



HIGH LEVEL DESING

ANALYZING AMAZON SALES DATA



BY- ANURAG KUMAR

15-OCT-2022

TABLE OF CONTENTS

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 Detail	
4. KPI	

1. Introduction

1.1 Why this High-Level Design Document?

This High-Level Design (HLD) document's goal is to provide the current project description with the extra depth it needs to reflect an appropriate coding model. This paper can be used as a reference guide to understand how the modules work at a high level and to help find conflicts before coding.

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 The Scope

The HLD documentation outlines the system's architecture, including the technology architecture, application architecture (layers), application flow, and database architecture. The HLD employs simple to somewhat complex concepts that system administrators should be able to understand.

2 General Description

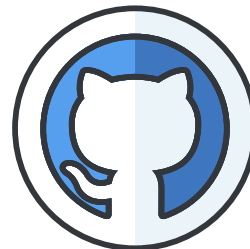
2.1 Product Perspective & Problem Statement

This Project aims to Analyze Amazon Sales from 2017 to 2019 to know more in-depth information about the items that are in high demand, items that are generating high profits which items should not be sold and how much stock we have to maintain for further Sales etc.

3. TOOLS USED



Power BI

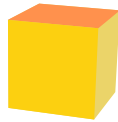


3 Design Details

3.1 Functional Architecture



step 1:
data from
source
system is
integrated
and loaded.



step 2:
data sets are
organized
into
analytics
data
modules



step 3:
bi analyst
and other
analytics
professionals
run
analytical
queries.



step 4:
results are
built into
data
visualizations,
dashboards
and reports



step 5:
business
executives
and workers
use the
information
for decision
making

HOW BI WORKS

ORGANIZATIONAL MEMORY

- Data Warehouse
- Enterprise resource planning (ERP)
- Knowledge Repository
- Content Management System (CMS)

INFORMATION INTEGRATION

- Business Analytical Tools
- Data Mining
- Real Time Decision

INSIGHT CREATION

- Text Mining Tool
- Web Mining Tool
- Environmental Scanning
- RFID

PRESENTATION

- Online Analytical Processing (OLAP) Tool
- Visualization Tool
- Digital Dashboard
- Score Card

3.2 Optimization

1. 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

2. 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.

3. 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 including 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 data filters.

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

4. 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.

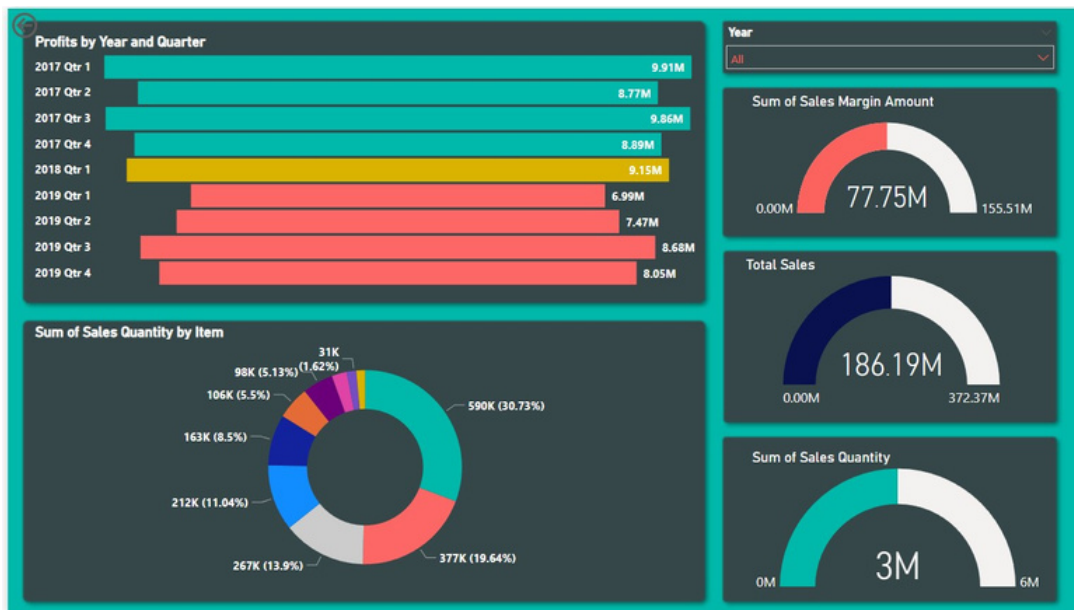
Where possible, use MIN or MAX instead of AVG. AVG requires more processing than MIN or MAX. Often rows will be duplicated and display the same result with MIN, MAX, or AVG.

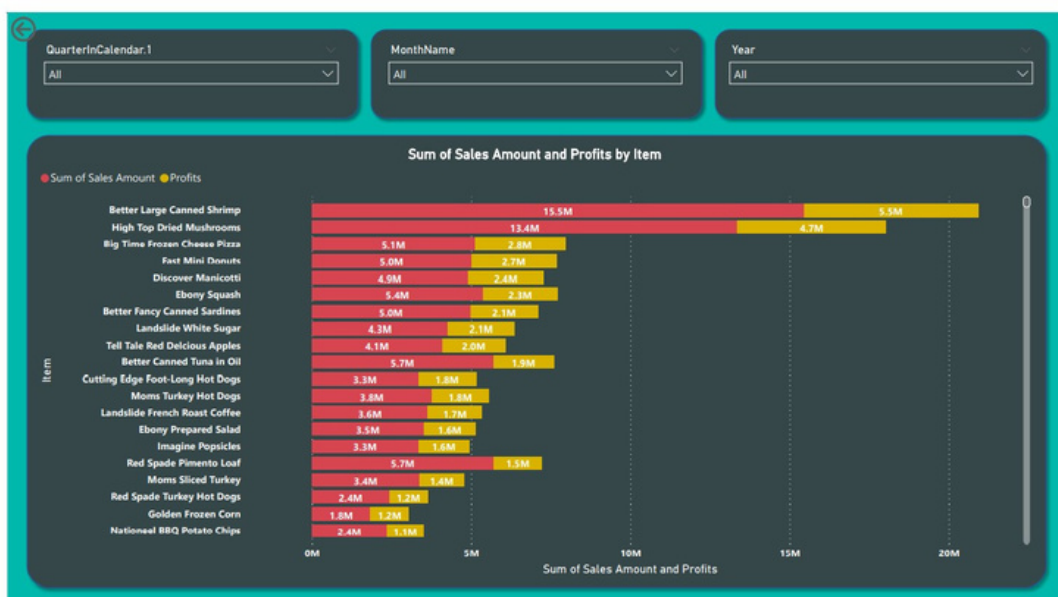
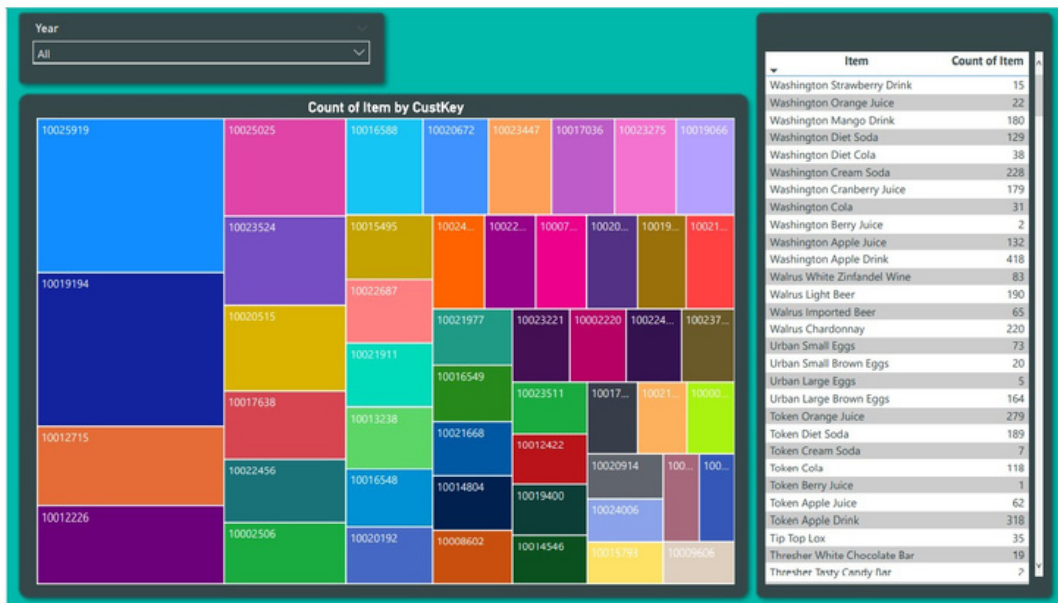
4. KPI

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









Year

All



Sum of Sales Amount by CustKey



Profits by CustKey





HIGH LEVEL DESING

THANKYOU



BY- ANURAG KUMAR

15-OCT-2022