Travel Insurance Data Analysis

Low-Level Design Document

Harsh Makwana,

iNeuron Intelligence Pvt. Ltd.

Document Control

Change Record:

Version	Date	Author	Comments
0.1	03/06/2022	Harsh Makwana	Introduction, Architecture, Architecture Description.

Reviews

Version	Date	Reviewer	Comments

Approval Status

Version	Review Date	Reviewed by	Approved by	Comments

Contents

1. Introduction

1.1 What is Low-Level Design Document?

The goal of the LDD or Low-level design document (LLDD) is to give the internal logic design of the actual program code for the House Price Prediction dashboard. LDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

1.2 What is Scope?

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

1.3 Project Introduction:

Travel insurance, as the name suggests, is an insurance product that is a kind of General Insurance Policy. that covers the risks associated with traveling. It can cover up the costs of trip cancellation, baggage loss, medical expenses, and other losses. It can be for losses during domestic or international trips. It is useful for

protection against unforeseen damage during traveling and associated financial costs. It is a beneficial plan for those traveling domestically or abroad, especially for those who are frequent travelers. In this Project I have analyzed and created meaningful insights and relationships between attributes and came up with interesting findings

1.4 Problem Statement :

Finance is used as a collective term to refer to a broad range of economic services provided by the finance industry, which encompasses a broad range of organizations that manage money, including credit unions, banks, credit card companies, insurance companies, consumer finance companies, stock brokerages, investment funds .

Do ETL : Extract-Transform-Load the dataset and find for me some information from this large data. This is form of data mining.

What all information can be achieved by mining this data, would be brainstormed by the interns

Find key metrics and factors and show the meaningful relationships between attributes.

Do your own research and come up with your findings

2. Architecture

MS Power BI architecture consists of four major steps that explain the whole process from data sourcing to the creation of reports and dashboards. Various technologies and processes work together to get the required results with extreme precision. Let's see those steps further.
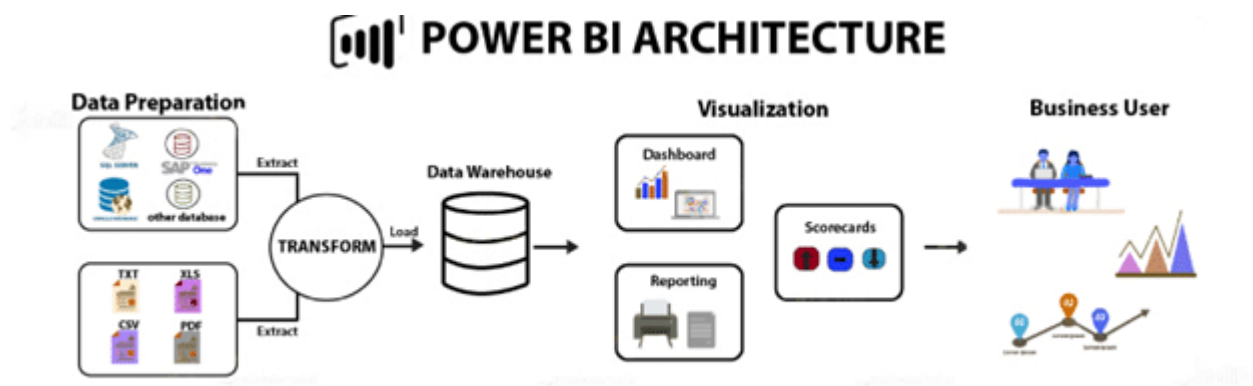
Sourcing data: Power BI extracts data from various servers, Excel sheets, CSV files, and databases. The extracted information can be directly imported to Power BI, or a live service link is established to receive it. If you directly import the data in Power BI, it will only be compressed up to 1 GB. Post that, you can only run live queries on your chunky datasets.

Transforming the data: Before visualizing the data, cleaning and preprocessing it should be done. This means removing useless or missing values from rows or

columns. Following that, certain rules will be applied to transform and load the datasets into the warehouse.

Report and publish: After cleaning and transforming the data, reports will be created based on requirements. A report is a visualization of the data with different filters and constraints presented in the form of graphs, pie charts, and other figures.

Creating dashboards: Power BI Dashboards are created by pinning individual elements or pages of live reports. Dashboards should be created after you have published your reports to the BI service. When the reports get saved, the visual maintains the filter settings chosen so that the user can apply filters and slicers.

3. Architecture Description

3.1 Data Description

Dataset Contains different travel insurance plans given with agencies for various different countries around the world, The attributes present in the dataset are as follows:

1. **Product Name**: It consists of a set of plans or packages, offered to clients based on their selection.

2. **Destination**: Where the client is interested in Travel.

3. **Distribution Channel**: The product delivered to the client, whether it should be in either one mode i.e. online or offline.

4. **Agency**: Which Agency company client adapts.

5. **Agency Type**: Airlines or Travel Agency, whether clients want to travel by plane or they want to personalize their tour.

6. **Gender**: Distinguish between Male or Female, who travels the most.

7. **Claim**: Whether they had claimed some condition in the format of yes/no.

8. **Duration**: How much time they had spent on tour.

9. **Net Sales**: Overall sold price of the company product.

10. **Commission values**: After-sales how much amount of profit was made.

3.2 Web Scrapping

Web scraping is a technique to automatically extract content and data from websites using bots. It is also known as web data extraction or web harvesting. Web scrapping is made simple now days, many tools are used for web scrapping. Some of python libraries used for web scrapping are Beautiful Soup, Scrapy, Selenium, etc.

3.3 Data Transformation

In the Transformation Process, we will convert our original datasets with other necessary attributes format. And will merge it with different dataset or change values by Cleaning and Preprocessing the data.

- Before building any model, it is crucial to perform data pre-processing to feed the correct data to the model to learn and predict
- Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

3.4 Deployment

The deployment process lets you clone content from one stage in the pipeline to another, typically from development to test, and from test to production.

During deployment, Power BI copies the content from the current stage, into the target one. The connections between the copied items are kept during the copy process. Power BI also applies the configured deployment rules to the updated content in the target stage. Deploying content may take a while, depending on the number of items being deployed. During this time, you can navigate to other pages in the Power BI portal, but you cannot use the content in the target stage

When you publish a Power BI Desktop file to the Power BI service, you publish the data in the model to your Power BI workspace. The same is true for any reports you created in Report view. You'll see a new dataset with the same name, and any reports in your Workspace navigator.

Publishing from Power BI Desktop has the same effect as using Get Data in Power BI to connect to and upload a Power BI Desktop file