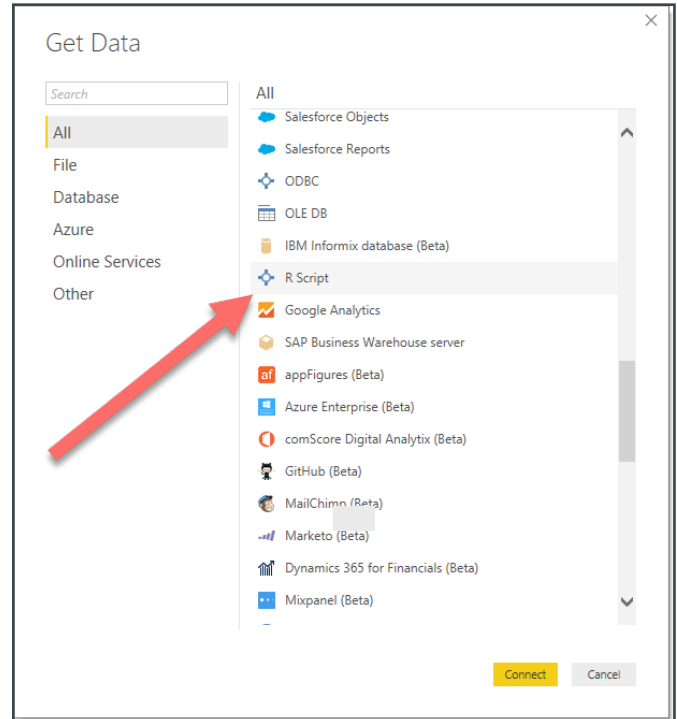
R, a programming language used by statisticians, data scientists, and data analysts, is the most widely used statistical language in the world. R integration in Power BI brings this language into all stages of generating insights. Using the R connector, users can run R scripts directly in Power BI and import the resulting data sets into a Power BI data model.

R in Power Query performs advanced data cleansing and preparation tasks, such as outlier detection and missing values completion. R visuals in Power BI allow you to visualize data by gaining endless flexibility and advanced analytics depth. Once the visuals are created, you can share the R visuals in your reports and on your dashboard, where they are interactive and cross-filterable.

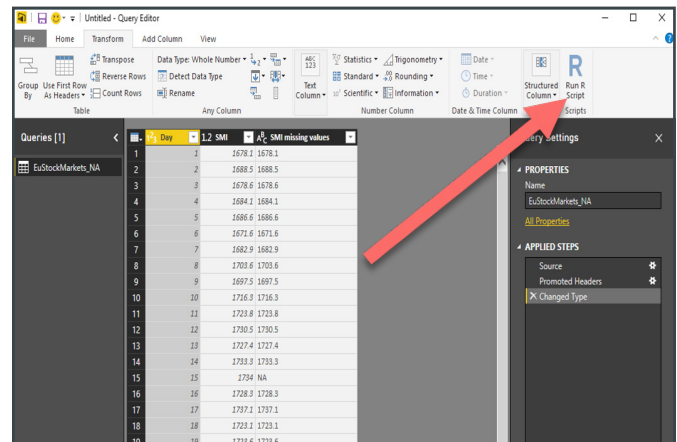
Check out the R showcase for amazing examples of what can be done with R in Power BI.

Power BI users do not need to have a background in working with R to leverage everything that R can do, such as prediction, clustering, association rules, and decision trees. R custom visuals allow users to apply the power of R without writing one line of R. Just import a custom R visual to your report, and drag your data to update your report.

Because R is run directly in the Power BI service, reports using R can be shared with and viewed by anyone—even if they don't have R installed.



From <https://powerbi.microsoft.com/en-us/documentation/powerbi-desktop-r-scripts/>



Learn more:

[R connector](#)

[R in Power Query](#)

[R visuals in Power BI](#)

[R showcase](#)

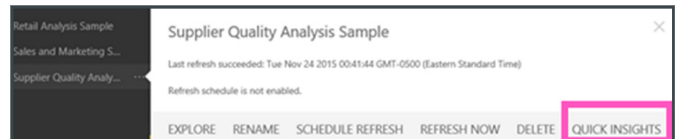
[R custom visuals](#)

Quick Insights feature

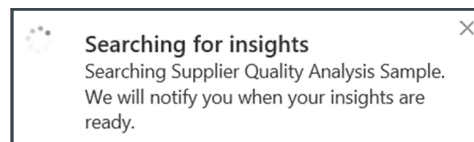
The Quick Insights feature in Power BI is built on a growing set of advanced analytical algorithms, developed in conjunction with Microsoft Research, which allows users to find insights in their data in new and intuitive ways. With a simple click, Quick Insights in Power BI searches different subsets of your data set while applying a set of sophisticated algorithms to discover potentially interesting insights. Power BI scans as much of a data set as possible in an allotted amount of time.

To use Quick Insights in Power BI, follow these steps.

1. In the left navigation pane under **Data sets**, select the ellipses (...), and then choose **Quick Insights**.



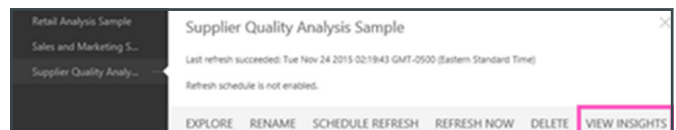
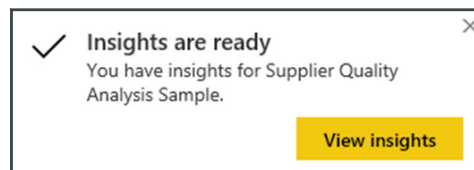
2. Power BI uses various algorithms to search for trends in your data set.



3. Within seconds, your insights are ready. Select **View Insights** to display visualizations.

Or, in the left navigation pane, select the ellipses (...) and then choose **View Insights**.

NOTE: Some data sets are unable to generate insights because the data isn't statistically significant. To learn more, see [Optimize your data for quick insights](#).

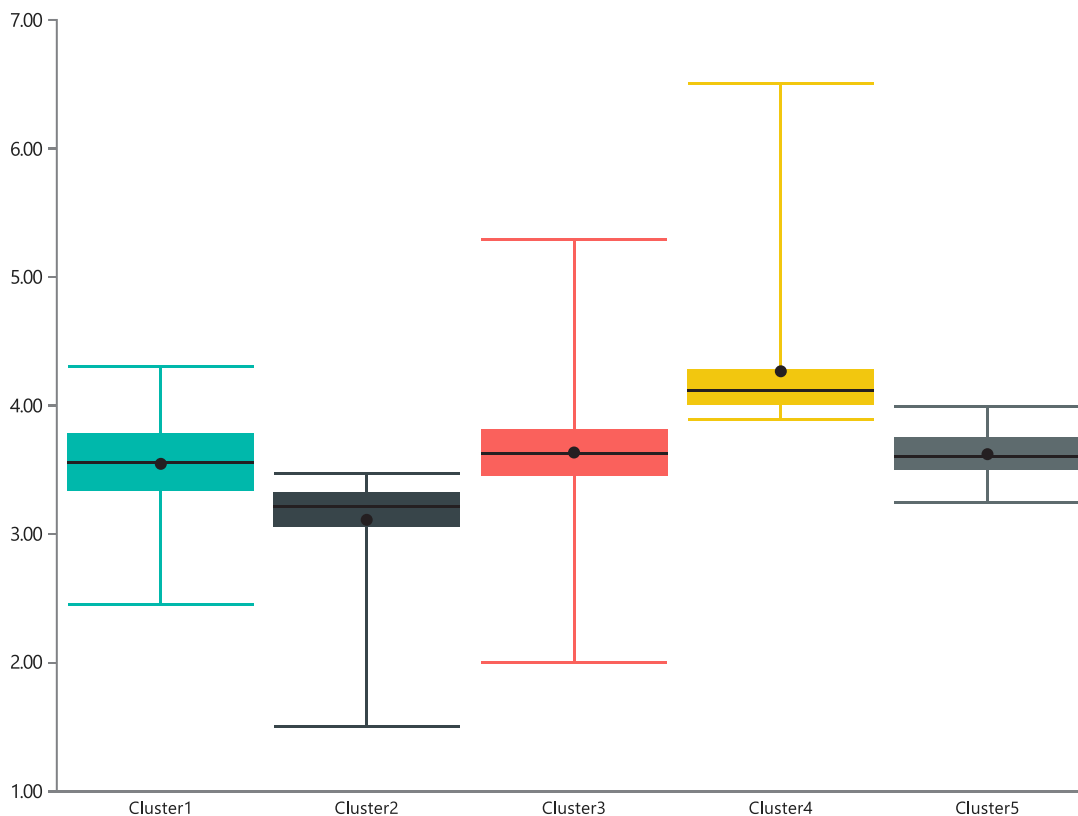


Segmentation and cohort analysis

Segmentation and cohort analysis is a simple, yet powerful, way to explore data and identify deviations from the norm. Segmentation and cohort analysis is simply the act of breaking down or combing data into meaningful groups, and then comparing those groups to identify meaningful relationships in your data. It is typically used to develop a hypothesis about your data and identify areas for further analysis. Power BI has several tools to help this process, including clustering, grouping, and binning.

Clustering allows you to use machine learning algorithms to quickly find groups of similar data points in a subset of your data. After you have created a cluster field of data, custom visuals in Power BI allow further analysis and evaluation of the clusters. For example, you could use the cluster column and each of the associated measures in a radar chart to see the aggregate of each measure for each cluster. You could also use the cluster column and one of the measures in a box and whiskers plot to see the distribution of values for that measure in each cluster. This can help you determine the minimum, maximum, and median values for that measure within each cluster.

Average of NSAT by ProdID (clusters) and ProdID



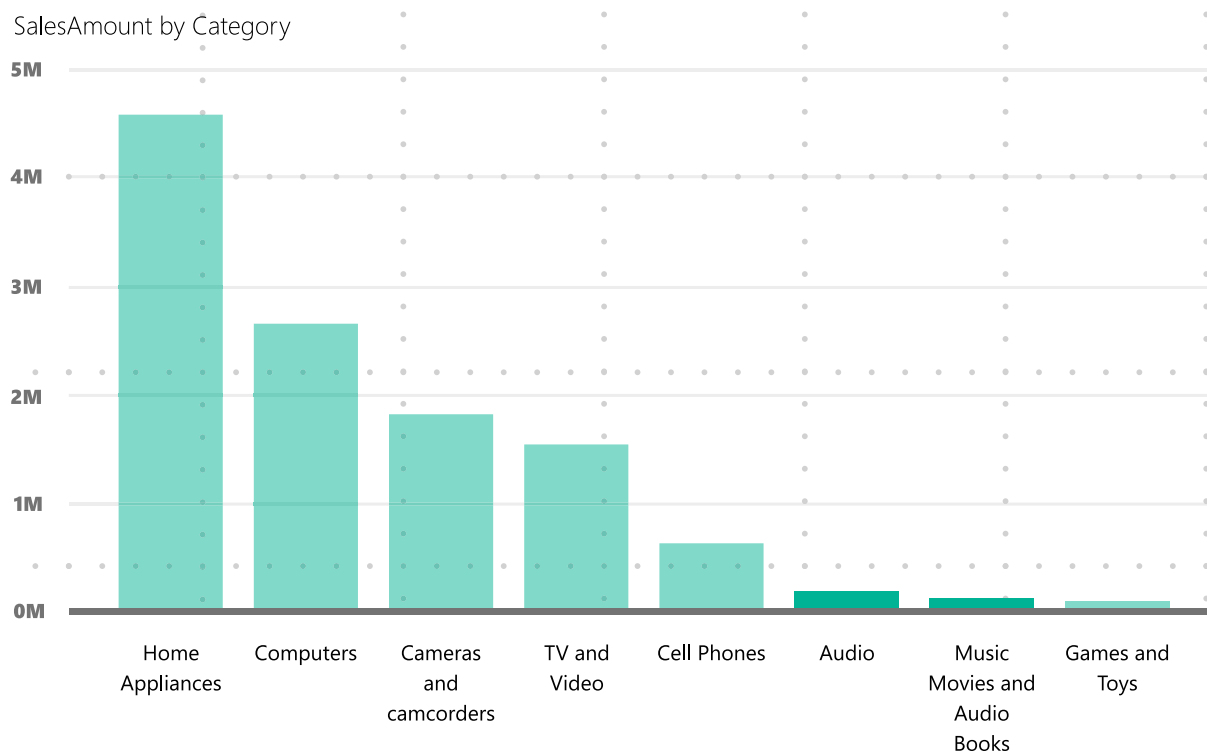
Learn more:

[Clustering](#)

Data grouping and Binning

Sometimes, two categories of data points within a field are better talked about together and should be grouped into a single category. Grouping data points this way can help more clearly view, analyze, and explore data and trends in visuals.

Typically applied in the explore phase of an analysis project, grouping manually aggregates data points into groups. These groups become part of the data model and automatically apply to new or refreshed data.



While grouping applies to category fields, binning applies to continuous fields such as date fields and numeric fields. Power BI recognizes these as continuous fields and brings up a dialog box that allows you to bin the results by setting the size of each bin.

Learn more:
[Grouping and binning](#)

Data streaming in Power BI

Real-time dashboards

Power BI lets you easily display and analyze your real-time data, empowering your organization to gain instant insights from time-sensitive information. Monitor social media campaigns as they go viral. Display streaming video on your dashboards. Bring your dashboards to life with IoT sensor readouts. The rich functionality of Power BI, combined with the velocity of real-time data, will transform the way you do business.

Real-time data is all around us, and Power BI features are designed to help you get to that data with minimum setup cost. Integration with streaming data providers such as Azure Event Hub, PubNub, and Temboo allows Power BI to get to your real-time data—no matter where it lives. In addition, Power BI allows you to dive deep into your real-time data through integration with Microsoft Azure Stream Analytics and Azure Machine Learning. Predictive intelligence can help you take proactive action to stay on the right course, and stream analytics can shape and aggregate your data. Data streams are stored in the cloud for historical analysis and visualization.

Setup of real-time streaming data sets

With Power BI real-time streaming, you can stream data and update dashboards in real time. Any visual or dashboard that can be created in Power BI can also be created to display and update real-time data and visuals. The devices and sources of streaming data can be factory sensors, social media, service usage metrics, and anything else from which time-sensitive data can be collected or transmitted. Streaming data can be consumed two ways in Power BI: as tiles with visuals from streaming data, or as data sets created from streaming data that persist in Power BI.

Learn more:

[Real-time streaming in Power BI](#)

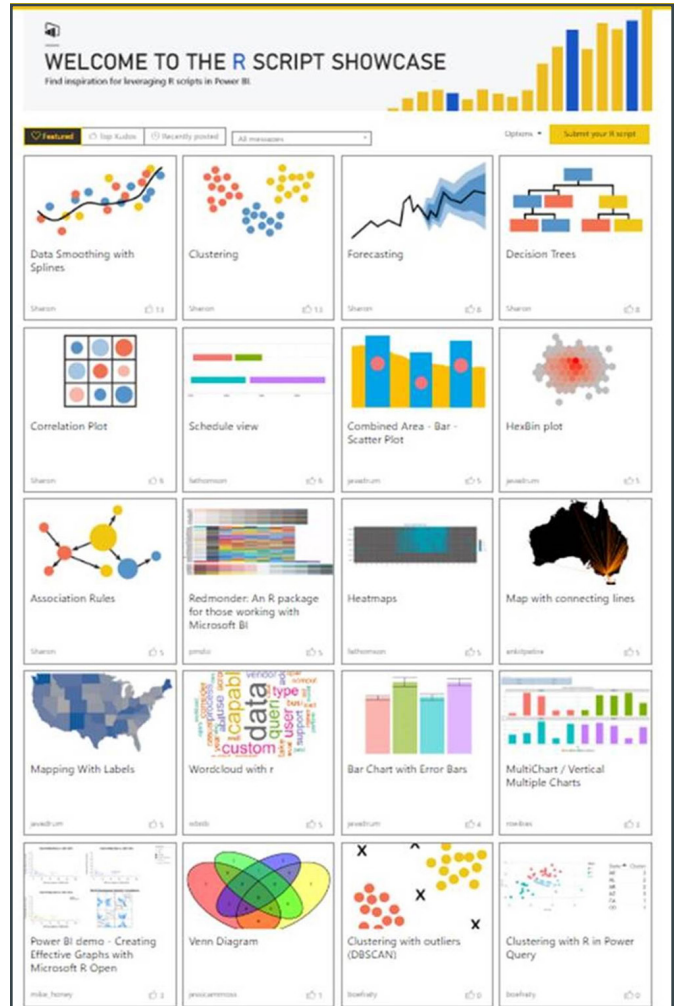
Visualizations in Power BI

Data visualizations allow you to interact with your data to find business insights. Power BI lets you choose from a list of available visualizations, add a custom visualization that you create yourself, or select from our expanding list of available visualizations in the community gallery.

Community-sourced visualizations

Power BI has a visuals gallery with many useful visualizations created by both the community and Microsoft, which you can download and use in your Power BI reports. To add a community-sourced visualization to your report, visit the visuals library on the Power BI site. On the Welcome to Power BI custom visuals page, browse the gallery. Select a visual tile to see more information about that visual, and download the visual that you want to use.

Learn more:
[Download a custom visual from the gallery](#)
[Add a custom visualization to a Power BI report](#)



R visualizations

The R Showcase in Power BI allows you to create new or use existing advanced analytics in R visualizations through the community R Script Showcase to leverage R scripts in Power BI.

Using the R Showcase, you'll be able to apply complex algorithms, visualizations, scripts, and more with just a click. You don't even need to know anything about R to use the R custom visuals. Just import an R custom visual to your report, drag the data, and the R power will be applied without writing one line of R code.

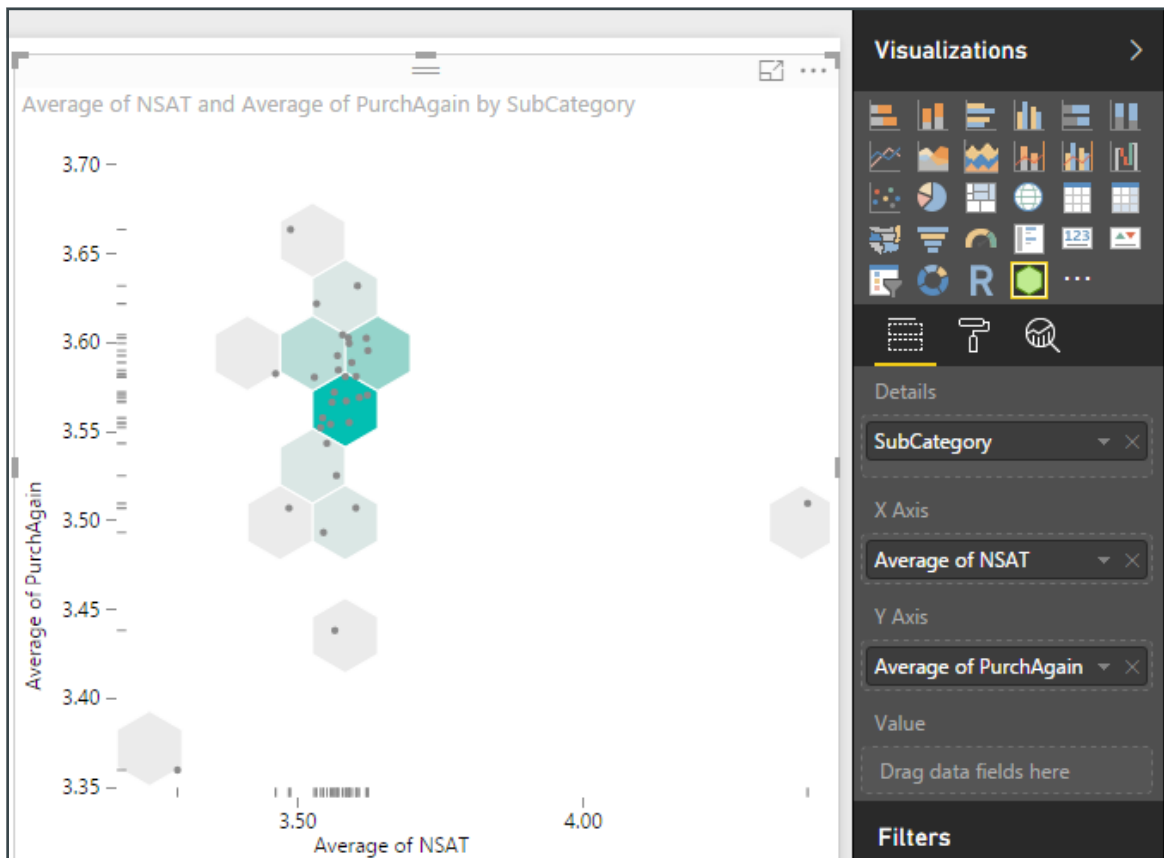
To use the R Script Showcase in your reports, visit the community page of the Power BI website. Click the **R Script Showcase** link, and then select and download the report you would like to use.

Learn more:
[R Script Showcase](#)

Custom visualizations

Custom visualization in Power BI allows you to create full custom visuals to add to reports or submit to the Power BI community for others to use. You can create a custom visual using Power BI developer tools, which let you design and test a custom visual by writing custom visual TypeScript code and creating CSS.

Once you've tested your custom visual, you can export it to your Power BI dashboard, or submit it to the Power BI visuals gallery.



Learn more:

[Power BI Custom Visuals Gallery](#)

Data connection and shaping

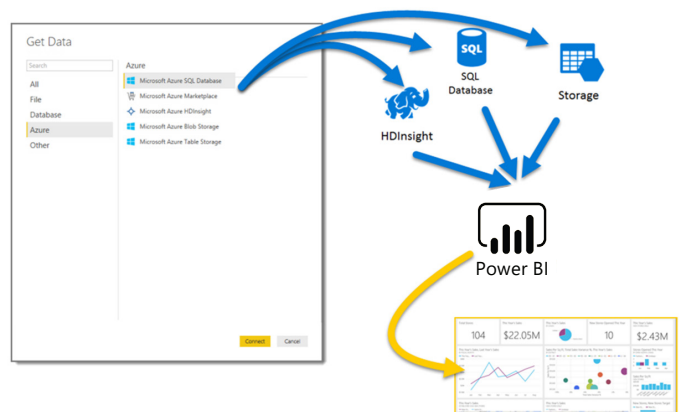
Azure services

With Azure services and Power BI, you can turn your data processing efforts into analytics and reports that provide real-time insights into your business. Whether your data processing is cloud-based or on-premises, straightforward or complex, single-sourced or massively scaled, warehoused or real-time, Azure and Power BI have the built-in connectivity

and integration to bring your business intelligence efforts to life. A multitude of Azure connections are available, and the business intelligence solutions you can create with Azure services are as unique as your business. You can connect just one Azure data source, or a handful, and then shape and refine your data to build customized reports.



To use an Azure connection in your Power BI dashboard, select an Azure data source in the **Get Data** dialog to bring in a wide variety of sources.



Learn more:
[Azure and PowerBI](#)

[Stream Analytics & Power BI](#)

DirectQuery

DirectQuery allows you to build visualization over very large data sets, where it would otherwise be unfeasible to import all of the data. Traditionally, underlying data changes can require a refresh of data. For some reports, the need to display current data can require large data transfers, making re-importing data unfeasible. Users can avoid this issue by using DirectQuery to query live against the data source.

When using DirectQuery, it is important to consider the performance and load of your data set, the security, and available supported features. All DirectQuery requests are sent to the source database, so the time required to refresh a visual depends on how long the back-end source takes to respond with the results from the query. Power BI creates queries that are as efficient as possible, but there are conditions under which a generated query may not be efficient enough and the performance may be impacted or the refresh would fail. This situation can be mostly avoided by using columns with a cardinality below 1 million.

Security should also be considered when using DirectQuery in the Power BI service. All users who consume a published report connect to the back-end data source using the credentials entered after publication to the Power BI service. This is the same situation as data that is imported: all users see the same data, irrespective of any security rules defined in the back-end source. Finally, not all Power BI Desktop features are supported in DirectQuery mode, or the feature may have some limitations. In addition, some capabilities in the Power BI service (such as Quick Insights) are not available for data sets using DirectQuery.

Learn more:

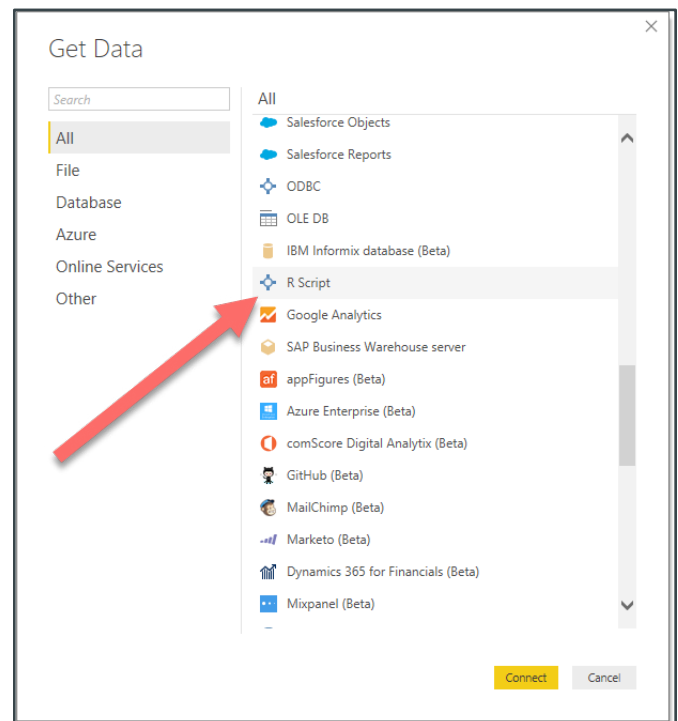
[Use DirectQuery in Power BI Desktop](#)

Data fetching with the R connector

Using the R connector, users can run R scripts directly in Power BI Desktop, and import the resulting data sets into a Power BI Desktop data model. With just a few steps, you can run R scripts and create a data model. To run an R script in Power BI Desktop, first create the script in your local R development environment, and make sure it runs successfully. Next, in Power BI Desktop, find the R Script data connector in **Get Data**. To run your R script, select **Get Data** and then **More**. Next, select **Other**, and then **R Script**.

Learn more:

[Visualizing and operationalizing R data in Power BI](#)

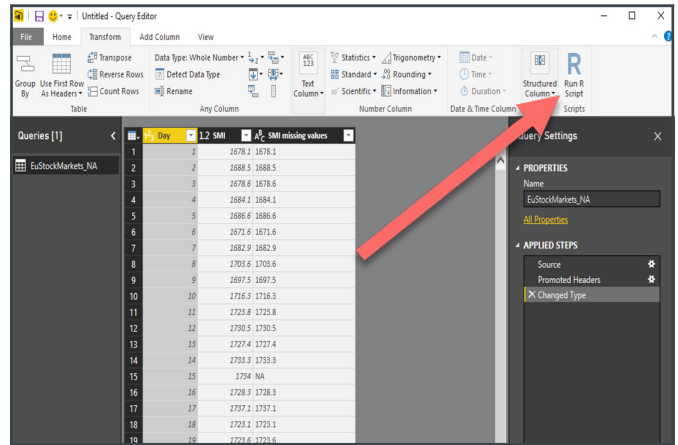


Data shaping in Power Query with R

The integration of R into Query Editor lets you perform a data cleansing using R, and then perform advanced data shaping and analytics in data sets. Examples include completion of missing data, predictions, and clustering,

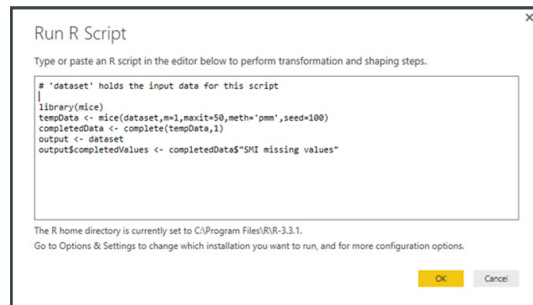
just to name a few. R is a powerful language, and can be used in Query Editor to prepare your data model and create reports. To use Query Editor, follow the steps below:

In the **Transform** tab, select **Run R Script**, and the Run R Script editor will appear:



When put into the Run R Script dialog, the code will look like this:

Note that for the R scripts to work properly in Power BI, all data sources need to be set to **public**.



Once we do so, we see a new column in the fields well called **completedValues**. With just five lines of R script, Query Editor will fill in the missing values (in rows 15 and 18 in the example) with a predictive model.

Data Analysis Expressions

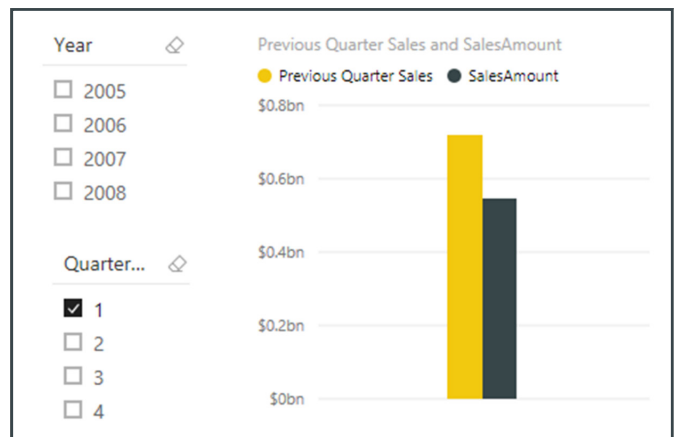
Data Analysis Expressions (DAX) is a collection of functions, operators, and constants that can be used in a formula or expression to calculate and return one or more values. DAX helps you create new information from data that is already in your model. Learning how to create effective DAX formulas will help you get the most out of your data. It's quite easy to create a new Power BI Desktop file and import data into it. You can even create reports that show valuable insights without using any DAX formulas at all.

But, what if you need to analyze growth percentages across product categories and for different date ranges? Or, you need calculate year-over-year growth compared to market trends? DAX formulas provide this capability and many other important capabilities as well. Learning how to create effective DAX formulas will help you get the most out of your data. When you get the information you need, you can begin to solve real business problems that affect your bottom line. This is the power in Power BI, and DAX will help you get there.

Here is an example of DAX in action.

```
Previous Quarter Sales =  
CALCULATE(SUM(Sales[SalesAmount]),  
PREVIOUSQUARTER(Calendar[DateKey]))
```

This formula will calculate the total sales for the previous quarter, depending on the filters applied in a report. For example, if we put **SalesAmount** and our new Previous Quarter Sales measure in a chart, and then added **Year** and **QuarterOfYear** as slicers, we'd get something like this:



Learn more:
[Introduction to DAX](#)

Conclusion

Power BI business analytics tools enable users at all levels of an organization to analyze data and share insights. Through dashboards, Power BI provides a 360-degree view of your most important metrics—in one place, updated in real time, and available on all of your devices. With one click, explore data using intuitive tools to quickly find answers and

uncover new insights. Power BI allows you to dig deep into your data, while being productive and creative with what you build. With more than 20 built-in visuals and a gallery of vibrant custom visualizations, Power BI makes it easy to use advanced analytics to effectively communicate your message and address business challenges.

Additional resources:

[Explore Power BI](#)

[Download for free](#)

[Guided learning](#)

[Join the community](#)

[Blog](#)

[Solutions](#)