

PRD: Healthcare Data Analysis

Product Engineering Project 2

by Nick, Alex, Amrita, Mario

Scrum Master: Amrita Kondeti

Change history:

- Document created 2/5/24

Overview:

- Our team aims to navigate through this dataset, utilizing data science techniques and AWS technologies to bring to light patterns and predictions that could shape future healthcare strategies related to cardiovascular disease.

Objectives:

- Creating S3 bucket to store dataset
- Set up an RDS
- Create a lambda function, which would clean data on upload and then input into the RDS
- Pull clean data from the RDS locally and create five statistical insights
- Created a EC2 website to view statistical analysis (option 1 extension)
- Additional analysis of the data in the form of visualizations or automation (if we have time)

Success Metrics:

- Created an S3 bucket that stores data
- Initialized a RDS with a proper schema
- Utilize a lambda function to clean data before entry into the RDS
- Created five statistical insights based on the cleaned data

Messaging:

- The deliverable of this product will be an RDS with cleaned data and some statistical insights.

Timeline and Release Planning:

- Release Date 2/8/24

Personas:

- Red Shield, a hospital in Los Angeles, wants to find patterns in patients with cardiovascular disease

Scenarios:

- Red Shield is looking a different factors to determine common features in patients with cardiovascular disease:
 - Physical features like Age, Height, Weight, and Gender
 - Habits like Smoking, Alcohol intake, and Physical activity

Features In:

[M] denotes minimum viable experience requirement

- RDS in a public VPC [M]
- Statistical Insights [M]
- EC2 Website

Features Out:

- The ability to make backups of the database

Designs:

- Presentation Slides
- ER Diagram
- Jira Board
- Five statistical insights on the data

Open Issues:

- None right now