Power BI
Jargon Buster

A handy guide by Chandoo

Chandoo.org
BECOME AWESOME

What is Power BI?

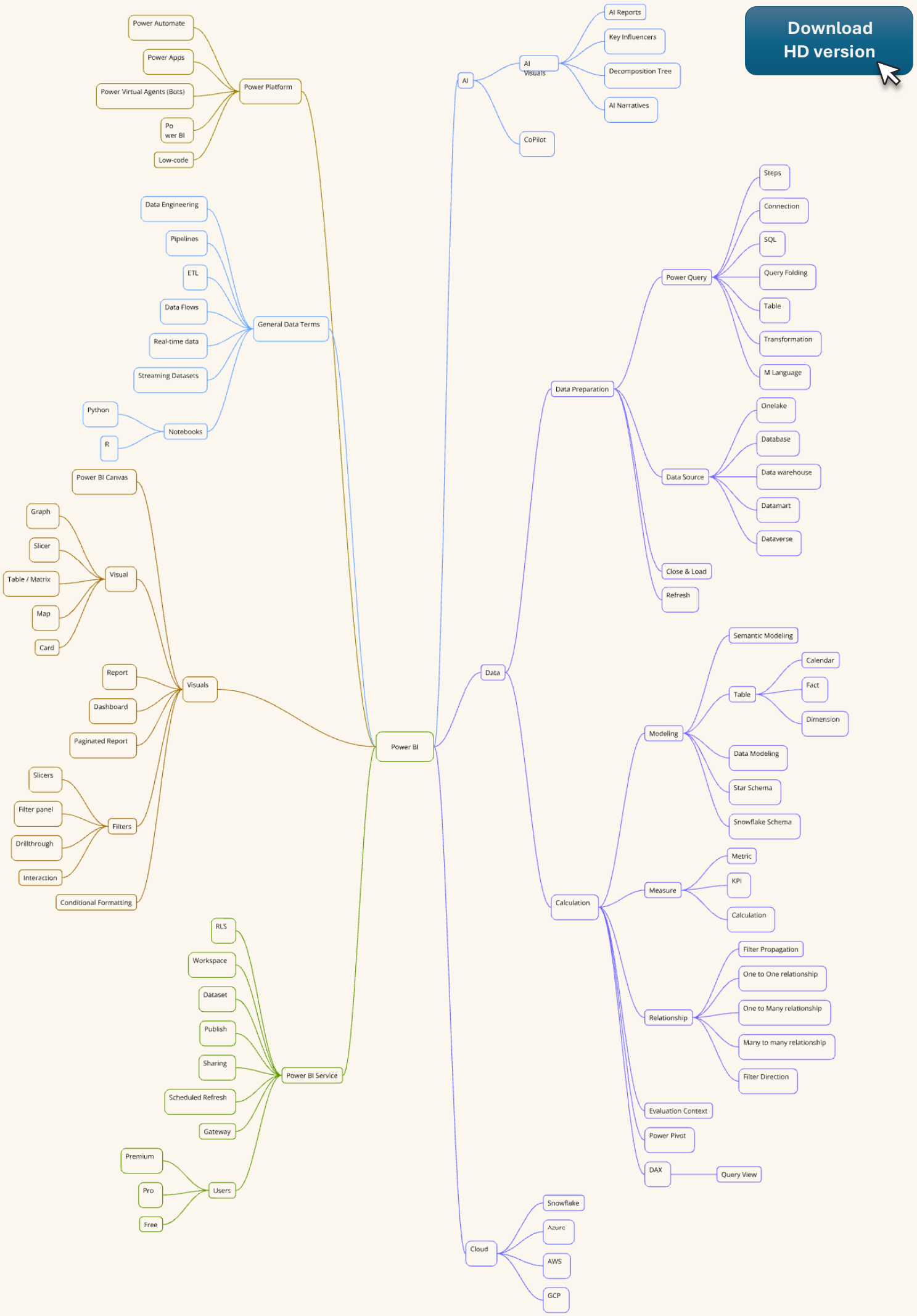
Power BI is a data analysis platform. We use it to create business intelligence (BI) reports, dashboards and share them easily with others in your organization (or outside).

A sample Power BI report of sales performance.

What are the key ingredients of Power BI technology

- Power Query for connections, data prep, querying, cleanup, and joins
- Power Pivot for data and semantic modeling (i.e. linking various tables, defining calculations)
- Power BI Visual layer for making visuals, setting up interactions and formats.
- Power BI Service to share the datasets and reports with your audience.
- Power BI Apps to share reports with mobile audience.

Power BI Mind-map



60+ Power BI Terminology

1. **Connections**: When you bring data from anywhere to Power Query, you are making a connection. For example, you can "connect" to SharePoint to bring CSV file into Power Query.
2. **Query**: The process of getting data from somewhere, cleaning / preparing it for analysis is called "querying".
3. **Table**: A table is a collection of rows & columns. Usually, each column contains one kind of data and each row is one *instance* of the data. For example, a shipment table contains list of all the shipments we make, each row with one shipment.
4. **OneLake**: A centralized data lake in Microsoft Fabric that provides a unified storage layer for all types of data. It allows seamless integration with various data services and tools within the Microsoft ecosystem.
5. **Warehouse**: A large, centralized repository that stores structured data from multiple sources for analysis and reporting. Data Warehouses are optimized for creating business intelligence reports, machine learning models and large-scale analysis.
6. **Database**: A place where we keep data. Examples include MySQL, Oracle, SQL Server and PostgreSQL. Power BI can connect to various types of databases to source data for analysis and reporting.
7. **Datamart**: A subset of a data warehouse focused on a specific business area or department. In Power BI, it's a feature that allows users to create managed semantic models for easier data sharing and reuse.
8. **Dataverse**: A cloud-based storage platform that securely stores and manages data used by business applications. In other words, another *kind of database*. These are primarily used to manage the data needs of a Power App (a simple no-code data application you can make with Power Platform).
9. **SQL**: The language of databases and warehouses. Structured Query Language, a standard language for managing and manipulating relational databases. Power BI uses SQL to query databases and retrieve data for analysis.
10. **Transformation**: The process of modifying, cleaning, or restructuring data to make it suitable for analysis. In Power Query, transformations are applied as steps to prepare data before loading it into Power BI. For example, splitting full name to first and last name columns is a transformation.
11. **Refresh**: The process of updating a Power BI report or dataset with the latest data from its source. This ensures that your visualizations and analyses are based on the most current information.
12. **Query Folding**: An optimization technique where Power Query pushes data transformation steps back to the source database. This improves performance by reducing the amount of data transferred and processed locally.

13. **M Language**: The formula language used in Power Query for data transformation and manipulation. It's a functional language designed for building queries that extract and transform data from various sources.
14. **Close & Load:** A command in Power Query that finalizes the query design and loads the resulting data into Power BI for analysis. It marks the transition from data preparation to data modeling and visualization.

15. **Query Parameter**: A variable used in Power Query to make queries more flexible and reusable. For example, you can have a “folder name” parameter in Power Query so it can point to a *dynamic* folder.
16. **Power Pivot**: This is the calculation engine of Power BI. We can create complex data models, define relationships between tables, and create advanced calculations using DAX with Power Pivot.
17. **Modeling**: The process of structuring and organizing data to create a logical representation of business information. In Power BI, modeling involves defining relationships, creating calculated columns and measures, and setting up hierarchies.
18. **Data vs. Semantic Modeling**: In older versions of Power BI the Modeling activity (described above) is referred to as data modeling. But now, Power BI (and Fabric) call this Semantic Modeling. They both refer to the same concept of modeling your table relationships, calculations and business rules.
19. **DAX**: Data Analysis Expressions, a formula language used in Power Pivot for creating custom calculations. DAX is used to define measures, calculated columns, and custom tables in Power BI models.
20. **Measure**: A DAX calculation that performs dynamic aggregations or complex calculations in Power BI. For example, a **`Total Sales`** measure defines the logic to calculate the total sales. The formula could look like =SUM(`transactions[amount]`)
21. **KPI, Metric & Calculation**: These are some other words for the ***measure*** in the context of Power BI.
22. **Relationship**: A connection between two tables in a data model, typically based on a common field. Relationships in Power BI enable the combination of data from multiple tables for analysis and reporting. Power BI supports one-to-many and many-to-many relationships.
23. **One to many relationship**: A type of relationship where one record in a table can be associated with multiple records in another table. This is the most common type of relationship in Power BI data models.
24. **Filter Direction**: Determines how filters are propagated between related tables in a Power BI model. The default filter direction is from the "one" side to the "many" side of a relationship. Usually the filter direction goes from **dimension table** (one side) to the **fact table** (many side).

25. **Star Schema**: A data modeling technique where a central fact table is connected to multiple dimension tables. This schema is commonly used in data warehousing and business intelligence for efficient querying and analysis.
26. **Snowflake Schema**: An extension of the star schema where dimension tables are normalized into multiple related tables. This can lead to more complex queries but can save storage space and improve data integrity.

27. **Evaluation Context**: This is the *context* or what_is_happening_on_the_screen that tells Power Pivot how to evaluate the measure. For example, you may see different values for the **Total Sales** measure based on the evaluation context.
    - With no context: overall total value
    - With product context: total sales by product
    - With date context: total sales by month
    - With cross-filtering on person and date: total sales by person in that month.
28. **Power BI Canvas**: The main working area in Power BI Desktop where you create and arrange visuals for your report. It's where you drag and drop fields, adjust layouts, and design the overall look of your report.
29. **Visual**: Another word for **chart** or graph in the context of Power BI. Visuals are the building blocks of Power BI reports and dashboards, used to present data in an easily understandable format.
30. **Report**: A collection of visualizations, text, and other elements that together provide insights into a dataset. Reports in Power BI are interactive and can be shared with others.
31. **Dashboard vs. Report**: A dashboard is a single page that displays key metrics at a glance, often using tiles that link to reports. A report is a multi-page, detailed view of data with various visualizations and interactivity options.
32. **Filters**: Tools in Power BI that allow users to narrow down the data displayed in visualizations. Filters can be applied at the visual, page, or report level to focus on specific subsets of data.
33. **Slicer**: A type of visual filter in Power BI that allows users to interactively filter other visualizations on a report page. Slicers can be configured as lists, dropdowns, or even as timelines for date-based filtering.
34. **Drillthrough**: A feature that allows users to navigate from one report page to a more detailed page, carrying over the context of the selected data point. For example, if you have a Power BI report with 2 pages – business summary and product detail, you can ***drillthrough*** to product detail page to see the result for a specific product.
35. **Interaction**: When you click on one visual's data point, other visuals / elements on the page respond to this and *usually* highlight / show the relevant portion on them. This is called interaction.

36. **Paginated Report**: A type of report in Power BI optimized for printing. Paginated reports are ideal for operational reporting, invoices, or any scenario requiring precise formatting and pagination.
37. **Conditional Formatting**: A feature that allows you to apply formatting to data points based on specific conditions or rules. It helps highlight important information or trends in your data visually.
38. **Power BI Service**: The cloud-based platform where Power BI reports and dashboards are published, shared, and accessed by users. It provides collaboration features, scheduled refreshes, and mobile access to reports.
39. **RLS (Row Level Security)**: A feature in Power BI that restricts data access for certain users based on their roles. It ensures that users can only view data they're authorized to see, enhancing data security and privacy.
40. **Publishing a Report**: The process of uploading a report from Power BI Desktop to the Power BI Service. Publishing makes the report available for sharing and collaboration in the cloud.
41. **Dataset**: A collection of data used to create reports and visualizations in Power BI. A dataset can include data from various sources and contains the data model, relationships, and calculations. In a way, dataset is same as *semantic model*.
42. **Sharing a report**: The act of granting access to a Power BI report to other users or groups. Sharing can be done directly or through workspaces, apps, or embedding, depending on the scenario.
43. **Workspace**: A collaborative environment in the Power BI Service where teams can work together on reports, dashboards, and datasets. Think of them as a folder where all the files, datasets go.
44. **Scheduled Refresh**: A feature that automatically updates a dataset in Power BI Service at specified intervals. This ensures that reports and dashboards always display the most up-to-date data.
45. **Users**: Individuals who interact with Power BI, either as creators (developing reports and dashboards) or consumers (viewing and interacting with published content). User management in Power BI includes assigning roles and permissions.
46. **Power BI Pro & Premium**: Licensing options for Power BI. Pro is for individual users who need to create and share content, while Premium provides dedicated capacity and advanced features for organizations.
47. **Gateway**: A bridge that facilitates secure data transfer between on-premises data sources and Power BI Service. Sometimes you need to install a gateway when you have data in a local computer / network that is required for the reports published on the service (cloud). For example, if you have an Excel file with data that is referred to in the published Power BI report, you need a gateway to *automatically refresh* the report.
48. **Power Platform**: A suite of Microsoft business applications that includes Power BI, Power Apps, Power Automate, and Power Virtual Agents. These tools work together to enable data analysis, app creation, process automation, and chatbot development.

49. **Power Automate**: This is a RPA (Robotic Process Automation) tool from Microsoft. We can use it to automate actions or steps on your computer or on cloud.
50. **Power Apps**: A platform for building custom business applications with little to no code. Power Apps can be integrated with Power BI to create data-driven apps.

51. **Low-code**: An approach to software development that requires minimal hand-coding. Power BI and the Power Platform use low-code principles to enable rapid development of business solutions.
52. **CoPilot**: An AI-powered assistant in Power BI that helps users create reports, write DAX formulas, and generate insights from their data using natural language interactions.
53. **AI visuals**: Specialized visualizations in Power BI that use artificial intelligence and machine learning to provide advanced analytics, such as key influencers, decomposition trees, and anomaly detection.
54. **DAX Query View**: A feature in Power BI Desktop that allows users to write and execute DAX queries directly against the data model. It's useful for testing complex DAX expressions and troubleshooting measures.
55. **Cloud Platforms**: Online computing services that provide scalable and flexible resources for data storage, processing, and analytics. Power BI integrates with various cloud platforms, including Microsoft Azure, for enhanced data capabilities.
56. **Data Engineering**: The practice of designing, building, and maintaining data infrastructures and pipelines. In the context of Power BI, data engineering involves preparing and optimizing data for analysis and reporting.
57. **ETL**: Extract, Transform, Load - a process of extracting data from various sources, transforming it to fit operational needs, and loading it into a target system. Power Query in Power BI is often used for ETL operations.
58. **Data Flows**: A self-service data preparation technology in Power BI that enables users to ingest, transform, and load data into Power BI. Dataflows can be reused across multiple datasets and reports.
59. **Pipeline**: A series of connected data processing elements, where the output of one element is the input of the next. For example, you can have a pipeline to fetch data from website → combine with internal CRM system → Remove any duplicates → Calculate the discount percentages → update the invoices.
60. **Real-time data**: Information that is delivered immediately after collection, allowing for instant analysis and decision-making. Power BI supports real-time data through streaming datasets and real-time dashboards.
61. **Streaming data**: A continuous flow of data that can be processed in real-time. Power BI can connect to streaming data sources to create live-updating dashboards and reports.
62. **Notebook**: An interactive computing environment that combines code execution, rich text, mathematics, plots, and rich media. Jupyter Notebooks can be used with Power BI for advanced data analysis and preparation.