Architecture Design

# Amazon Sales Data Analysis

|  |  |
| --- | --- |
| **Written By** | Arnab De, Siddhartha Shankar Dhibar |
| **Document Version** | 0.3 |
| **Last Revised Date** |  |

**DOCUMENT CONTROL**

## Change Record:

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **AUTHOR** | **COMMENTS** |
| 0.1 | 14- Oct -  2022 | Arnab De | Introduction and architecture defined |
| 0.2 | 15 - May -  2021 | Siddharta Shankar Dhibar | Final revision |

**Reviews:**

|  |  |  |  |
| --- | --- | --- | --- |
| **VERSION** | **DATE** | **REVIEWER** | **COMMENTS** |
| 0.2 | 21- May -  2021 | Author 3 | Unit test cases to be added |

**Approval Status:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **VERSION** | **REVIEW**  **DATE** | **REVIEWED BY** |  | **APPROVED BY** | **COMMENTS** |
|  |  |  |  |  |  |

# Introduction

## What is Architecture design document?

Any software needs the architectural design to represents the design of software. IEEE defines architectural design as “the process of defining a collection of hardware and software components and their interfaces to establish the framework for the development of a computer system.” The software that is built for computer-based systems can exhibit one of these many architectures.

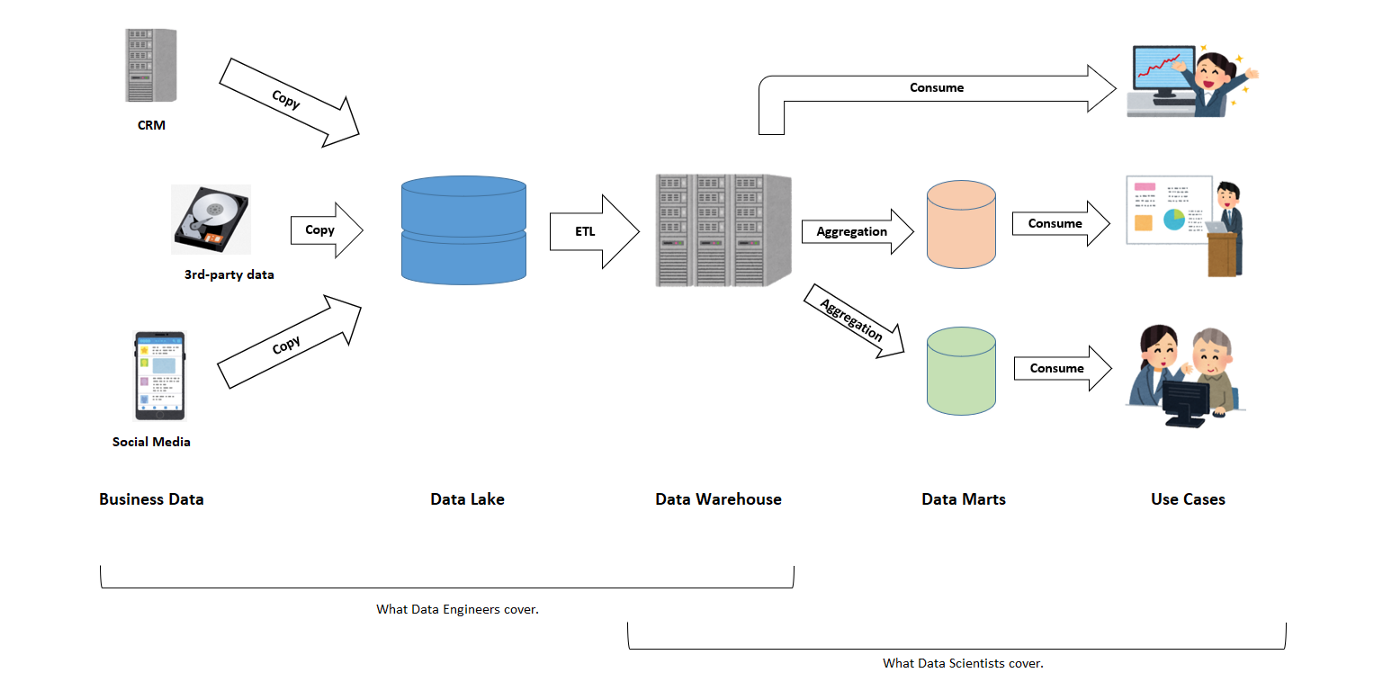
Each style will describe a system category that consists of:

* A set of components (e.g.: a database, computational modules) that will perform a function required by the system.
* The set of connectors will help in coordination, communication, and cooperation between the components.
* Conditions that how components can be integrated to form the system.
* Semantic models that help the designer to understand the overall properties of the system.

## Scope

Architecture Design Document (ADD) is an architecture design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the design principles may be defined during requirement analysis and then refined during architectural design work.

# Architecture

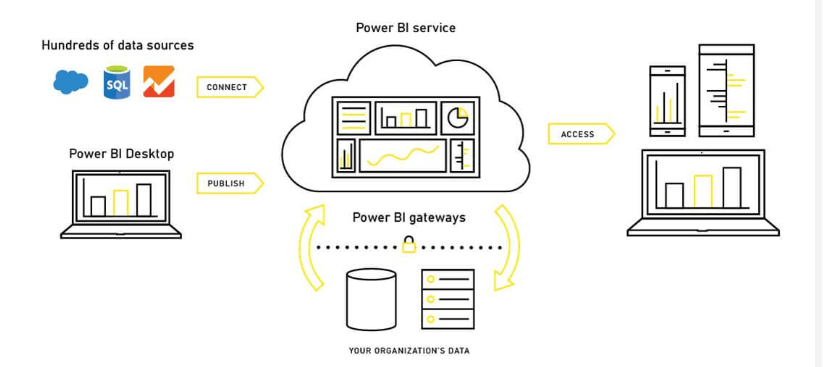


**Power BI Architecture**

Power BI has a highly scalable, n-tier client-server architecture that serves mobile clients, web clients and desktop-installed software. Power BI architecture supports fast and flexible deployments.

### ARCHITECTURE DESIGN

The following diagram shows Power BI’s architecture:



Power BI is internally managed by the multiple server processes.

1. Sourcing data:

Power BI extracts data from various servers, Excel sheets, CSV files, and databases. The extracted information can be directly imported to Power BI, or a live service link is established to receive it. If you directly import the data in Power BI, it will only be compressed up to 1 GB. Post that, you can only run live queries on your chunky datasets.

1. Transforming the data:

Before visualizing the data, cleaning and pre-processing it should be done. This means removing useless or missing values from rows or columns. Following that, certain rules will be applied to transform and load the datasets into the warehouse.

1. Report and publish:

After cleaning and transforming the data, reports will be created based on requirements. A report is a visualization of the data with different filters and constraints presented in the form of graphs, pie charts, and other figures.

1. Creating dashboards:

are created by pinning individual elements or pages of live reports. Dashboards should be created after you have published your reports to the BI service. When the reports get saved, the visual maintains the filter settings chosen so that the user can apply filters and slicers.

5.Power BI Desktop:

Power BI Desktop is a free software used to convert, connect, and visualize datasets on a PC or laptop. It’s one of the most important Power BI components where you can integrate distinct information sources and combine them to form a data model. Then, you can create graphics or image collections to share them as records with other individuals in your organization.

1. Power BI Service:

After the reports are created on Power BI Desktop, you can publish them on the cloud using Power BI Service. The service connects users and allows them to create dashboards known as Power BI Workspace. It offers natural language Q&A and alerts, and it is available in both Power BI free and Power BI Pro versions,

1. Power BI Mobile Apps:

The mobile apps of Power BI keep you connected with the data no matter where you are. You can see live reports and dashboards on your iOS and Android smartphones and make better market decisions on the go. Only pro Power BI architecture provides the feature of Mobile reports and dashboards.

1. Power BI Query:

Power Query allows users to connect distinct information from multiple sources and convert them to satisfy their business requirements. Power Query is included in the Power Query Editor of Power BI Desktop.

1. Power Q&A:

Power Q & A allows business users to explore information in their own words and phrases. This natural language question and reply engine is the fastest way to get the response from your data.

1. Power Map:

Power BI queries offer a 3D visualization tool, Power Map, that shows differences in your datasets with shadings ranging from dark to light.

1. Power Pivot:

Power Pivot allows data storage with high compression, quick aggregation, and calculation. With Power Query, users can load information into it, or the pivot can load information on its own.

1. Power View:

For a quick and effective visualization in your Excel workbooks, you can try Power View’s drag-n-drop feature and save your time. It’s an important part of MS Power BI architecture that enables the user to quickly visualize the data in a few clients.

**Deployment Description**We should be able to address these questions satisfactorily for a successful sales pitch. To answer these questions we should be aware of the deployment models of Power BI. In this blog let us go through these aspects of Power BI.

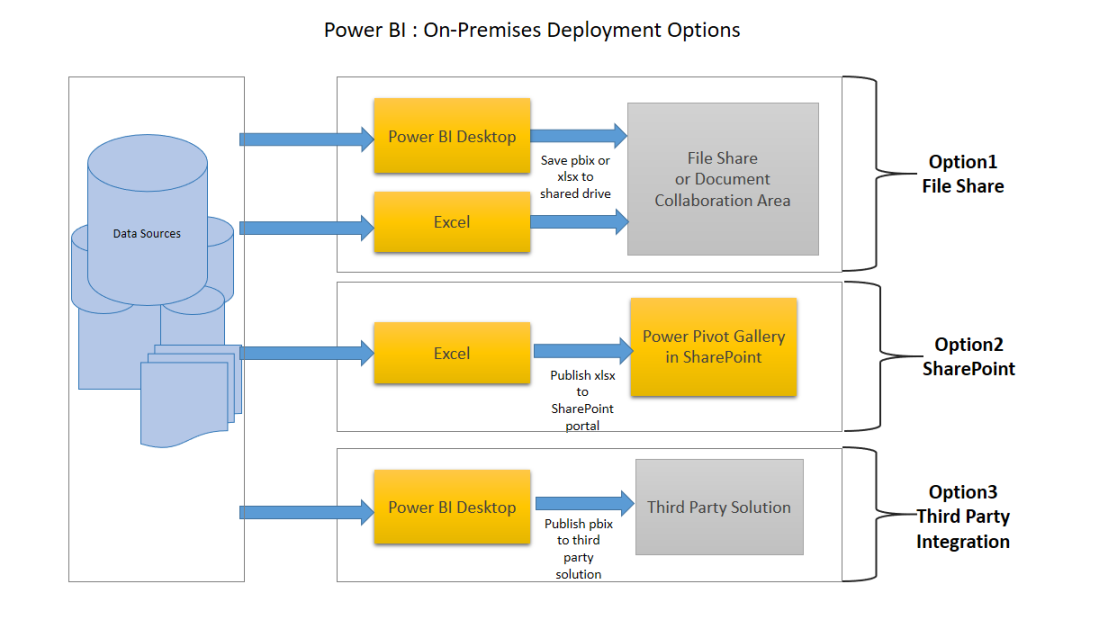
We need to first understand generic  deployment models

* **On-Premises:** Refers to data, applications and infrastructure entirely owned by client at client data center and client has complete control over it.
* **Cloud**: Refers to data, infrastructure and/or services residing in a public cloud environment and completely managed /controlled by third party. Microsoft Azure and web based Power BI service are examples of the cloud offerings.
* **Hybrid**: This denotes to the implementation which  spans both on premises and cloud sources which can be services, infrastructure and data sources

Power BI supports all three types of models. Let us see how this is possible and exactly what is to be  done.

**On-Premises Deployment :**

Power BI can be deployed on premise three different options . Kindly refer below diagram.



**Option 1 : File Share**

The first on-premises option involves usage of a file share:

* Data preparation and report creation is done in client tools: Power BI Desktop and/or Excel.
* The completed Power BI Desktop and/or Excel file is published to a file share or a document collaboration area / repository.
* To view the reports, Excel or Power BI desktop has to be installed on the viewer’s machine

**Option 2 : SharePoint**

The second on-premises option involves a specialized document library in SharePoint called the Power Pivot Gallery. Due to my limited knowledge, I am not going in details of this option

* Data preparation and report creation occurs in Excel.
* The completed Excel file is published to SharePoint within a Power Pivot Gallery.

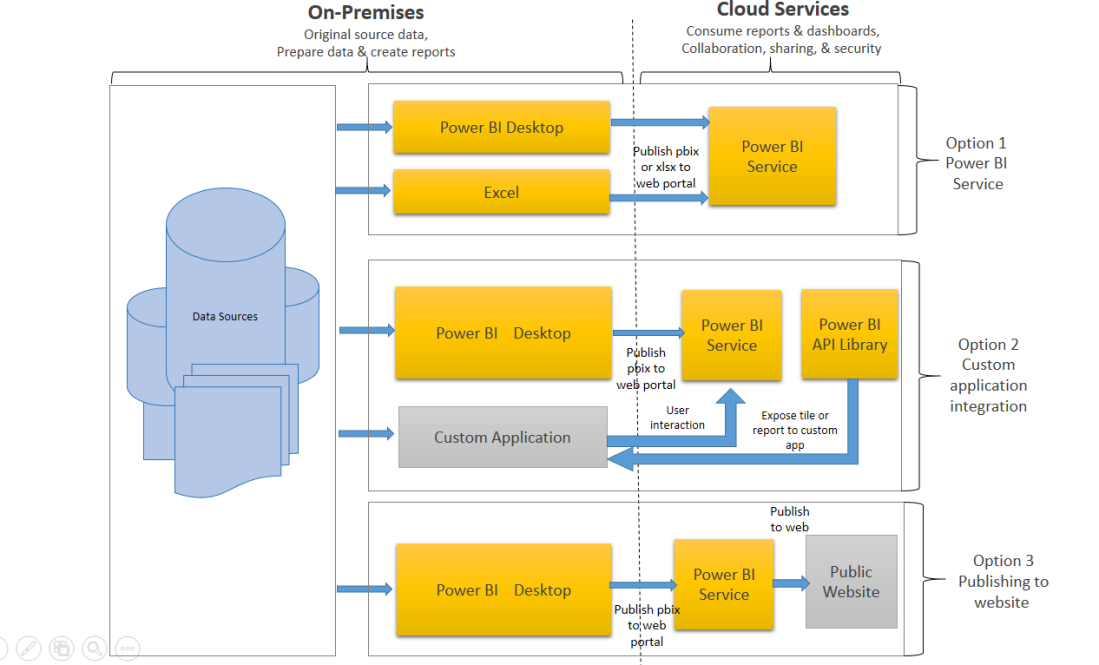
**Option 3: Third Party Integration**

The third on-premises option involves a third party which integrates with Power BI.

* Data preparation and report creation occurs in Power BI Desktop.
* The completed Power BI Desktop file is published to the third party server

**Hybrid Deployment**

Power BI can be deployed in hybrid mode in three different options . Kindly refer below diagram.



**Option 1: Power BI Service**

* Data is either from the on premises corporate applications or it might be born in cloud. It can even mix of these two
* Data preparation and report creation occurs in Power BI Desktop or excel
* Completed Power BI reports are then published to Power BI service
* Report consumption, sharing, security, collaboration, data refresh happens in Power BI service
* Dashboards are created in Power BI service and reports can also be edited or created in Power BI service

**Option 2: Custom Application Integration**

* Data is either from the on premises corporate applications or it might be born in cloud. It can even mix of these two
* Data preparation and report creation occurs in Power BI Desktop or excel
* Completed Power BI reports are then published to Power BI service
* With Power BI API , these reports can be published in custom web application or mobile app  within iFrame
* If user interacts with this report, he/she will be redirected to Power BI service
* Application can be on premise or cloud application

**Option 3: Public Website**

* Data is either from the on premises corporate applications or it might be born in cloud. It can even mix of these two
* Data preparation and report creation occurs in Power BI Desktop
* Completed Power BI reports are then published to Power BI service
* An embed code is generated by Power BI service for selected report and this code can be embedded in web page of the website within iFrame
* Here no security is maintained as its public website, hence suitable for the data which can be made publicly available