Architecture Design

Amazon Sales Analysis

Written By	Shubham Sharma
Document Version	0.1
Last Revised Date	24-08-2023

DOCUMENT CONTROL

Change Record:

VERSION	DATE	AUTHOR	COMMENTS
0.1	24-08-23	Shubham	Introduction and architecture defined

Reviews:

VERSION	DATE	REVIEWER	COMMENTS

Approval Status:

VERSION	REVIEW DATE	REVIEWED BY		APPROVED BY	COMMENTS

Contents

1. Introduction

1.1 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.

1.2 Scope

Architecture Design Document (ADD) is an architecture design process that follows a step-by-steprefinement 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.

2. Architecture

2.1 Python Pandas Architecture

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

The following diagram shows Python panda's architecture:
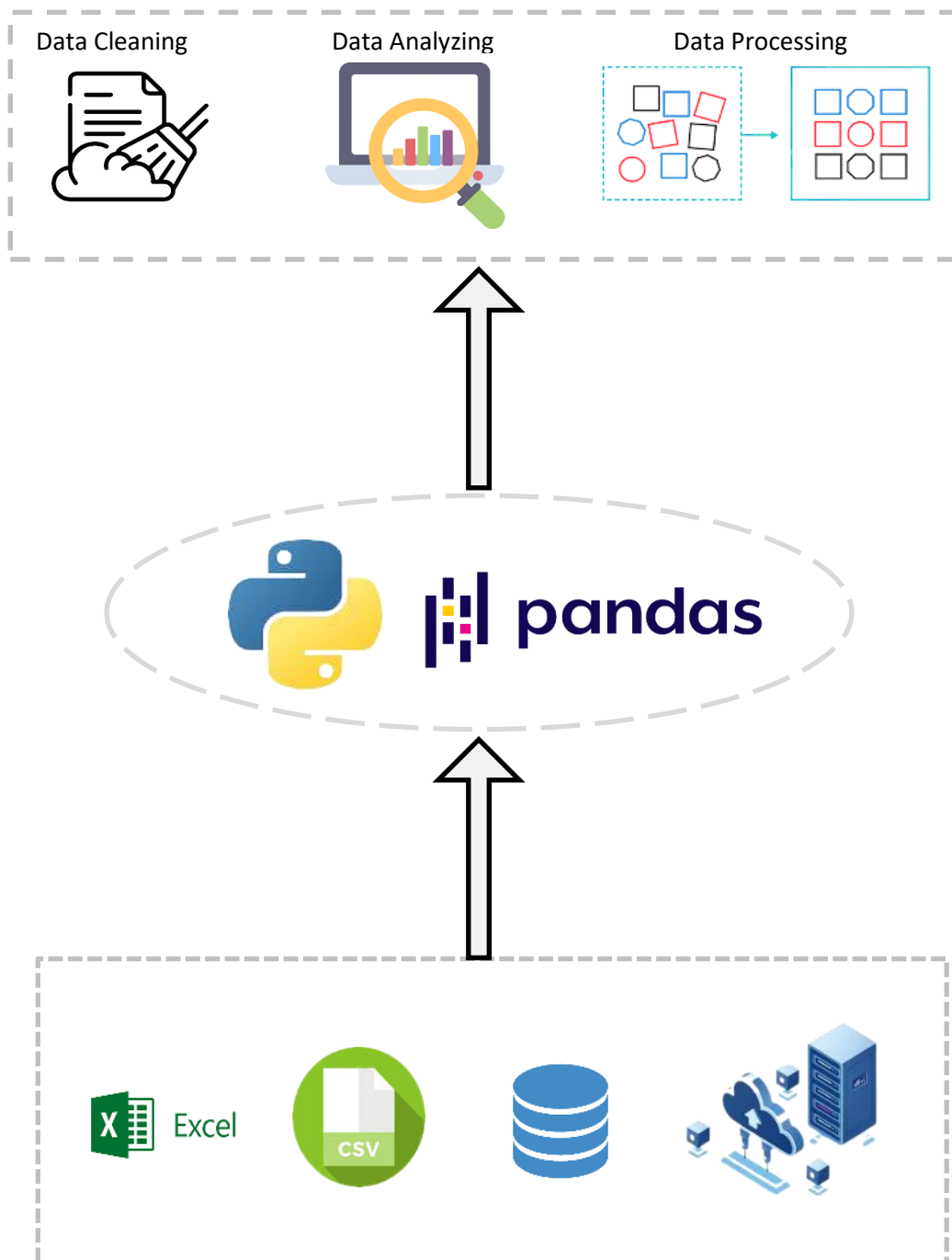
Pandas is internally managed by python for handling and manipulating the data from the databases.

1. Reading the dataset

The dataset has to be read using the pandas.read_csv(). This function enables us to read the csv file into out pandas dataframe. Which is the 2-D representation of our data.

2) Pandas DataFrame & Pandas functions:-

The pandas dataframe is a table like structure which represents the data in a two dimension manner. Here the each data are represented in means of columns and rows. The number of columns is of the number of features the dataset has. The number of rows is the number of data that the dataset contain. This pandas contains many inbuild functions which are more than enough to clean the data, analyze the data, and to process the data based on the requirements.

2.2 Power BI Architecture

Power BI is a service built on top of Azure. There are multi data sources that Power BI can connect to. Power BI gateway is connected to on-premise data sources to get continuous data for reporting and analytics.

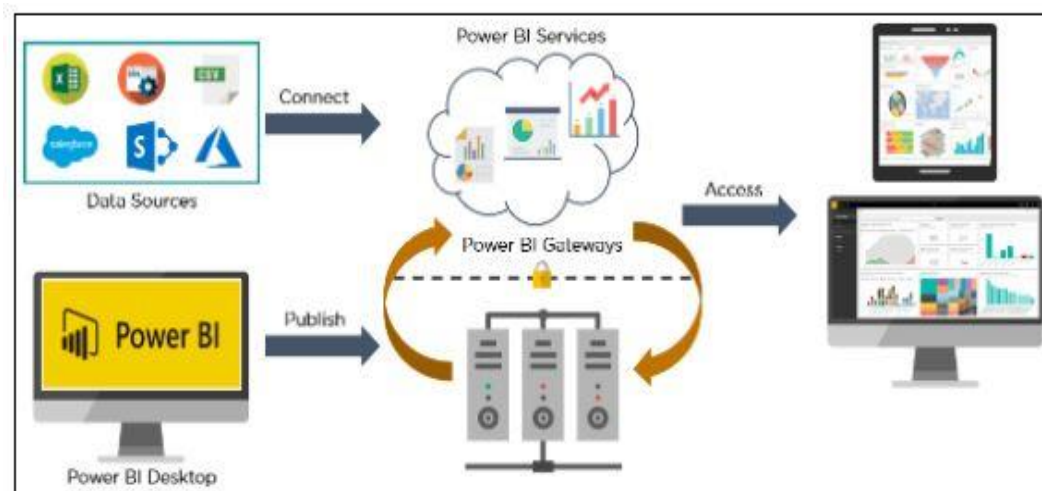The following diagram shows Tableau Server's architecture:

Tableau Server is internally managed by the multiple server processes.

2.3 Power Query

Power Query is the data transformation and mash up the engine. It enables you to discover, connect, combine, and refine data sources to meet your analysis need. It can be downloaded as an add-in for Excel or can be used as part of the Power BI Desktop.

2.4 Power Pivot:-

Power Pivot is a data modeling technique that lets you create data models, establish relationships, and create calculations. It uses Data Analysis Expression (DAX) language to model simple and complex data.

2.5 Power View:-

Power View is a technology that is available in Excel, Sharepoint, SQL Server, and Power BI. It lets you create interactive charts, graphs, maps, and other visuals that bring your data to life. It can connect to data sources and filter data for each data visualization element or the entire report.

2.6 Power Map:-

Microsoft's Power Map for Excel and Power BI is a 3-D data visualization tool that lets you map your data and plot more than a million rows of data visually on Bing maps in 3-D format from an Excel table or Data Model in Excel. Power Map works with Bing maps to get the best visualization based on latitude, longitude, or country, state, city, and street address information.

2.7 Multiple Data Sources:-

Support for various data sources is one of the vital features of Power BI. You can access various sources of data such as Excel, CSV, SQL Server, Web files, etc. to create interactive visualizations.

2.8 Steam Analytics:-

Power BI's primary advantage is its support stream analytics. From factory sensors to social media sources, Power BI assists in real-time analytics to make timely decisions.   ## 2.9 Custom Visualization:-
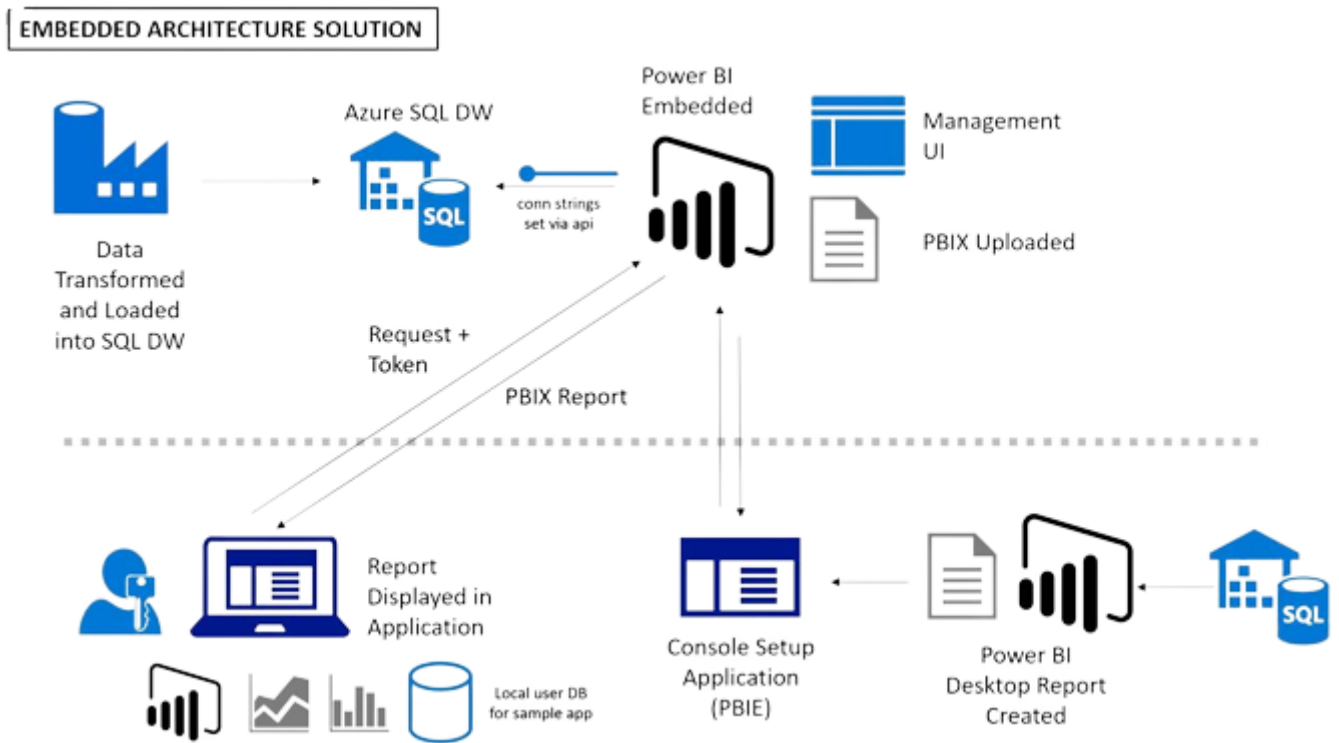
While dealing with complex data, Power BI's default standard might not be enough in some cases. In that case, you can access the custom library of visualization that meets your needs.

2.10 Power BI Communication Flow

The above picture shows the communication flow of the power bi application.