## Experiment No. 1

**Aim:**

Practical 1: Import the legacy data from different sources such as (Excel, SqlServer, Oracle etc.) and load it into the target system.

**Objective:**

In this experiment, we will be able to

● Gather data from various sources and load it in the targeted system, Data Extraction and Loading operation.

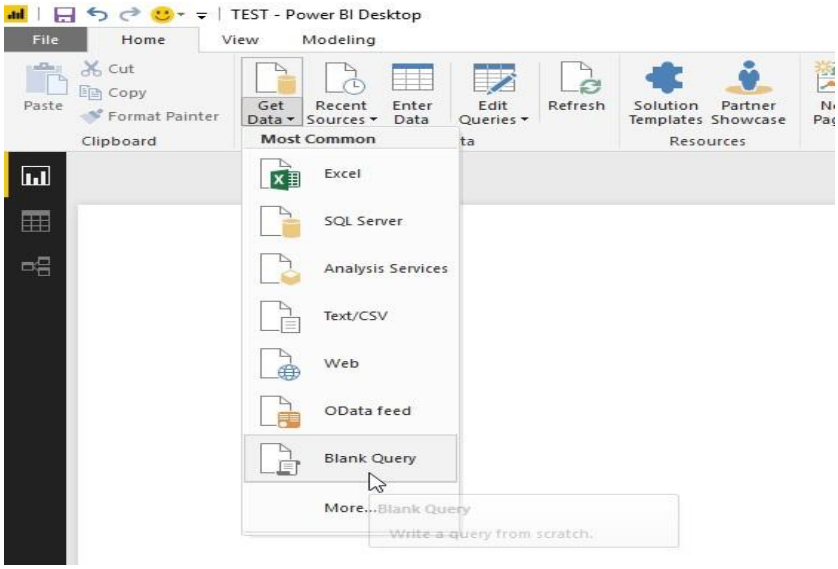
**Theory:**

Importing Excel Data and loading in BI Tool (Power BI)

1) Launch Power BI Desktop.

2) From the Home ribbon, select Get Data.

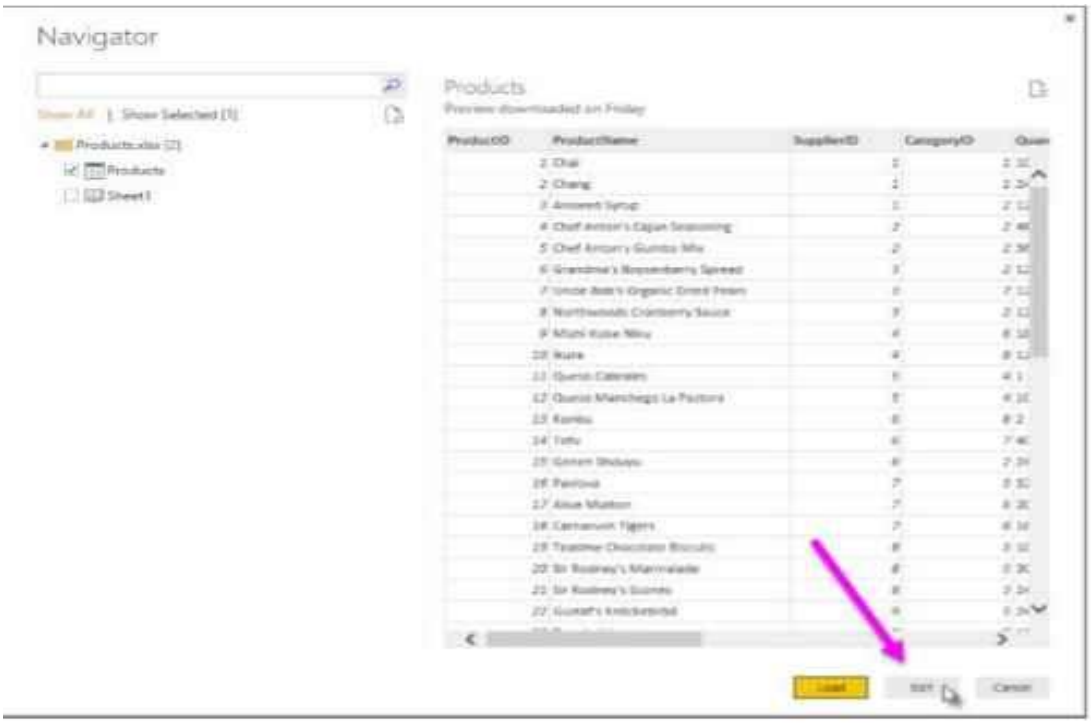
Excel is one of the Most Common data connections, so you can select it directly from the Get Data menu.

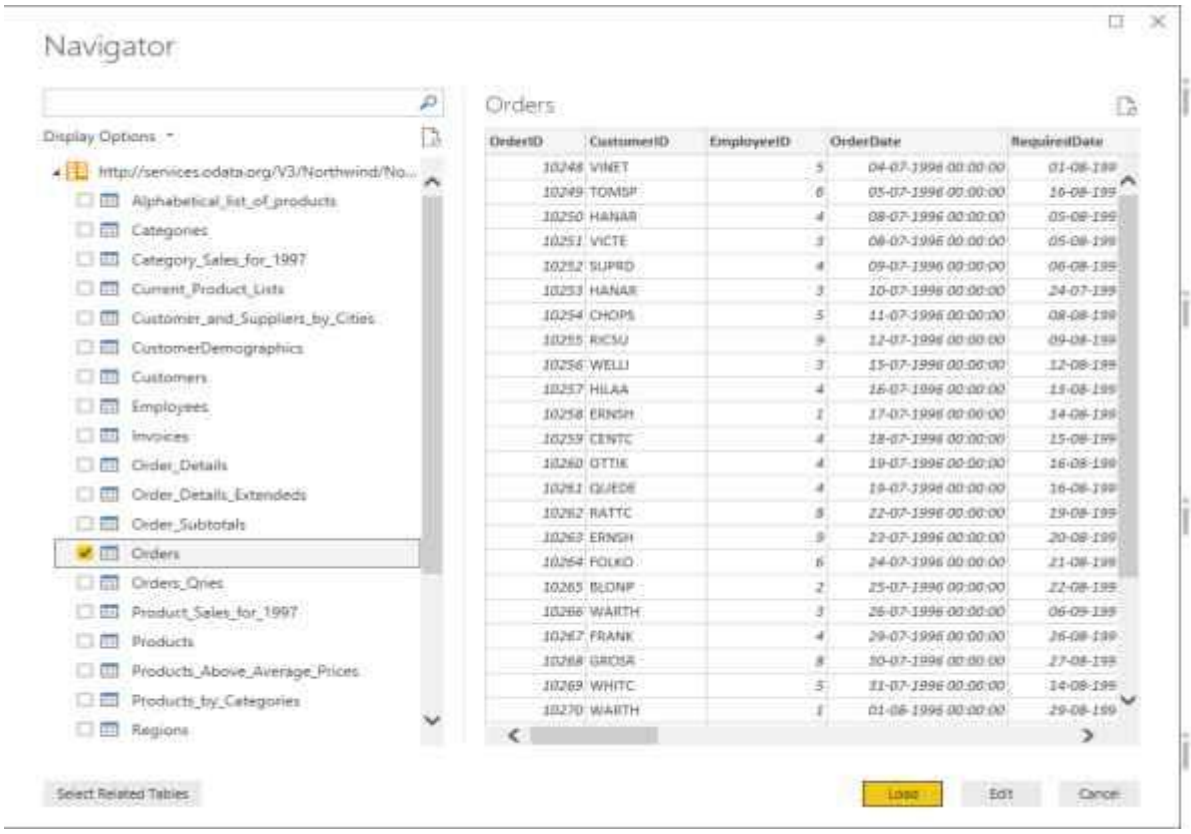



3) If you select the Get Data button directly, you can also select File > Excel and select Connect.

4) In the Open File dialog box, select the Products.xlsx file.

5) In the Navigator pane, select the Products table and then select Edit**.**




**Importing Data from OData Feed**

In this task, you'll bring in order data. This step represents connecting to a sales system. You import data into Power BI Desktop from the sample Northwind OData feed at the following

URL, which you can copy (and then paste) in the steps below:

http://services.odata.org/V3/Northwind/Northwind.svc/

Connect to an OData feed:

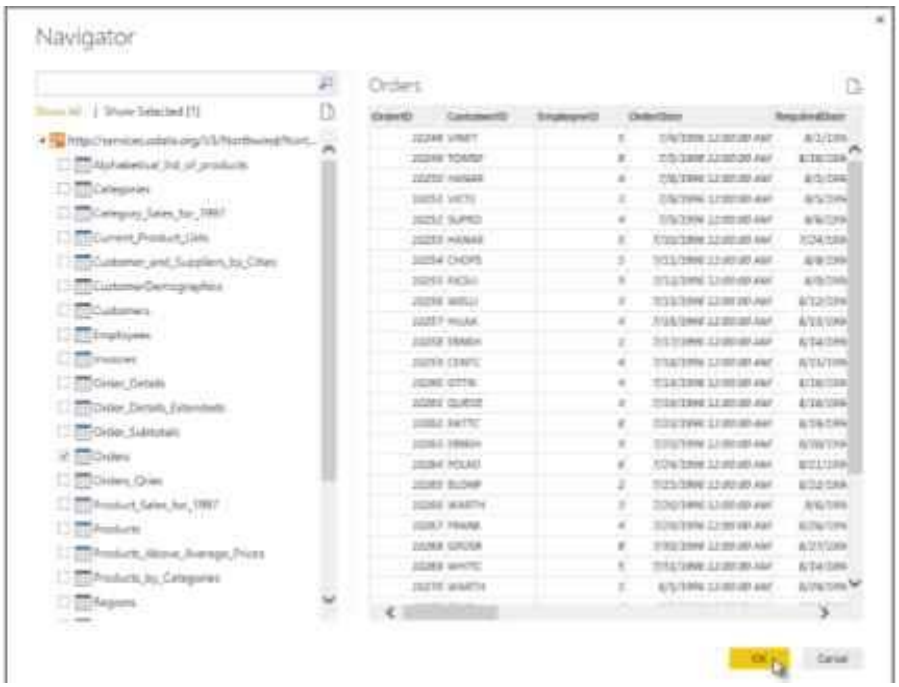
1) From the Home ribbon tab in the Query Editor, select Get Data.

2) Browse to the OData Feed data source.

3) In the OData Feed dialog box, paste the URL for the Northwind OData feed.

4) Select OK.

5) In the Navigator pane, select the Orders table, and then select Edit.

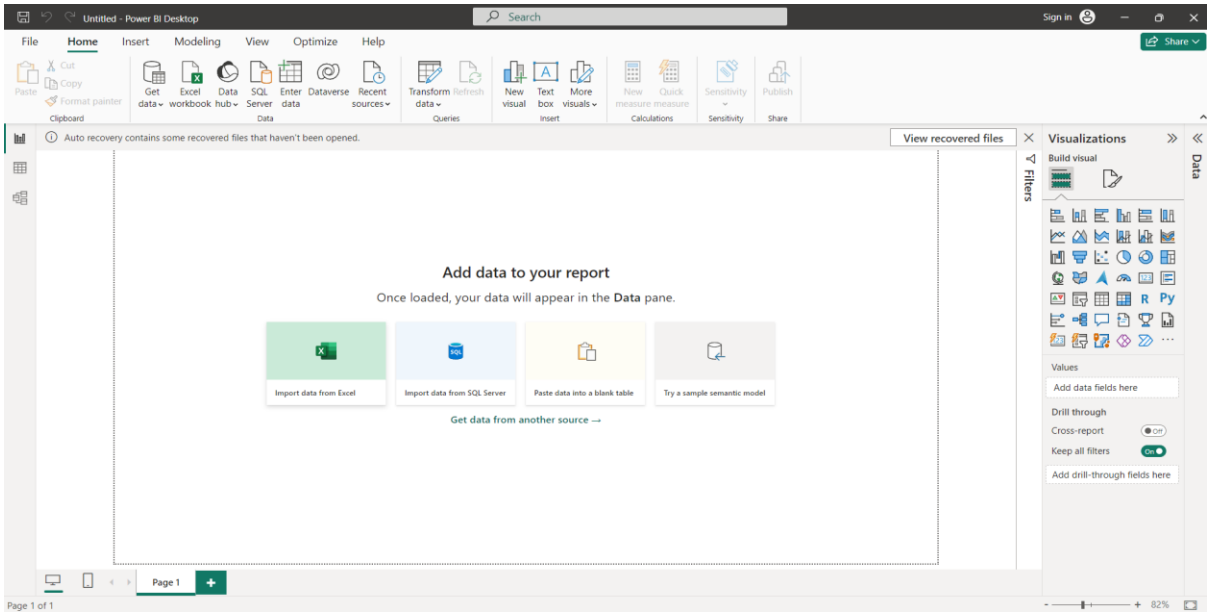


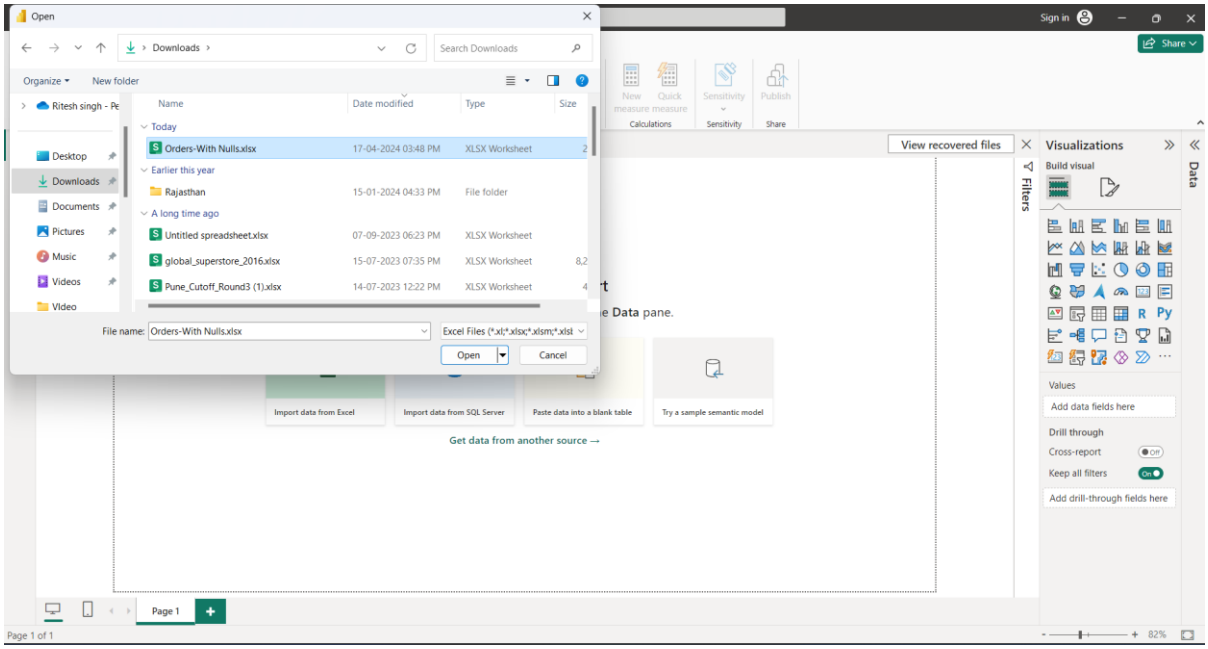

Note - You can click a table name, without selecting the checkbox, to see a preview

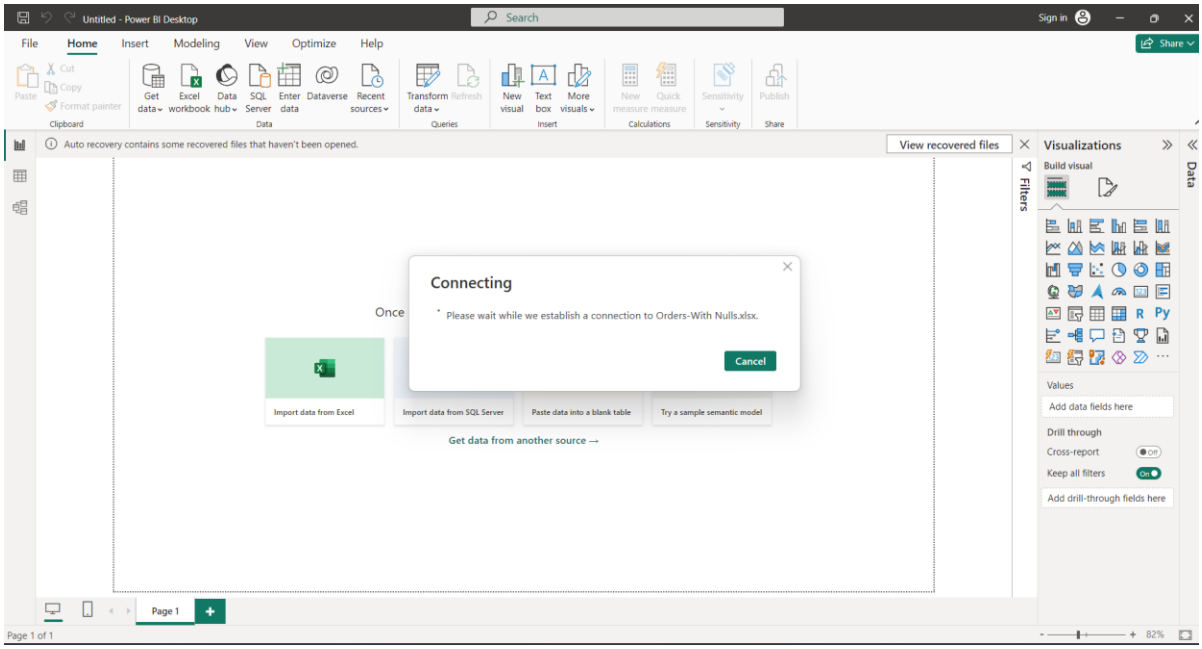
**Conclusion:** In This way we have studied data loading operation from various sources into power BI.

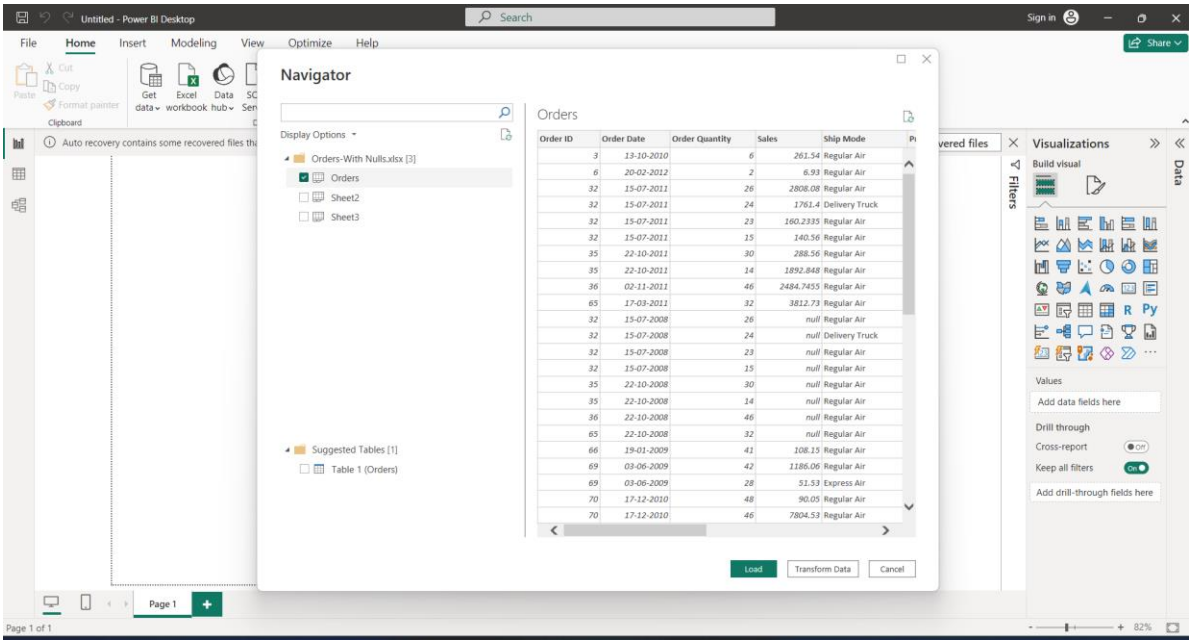
**Input:** Importing Excel file and loading in the target system

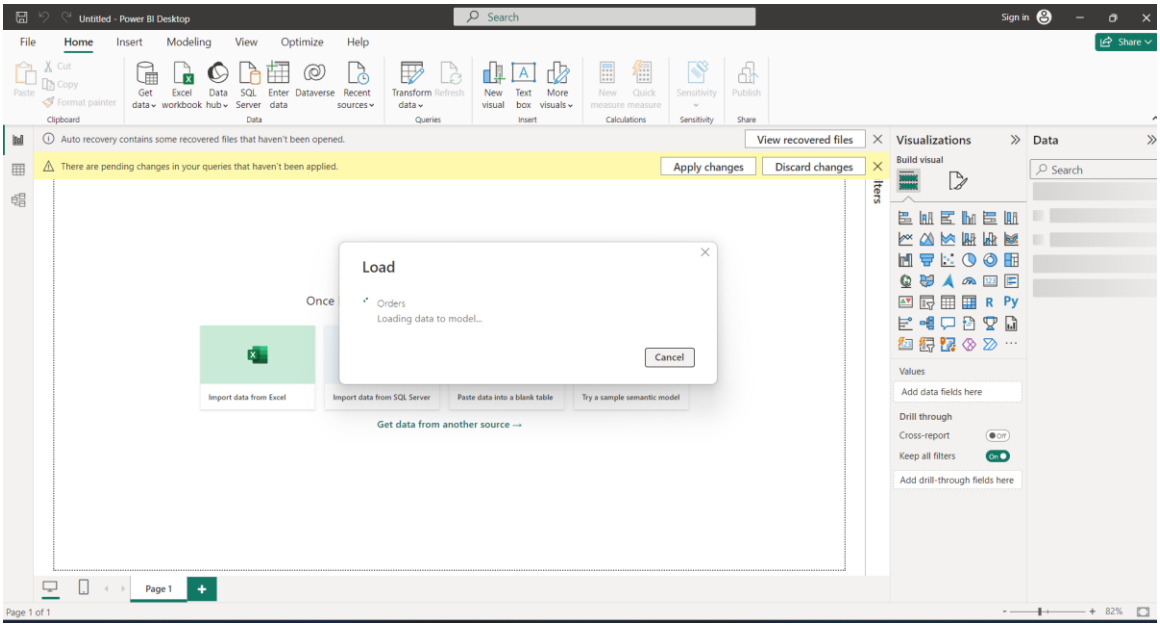
**Outcome:**

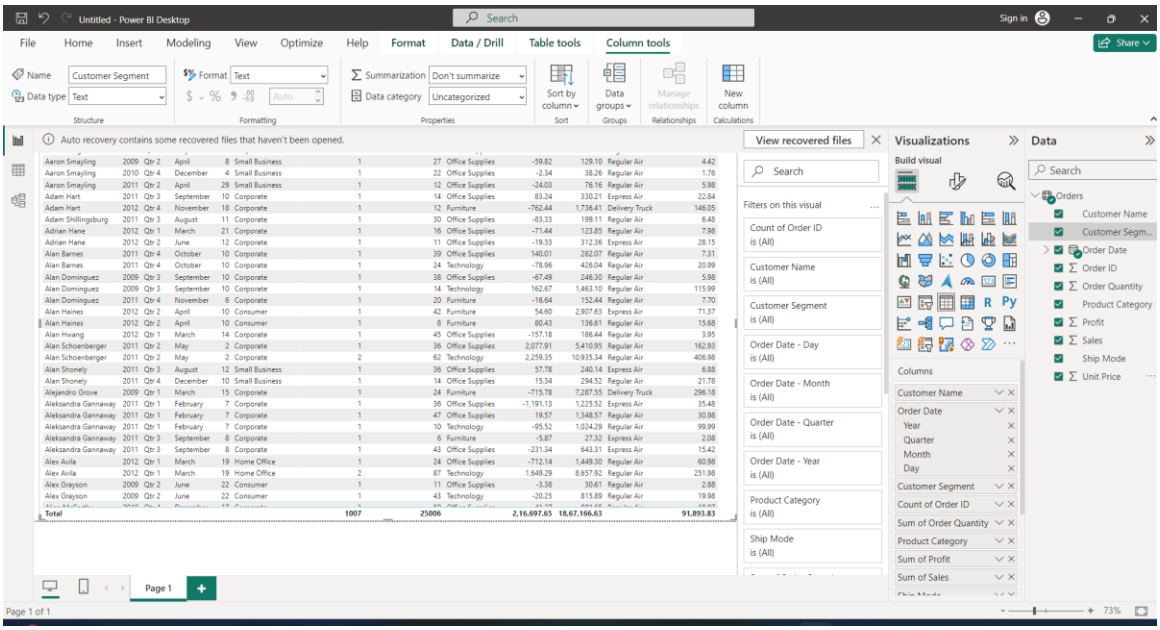

Untitled - Power BI Desktop
Search
Sign in
File
Home
Insert
Modeling
View
Optimize
Help
Share
Auto recovery contains some recovered files that haven't been opened.
View recovered files
Add data to your report
Once loaded, your data will appear in the Data pane.
Import data from Excel
Import data from SQL Server
Paste data into a blank table
Try a sample semantic model
Get data from another source →
Visualizations
Build visual
Values
Add data fields here
Drill through
Cross-report
Keep all filters
Add drill-through fields here
Filters
Data
Page 1
Page 1 of 1
82%


Open
Downloads
Search Downloads
Organize
New folder
Name
Date modified
Type
Size
Today
Orders-With Nulls.xlsx
17-04-2024 03:48 PM
XLSX Worksheet
Earlier this year
Rajasthan
15-01-2024 04:33 PM
File folder
A long time ago
Untitled spreadsheet.xlsx
07-09-2023 06:23 PM
global_superstore_2016.xlsx
15-07-2023 07:35 PM
Pune_Cutoff_Round3 (1).xlsx
14-07-2023 12:22 PM
Desktop
Downloads
Documents
Pictures
Music
Videos
File name: Orders-With Nulls.xlsx
Excel Files (*.xl;*.xlsx;*.xlsm;*.xlsb
Open
Cancel
Sign in
Share
View recovered files
Visualizations
Build visual
Values
Add data fields here
Drill through
Cross-report
Keep all filters
Add drill-through fields here
Filters
Data
Page 1
Page 1 of 1
82%


Untitled - Power BI Desktop
Search
Sign in
File
Home
Insert
Modeling
View
Optimize
Help
Share
Auto recovery contains some recovered files that haven't been opened.
View recovered files
Connecting
Please wait while we establish a connection to Orders-With Nulls.xlsx.
Cancel
Once
Import data from Excel
Import data from SQL Server
Paste data into a blank table
Try a sample semantic model
Get data from another source →
Visualizations
Build visual
Values
Add data fields here
Drill through
Cross-report
Keep all filters
Add drill-through fields here
Filters
Data
Page 1
Page 1 of 1
82%


Untitled - Power BI Desktop
Search
Sign in
File
Home
Insert
Modeling
View
Optimize
Help
Share
Navigator
Display Options
Orders-With Nulls.xlsx [3]
Orders
Sheet2
Sheet3
Suggested Tables [1]
Table 1 (Orders)
Orders
Order ID
Order Date
Order Quantity
Sales
Ship Mode
3 13-10-2010 6 261.54 Regular Air
6 20-02-2012 2 6.93 Regular Air
32 15-07-2011 26 2808.08 Regular Air
32 15-07-2011 24 1761.4 Delivery Truck
32 15-07-2011 23 160.2335 Regular Air
32 15-07-2011 15 140.56 Regular Air
35 22-10-2011 30 288.56 Regular Air
35 22-10-2011 14 1892.848 Regular Air
36 02-11-2011 46 2484.7455 Regular Air
65 17-03-2011 32 3812.73 Regular Air
32 15-07-2008 26 null Regular Air
32 15-07-2008 24 null Delivery Truck
32 15-07-2008 23 null Regular Air
32 15-07-2008 15 null Regular Air
35 22-10-2008 30 null Regular Air
35 22-10-2008 14 null Regular Air
36 22-10-2008 46 null Regular Air
65 22-10-2008 32 null Regular Air
66 19-01-2009 41 108.15 Regular Air
69 03-06-2009 42 1186.06 Regular Air
69 03-06-2009 28 51.53 Express Air
70 17-12-2010 48 90.05 Regular Air
70 17-12-2010 46 7804.53 Regular Air
Load
Transform Data
Cancel
Visualizations
Build visual
Values
Add data fields here
Drill through
Cross-report
Keep all filters
Add drill-through fields here
Filters
Data
Page 1
Page 1 of 1
82%




Successfully able to Import legacy data from different sources such as (Excel, SqlServer, Oracle etc.) and load in the targeted system.

**Application:**

Data Integration from Multiple Sources: ETL can be employed to extract data from these diverse sources such as different databases, spreadsheets, and external APIs., transform it into a unified format, and load it into the SQL Server database. This integration ensures a consolidated and consistent dataset for analysis in Power BI.

**Questions:**

Q 1: How do you ensure data integrity and consistency during the import process from different sources?

Q 2: What strategies would you employ to handle large volumes of data during the import process from various sources?

Q 3: Can you describe the role of data mapping and transformation in the context of importing data from different sources into the target system?

Q 4: What security measures would you implement to safeguard the integrity and confidentiality of the imported data during the transfer process?

**Answers:**

1. To ensure data integrity and consistency during import from different sources, employing techniques like data validation and enforcing standards is crucial. Data validation involves checking data for accuracy and completeness, while standards ensure that imported data meets predefined criteria, minimizing inconsistencies.

2. Handling large volumes of data during import from various sources requires strategies like parallel processing and distributed computing. Parallel processing involves breaking down tasks into smaller units and executing them simultaneously, while distributed computing utilizes multiple interconnected computers to process data in parallel, enabling efficient handling of large datasets.

3. Data mapping and transformation play vital roles in importing data from different sources into the target system. Data mapping involves aligning source data structures with the target system's schema to ensure compatibility, while transformation involves converting data formats, cleaning data, and resolving discrepancies to ensure accurate and meaningful data integration.

4. Implementing security measures is essential to safeguard the integrity and confidentiality of imported data during the transfer process. This includes encryption of data during transmission, using secure connections such as SSL/TLS, implementing access controls to restrict unauthorized access to data, and regularly auditing and monitoring data transfer processes for potential security breaches.