Template - Requirements Specifications Document

# **Introduction** - *This introduction is very important as it sets expectations that we will come back to throughout the SRS.*

## **Purpose** - The primary aim of this project is to equip the health insurance company with the necessary tools for making data-driven decisions. The project aims to uncover patterns, trends, and correlations within the data acquired from competitors through diverse sources, including web scraping and third-party platforms. The analysis of this data is intended to inform and shape the company's business strategies effectively.

## **Intended Audience and Use** – Marketing and Sales team, Product development team, Data Analyst and Data Scientist. Developers, QA, Customer Service team*.*

## **Product Scope** - Using a combination of AWS S3, AWS Redshift, Databricks, AWS EMR Studio, and Pyspark technologies, this project entails the creation of strong data pipelines. The pipeline's scope will include Data Collection, Data Analysis, Result Creation, and Targeted Offers for customers. Our aim is to gather information that will help us better understand our client's individual needs to adapt our services and increase income.

## **Definitions and Acronyms** –

* AWS – Amazon Web Service
* S3- Amazon Simple Storage Service to store data
* AWS Redshft – Amazon Service to analyze data
* Databricks: Software we can use it to process, clean, share, analyze, read, and write data into d

# **Overall Description** – Using a combination of AWS S3, AWS Redshift, Databricks, AWS EMR Studio, and Pyspark technologies, this project entails the creation of strong data pipelines. The pipeline will include the following significant phases such as data collection. data analysis, result creation, and targeted offers.

# By focusing on these stages, this project aims to establish a comprehensive data-driven framework that gives decision-makers the information they need to encourage revenue development, customer engagement, and operational excellence inside a health insurance company.

# **User Needs** –

* Executive leadership: To make strategic decisions based on a comprehensive understanding of market dynamics and customer behaviors.
* Marketing and Sales Teams: Create targeted marketing campaigns and sales strategies.
* Product Development Team: Designing and refining insurance that carter to the evolving needs of customers
* Data Analyst and Scientist: Use it to further analyses.
* Customer Service team: use it to focus on optimizing customer experiences
* Developers: Can use it to modify witsh new needs or other similar project
* Tester: Can use this to make sure everything is where it suppose to be

## **Assumptions and Dependencies** –

## **Assumptions**

* **Data Quality**: We are assuming the data we are collecting is of reasonable quality, without inaccuracies and inconsistencies.
* **Resource Availability**: We are assuming there are enough resources for the development of the pipeline.

**DEPENDENCIES:**

**Data collection**: Project success depends on timely and accurate data collection from competitors’ sources.

**Data Cleaning**: The accuracy and reliability of analysis depend on comprehensive data cleaning and formatting. Any problems in this phase will cause inaccurate insights.

**Technology infrastructure**: Dependence on the availability and proper functioning of AWS services

Collaboration and Communication: Effective communication and collaboration among different teams are essential to ensure alignment and success of this project

# **System Features and Requirements –**

# *.*

## **Functional Requirements -**

* Which disease has a maximum number of claims.
* Find those Subscribers having age less than 30 and they subscribe any subgroup
* Find out which group has maximum subgroups.
* Find out hospital which serve most number of patients
* Find out which subgroups subscribe most number of times
* Find out total number of claims which were rejected
* From where most claims are coming (city)
* Which groups of policies subscriber subscribe mostly Government or private
* Average monthly premium subscriber pay to insurance company.
* Find out Which group is most profitable
* List all the patients below age of 18 who admit for cancer
* List patients who have cashless insurance and have total charges greater than or equal for Rs. 50,000.
* List female patients over the age of 40 that have undergone knee surgery in the past year

## External Interface Requirements –

## System Features - *System features are a type of functional requirements. These are features that are required in order for a system to function.*

## Nonfunctional Requirements - *Nonfunctional requirements, which help ensure that a product will work the way users and other stakeholders expect it to, can be just as important as functional ones. These may include:*

### Performance requirements

### Safety requirements

### Security requirements

### Usability requirements

### Scalability requirements

## 