High level design (HLD)



Analyzing world's best wine reviews dataset

**Kashish malviya**

**Document Version Control**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date Issued** | **Version** | **Description** | **Author** |
| 14.02.2025 | 1.0 | First Version of Complete HLD | Kashish Malviya |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents**

Document Version Control…………..

Abstract…………………………………………………………..

1. Introduction …………………………………………………..

1.1 Why this High-Level Design Document?

1.2 Scope

2. General Description

2.1 Product Perspective & Problem Statement

2.2 Tools used

3. Design Details

3.1 Functional Architecture

3.2 Optimization

4. KPIs

4.1 KPIs (Key Performance Indicators)

5. Deployment

Abstract:

Analyzing World's Best Wine Reviews Dataset using Power BI

Wine is a globally appreciated beverage with diverse varieties, price ranges, and quality levels. This project aims to analyze the "World's Best Wine Reviews" dataset using Power BI, providing insights into wine ratings, pricing trends, regional variations, and reviewer preferences.

The analysis will focus on:

Wine Ratings & Quality: Identifying the highest-rated wines and trends over time.

Price vs. Quality Correlation: Exploring whether higher prices guarantee better quality.

Geographical Insights: Analyzing top wine-producing countries and regions.

Reviewer Trends: Understanding how expert reviews influence consumer choices.

Using interactive dashboards and data visualization, Power BI will present findings in a user-friendly format, helping wine enthusiasts, producers, and businesses make informed decisions. The study aims to uncover patterns that define top-quality wines and their market trends.

1 Introduction

* 1. Why this High-Level Design Document?

A high-level document is a document that provides a high-level overview of a system, product, or process. It is used to help stakeholders understand the design approach and make decisions.

What does a high-level document include?

* A high-level architecture diagram
* A brief description of the solution
* A reference to requirements
* Business processes, use cases, and user stories
* Data flow diagrams
* A logical breakdown of the solution
* A consideration of risks, issues, and assumptions

SCOPE

The scope of a design project refers to the specific goals, deliverables, and tasks that need to be completed to successfully achieve the project objectives. It outlines the boundaries and requirements of the project, and typically includes details on the project timeline, budget, and resources.

2 General Description

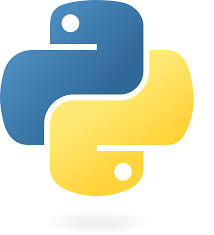
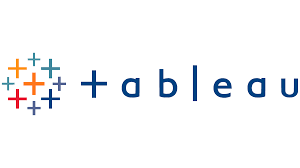
2.1 Product Perspective & Problem Statement

The wine industry is vast, with thousands of varieties, producers, and price points, making it difficult for consumers and businesses to navigate effectively. This project seeks to solve these issues by analyzing a large dataset of wine reviews, using Power BI to uncover trends, correlations, and key insights that benefit both consumers and industry stakeholders.

The objective of the project is to perform data visualization techniques to understand the insight of the data. This project aims apply various Business Intelligence tools such as Tableau or Power BI to get a visual understanding of the data.

Tools Used:

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

3 Design Details

3.1 Functional Architecture

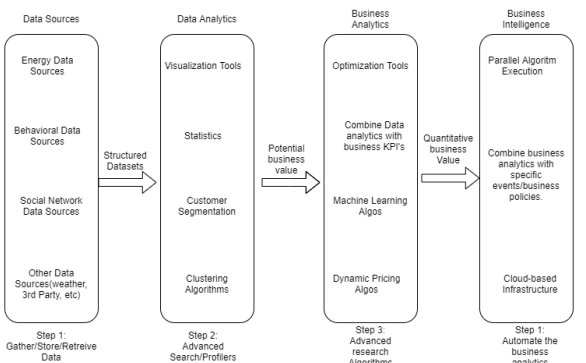
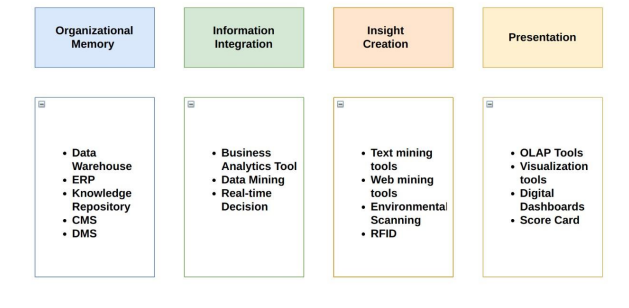


Figure 1: Functional Architecture of Business Intelligence

HOW BI REALLY WORKS:



**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.

• 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.

o LODs - Look at the number of unique dimension members in the calculation.

o 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.

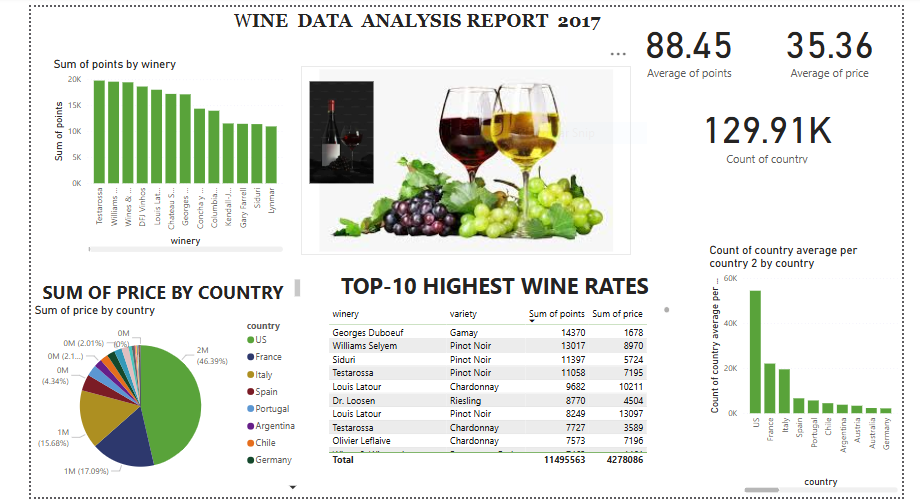
**KPIs for wine analysis**

* Average order value (AOV): The amount of money customers spend on a single order of wine. AOV helps wineries adjust prices and improve offerings.
* Distribution costs: A significant part of a winery's budget.
* Tasting room performance: KPIs can help wineries understand how their tasting room is performing against its goals. This can help wineries identify opportunities for growth.

5 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 analysing 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 Power BI at scale, as well as organizing, orchestrating, and unifying disparate sources of data for business users and experts alike to author and consume content.

Power BI prioritizes choice in flexibility to fit, rather than dictate, your enterprise architecture. Power BI Desktop and Power BI Service leverage your existing technology Investments and integrate them 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 Power BI to match your requirements.



DASHBOARD OF WINE ANALYSIS