High Level Design (HLD)

Wine Data Analysis

Aditya Phulallwar
Priyam Bhalla
Tanmay Dharmadhikari

Document Version Control

Date Issued	Version	Description	Author
8th August 2021	1.0	First Version of Complete HLD	Aditya Phulallwar Priyam Bhalla Tanmay Dharmadhikari

Contents

Abstract

Business analytics (BA) refers to the skills, technologies, and practices for continuous iterative exploration and investigation of past business performance to gain insight and drive business planning. Business analytics focuses on developing new insights and understanding of business performance based on data and statistical methods. In contrast, business intelligence traditionally focuses on using a consistent set of metrics to both measure past performance and guide business planning. In other words, business intelligence focusses on description, while business analytics focusses on prediction and prescription. Here, we will discuss about different types of wines which are best according to the tasters and wines which are expensive and the wines which are above average and below average and the taster who has done most reviews.

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.

2 General Description

2.1 Product Perspective & Problem Statement

While the wine industry generally has been very proactive about dealing with climate change, from capturing fermentation carbon to trialling new varieties, this is ultimately a political problem. To find an effective solution for this problem you were asked to help in ETL process.

2.2 Tools used

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

In this project we have used Tableau Public to work on the data and derive the insights that can be useful.

3 Design Details

3.1 Functional Architecture

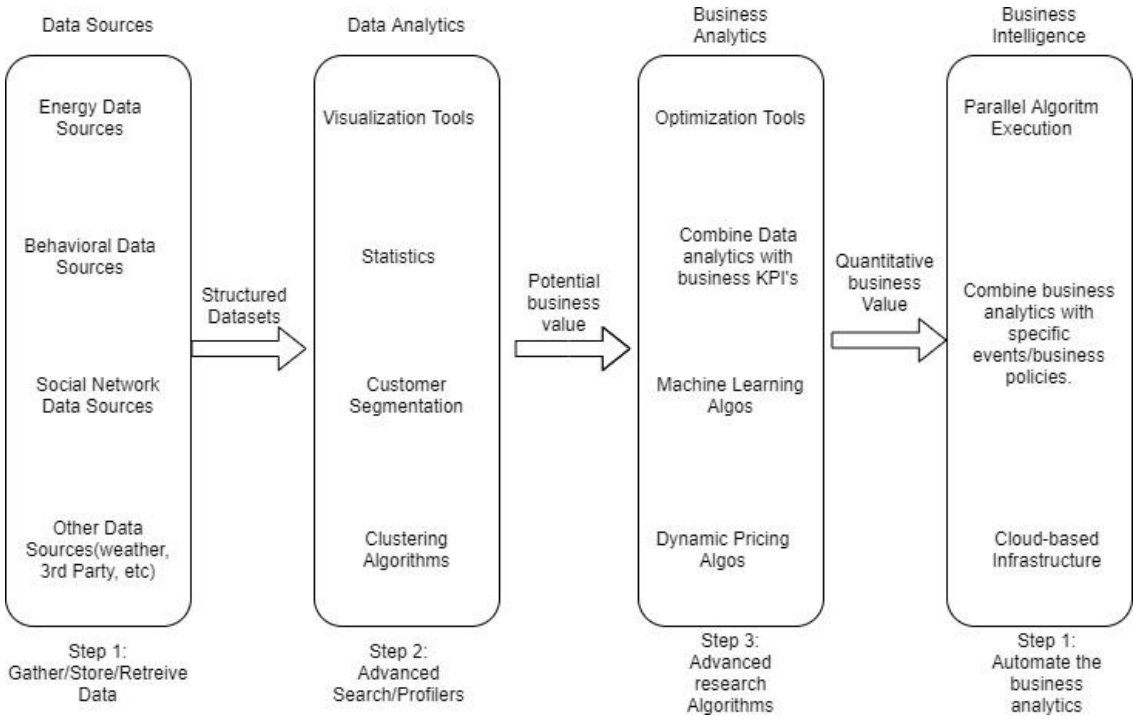
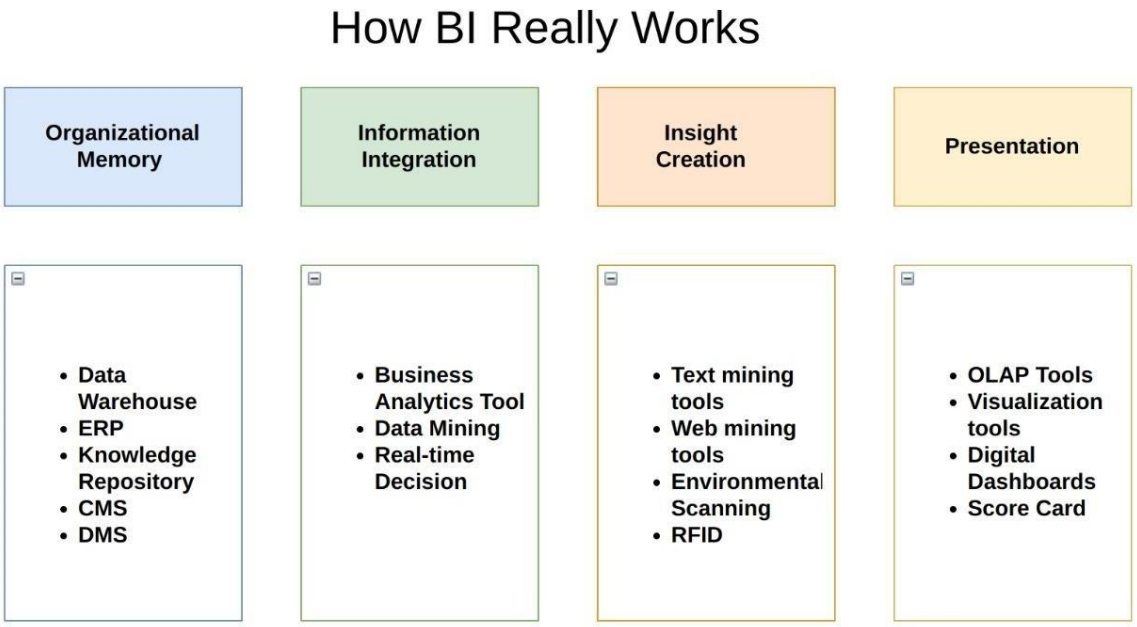
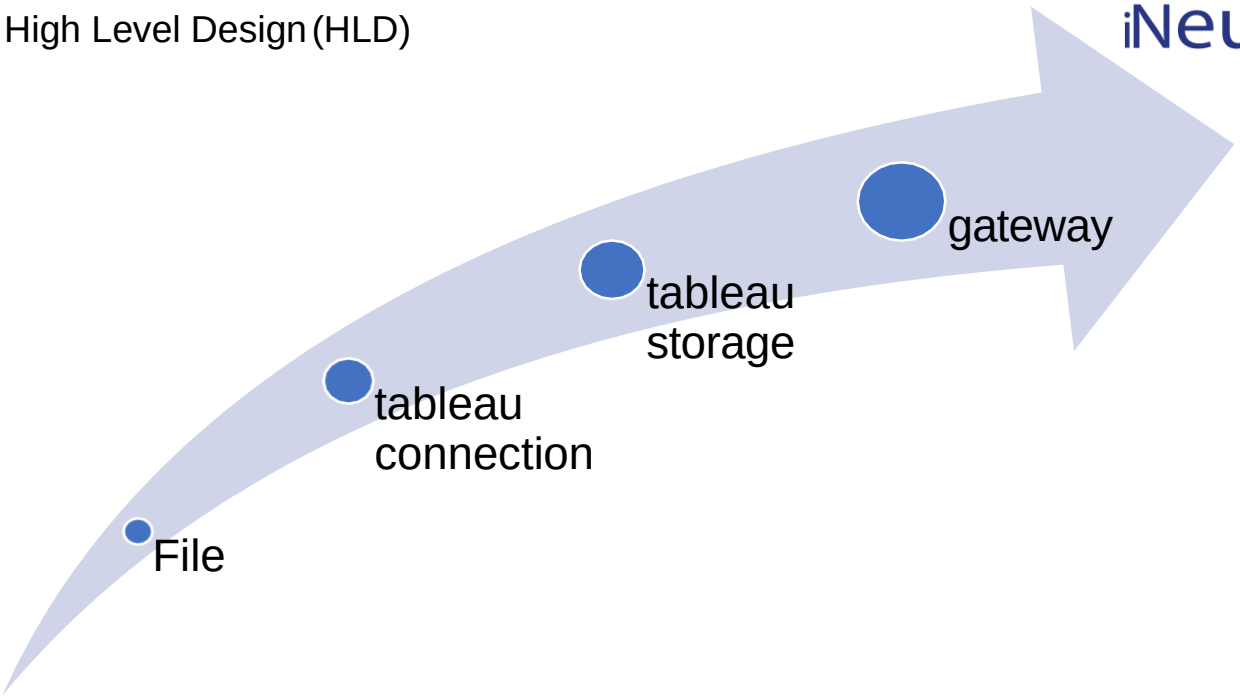


Figure 1: Functional Architecture of Business Intelligence





3.2 Optimization

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 Countries

- 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.

Optimize and materialize your calculations

- Perform calculations in the data
- 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 of Countries and Province so that Province can be dependent on Countries.

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

3.3 KPIs (Key Performance Indicators)

Key indicators displaying Analysis of Wine Review Data

- Relation between Countries and Average Wine Points
- Relation between Variety and Maximum Points
- Relation between Countries selling most wines
- Relation between Taster and Reviews Done
- Best Winery with maximum price
- Relation between Variety and Price
- Relation between Province and wines sold
- Province selling most wines
- Relation between Winery and points
- Relation between Title and Maximum points
- Relation between Country and most wines available
- Wine Data with respect to Country

4 Deployment

Prioritizing data and analytics couldn't come at a better time. Your company, no matter what size, is already collecting data and most likely analyzing just a portion of it to solve business problems, gain competitive advantages, and drive enterprise transformation. With the explosive growth of enterprise data, database technologies, and the high demand for analytical skills, today's most effective IT organizations have shifted their focus to enabling self-service by deploying and operating Tableau at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users and experts alike to author and consume content.

Tableau prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. Tableau Server and Tableau Online leverage your existing technology investments and integrate into your IT infrastructure to provide a self-service, modern analytics platform for your users. With on-premises, cloud, and hosted options, there is a version of Tableau to match your requirements. We have used Tableau Public in this **WINE REVIEW DATA ANALYSIS** project.

Tableau Public

Tableau Public is a free platform to publicly share and explore data visualizations online. Anyone can create visualizations using either Tableau Desktop Professional Edition or the free Public Edition. With millions of inspiring data visualizations, or "vizzes" as we affectionately call them, anyone can see and understand vizzes about any public data topic under the sun, making data part of everyday life and supporting a community.

With the help of tableau public, a link is created which can be used to share and the person can be able to view the visualizations in different screens and the visualization project can also be downloaded and can also edited as per the user needs.