

Self-Service Data Analytics

Bharath Sunkavalli, Bayer U.S. LLC, Whippany, USA

ABSTRACT

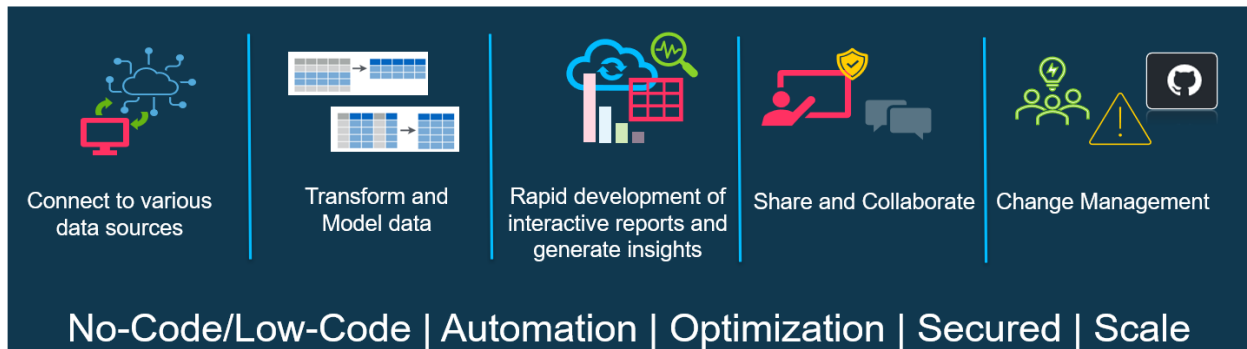
Data Analytics has become an integral part of all functions working with clinical data. Several functions such as Data Management, Data Analytics, Medical Review etc. are taking advantage of interactive applications to perform their tasks. However, the development and maintenance of such applications requires dependency on other groups such as IT/Infrastructure teams. This paper will showcase the possibilities of self-service data analytics with Power BI that will allow functional groups to develop business solutions on their own. With Self-service data analytics, functional groups will be able to connect and fetch data from various sources, transform and visualize data, build interactive dashboards, and communicate results securely. Power BI caters these services in one no-code/low-code environment and in addition it supports R/Python scripts to enhance its analytical and visualization capabilities. Organizations of any size can leverage domain expertise combined with Power BI capabilities to gain real-time actionable insights.

INTRODUCTION

It has been an emerging trend in all functional groups of Clinical Operations to leverage Data Analytics for actionable insights and take data-driven decisions. Data Analytics with emerging technologies involves provision of a web application with user focused interactive reports and data-driven narratives. These applications are developed and maintained by Data Analyst/Data Scientists and are delivered to Subject Matter Experts/End users of that functional group. There are several tools available to develop these applications however to deploy and scale in a secured environment has been a challenge for data analysts due to dependencies on external teams such as IT Management. This paper will illustrate on how to self-serve development and maintenance of applications within functional group by leveraging Business Intelligence (BI) tools. There are several BI tools with self-service capabilities however for illustration purposes Power BI will be used as reference in this paper.

CORE FUNCTIONALITIES

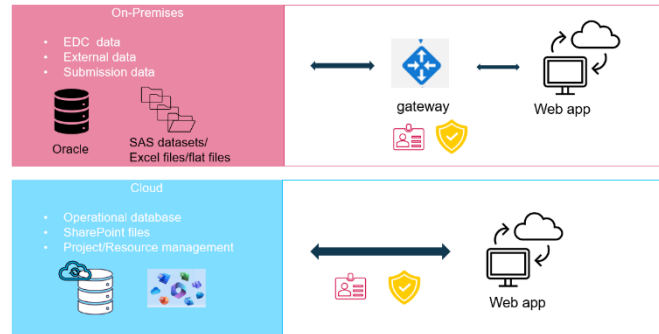
Self-Service Data Analytics is capable of below core functionalities to support business use case within a functional group. It starts with connections to variety of data sources, reshape and format data as per reporting needs, build interactive reports, and finally share the results and collaborate to make informed and collective decisions. Details of each functionality are explained in below sections, they are limited to the experience acquired in recent years working with distinct use cases across functional groups within our organization. For in depth technical details and solutions specific to your use case, it is recommended to refer to BI tool's documentation or online resources and adhere to your company's SOP's and best practices.



DATA CONNECTIONS

BI tools come with inbuilt data connectors that will allow you to connect to databases, files, online services, and several other data sources.

To connect to on-premises data sources such as Oracle database and file servers with SAS® datasets and excel files, we need to set up a gateway on the server where our application is hosted. Gateway will need the server details and authentication method for connection. Once the Gateway is set-up the web app can securely interact with source data. For data sources that are in cloud will not need a gateway and can still support secure connection between data source and web app.



DATA WRANGLING

BI tools offer No-code/Low-code data transformations with graphical interface that will allow to reshape and combine data as per your reporting needs. Data transformations within Power BI are done with Power Query engine which runs on M language and Data Analysis Expressions (DAX) language, it also allows you to execute R and Python scripts for data transformation. In certain use cases where data wrangling is complex, best practice is to pre-process the data and use Power BI for visualization. This will make the data transformations transparent for other Analysts, supports validation and code maintenance.

DATA INSIGHTS

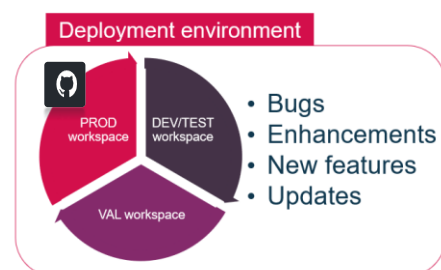
Interactive reports and data-driven insights can be created quickly with drag & drop gestures. Visual elements (visuals) include graphs, tabular displays, filters, and data-driven narratives. If the default visuals do not fit your purpose, then custom visuals can be imported from third party developers. To support advanced analytics and visualizations R/Python scripts can be used to create visuals however they can't be used to filter other visuals. All visuals are interconnected allowing users to highlight, filter and drill down to other visuals. Templates can be created to scale the reports, across clinical studies or functions.

SHARE & COLLABORATE

Reports that are ready to be shared can be published from Power BI desktop to a shared workspace. Once published report authors can grant access to users and control the read/edit permission. The access can be granted to individuals and Microsoft 365 groups and users will be authenticated with Azure Active Directory. The report can then be shared with a web link where end users can consume the reports, save the state of report with personal bookmarks and collaborate with others with inbuilt chat functionalities. Reports can be refreshed on-demand by end user, or it can be scheduled. Datasets that failed to refresh for any reason will send you an email notification. All these features are built-in to Power BI services and can be triggered with click of button.

MAINTAINENCE

Applications need regular maintenance to fix bugs, enhancements, new features etc. To handle post-production changes a change management process must be defined as per company SOP's. It is common practice to have DEV/VAL/PROD workspaces to allow validation and UAT before the new version is released in Production. Version control and change history can be handled by GIT repositories or for a less critical application Microsoft SharePoint may suffice.

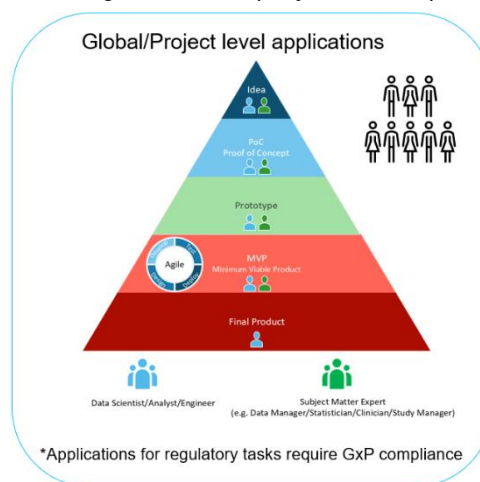


APPLICATIONS

BA tools can support the needs of a single individual, functional group, department, and Organization. The life cycle to develop an application depends on the criticality of reports and should be aligned with company standard operating procedures.

APPLICATION DEVELOPMENT LIFE CYCLE

The development lifecycle is as shown in the right. It is very important that end users are involved in every step of application development life cycle to ensure application meets end user requirements. We start with ideation phase together with end user and create a PoC - to check the technical feasibility of concept, followed by prototype - which is a working model but not validated. During these steps we develop quick and dirty reports to ensure the prototype meets core user requirements. It is possible we may not meet user requirements and may discard the idea and we start over with new concept. This is the benefit of rapid prototyping; it will allow us to fail fast at an early stage of development. Assuming the prototype is approved by end users you can apply agile methodologies to develop Minimum Viable Product (MVP). MVP is a thoroughly tested by peer programmers (Testing) and end users (UAT). New features are gradually added to MVP in every iteration (sprints) until it meets all user requirements of final product. In addition, if the application supports regulatory activities it needs to be published to GxP compliant environment all other applications need not be GxP compliant but should follow validation Standard Operating Procedures to ensure plausible results.



USE CASES:

- Project/Study lead can monitor status of deliverables, proactively identify risks and plan mitigation strategies
 - Create reports for RWD/RWE teams by fetching data from various data sources
 - Graphical patient profiles
 - Study Analyst can automate data quality and compliance checks on EDC and external data
 - Create dashboards that automate manual Clinical data review and cleaning process.
- Check out [PHUSE EU connect 2022 presentation](#)

CONCLUSION

Business intelligence (BI) tools such as Power BI will allow functional groups to create fast and reliable end-to-end solutions to business problems. The non-functional requirements such as publishing, scaling, data security and user access management services are built-in to application servers, this will allow Data analysts/Data scientists to focus on functional requirements, such as end-user requirements, user interface and user experience. BI tools are easy to learn, ample online resources and reliable support from user communities. Functional groups can significantly benefit by upskilling data analysts within the group as they can leverage existing domain expertise.

Self-Service Data Analytics plays significant role in achieving company business objectives of Optimization, Automation, and Data Democratization that promotes real-time, data-driven, and collective decision-making.

RECOMMENDED READING

Tumor assessment evaluation using MS Azure and Power BI: [PHUSE EU connect 2022 presentation](#)

Power BI get started: <https://learn.microsoft.com/en-us/power-bi/fundamentals/>

On-premises data gateway: <https://learn.microsoft.com/en-us/power-bi/connect-data/service-gateway-onprem>

R visuals in Power BI: <https://learn.microsoft.com/en-us/power-bi/visuals/service-r-visuals>

Power BI licenses: <https://learn.microsoft.com/en-us/power-bi/fundamentals/service-features-license-type>

Power BI templates: <https://github.com/bharath-viz/PBI>

CONTACT INFORMATION

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

Author Name: Bharath Sunkavalli

Company: Bayer U.S. LLC

Address: 100 Bayer Boulevard
Whippany NJ 07981

Email: bharath.sunkavalli@bayer.com

Web: www.linkedin.com/in/bharath-sunkavalli