Project Report

|  |  |
| --- | --- |
| Date | 17 November 2022 |
| Team ID | PNT2022TMID14568 |
| Project Name | Visualizing and predicting heart  diseases with an interactive |

# INTRODUCTION

* 1. Project Overview

Improving the precision of heart diseases detection has been investigated by many researchers in the literature. Such improvement induced by the overwhelming health care expenditures and erroneous diagnosis. As a result, various methodologies have been proposed to analyse the disease factors aiming to decrease the physicians practice variation and reduce medical costs and errors. Heart disease is one of the biggest causes of morbidity and mortality among the population of the world. Prediction of cardiovascular disease is regarded as one of the most important subjects in the section of clinical data analysis. The amount of data in the healthcare industry is huge. Data mining turns the large collection of raw healthcare data into information that can help to make informed decisions and predictions.

# LITERATURE SURVEY

1.1 Existing problem

# Effective heart disease prediction system using data mining techniques

In order to predict the probability of patients having heart disease, a confusion matrix was created, where A denotes patients with heart disease, and B denotes patients with no heart disease.The confusion matrix contains the following four entries:

\*TP (true positive): The number of records classified as true while they were actually true.

\*FP (false positive): The number of records classified as true while they were actually false.

\*FN (false negative): The number of records classified as false while they were actually true.

\*TN (true negative): The number of records classified as false while they were actually false.

The overall process of effective heart disease prediction system (EHDPS) is based on the following three steps:

1. Data collection
2. Data pre-processing and
3. The classification of data.

# Predicting Heart Diseases In Logistic Regression Of Machine Learning Algorithms

The aim of this study is to evaluate the risk of 10-year CHD using 14 IVs. The attributes are selected after the backward elimination process considering the P values which are lower than 5%. Therefore, the logistic regression model is derived through P values of the variables <0.05 (sex, age,cigsPerDay, totChol, sysBP, glucose). According to the logistic regression outcome, men are more suspectable to heart disease than women. Age, number of cigarettes per day and systolic blood pressure are the odds of CHD. However, There is no significance change in the total cholesterol level and the glucose level. But, the level of

glucose has a negligible change in odds. The model is more specific than sensitive. Further, the accuracy of the model is 0.87. The value under the ROC curve is 73.5 which is somewhat satisfactory. Moreover, the model could be improved by using more data.

# Evolutionary algorithm-based convolutional neural network for predicting heart diseases

Convolutional neural networks (CNNs) have been commonly used in medical decision support systems to predict and diagnose different diseases with good precision. CNNs are extremely successful in developing health support systems because of their ability to identify relationships and hidden patterns in healthcare data. One of the most important and useful applications of such systems is in the prediction of heart diseases by observing cardiac anomalies. Fundamentally, CNNs have multiple hyperparameters and various specific architectures, which are costly and impose challenges in selecting the best value among possible hyperparameters**.**

1.1 Problem Statement Definition

Day by day the cases of heart diseases are increasing at a rapid rate and it’s very Important and concerning to predict any such diseases beforehand. This diagnosis is a difficult task i.e. it should be performed

precisely and efficiently. The research paper mainly focuses on which patient is more likely to have a heart disease based on various medical attributes. We prepared a heart disease prediction system to predict whether the patient is likely to be diagnosed with a heart disease or not using the medical history of the patient.

# IDEATION & PROPOSED SOLUTION

* + 1. Empathy Map Canvas

t



|  |  |  |
| --- | --- | --- |
|  | | |
|  | To foresee To know if To gain the health issues the result is trust of prior to its accurate doctors and  occurence patients  To compare To abide by People choose to  Look for Random Promote their way the results of the traditional believe doctor's  alternative suggestions of problem solving people with medical opinion rather  solutions for for with out knowing similar procedures than computer  diagonisis counsulting the the person's diagnosis programs  health status  Get their Thinking about Pretend as if  results the cost rather the health  double than the condition is  effectiveness of good though its  checked the solution not    fear of getting getiing a redeuces  not getting osilating accurate  proper thoughts about predictions waiting time  the result being  treatment right or not in hospitals |  |

* + 1. Ideation & Brainstorming



**1**

**Define your problem statement**

What problem are you trying to solve? Frame your problem as a How Might We statement. This will be the focus of your brainstorm.

**2**

**Brainstorm**

Write down any ideas that come to mind that address your problem statement.

**3**

**Group ideas**

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

**4**

**Prioritize**

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

**10 minutes**

**5 minutes**

**20 minutes**

**20 minutes**

**Swetha**

**Pradyuth**

**Gokul**

**Swathi**

**PROBLEM**

**Visualising and predicting heart diseases with an interactive dashboard.**

Monitoring h Recording

eart activities

and blood and before

sugar. and after excercise.

use ml

techniques(naive

bayers, random

forest, svm,

decsion tree) to

develop the model

prepare

empathy map

design and build an UI

the results

to display

analyse dataset and prediction methods

**Key rules of brainstorming**

To run an smooth and productive session

Stay in topic. Encourage wild ideas.

preprocess the dataset

Defer judgment. Listen to others.

**Importance**

If each of these tasks could get done without any difficulty or cost, which would have the most positive impact?

collect the necessary

required to create dataset

details

Go for volume. If possible, be visual.

**TIP**

Participants can use their cursors to point at where sticky notes should go on the grid. The facilitator can confirm the spot by using the laser pointer holding the **H key** on the keyboard.

**Feasibility**

Regardless of their importance, which tasks are more feasible than others? (Cost, time, effort, complexity, etc.)

**Need some inspiration?**

See a finished version of this template to kickstart your work.

**Open example**

**Share template feedback**

**Keep moving forward**

**Strategy blueprint**

Define the components of a new idea or strategy.

**Open the template**

**Customer experience journey map**

Understand customer needs, motivations, and obstacles for an experience.

**Open the template**

**Strengths, weaknesses, opportunities & threats**

Identify strengths, weaknesses, opportunities, and threats (SWOT) to develop a plan.

**Open the template**

**Quick add-ons**

1. **Share the mural**

**Share a view link** to the mural with stakeholders to keep them in the loop about the outcomes of the session.

1. **Export the mural**

Export a copy of the mural as a PNG or PDF to attach to emails, include in slides, or save in your drive.

**After you collaborate**

You can export the mural as an image or pdf to share with members of your company who might find it helpful.

1. **Team gathering**

Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.

1. **Set the goal**

Think about the problem you'll be focusing on solving in the brainstorming session.

1. **Learn how to use the facilitation tools**

Use the Facilitation Superpowers to run a happy and productive session.

**Open article**

**Before you collaborate**

A little bit of preparation goes a long way with this session. Here’s what you need to do to get going.

**10 minutes**

**Share template feedback**

**Brainstorm**

**& idea prioritization**

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

**10 minutes** to prepare

**1 hour** to collaborate

**2-8 people** recommended

various

Read

heart patient histories

age,sex,blood

profile that includes

sugar level,blood oxygen level, murmur in heart

Make a patient

Involve ml to predict heart dieases

make use of neural networks

Keep track of pulse rate

Check blood sugar level

design a UI that is apt to display the key requirements

organise the dataset and input the patient characteristics used as features in the predictive model.

identifying the crucial requirements

collect and explore different datasets

conduct a survey based on cvd

**TIP**

Add customizable tags to sticky notes to make it easier to find, browse, organize, and categorize important ideas as themes within your mural.

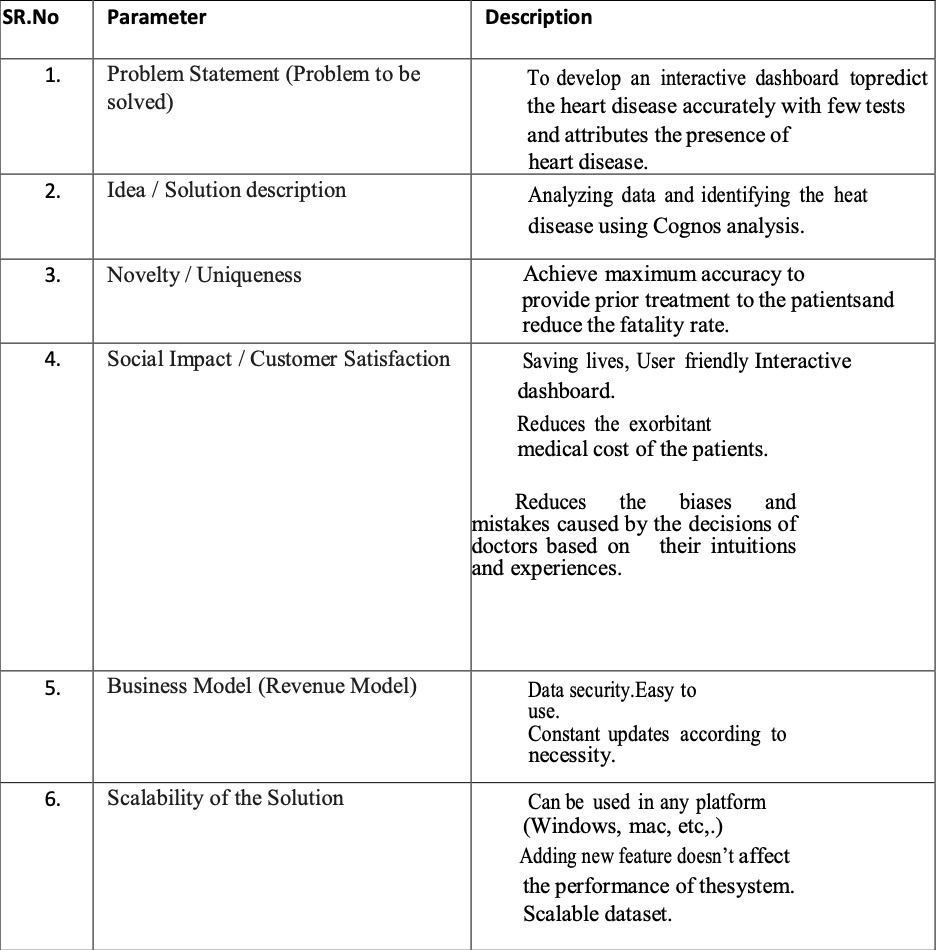
conduct survey

**TIP**

You can select a sticky note and hit the pencil [switch to sketch] icon to start drawing!

**Template**

* + 1. Proposed Solution



* + 1. Problem Solution fit



**1. CUSTOMER SEGMENT(S)**

Who is your customer?

i.e. working parents of 0-5 y.o. kids

**CS**

**6. CUSTOMER CONSTRAINTS**

**CC**

**AS**

What constraints prevent your customers from taking action or limit their choices

of solutions? i.e. spending power, budget, no cash, network connection, available devices.

**5. AVAILABLE SOLUTIONS**

Which solutions are available to the customers when they face the problem

People suffering from cardiac diseases

i. Insufficient money for health checkups

or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking

i. Medical checkup regarding the symptoms

1. Lack of time ii. X-rays
2. Ignorance iii. Regular general checkups
3. Unaware about medical checkups

Problem-Solution it canvas is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 license Created by Daria Nepriakhina / Amaltama.com

Purpose / Vision

**Explore AS, differentiate**

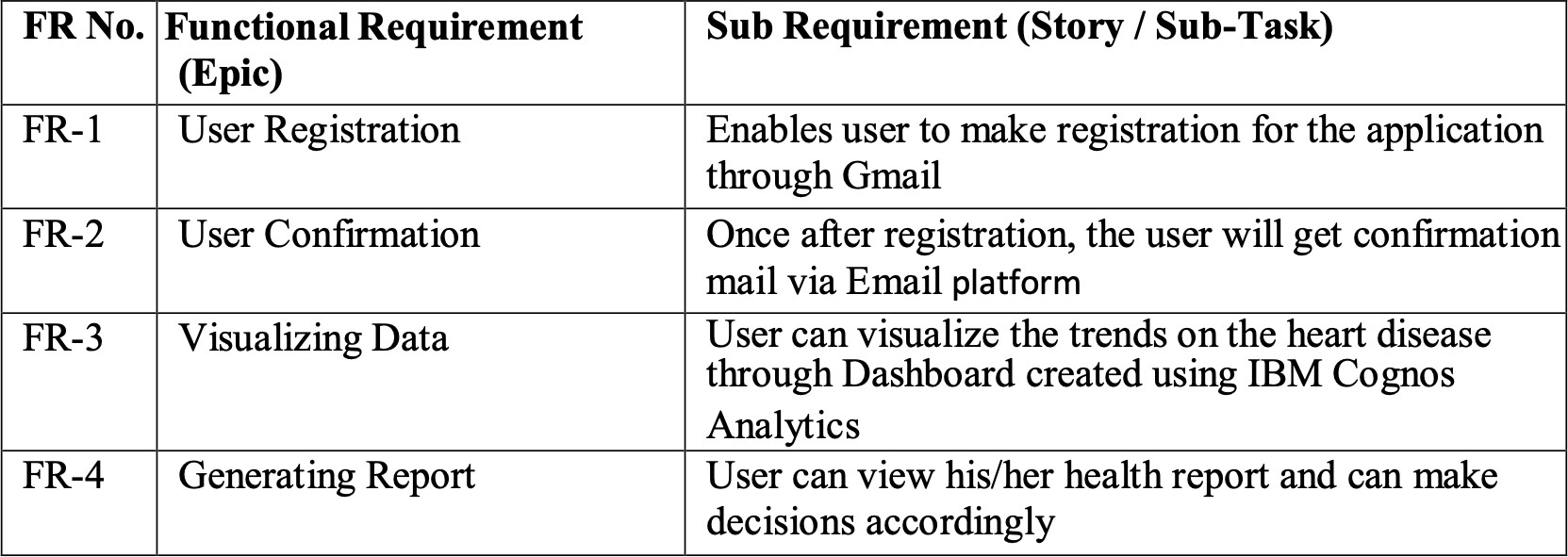
**Deﬁne CS, ﬁt into CC**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Focus on J&P, tap into BE, understand RC** | 1. **JOBS-TO-BE-DONE / PROBLEMS J&P**   Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.   * 1. Medical checkups are expensive   2. Delayed medical reports   3. Complex test results | **9. PROBLEM ROOT CAUSE RC**  What is the real reason that this problem exists? What is the back story behind the need to do this job?  i.e. customers have to do it because of the change in regulations.  Lack of cost effective, reliable, speedy and accurate methods for predicting the heart disease | **7. BEHAVIOUR BE**  What does your customer do to address the problem and get the job done?  i.e. directly related: ﬁnd the right solar panel installer, calculate usage and beneﬁts; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)   * 1. Getting stressed about their health   2. Finding out the seriousness of the disease | **Focus on J&P, tap into BE, understand RC** |

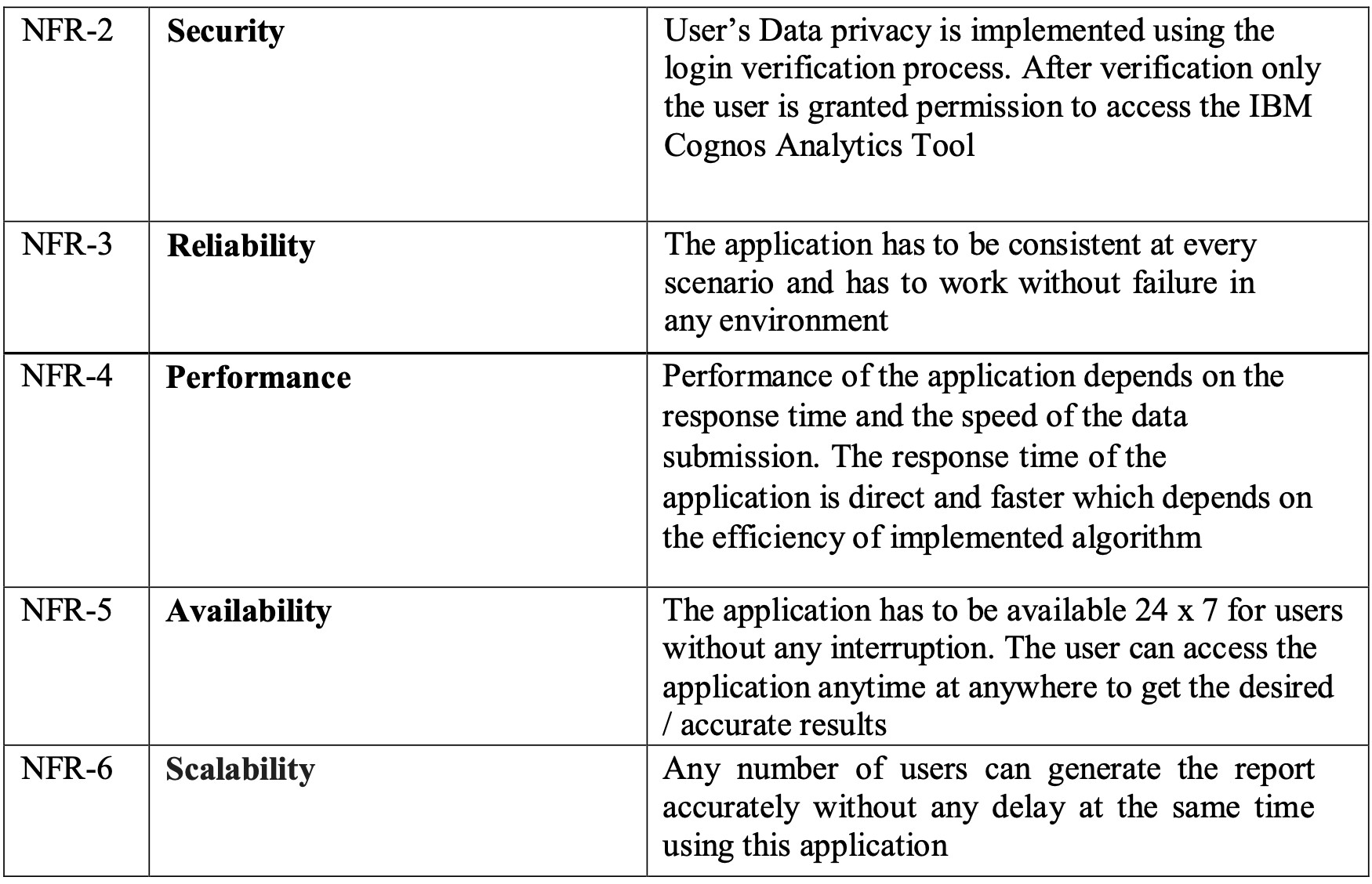
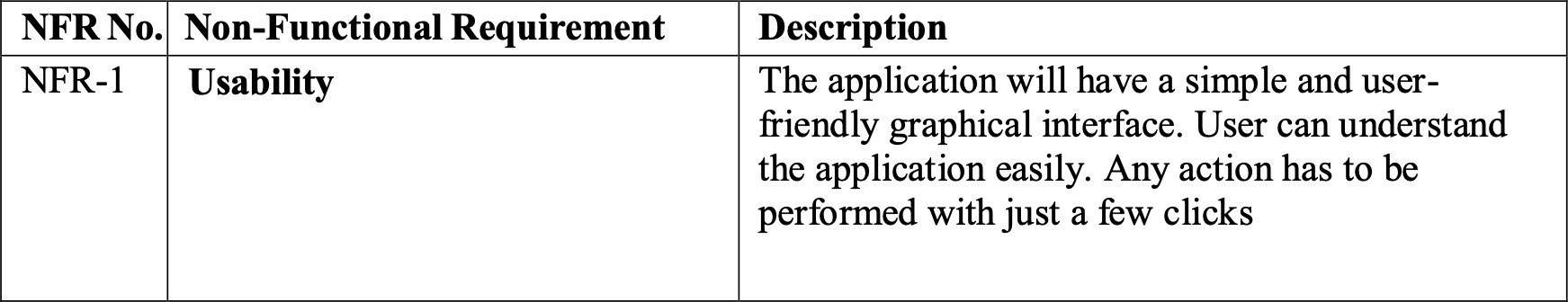
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Identify strong TR & EM** | 1. **TRIGGERS TR**   What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efﬁcient solution in the news.   * 1. Having symptoms   2. Biological cycle and lifestyle getting affected | **10. YOUR SOLUTION SL**  If you are working on an existing business, write down your current solution ﬁrst, ﬁll in the canvas, and check how much it ﬁts reality.  If you are working on a new business proposition, then keep it blank until you ﬁll in the canvas and come up with a solution that ﬁts within customer limitations, solves a problem and matches customer behaviour.  Developing an interactive dashboard for predicting and visualizing heart diseases | 1. **CHANNELS of BEHAVIOUR CH**     1. **ONLINE**   What kind of actions do customers take online? Extract online channels from #7  Consulting doctors online and searching about the symptoms in the internet   * 1. **OFFLINE**   What kind of actions do customers take ofﬂine? Extract ofﬂine channels from #7 and use them for customer development.  Consulting doctors offline and asking friends and family about the disease | **Extract online & ofﬂine CH of BE** |
| **4. EMOTIONS: BEFORE / AFTER EM**  How do customers feel when they face a problem or a job and afterwards?  i.e. lost, insecure > conﬁdent, in control - use it in your communication strategy & design.   * 1. Feeling unsecured and depressed   2. Spending a lot of money |

# REQUIREMENT ANALYSIS

* + 1. Functional requirement

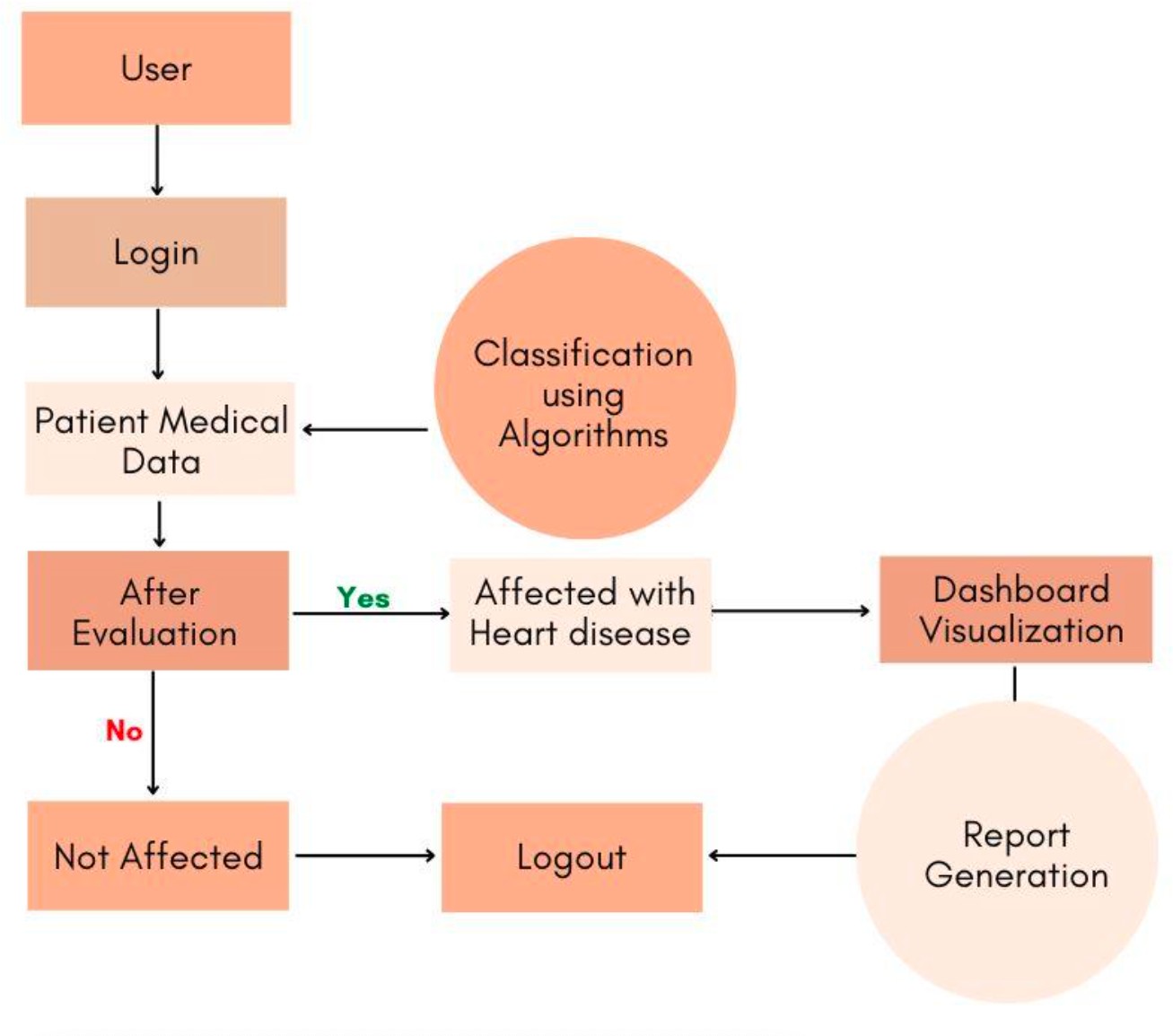


* + 1. Non-Functional requirements

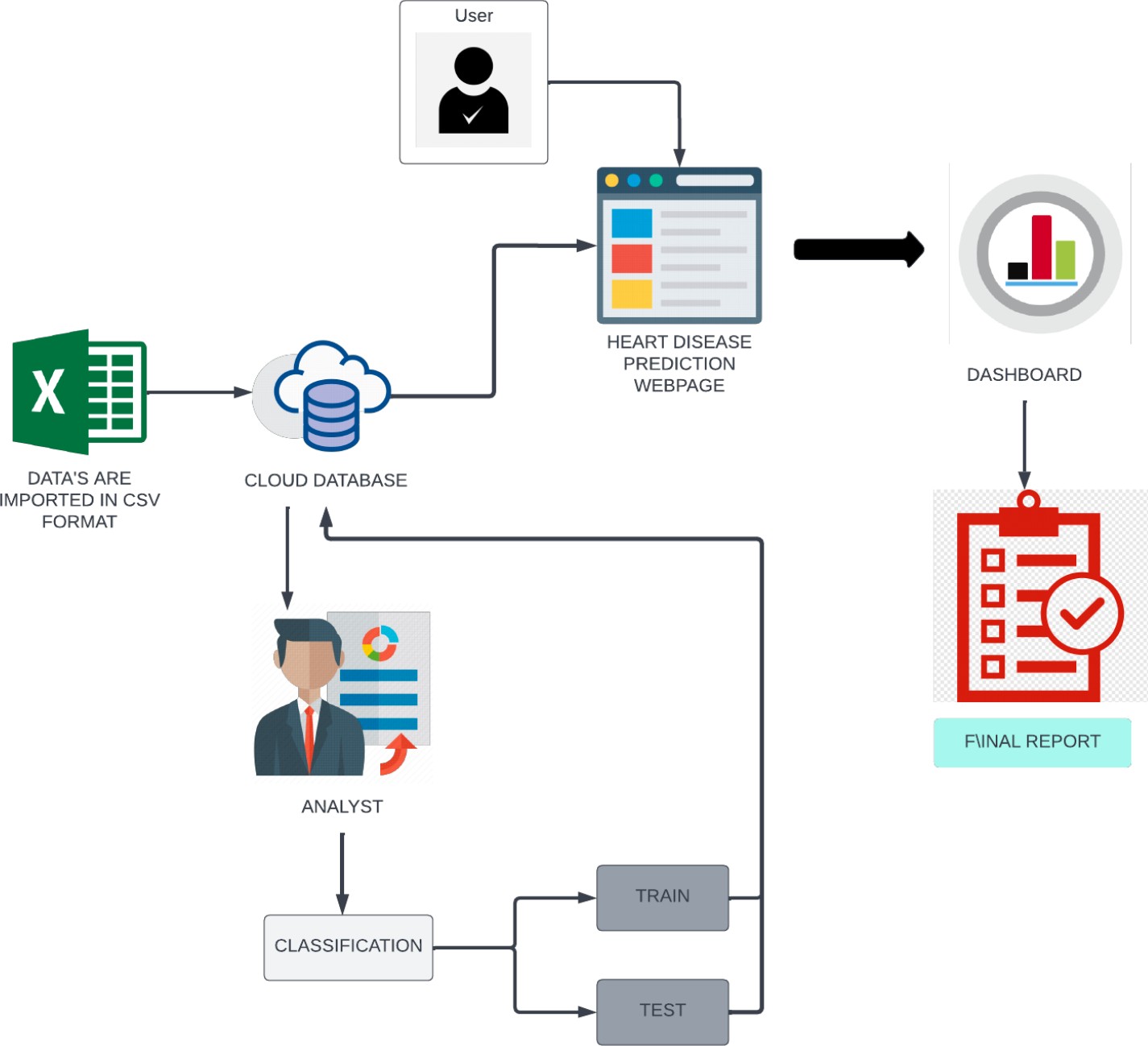


# PROJECT DESIGN

* + 1. Data Flow Diagrams

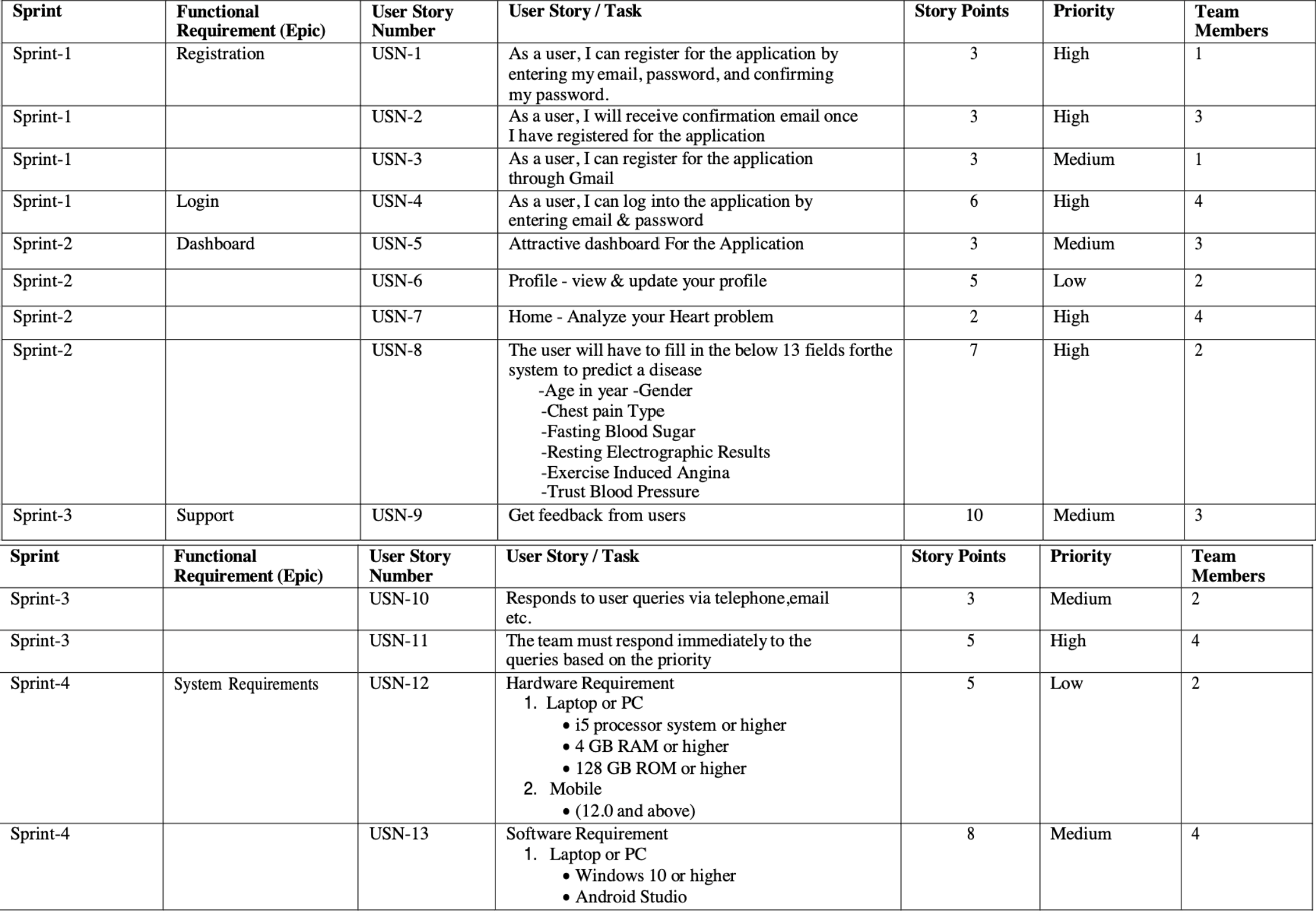


* + 1. Solution & Technical Architecture

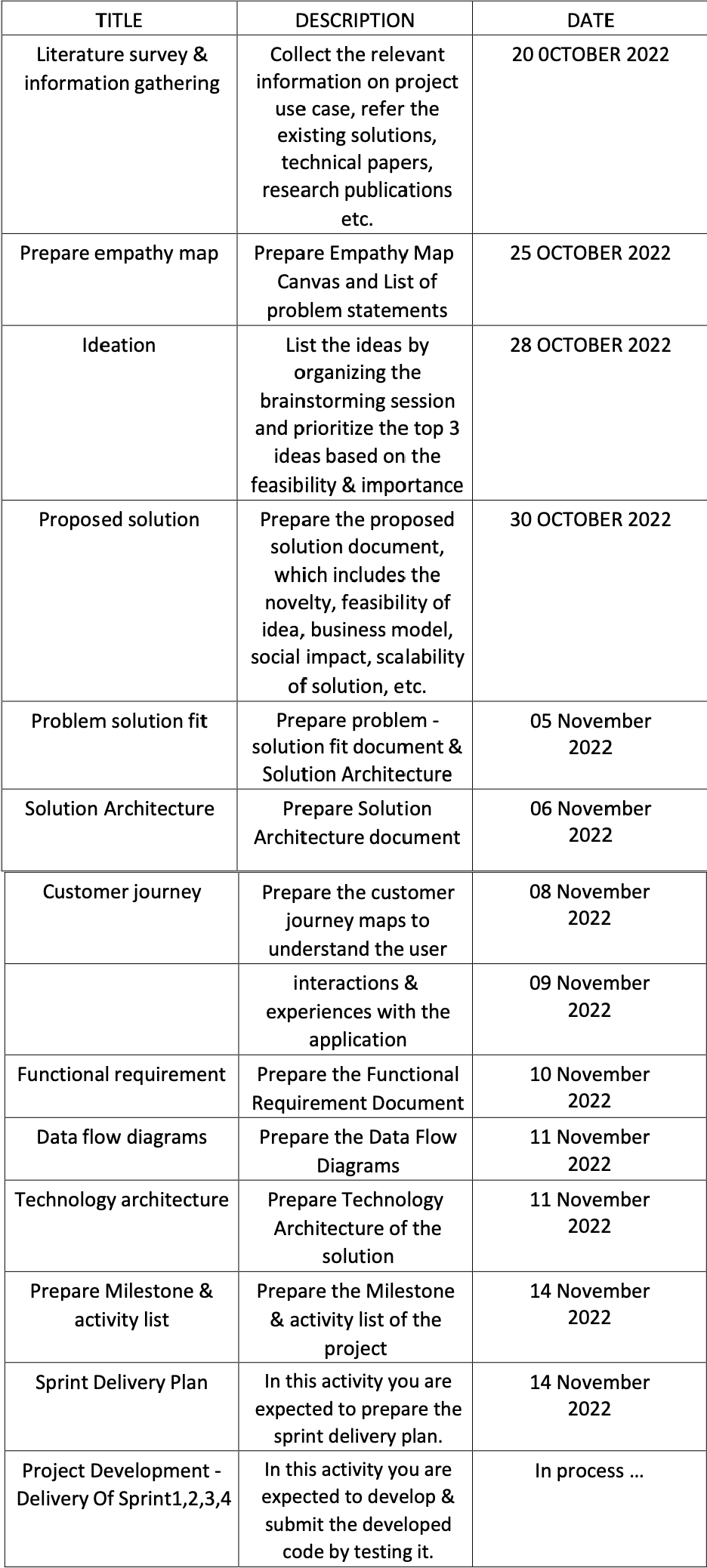


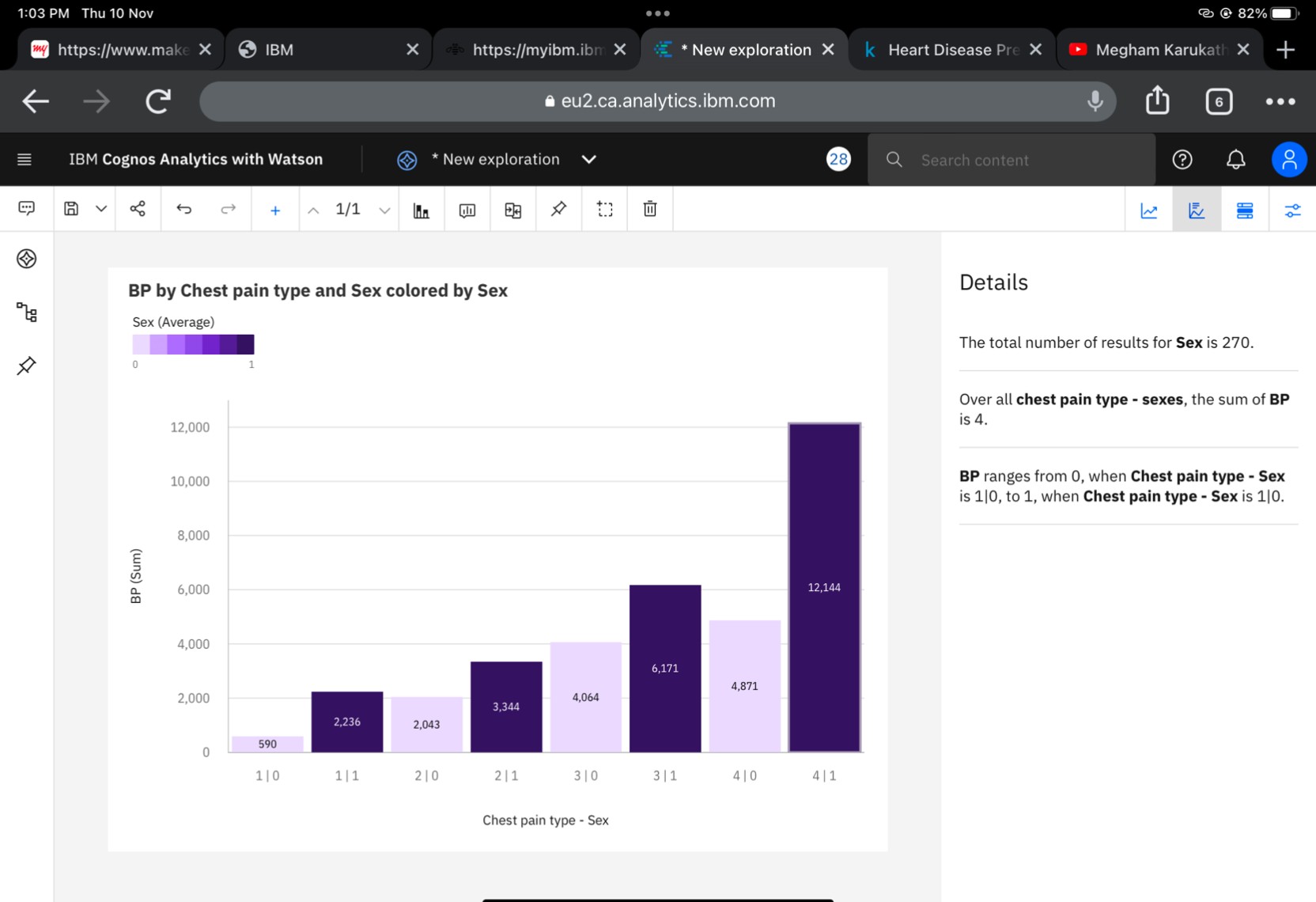
# PROJECT PLANNING & SCHEDULING

* + 1. Sprint Planning & Estimation



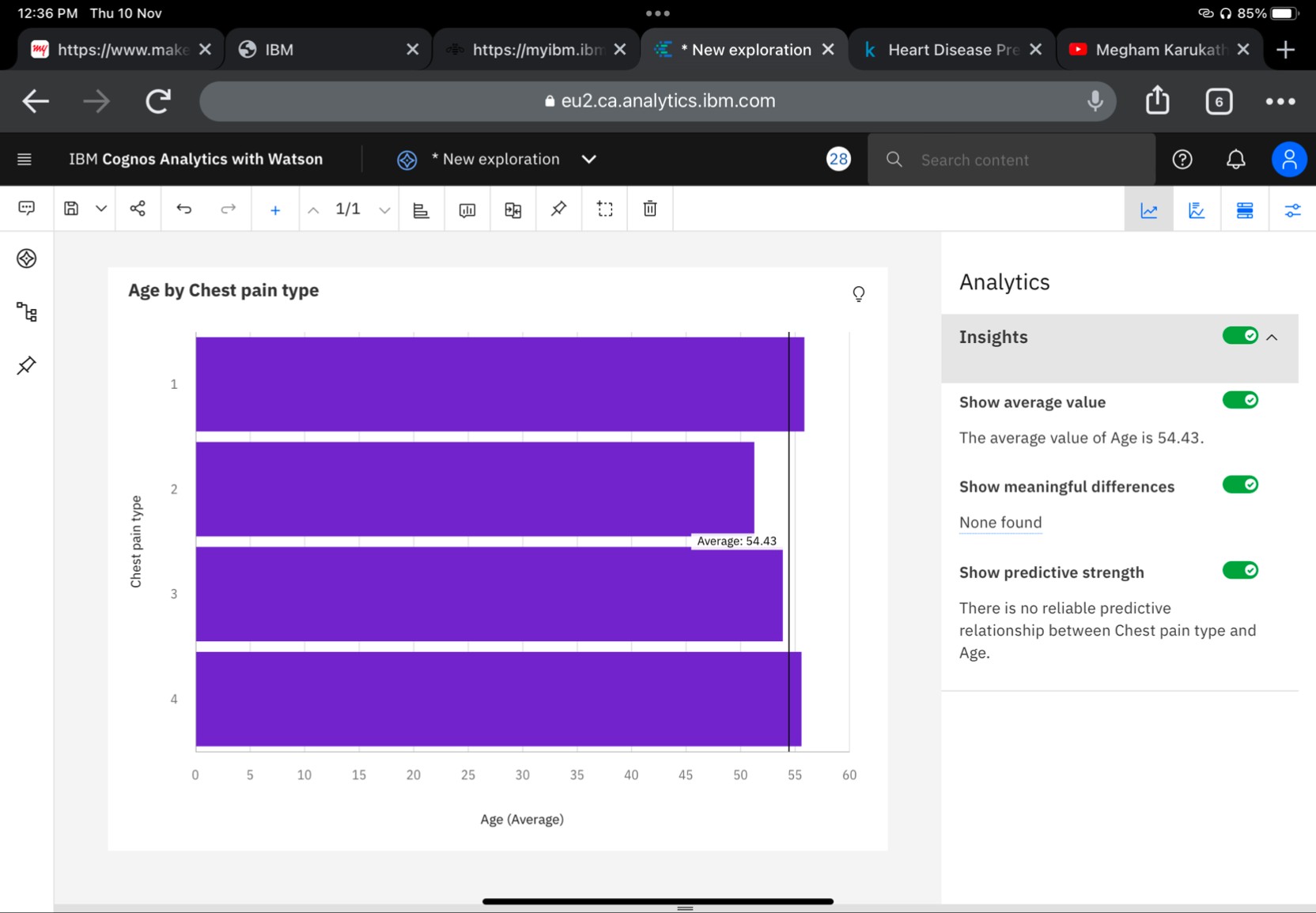
* + 1. Sprint Delivery Schedule

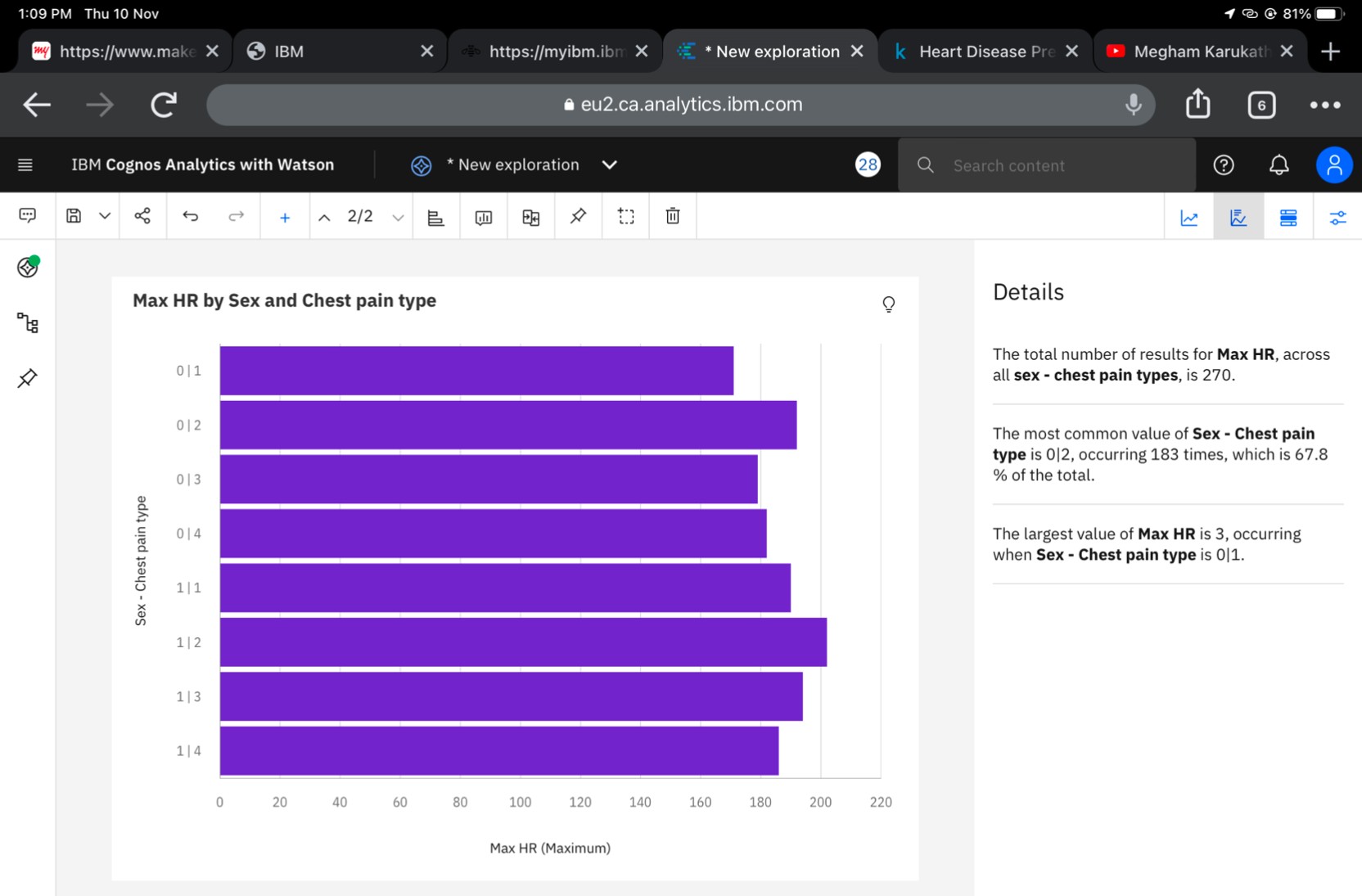


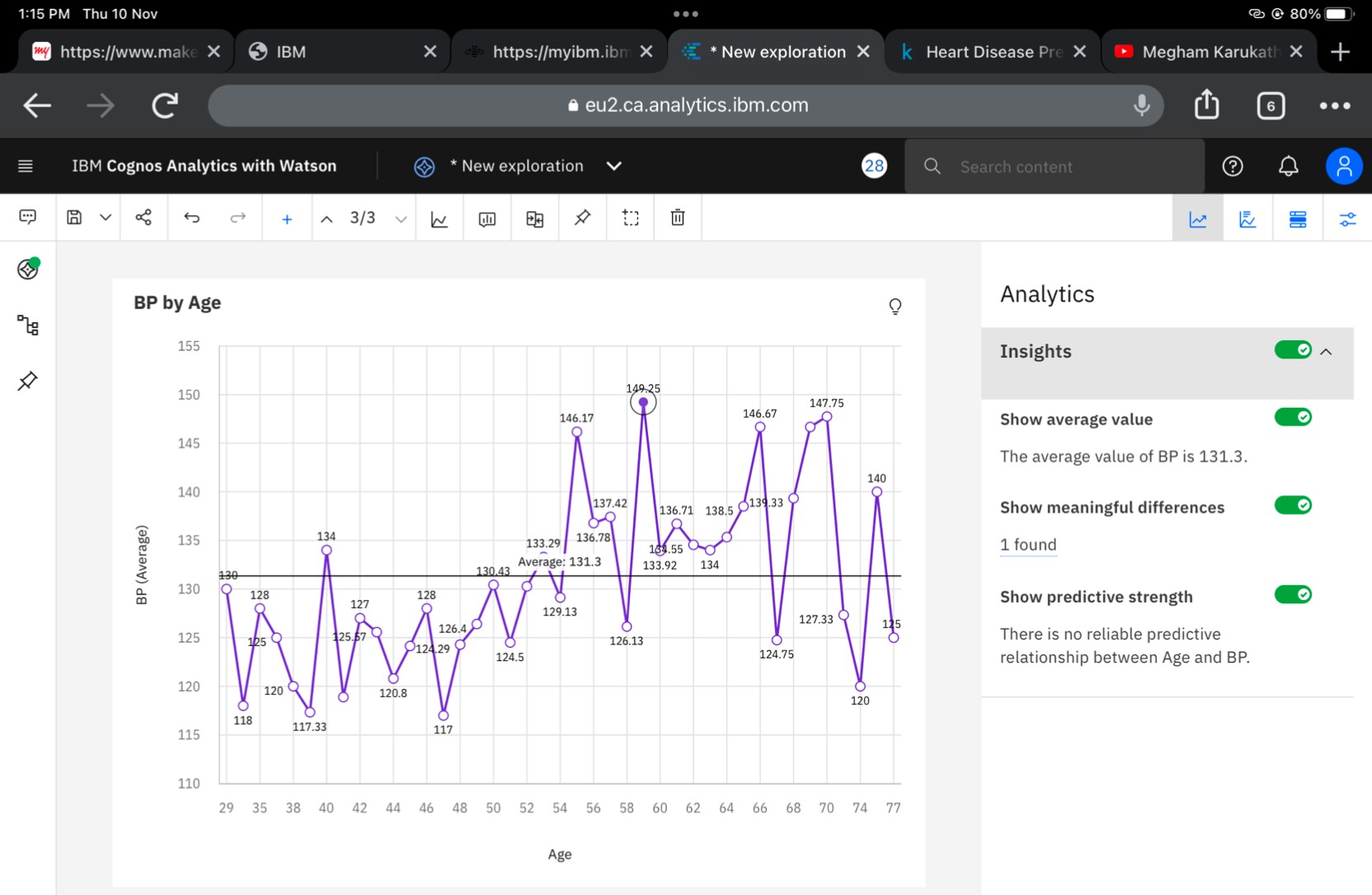


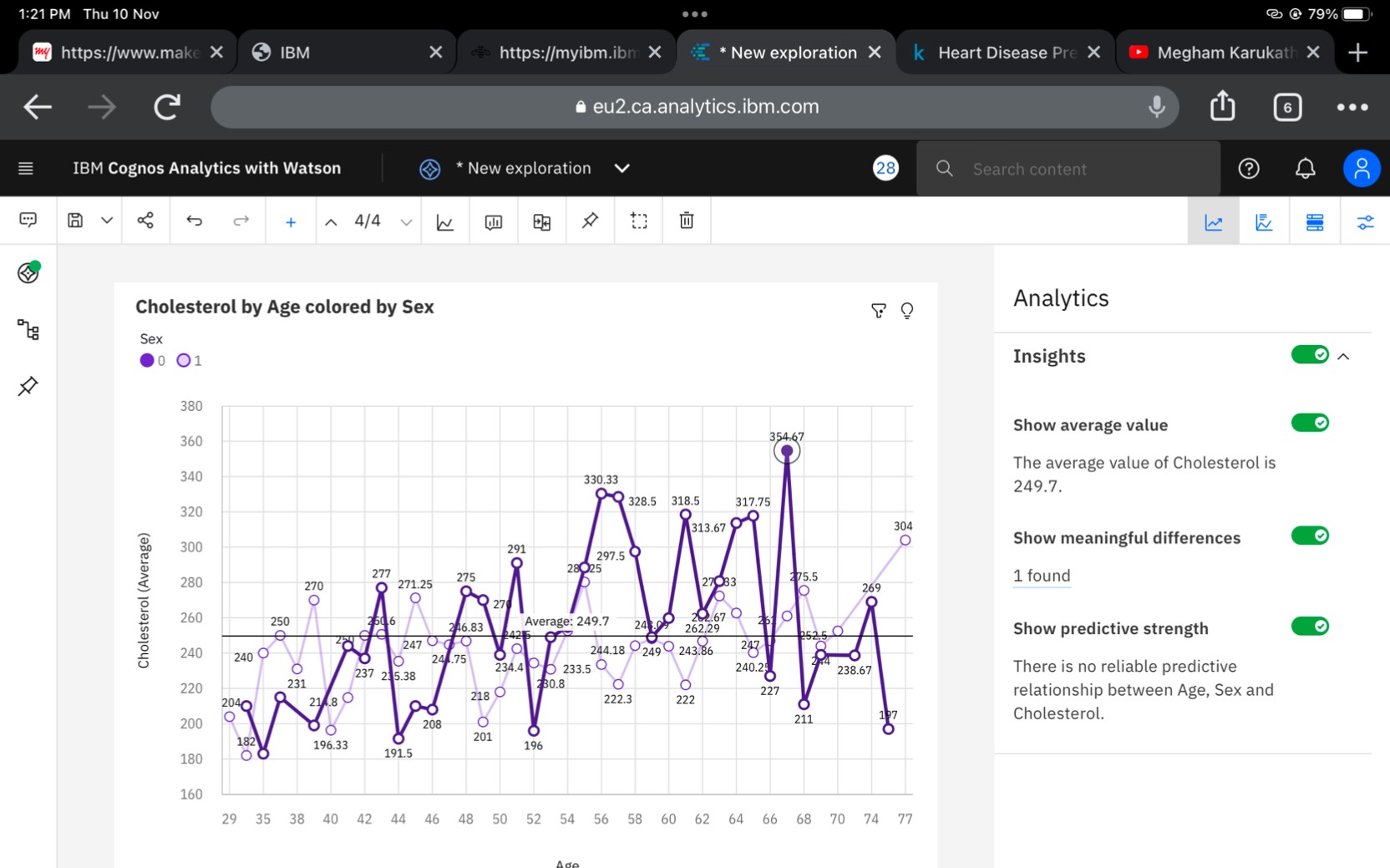
# WORKING WITH THE DATASET & DATA VISUALISATION

* + 1. Understanding the dataset



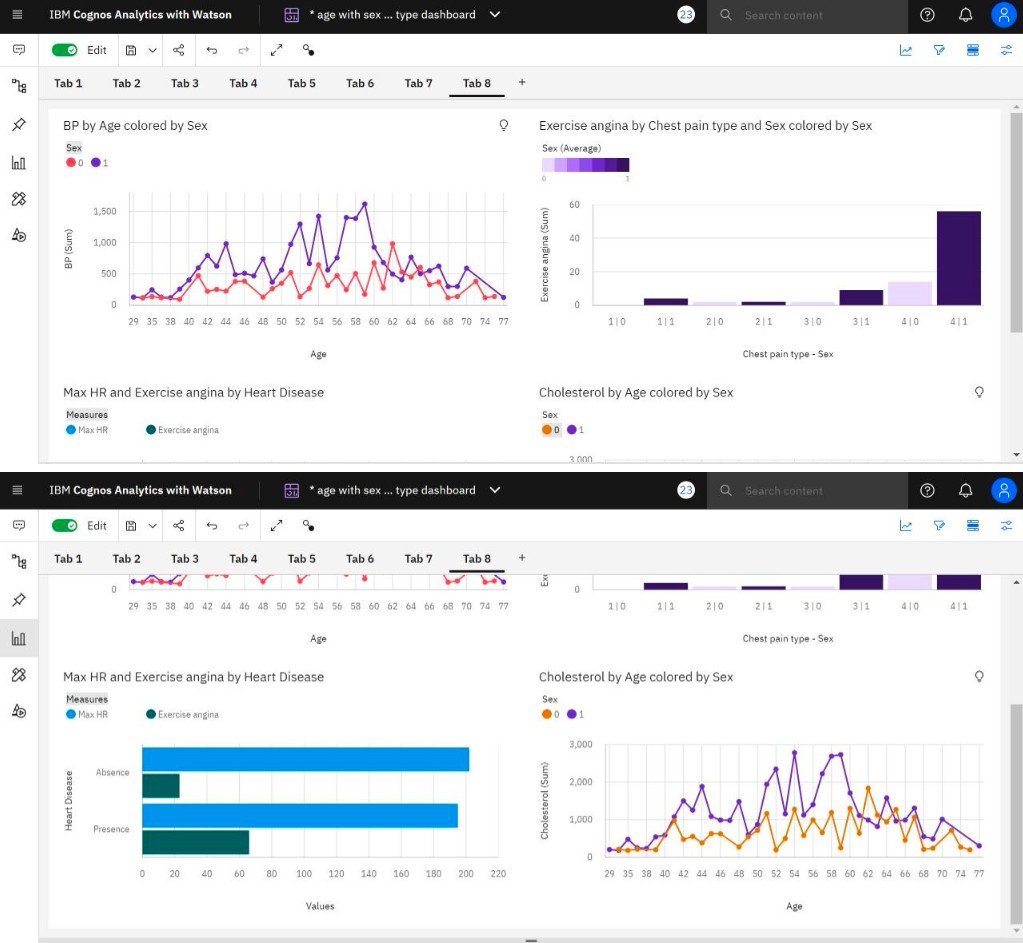


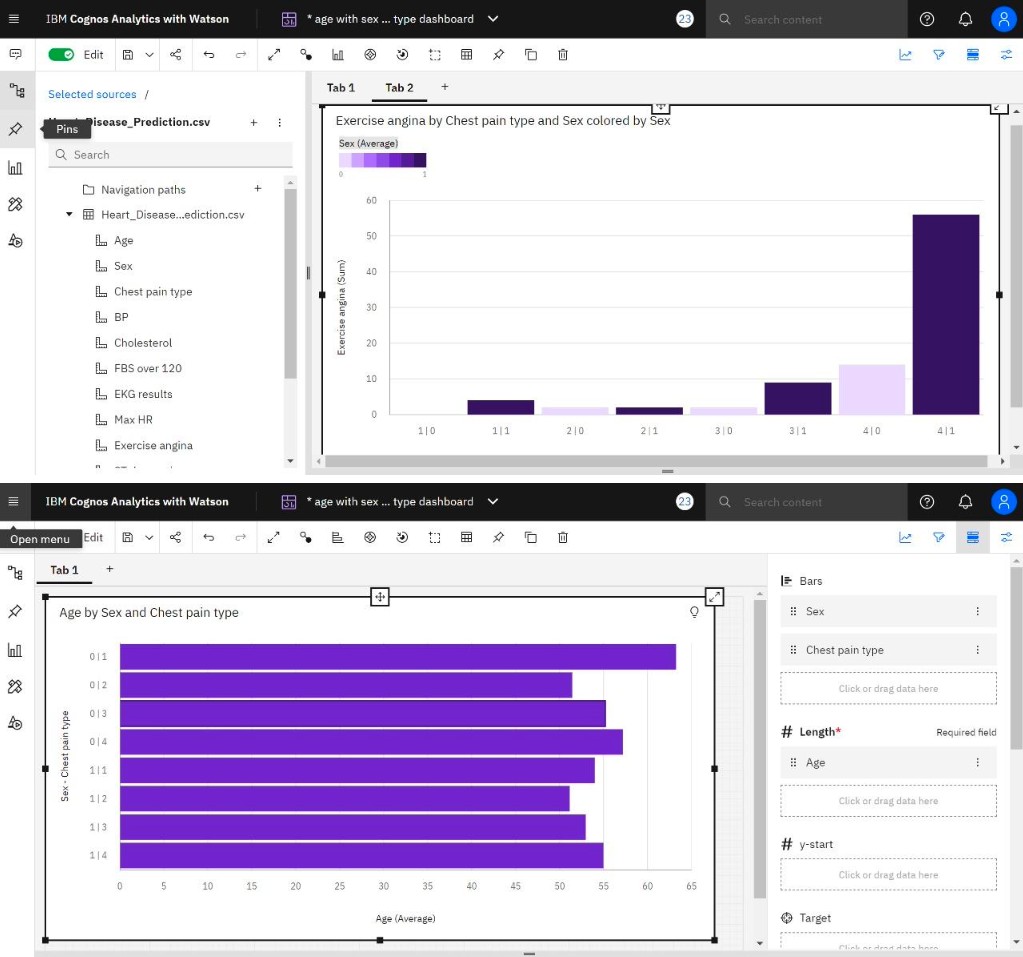


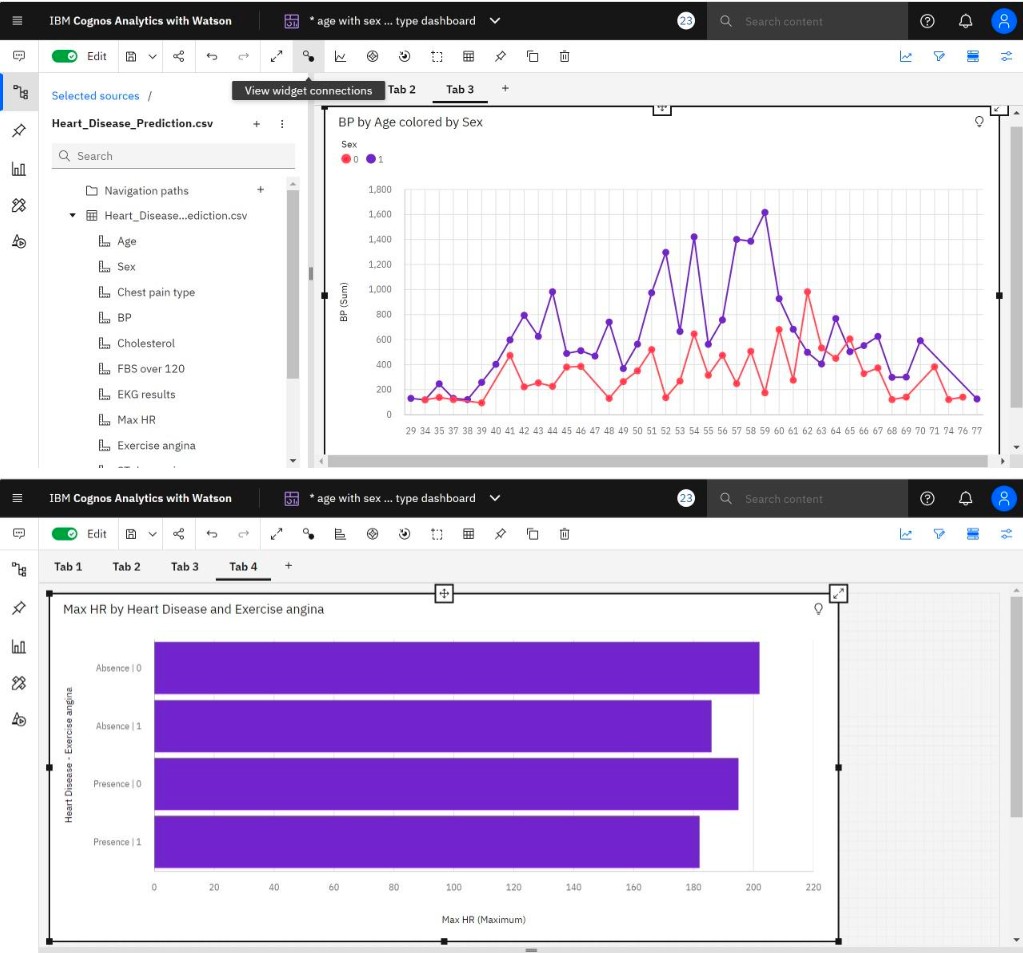


# CREATING THE DASHBOARD AND EXPORT THE ANALYTICS

* + 1. Creating the Dashboard







# ADVANTAGES & DISADVANTAGES

* + 1. Advantages:

1. Increased accuracy for effective heart disease diagnosis.
2. Handles roughest(enormous) amount of data using random forest algorithm and feature selection.
3. Reduce the time complexity of doctors.
4. Cost effective for patients.
   * 1. Advantages:
5. Prediction of cardiovascular disease results is not accurate.
6. Data mining techniques does not help to provide effective decision making.
7. Cannot handle enormous datasets for patient records.

# 10.CONCLUSION

With the increasing number of deaths due to heart diseases, it has become mandatory to develop a system to predict heart diseases effectively and accurately. In future the work can be enhanced by developing a web application based on the Random Forest algorithm as well as using a larger dataset as compared to the one used in this analysis which will help to provide better results and help health professionals in predicting the heart disease effectively and efficiently.