High Level Design (HLD)

HEALTH CARE ANALYSIS

Revision Number: 1.0
Last date of revision: 23-09-2024

Pawan Nagar

Document Version Control

Date Issued	Version	Description	Author
23-09-2024	0.1	First Version of Complete HLD	Pawan Nagar
23-09-2024	1.0	Final Version of Complete HLD	Pawan Nagar

Contents

Abstract

HR is not just about hiring people it is an ocean of its own. HR department goes through a constant journey of finding, selecting, on boarding and monitoring the right talent. You are required to use analytics concept to provide a smooth monitoring of workforce for the HR department. Edward Babushkin is a Russian people analyst and prolific writer. Through his Russian blog he has built a large community of people analytics practitioners and has become the face of people analytics in the East. In one of his translated posts he poses the question: Which employee will be most likely to stay the longest, Johnson, Peterson, or Sid Orson? In his support article, he than shows how to predict this using survival analysis. According to Edward, the data set is real – which is exciting! For the rest, the data is pretty straight forward. The only thing to keep an eye on is that some terms got lost in translation from Russian to English. As an example, 'independ' translates to a reversed scale of agreeableness, 'self-control' is conscientiousness, 'anxiety' is neuroticism, and 'novator' stands for openness.

1 Introduction

1.1 Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project •
- List and describe the non-functional attributes like:
  Security ○ Reliability

    - Maintainability
    - Portability
    - Reusability
    - Application compatibility
    - Resource utilization
    - Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

iNeuron

2 General Description

2.1 Product Perspective & Problem Statement

HR is not just about hiring people it is an ocean of its own. HR department goes through a constant journey of finding, selecting, on boarding and monitoring the right talent. You are required to use analytics concept to provide a smooth monitoring of workforce for the HR department. Edward Babushkin is a Russian people analyst and prolific writer. Through his Russian blog he has built a large community of people analytics practitioners and has become the face of people analytics in the East. In one of his translated posts he poses the question: Which employee will be most likely to stay the longest, Johnson, Peterson, or Sid Orson? In his support article, he than shows how to predict this using survival analysis. According to Edward, the data set is real – which is exciting! For the rest, the data is pretty straight forward. The only thing to keep an eye on is that some terms got lost in translation from Russian to English. As an example, 'in depend' translates to a reversed scale of agreeableness, 'self-control' is conscientiousness, 'anxiety' is neuroticism, and 'Novato' stands for openness.
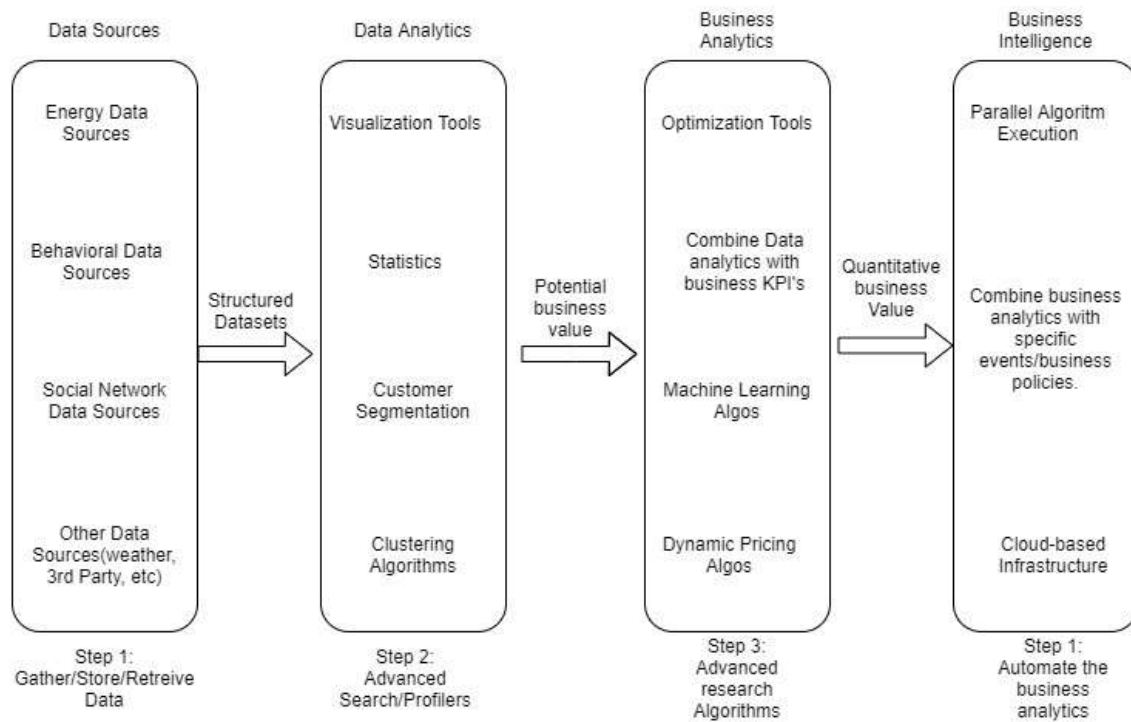
2.2 Tools used

Business Intelligence tools and libraries works such as Excel, Power BI are used to build the whole framework.

Power BI Desktop

3 Design Details

3.1 Functional Architecture

Figure 1: Functional Architecture of Business Intelligence

3.2 Optimization

Your data strategy drives performance

- Minimize the number of fields
- Minimize the number of records
- Optimize extracts to speed up future queries by materializing calculations, removing columns and the use of accelerated views

Reduce the marks (data points) in your view

- Practice guided analytics. There's no need to fit everything you plan to show in a single view. Compile related views and connect them with action filters to travel from overview to highly-granular views at the speed of thought.
- Remove unneeded dimensions from the detail shelf.
- Explore. Try displaying your data in different types of views. **Limit your filters by number and type**

- Reduce the number of filters in use. Excessive filters on a view will create a more complex query, which takes longer to return results. Double-check your filters and remove any that aren't necessary.
- Use an include filter. Exclude filters load the entire domain of a dimension, while include filters do not. An include filter runs much faster than an exclude filter, especially for dimensions with many members.

- Use a continuous date filter. Continuous date filters (relative and range-of-date filters) can take advantage of the indexing properties in your database and are faster than discrete date filters.

- Use Boolean or numeric filters. Computers process integers and Booleans (t/f) much faster than strings.

How BI Really Works

Organizational Memory	Information Integration	Insight Creation	Presentation
• Data Warehouse • ERP • Knowledge Repository • CMS • DMS	• Business Analytics Tool • Data Mining • Real-time Decision	• Text mining tools • Web mining tools • Environmental Scanning • RFID	• OLAP Tools • Visualization tools • Digital Dashboards • Score Card

- Use parameters and action filters. These reduce the query load (and work across data sources).

Optimize and materialize your calculations
- Perform calculations in the database
- Reduce the number of nested calculations.
- Reduce the granularity of LOD or table calculations in the view. The more granular the calculation, the longer it takes.
  - LODs - Look at the number of unique dimension members in the calculation.
  - Table Calculations - the more marks in the view, the longer it will take to calculate.
- Where possible, use MIN or MAX instead of AVG. AVG requires more processing than MIN or MAX. Often rows will be duplicated and display the same result with MIN, MAX, or AVG.
- Make groups with calculations. Like include filters, calculated groups load only named members of the domain, whereas Tableau's group function loads the entire domain.
- Use Booleans or numeric calculations instead of string calculations. Computers can process integers and Booleans (t/f) much faster than strings. Boolean>Int>Float>Date>DateTime>String

4 KPIs

Dashboards will be implemented to display and indicate certain KPIs and relevant indicators for the disease.

As and when, the system starts to capture the historical/periodic data for a user, the dashboards will be included to display charts over time with progress on various indicators or factors

4.1 KPIs (Key Performance Indicators)

Key indicators displaying a summary of the HR Analytics - Turnover and its relationship with different metrics

1. **Some Tiles for Quick Analysis:**
   a. First tile show that Patient Admit Date .
   b. Second tile show that Doctor Following Date .
   c. Third tile show sum of Billl.
   d. Fourth tile show total Patient Id .
   e. Fifth tile show total Test. .

2. Billing Amount Analysis

Here, we have our Dashboard interpret as follow: -

➔ Billing Amount by Bed_occuopancy
➔ Billing and Heath insurance Amount By Year Month
➔ Sum Of Billing Amount and health Insurance.
➔ Count Digenesis

3. Patient Based Analysis Visuals

- Relation between Admit date and Following Date.
- Relation between Sum of Billing amount and Health insurance.
- Relation between Health Insurance every mount and Year.
-

5 Deployment

Power BI is a business suite that includes several technologies that work together. To deliver outstanding business intelligence solutions, Microsoft Power BI technology consists of a group of components such as:

Data Sources

An important component of Power BI is its vast range of data sources. You can import data from files in your system, cloud-based online data sources or connect directly to live connections.

- Excel Text/CSV XML
- JSON
- Oracle Database
- IBM DB2 Database
- MySQL Database
- PostgreSQL Database
- Sybase Database
- Teradata Database etc.

Power BI Service

Power BI Service is a web-based platform from where you can share reports made on Power BI Desktop, collaborate with other users, and create dashboards.

- Free version
- Pro version
- Premium versionaca

Power BI Desktop

Power BI Desktop is a client-side tool known as a companion development and authoring tool. This desktop-based software is loaded with tools and functionalities to connect to data sources, transform data, data modelling and creating reports.

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

- **<u>On-Premises</u>:** Refers to data, applications and infrastructure entirely owned by client at client data centre and client has complete control over it.

- **<u>Cloud</u>:** 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.

- **<u>Hybrid</u>:** 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.

<u>On-Premises Deployment:</u>
Power BI can be deployed on premise three different options. Kindly refer below diagram.

Power BI : On-Premises Deployment Options
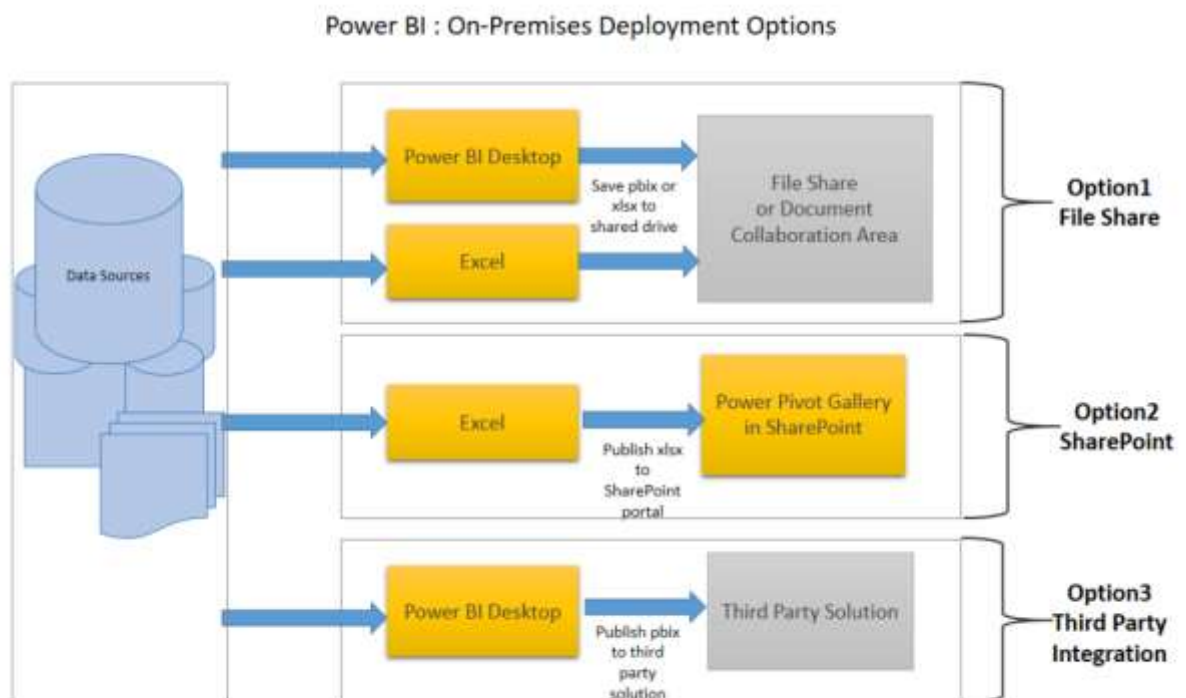
Option 1 : File Share
The first on-premises option involves usage of a file share: Data preparation and report creation is done in client tools: Power BI Desktop and/or Excel.
The completed Power BI Desktop and/or Excel file is published to a file share or a document collaboration area / repository.
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.
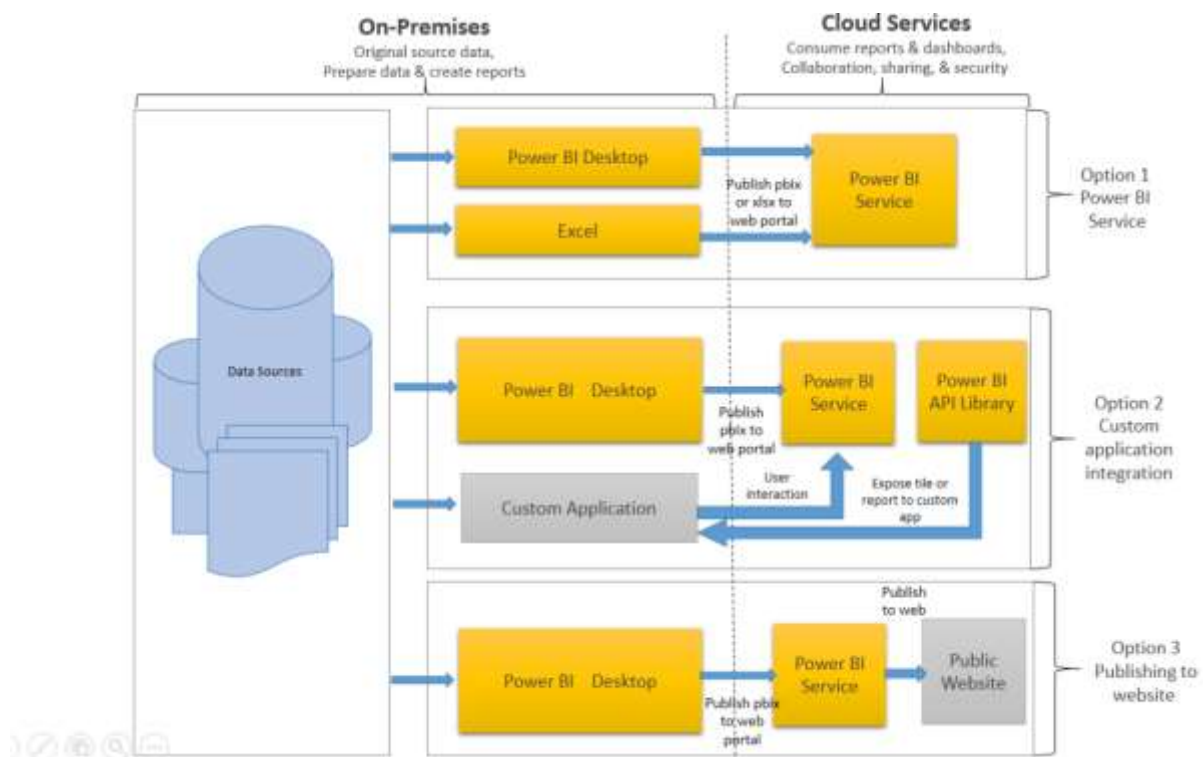
Option 1: Power BI Service

- 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
- 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