Microsoft Power BI
Certification Training (DA-100)

Module 2 – Data Extraction

Agenda

01 Overview of Power BI Desktop

02 Data Sources in Power BI

03 Using Files as a Data Source

04 Updating Files in Power BI

05 Using a Database as a Data Source

06 Using SaaS Connectors

Agenda

07 Other Data sources

08 Python Script

09 R Script

10 Power Query Editor

11 Areas in Power Query Editor

12 Advanced Editor

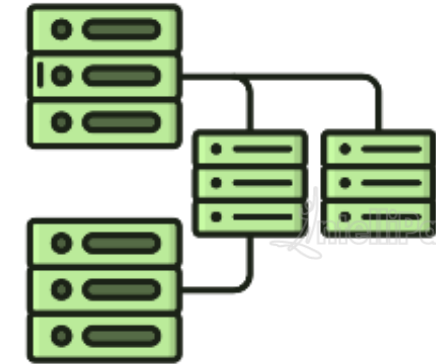# Overview of Power BI Desktop

Overview of Power BI Desktop

Power BI Desktop features a workspace for creating reports and comprises three key views in which we can work

Report View

Data View

Model View

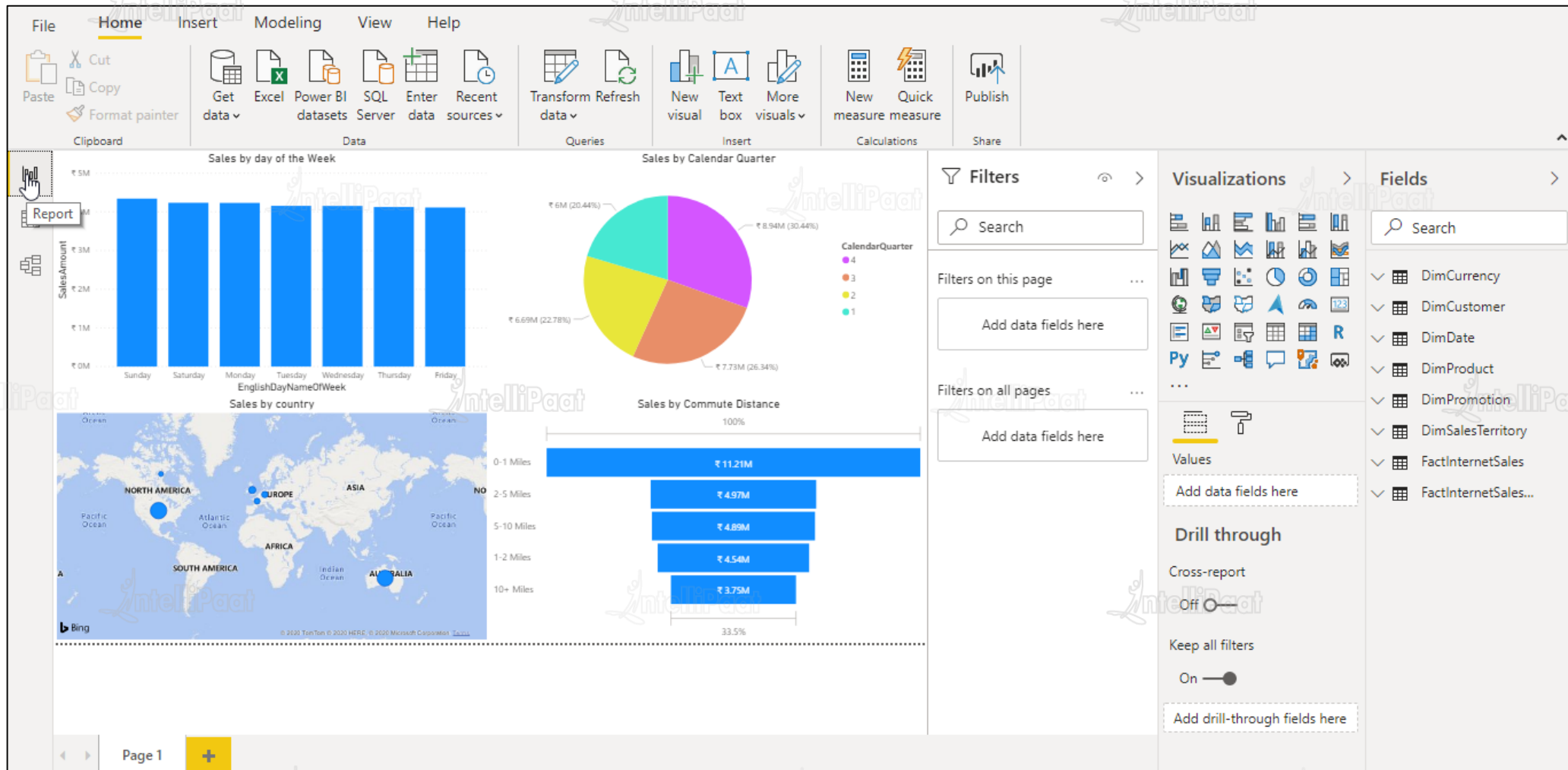# Overview of Power BI Desktop

Report View

Data View

Model View

This is the main workspace where we add report items, such as bar charts, maps, and pie charts, and display data using these report items

Overview of Power BI Desktop

Report View

Overview of Power BI Desktop

Report View

Data View

Model View

The data view is used to see imported datasets, in addition to shaping the data using transformations and M expressions

Overview of Power BI Desktop

Data View

Overview of Power BI Desktop

Power BI has an autodetect relationships feature, and we can also create relationships manually and set relationship properties

Report View

Data View

Model View

Overview of Power BI Desktop

Model View
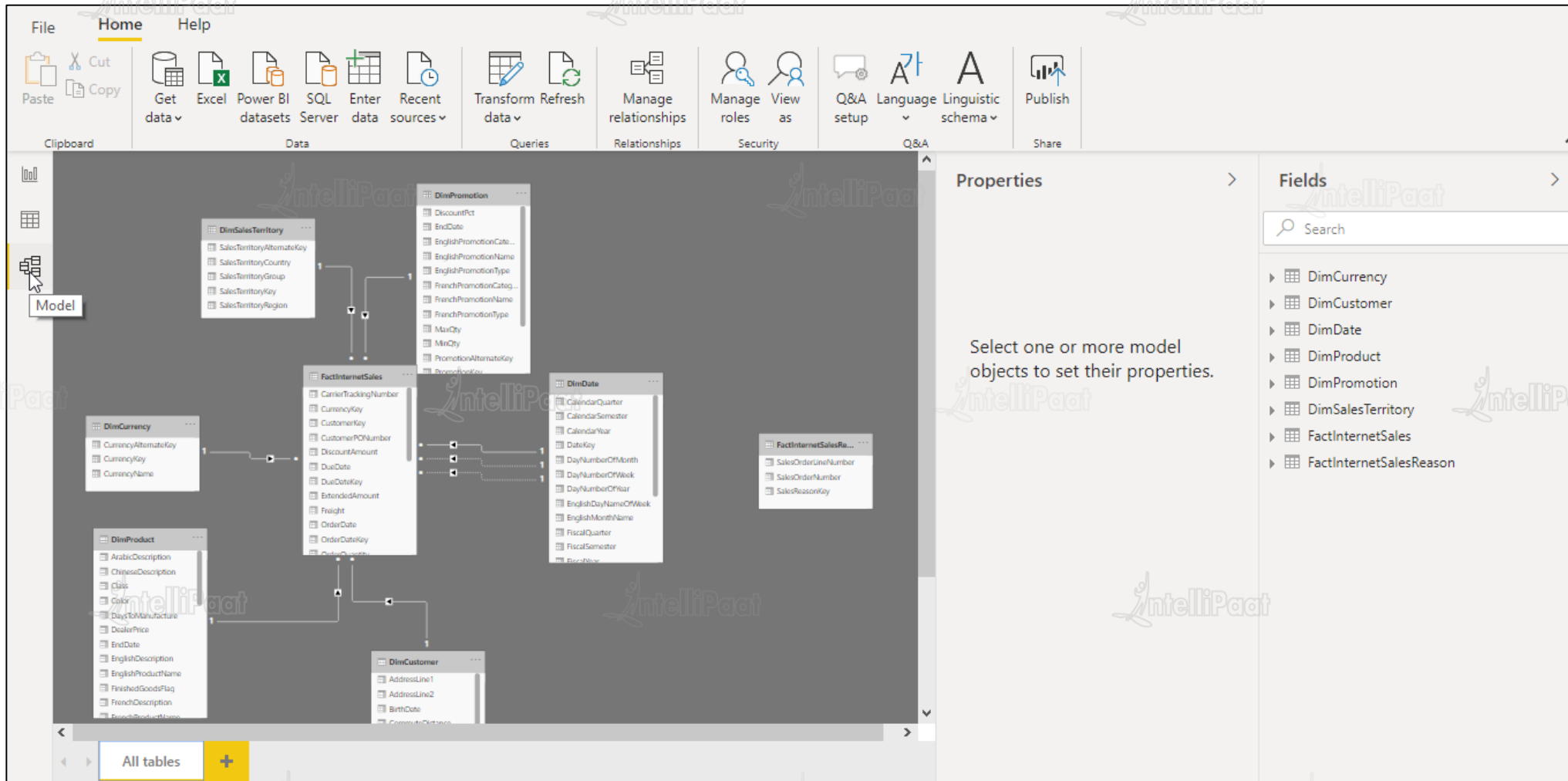
Data Sources in Power BI

Data Sources in Power BI

In Power BI Desktop, we have a wide choice of sources to import data from, including database, file, Azure, and SaaS connectors

Database Connectors
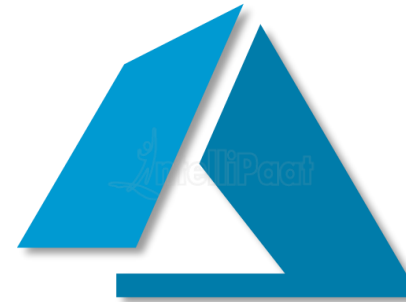
File Connectors

Azure Connectors

SaaS Connectors

Data Sources in Power BI

Data Sources

Database Connectors

File Connectors

Azure Connectors
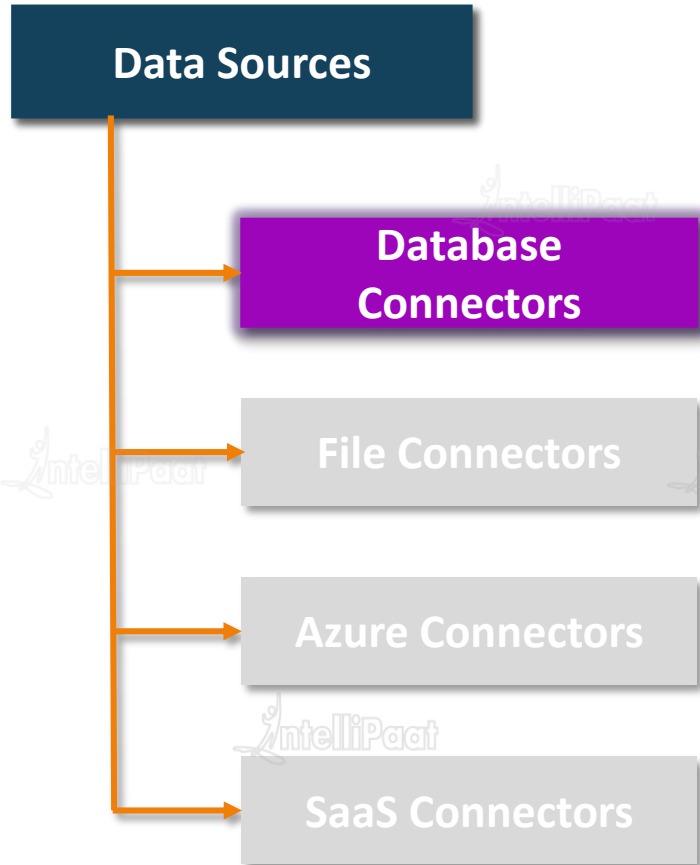
SaaS Connectors

Power BI supports the database connections for importing data from on-premises sources

Database connectors include:

MySQL

SQL Server

IBM Db2

Oracle

Teradata

Data Sources in Power BI

Data Sources

Database Connectors

File Connectors
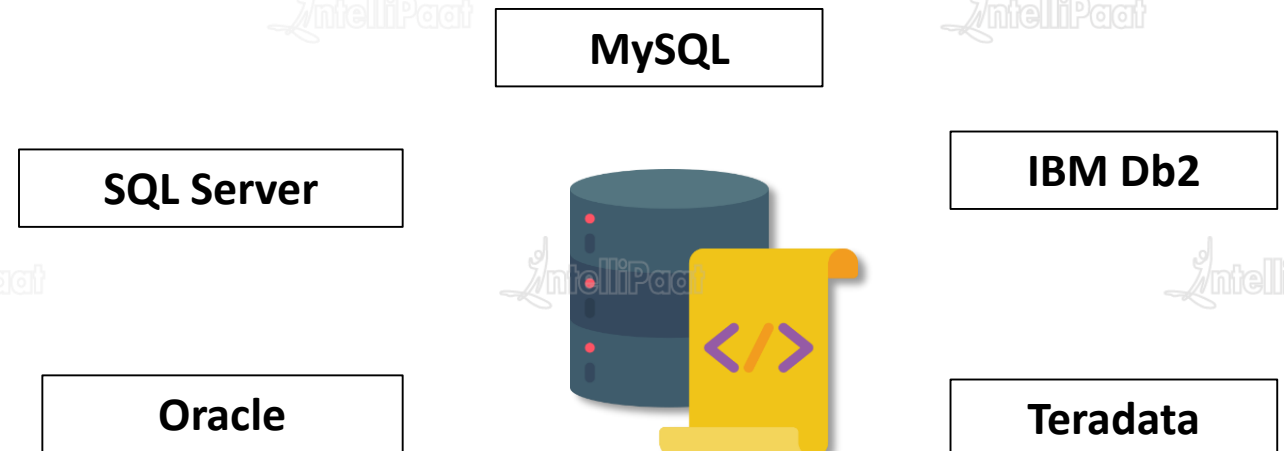
Azure Connectors
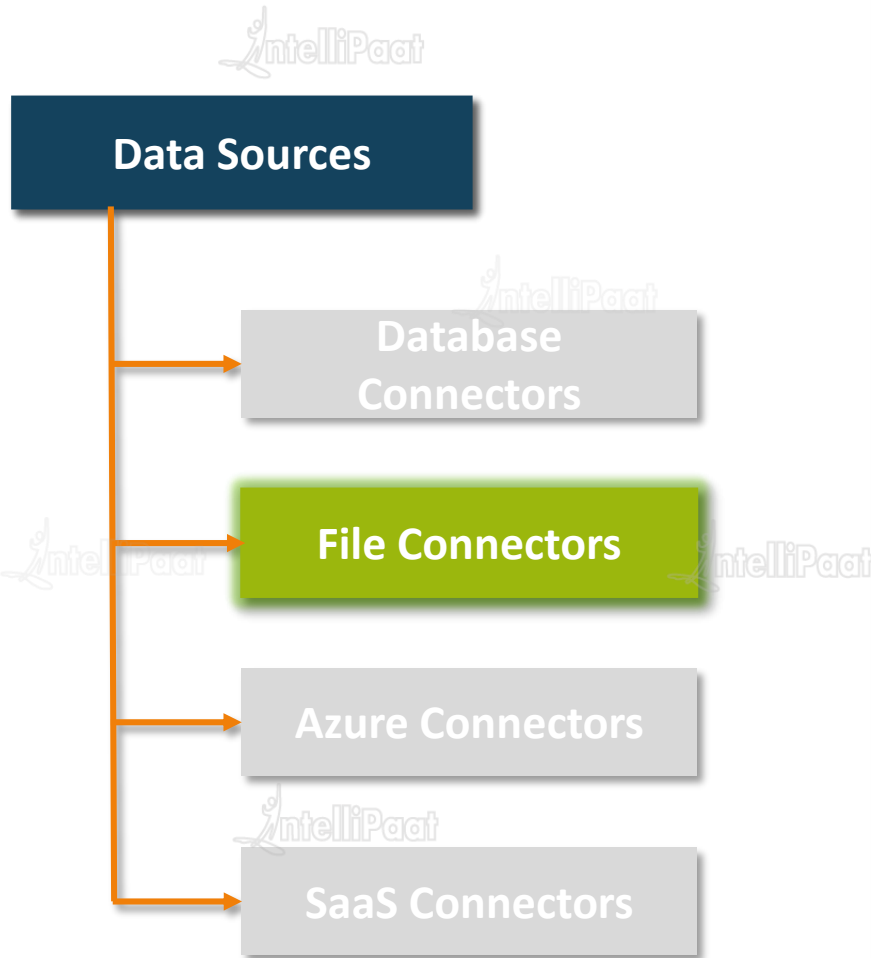
SaaS Connectors

Using a file connector, we can import either a single file or a folder to select multiple files. This is useful when we have a folder location used to store files created on a schedule

File connectors include:

CSV

MS Excel

Text

XML

JSON

Data Sources in Power BI

Data Sources

Database Connectors

File Connectors

Azure Connectors
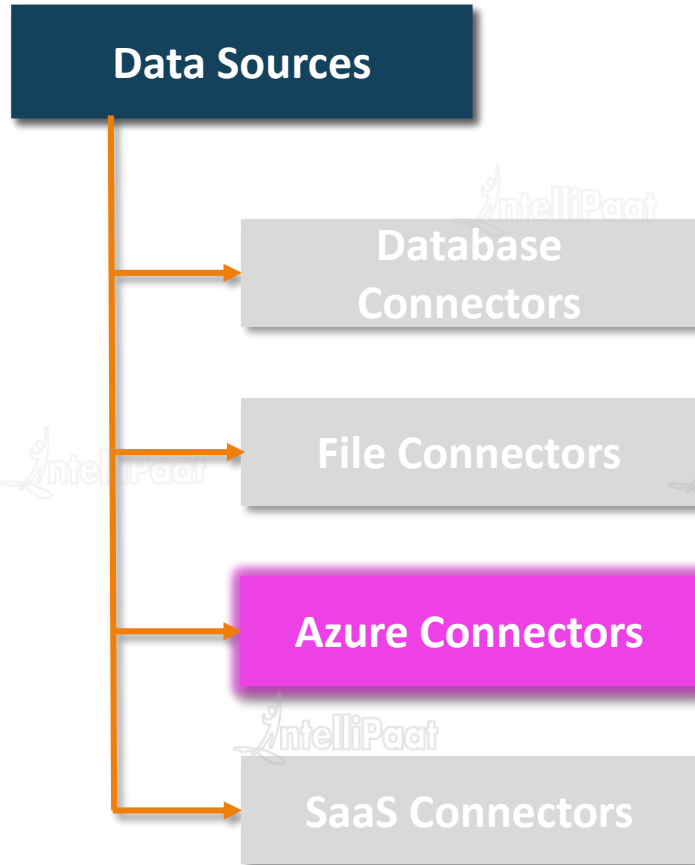
SaaS Connectors

Using Azure connectors in Power BI, we can connect to the database in the Azure cloud

Azure connectors include:

Azure SQL Data Warehouse

Azure Blob Storage

Azure Table Storage

Azure SQL Database

Data Sources in Power BI

Data Sources

- Database Connectors
- File Connectors
- Azure Connectors
- **SaaS Connectors**
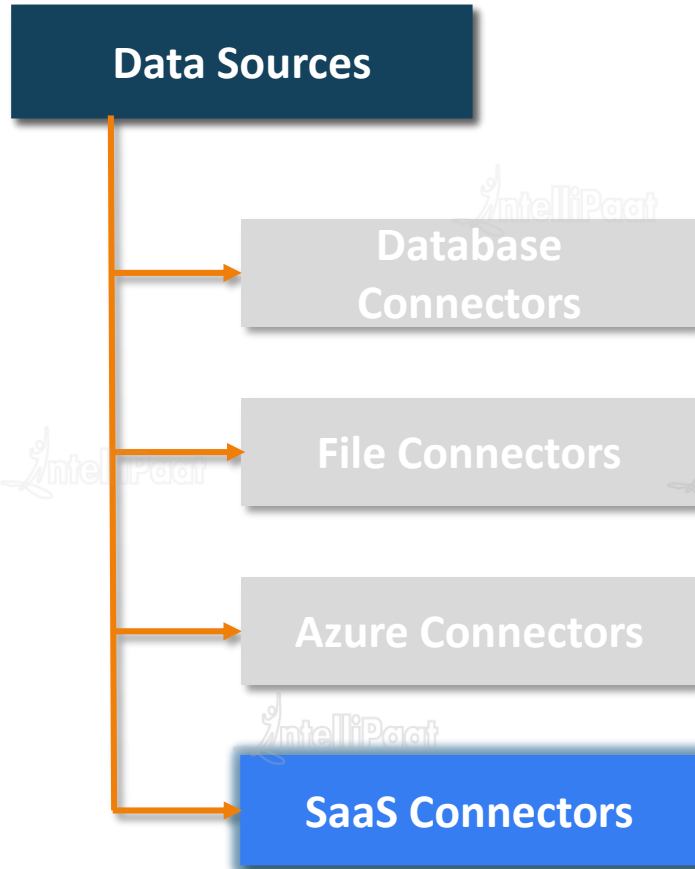
Power BI Desktop makes it easy to connect to external SaaS applications for analyzing data related to fields such as sales, marketing, finance, and social media

SaaS connectors include:

- Salesforce
- Google Analytics
- GitHub
- Facebook
- QuickBooks Online
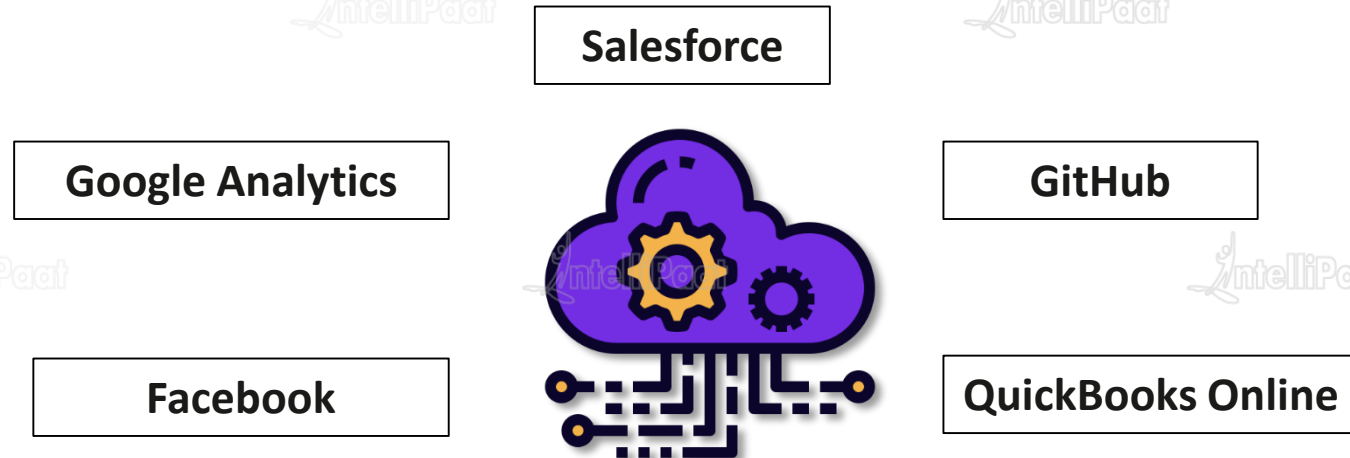
Using Excel as a Data Source

Using Excel as a Data Source

In Power BI, we can connect to files from Power BI Desktop or from Power BI Service

Using Power BI, we can also connect to various file formats, folders, etc. and even import Power BI Report files with the .pbix extension

When we import data directly into Power BI Service, the maximum file size should be 250 MB

Using Excel as a Data Source

There are two approaches in extracting an Excel file as a Data Source:

➡ Connect to an Excel workbook and use its contents as a dataset for Power BI reports and dashboards. We can edit our data using Power Query Editor

Connect the contents of Excel to Power BI Desktop

➡ Import a whole Excel workbook and explore it in the same way we would do when using Excel Online, but here we cannot edit our data

Connect the whole Excel workbook to Power BI Desktop

Using Excel as a Data Source

1 Click on Get Data

2 Different data sources appears

3 Click on Files

4 Select the file format of our choice
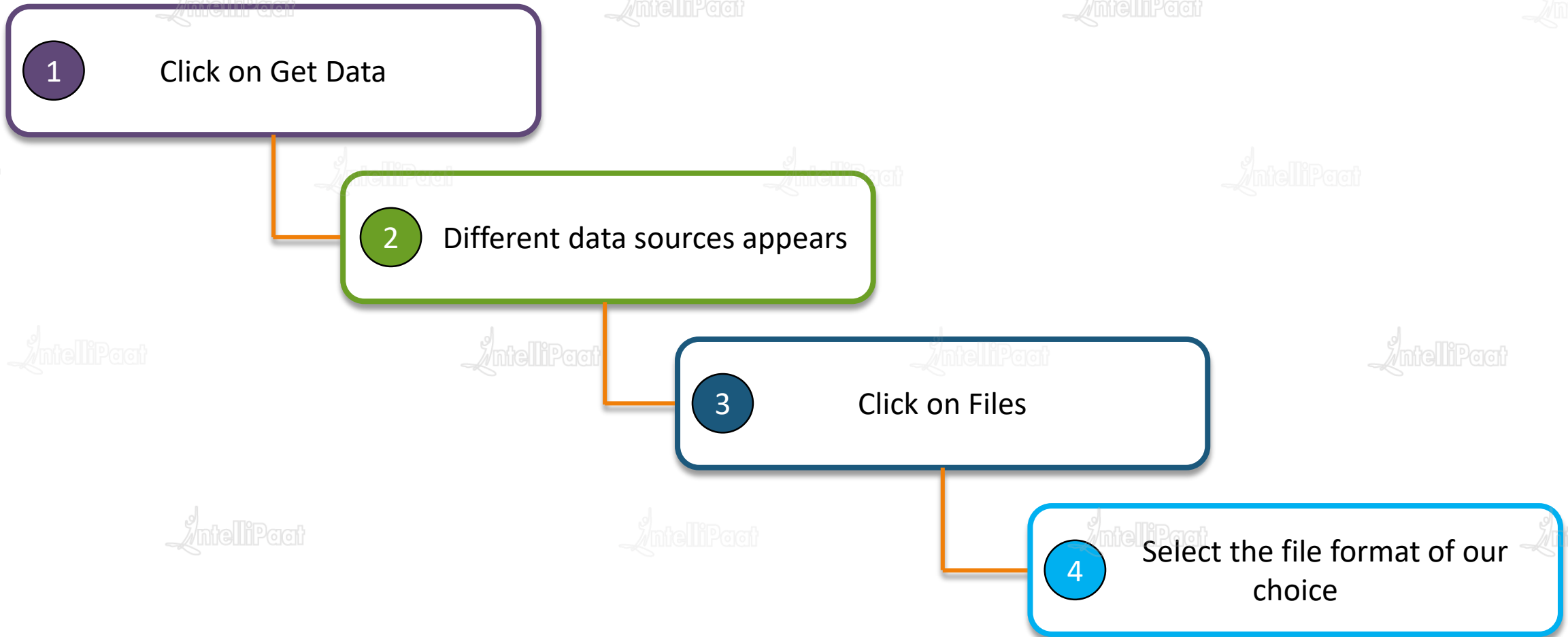
Demo: Importing Data from Excel

Updating Files in Power BI

Updating Files in Power BI

Uploading a File

Data Refresh

This applies to CSV, Power BI Desktop files, and Excel

Data Refresh in Power BI works based on the subscription types and on the data sources

We upload a local file to Power BI to build reports and dashboards; then, we make changes to the file and upload it again, providing the file name as the same, and then Power BI will update the file
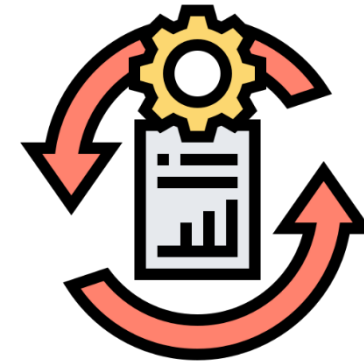
- For free users, data can be refreshed daily with a maximum of 10,000 rows/hour
- For pro users, data can be refreshed hourly up to 1 million rows/hour

Using a Database as a Data Source

Using a Database as a Data Source

Microsoft SQL Server is one of the database sources available in Power BI, which is a popular relational database management system (RDBMS)

It can handle multiple user connections and a high volume of data transactions, and it can run both in the cloud and on-premises

The steps for connecting SQL Server to Power BI is the same as connecting other databases such as Oracle, MySQL, or IBM Db2

Using a Database as a Data Source

1 — Power BI Desktop --> Get Data

2 — Click on SQL Server Database

3 — Enter the server name and the database name

4 — Use a query or select tables and views

5 — Load into Power Query Editor or directly into a data model
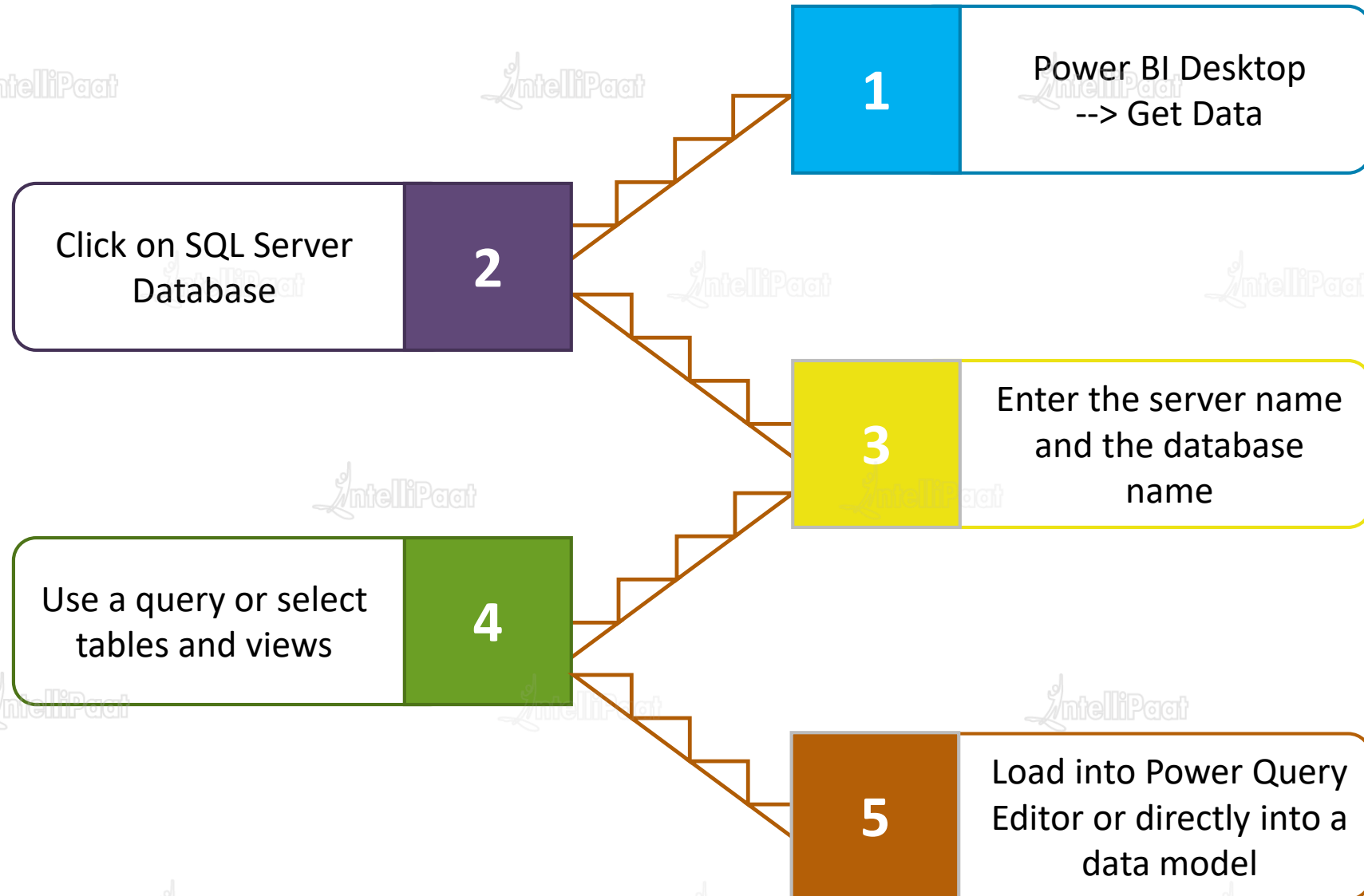
Demo: Importing Data from SQL Server

Using SaaS Connectors

Using SaaS Connectors

From Power BI Desktop, we can import data from third-party online solutions and combine the data in reports and dashboards

For example, we can create a report from Facebook that shows marketing data combined with the sales data that is from Salesforce
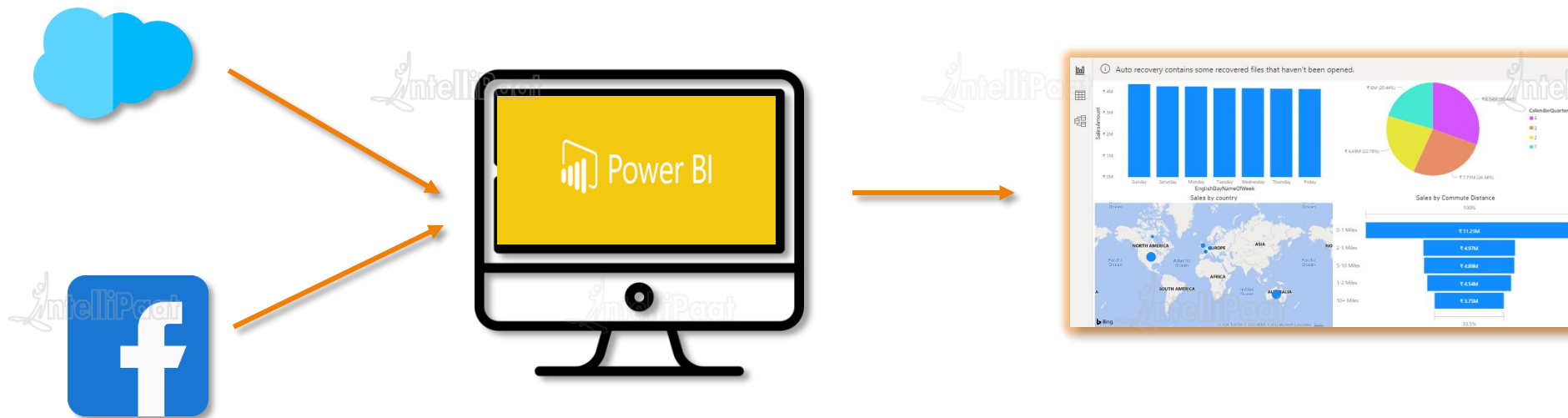
Other Data Sources

Other Data Sources

Web

Python Script

R Script Connector

OData Feed
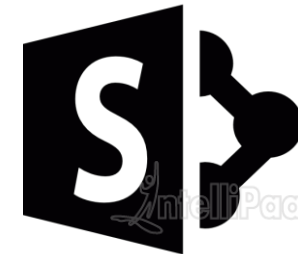
SharePoint

Python Script

Python Script

We can directly run Python scripts in Power BI Desktop by importing the datasets into the Power BI data model and adding visualizations to them
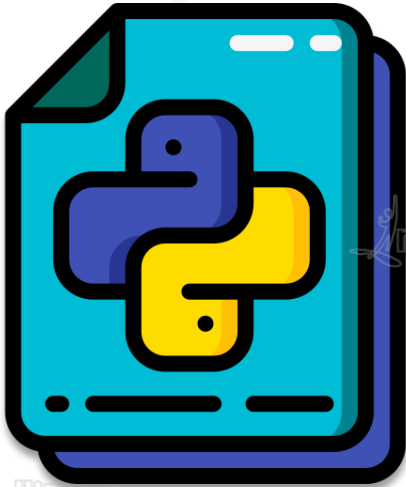
1 To run Python scripts in Power BI Desktop, install Python and the required Python packages

2 In Power BI Desktop, enable Python scripting

3 Run our Python script and import data from Power BI Desktop
Get data > Others > Python Script

4 Click on Connect; copy the Python script to the Python script dialog box, and click on OK

Python Script

Python in Query Editor

Python can also be used for data cleansing, performing advanced data shaping, analytics, etc. in Power BI Query Editor
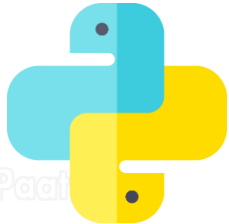
Python IDE

We can also use an external Python IDE to create Python scripts and use those scripts in Power BI Desktop

Python Visuals

We can also create Python visuals in Power BI Desktop

R Script

R Script

The R programming language has been integrated with the Transact-SQL language that helps Data Scientists develop applications in R and deploy them in a SQL Server production environment

To integrate R and SQL, a service was introduced known as SQL Server R Services

This service helps us run R scripts, create data models, and import results to Power BI Desktop

Power Query Editor

Power Query Editor

Power Query Editor helps us load data from a wide number of data sources and apply transformations on it
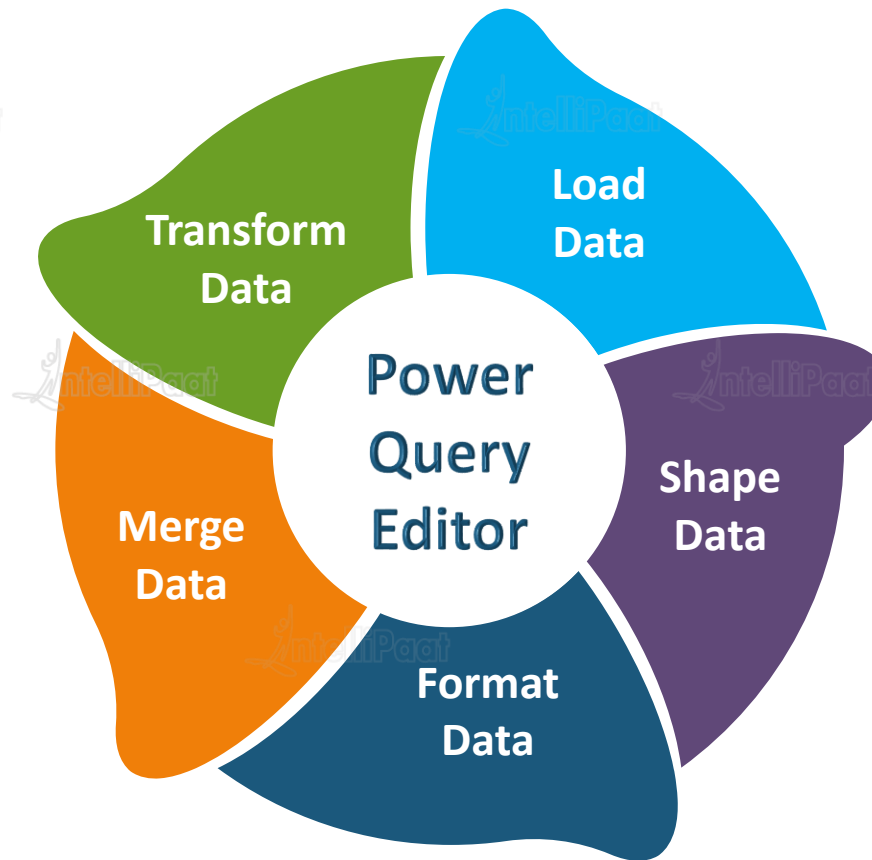
Areas in Power Query Editor

Areas in Power Query Editor

Four areas in Power Query Editor are:

01	Ribbon
02	Queries Pane
03	Data View Pane
04	Query Settings Pane

Areas in Power Query Editor

01	Ribbon
02	Queries Pane
03	Data View Pane
04	Query Settings Pane

Ribbon

The ribbon in Power Query Editor has six tabs:

- Home
- Transform
- Add Column
- View
- Tools
- Help
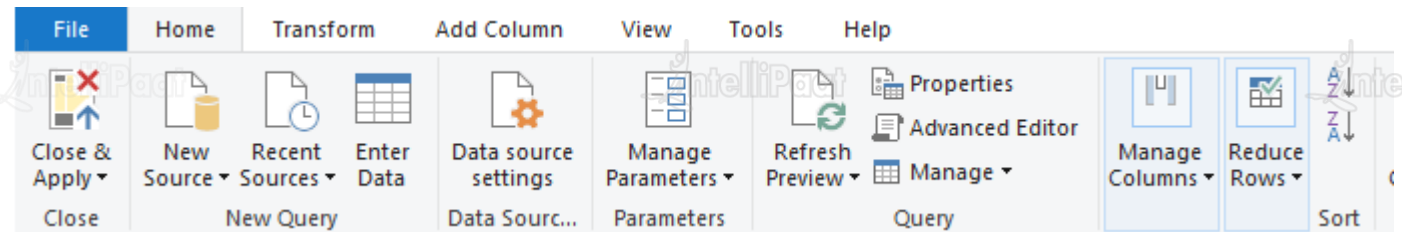
Areas in Power Query Editor

01 Ribbon

02 Queries Pane

03 Data View Pane

04 Query Settings Pane

Queries Pane

The Queries pane will show the active queries and their names. When we select a query from this pane, data will be displayed on the Data View pane

Queries [8] <

- FactInternetSales
- DimCurrency
- DimCustomer
- DimDate
- DimProduct
- DimPromotion
- DimSalesTerritory
- FactInternetSalesReason

Areas in Power Query Editor

Data View Pane

01	Ribbon
02	Queries Pane
03	Data View Pane
04	Query Settings Pane

In this pane, the data from the selected query is displayed where we can shape, transform, or merge the data to meet our needs

	1²₃ ProductKey	1²₃ OrderDateKey	1²₃ DueDateKey	1²₃ ShipDateKey	1²₃ Cus
1	310	20101229	20110110	20110105	
2	346	20101229	20110110	20110105	
3	346	20101229	20110110	20110105	
4	336	20101229	20110110	20110105	
5	346	20101229	20110110	20110105	
6	311	20101230	20110111	20110106	
7	310	20101230	20110111	20110106	
8	351	20101230	20110111	20110106	
9	344	20101230	20110111	20110106	
10	312	20101231	20110112	20110107	
11	312	20101231	20110112	20110107	
12	330	20101231	20110112	20110107	
13	313	20101231	20110112	20110107	
14	314	20101231	20110112	20110107	
15	314	20110101	20110113	20110108	
16	311	20110101	20110113	20110108	

Areas in Power Query Editor

01 Ribbon

02 Queries Pane

03 Data View Pane

04 Query Settings Pane

Query Settings Pane

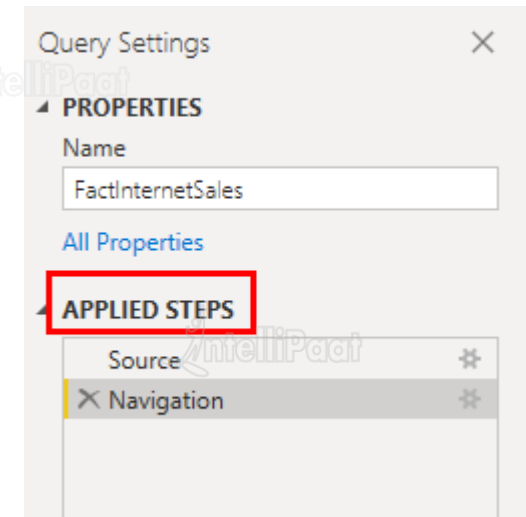This pane shows the steps associated with a query. All the operations that we perform on the data are captured by APPLIED STEPS in the Query Settings pane

Query Settings ✕

▲ **PROPERTIES**

Name

FactInternetSales

All Properties

▲ **APPLIED STEPS**

Source ⚙

✕ Navigation ⚙

Advanced Editor

Advanced Editor

Using Advanced Editor, we can see the code that Power Query Editor uses with each step, and we can also create our own code

To launch Advanced Editor, go to View > Advanced Editor

Demo: Overview of Query Editor

India: +91-7847955955

US: 1-800-216-8930 (TOLL FREE)

sales@intellipaat.com

24/7 Chat with Our Course Advisor