

EXTEND POWER BI CAPABILITIES WITH PYTHON & R

Christopher Hurley; Shriya Prasanna; Yara Seif



Expand your insights by utilizing open-source data visualizations within Power BI.

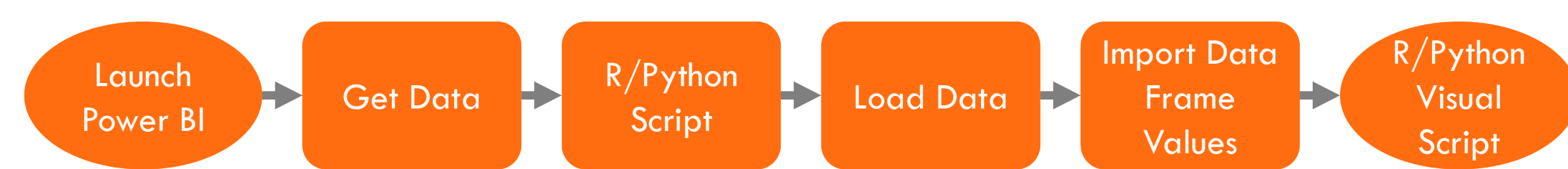
The aim of this poster is to demonstrate the utilization of open source visualizations within Power BI. Open-source programming packages power your data analysis through the ability to create rich visuals and incorporate advanced statistical methods into your plots and charts. By combining the sophistication and depth of the open source packages with the simplicity and ease of Power BI, there are no limits to the insights you can quickly derive from your data.

R and Python are two of the most popular open-source programming languages. We will show visualizations created with each. Finally, we will show an example of an open-source output integrated within a Power BI dashboard. The inherent flexibility of Power BI enables a wide variety of data functionality such as automation and filtering which give the ability to dive deeply into the data.

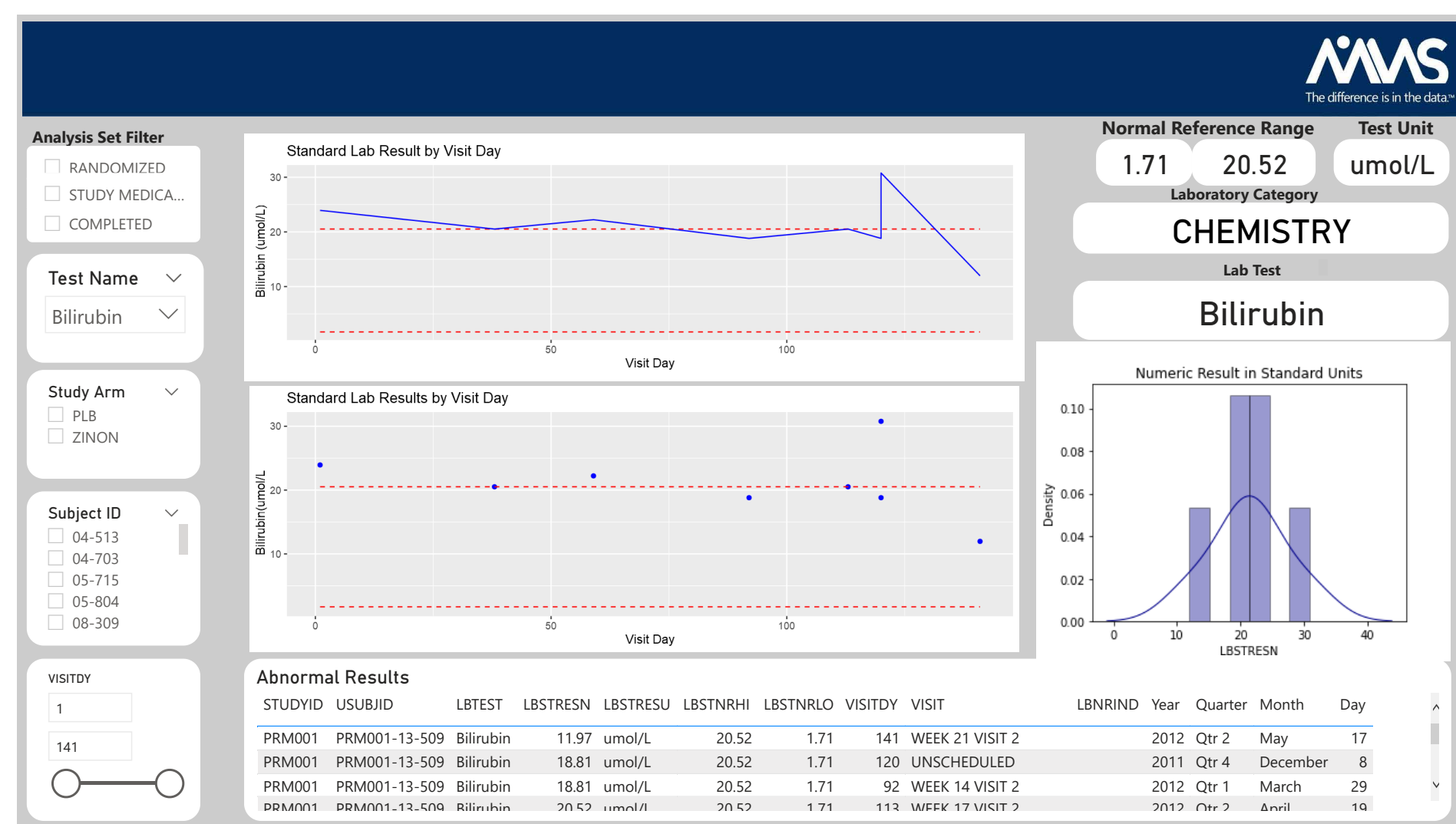
Using Python and R in Power BI can allow quick experimentation, data/calculation validation, statistical analysis/inferencing of the data demonstrated in a single interactive Power BI dashboard.

One of the first steps required to work with data in Power BI Desktop is to **import** it by clicking on Get Data, and here you choose R/Python script.

After making sure Python/R are installed on the machine, import that data frame. A window will pop up with a preview of the data to be imported; load the data onto Power BI. Finally, your own Python or R scripts can be added to Power BI to create new interactive visualizations that also can be filtered by using appropriate slicers and filters.



R & Python Integrated in Power BI



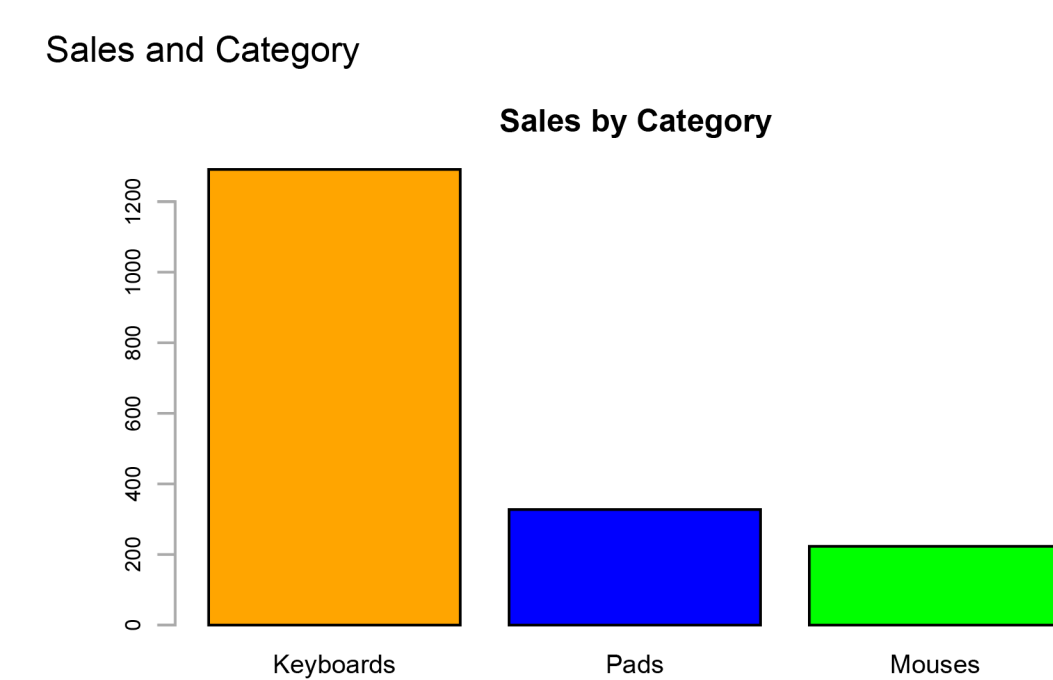
Integrating R & Power BI

With Power BI Desktop, you can use R to visualize your data. R is a language and environment for statistical computing and graphics.

The first official R release came in 1995. The Comprehensive R Archive Network (CRAN) was officially announced 23 April 1997 with 3 mirrors and 12 contributed packages. The first official “stable beta” version (v1.0) was released on 29 February 2000.

The PHUSE Education website has some excellent resources to learn more about Data Science and open-source programming languages.

<https://phuse.global/Education>



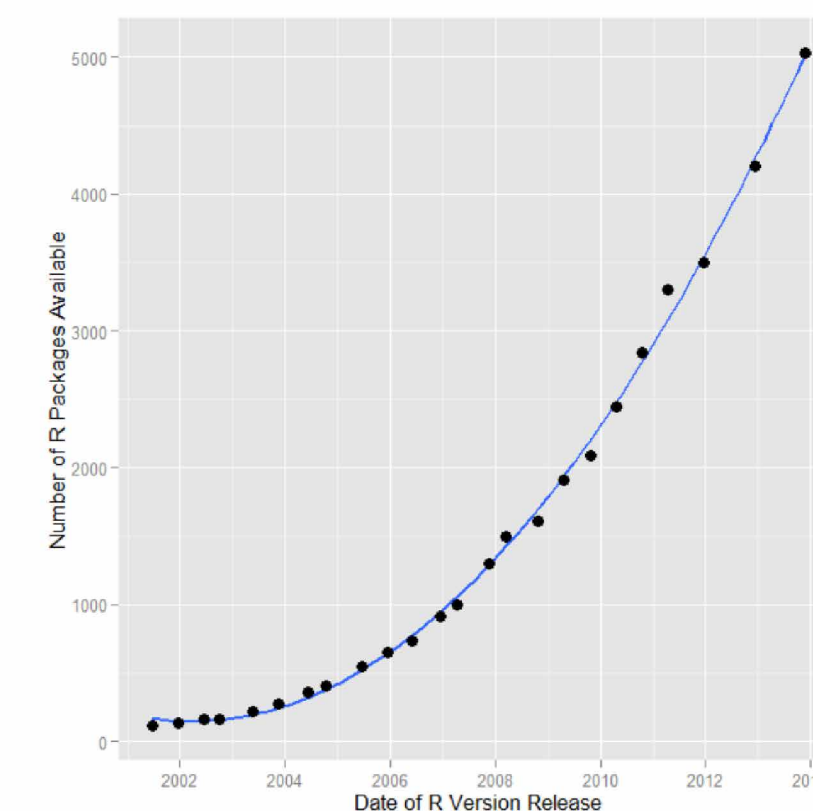
Power BI in the Modern Business Intelligence

The use of Power BI has greatly exceeded predictions. There are now over half a million unique users, across more than 45,000 companies spanning 185 countries have signed up for Power BI.

Beyond the inherent benefits of a SaaS-based approach, Power BI sets a new standard for “modern” business intelligence with a collection of capabilities completely unique in the industry.



Figure 2.



R is the most popular open-source language for Data Science. There are thousands of packages and libraries provided for doing different tasks. For example, dplyr and data.table are used for data manipulation, ggplot2 is used for data visualization, tidyr enables data cleaning, ‘Shiny’ gives the ability to create a Web application, knitr provides report generation and finally mlr3, xgboost, and caret are used in Machine Learning.

Integrating Python & Power BI

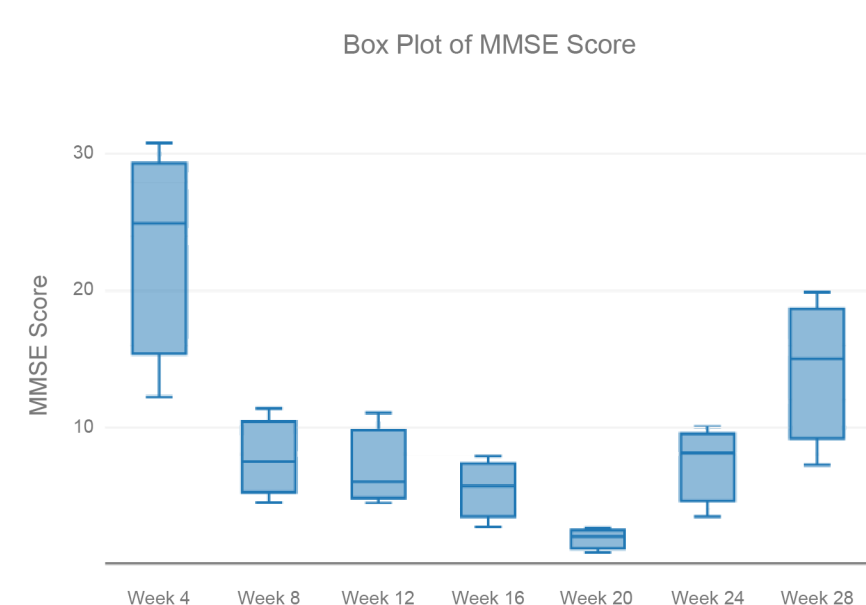
Power BI is Microsoft’s proprietary product for performing business intelligence tasks. Since 2018, integrating statistical and general-purpose software like R and Python has been made possible by Power BI.

Using Python in Power BI can:

1. Automate and optimize processes
2. Give the ability to create amazing customized visualizations
3. Enable the inclusion of machine learning modules and plot the associated data visualizations based on the predicated result.

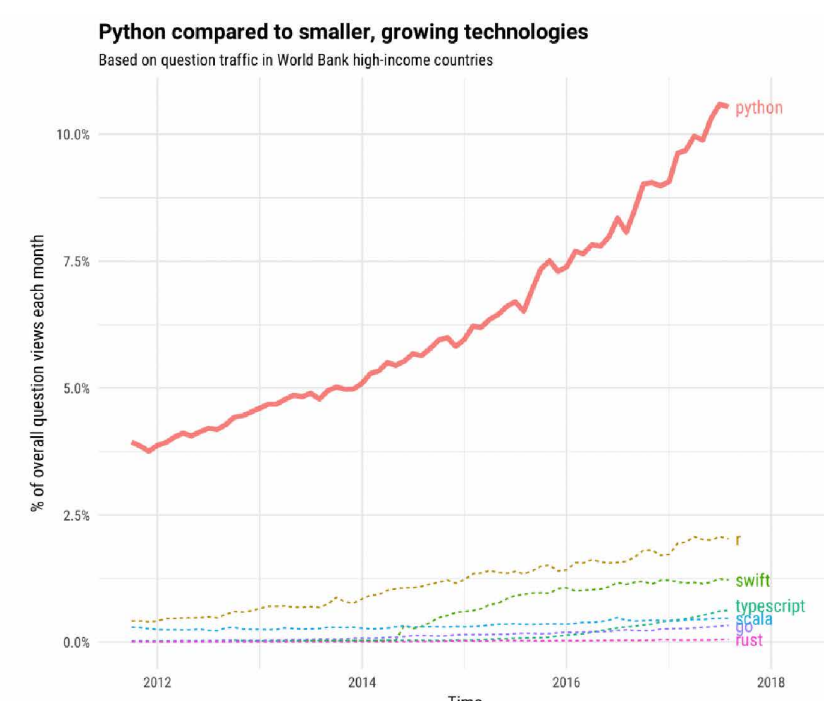
Most used libraries for data visualization include **Matplot**, **Plotly** and **Seaborn**.

Using Plotly: Here is an example of a box plot for a statistical representation of numerical data through their quartiles. The ends of the box represent the lower and upper quartiles, while the median (second quartile) is marked by a line inside the box.



Plotly has hover tool capabilities that allow us to detect any outliers or anomalies in many data points.

Figure 1.



There are over 137,000 python libraries and 198,826 python packages ready to ease developers’ regular programming experience. These libraries and packages are intended for a variety of modern-day solutions.¹

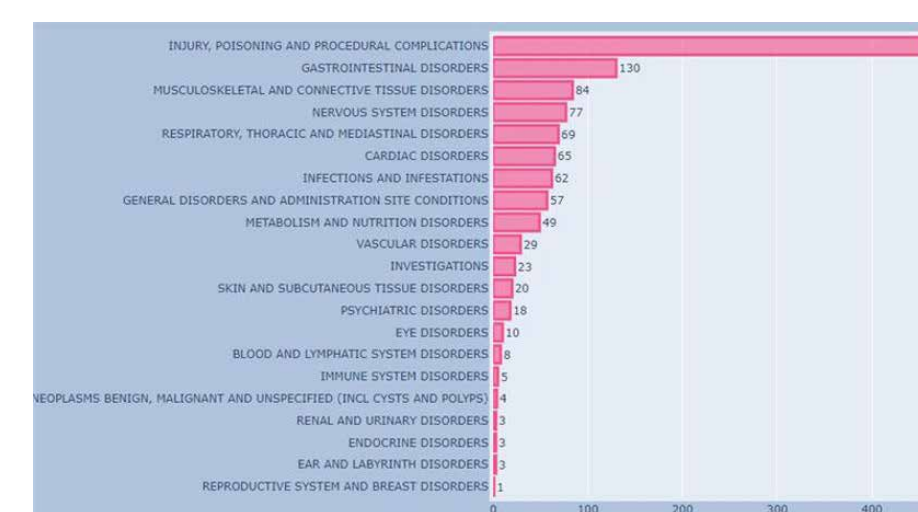
The following Python packages (non-Intel MKL) are currently supported for use in your Power BI reports: matplotlib; numpy; pandas; scikit-learn; scipy; seaborn; statsmodels.

Python & R Visualizations in Power BI

Python Visualization in Power BI

Data Description

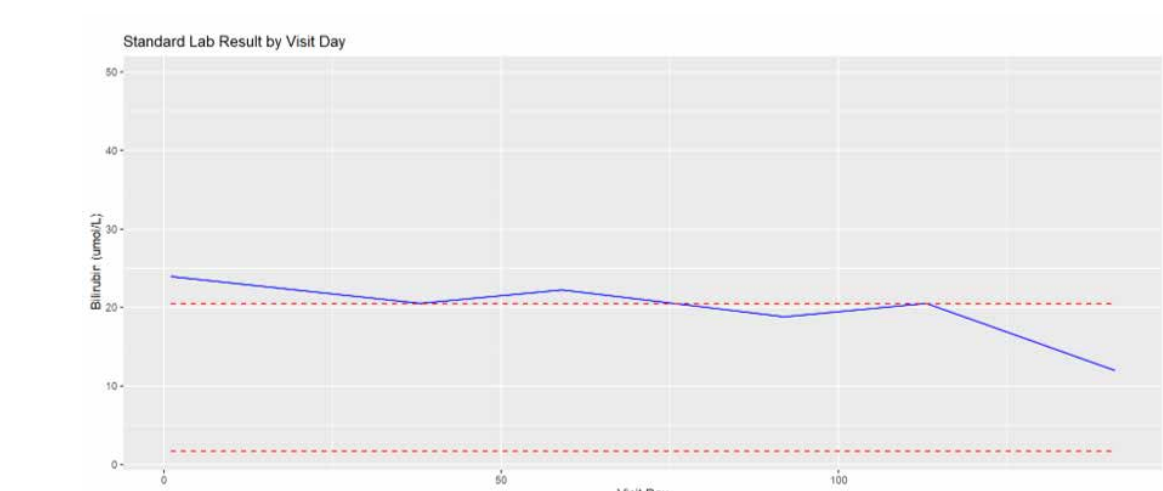
The data visualized in the stacked bar chart for all the lab test shown by visit day: visualized by using Python script using scatter plot feature from matplotlib library.



R Visualization Line Chart

Data Description

The data visualized in the line chart for all standard lab results shown by visit day: visualized by using R script using line geom_line feature from ggplot library.



References

1. Bichiashvili et al. Create power bi visuals using python in power bi desktop - power bi. Create Power BI visuals using Python in Power BI Desktop - Power BI | Microsoft Docs. <https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-python-visuals>. Published November 1, 2021. Accessed April 11, 2022.
2. Elizabeth J. Python is the fastest-growing major programming language, says stack overflow. JAXenter. <https://jaxenter.com/python-explosive-growth-stack-overflow-136943.html>. Published September 7, 2017. Accessed April 11, 2022.

The Data Science experts at MMS open new possibilities for Sponsors searching for data-driven insights for pharmaceutical, biotech, and healthcare organizations. The team employs a wealth of regulatory expertise, strong processes, and purpose-built, flexible technology to drive efficiency using real world data (RWD). With a global footprint across four continents, MMS maintains a 97 percent customer satisfaction rating, and the company has been recognized recently as Most Innovative CRO.

Learn more about MMS and our services by visiting www.mmsholdings.com or reach out to one of our poster authors:

- Christopher Hurley - churley@mmsholdings.com
- Shriya Prasanna - sprasanna@mmsholdings.com
- Yara Seif - yseif@mmsholdings.com

