Care



*Analytics*

*for*



*Hospitals'*

*Health*

*Data*

IBM – DOCUMENTATION

UNDER THE GUIDANCE OF

Industry Mentor(s) Name Shivam Shivare, Hari Prabu Faculty Mentor(s) Name Dr. S. RAMESH

TEAM ID : PNT2022TMID38936

SUBMITTED BY:

NAMES and reg.no



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

KRISHNASAMY COLLEGE OF ENGINEERING AND TECHNOLOGY 2019-2023

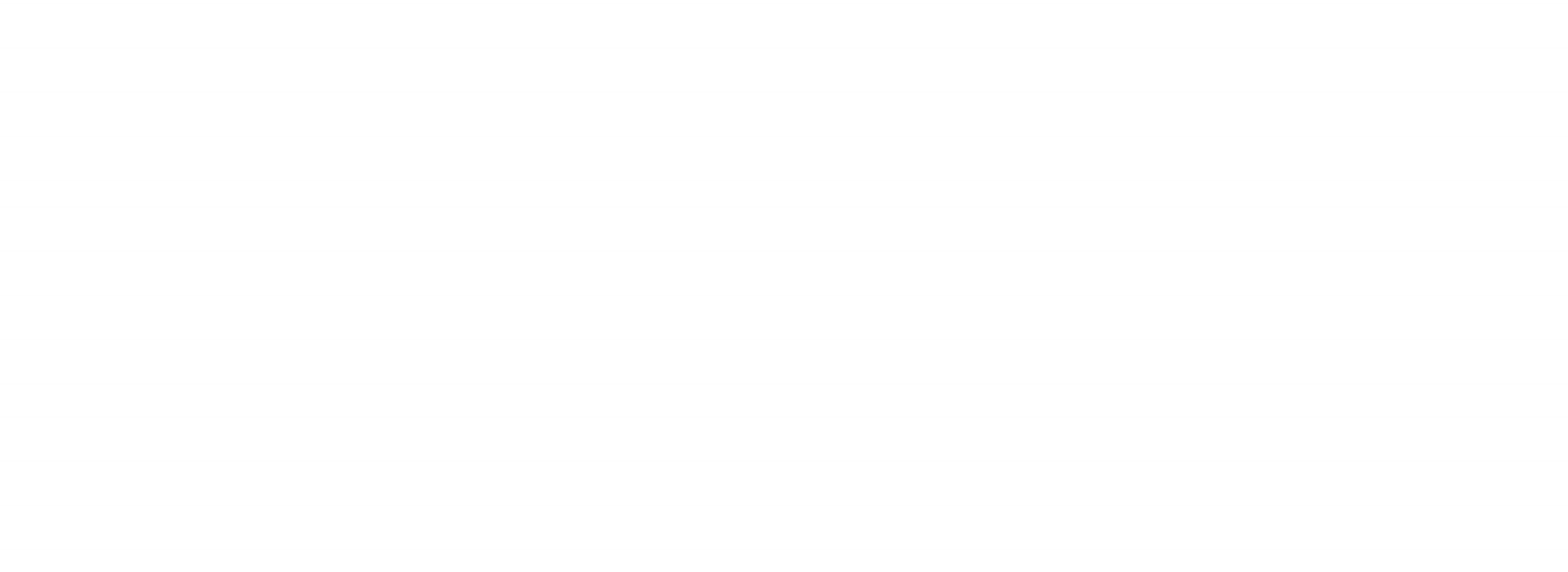
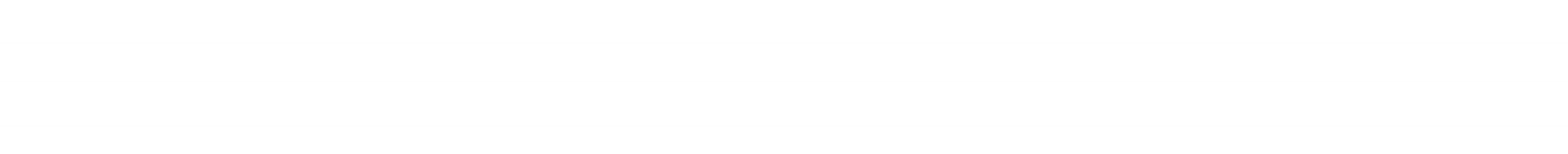
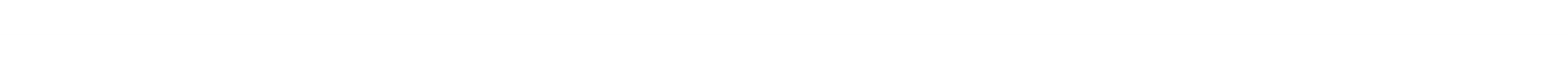
ANNA UNIVERSITY :: 2019-2023



|  |  |  |
| --- | --- | --- |
| *S.NO* | *Table of Content* | *Pg.no* |
| *1* | *Introduction* |  |
| *1.1* | *Project Overview* |  |
| *1.2* | *Purpose* |  |
| *2* | *Literature Survey* |  |
| *2.1* | *Existing Problem* |  |
| *2.2* | *References* |  |
| *2.3* | *Problem Statement Definition* |  |
| *3* | *Ideation & Proposed Solution* |  |
| *3.1* | *Empathy Map Canvas* |  |
| *3.2* | *Ideation & Brainstorming* |  |
| *3.3* | *Proposed Solution* |  |
| *3.4* | *Problem Solution Fit* |  |
| *4* | *Requirement Analysis* |  |
| *4.1* | *Functional Requirement* |  |
| *4.2* | *Non-Functional Requirement* |  |
| *5* | *Project Design* |  |
| *5.1* | *Data Flow Diagram* |  |
| *5.2* | *Solution & Technical Architecture* |  |
| *5.3* | *User Stories* |  |
| *6* | *Project Planning & Scheduling* |  |
| *6.1* | *Sprint Planning & Estimation* |  |

|  |  |  |
| --- | --- | --- |
| *6.2* | *Sprint Delivery Schedule* |  |
| *7* | *Testing* |  |
| *7.1* | *Test Cases* |  |
| *7.2* | *User Acceptance Testing* |  |
| *8* | *Results* |  |
| *8.1* | *Performance Metrics* |  |
| *9* | *Advantages & Disadvantages* |  |
| *10* | *Conclusion* |  |
| *11* | *Future Scope* |  |
| *12* | *Appendix* |  |
| *12.1* | *GitHub & Project Demo Link* |  |





*1. INTRODUCTION*

*1.1 PROJECT OVERVIEW*

*Recent Covid-19 Pandemic has raised alarms over one of the most overlooked areas to focus: Healthcare*

*Management. While healthcare management has various use cases for using data science,*

*patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiency of the healthcare management in a hospital.*

*This parameter helps hospitals to identify patients of high LOS-risk (patients who will stay*



*longer) at the time of admission. Once identified, patients with high LOS risk can have their*

*treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection. Also, prior knowledge of LOS can aid in logistics such as room and bed allocation planning.*

*Suppose you have been hired as Data Scientist of Health Man – a not for profit organization*

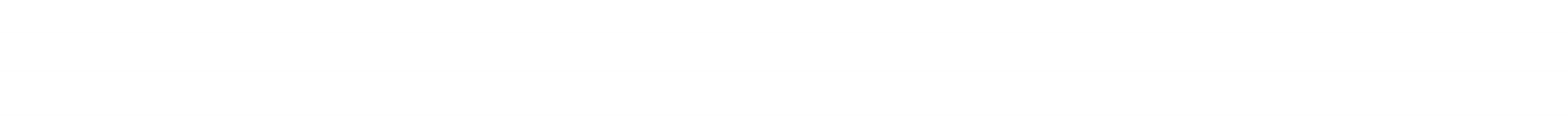
*dedicated to manage the functioning of Hospitals in a professional and optimal manner.*

*1.2 PURPOSE*

The purpose is to accurately predict using the given dataset, we plan to create various graphs and charts to highlight the insights and visualizations.

1. *Length of stay for each Case of Patients.*
2. *Stay by patient ID using Column Chart.*
3. *Severity of illness by Patient-ID using Tree Map.*
4. *Age, Department wise patient using Table.*
5. *Room availability by Pie Chart.*
6. *Dashboard Creation.*
7. *Department wise no. of Admissions by Waterfall Chart.*

2. LITERATURE SURVEY



*2.1 EXISTING PROBLEM*

*The most and common problems faced by every hospital organization is collecting the right data , cleaning the data, data storage , securing patients data, querying the data. The most existing problem is to update real time data.*



* 1. *REFERENCES*

Development of the Health Information Analytics Dashboard Using Big Data Analytics

The development of digital technology has the impact on healthcare facilities in Indonesia, one of which is the digitization of medical records. This will generate abundant clinical data from various sources including electronic medical records. Therefore, a large infrastructure is needed to store data from various sources that can facilitate the process of data aggregation to then be processed into information. Health Information Analytics Dashboard is the solution to get accurate, complete, and real-time insight from big data in healthcare. Data collection is carried out from various sources of health service facilities in Indonesia that are integrated into the system. With a user-friendly display, the analytic dashboard can be used to create monitoring reports with just one click. The method of this study uses big data analytics. The data analysis results are visualized through display charts/graphs that make it easier for users to understand the data analysis results and interpretation. This dashboard is useful to facilitate decision making so that stakeholders can find out more quickly to be able to respond appropriately and also improve the quality of health services so as to improve the degree of public health.

Reference link:

https://[www.researchgate.net/publication/348834045\_Development\_of\_the\_Health\_Informatio](http://www.researchgate.net/publication/348834045_Development_of_the_Health_Informatio) n\_Analytics\_Dashboard\_Using\_Big\_Data\_Analytics

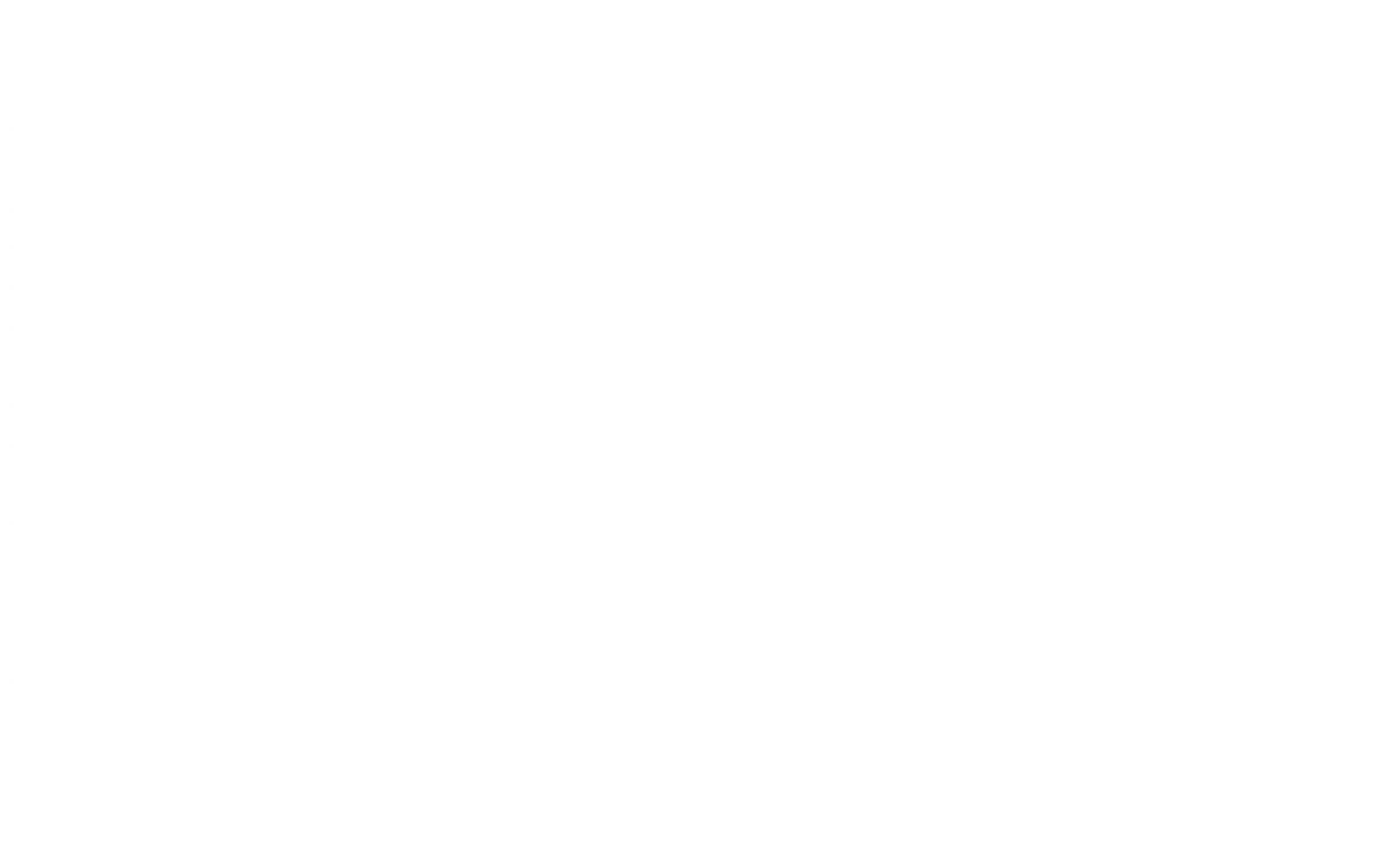
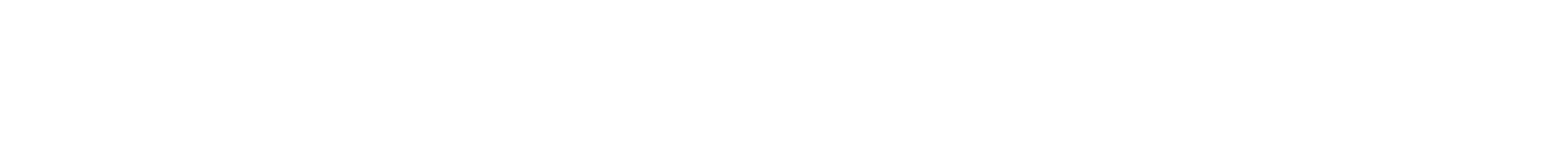
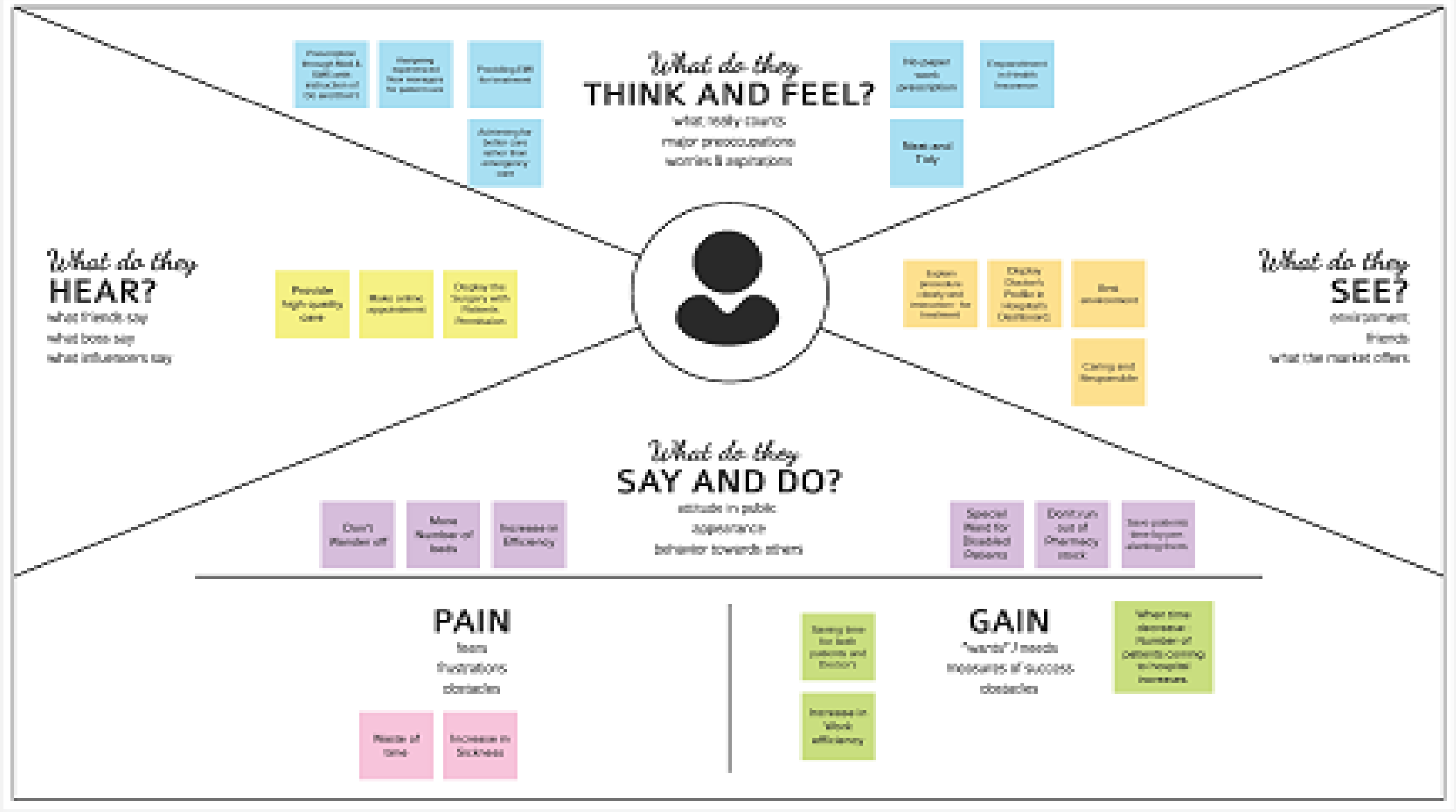
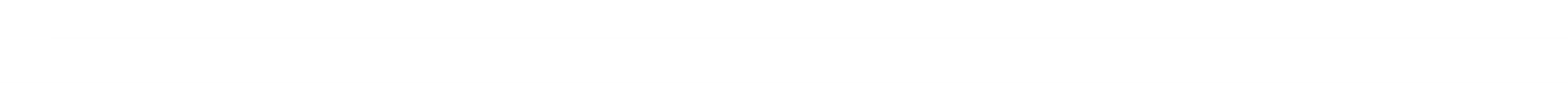
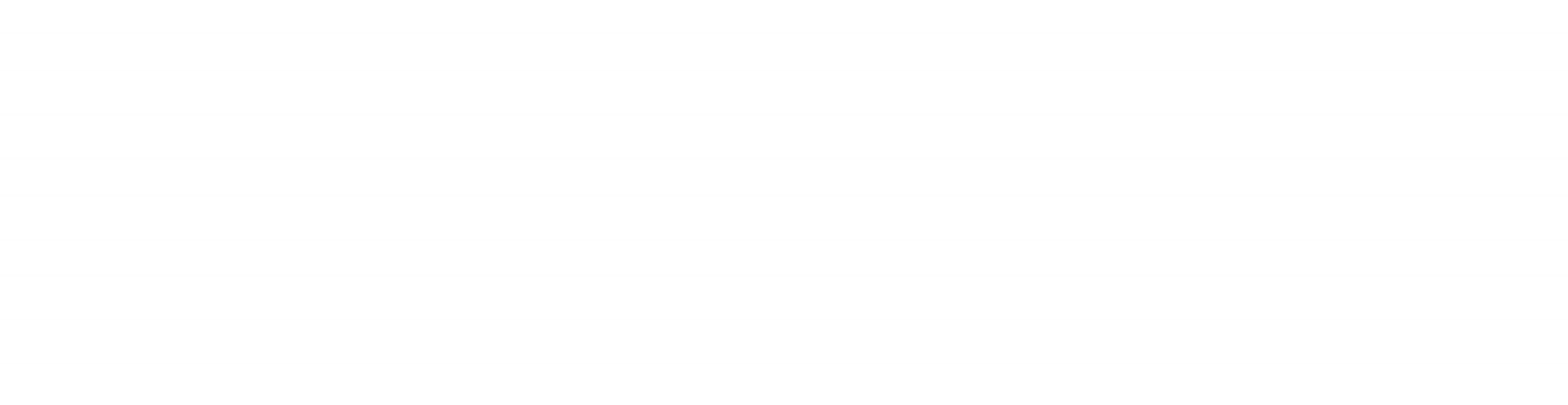
* 1. *Problem Statement Definition*
     1. *A Patient is a Customer who needs a better treatment because they don’ t need to*

suffer.

* + 1. *A Patient is a Customer who needs doorstep treatment because it’ s helpful for the*

patients, who are aged and can’ t able to travel.

* + 1. *A User needs a way to do something that addressed to the respective patients so that the patient benefits directly.*
    2. *A Patient needs a way to analyze and choose Doctor by their profile because of their experience and comfortable.*

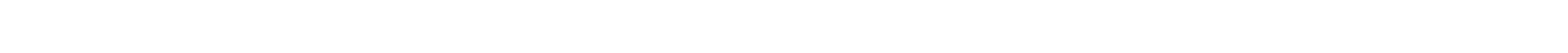
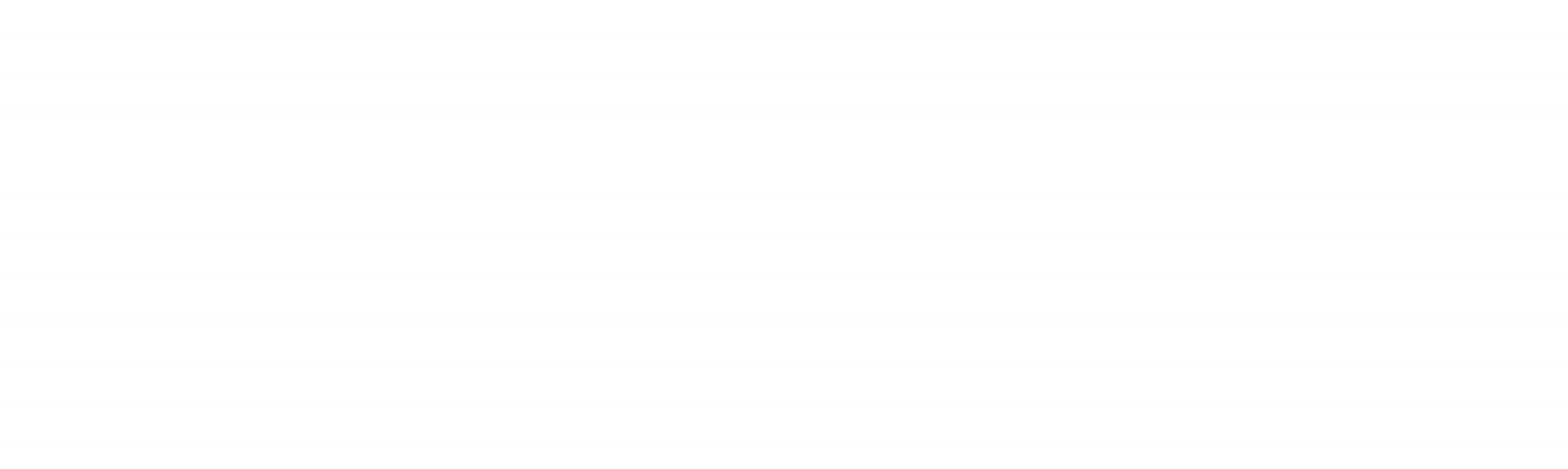


3. IDEATION & PROPOSED SOLUTION

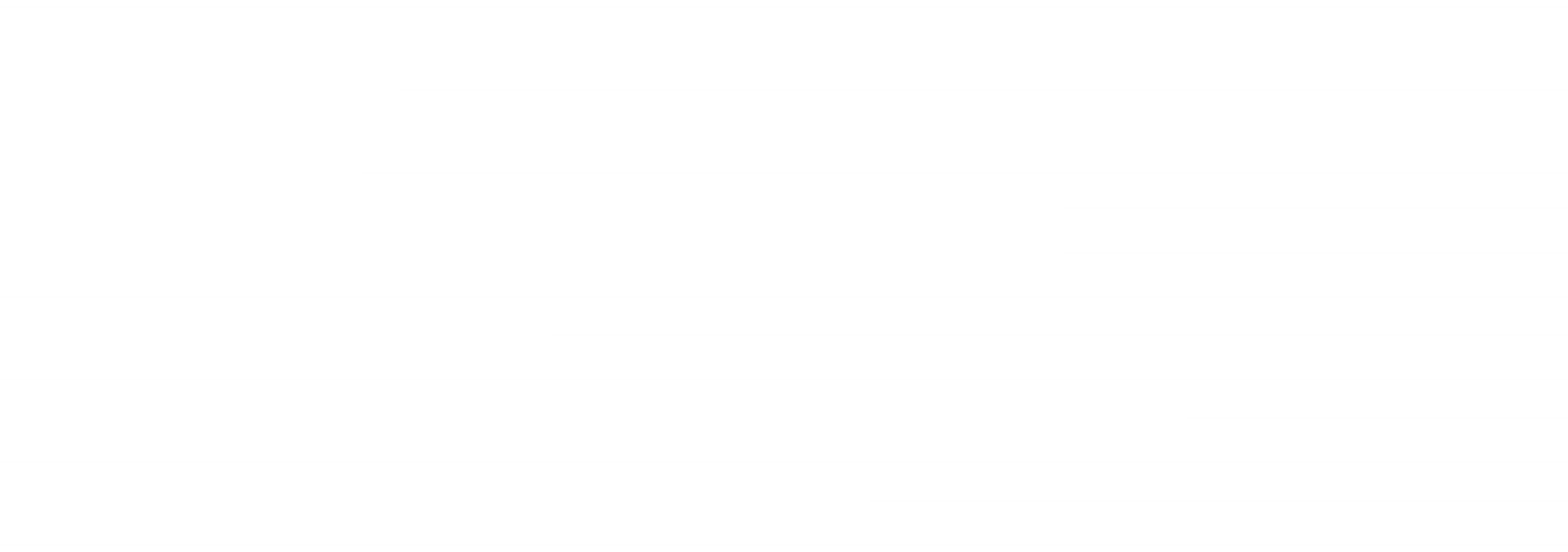
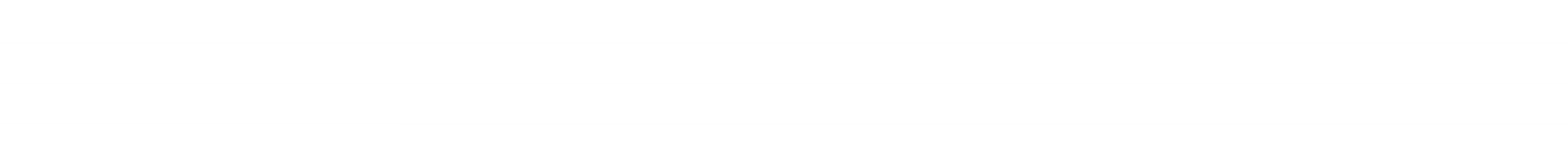
* 1. *Empathy Map Canvas*

3.2.Ideation & Brainstorming





*3.3 Proposed Solution*



*3.4 Problem Solution Fit*

*1. CUSTOMER SEGMENT*

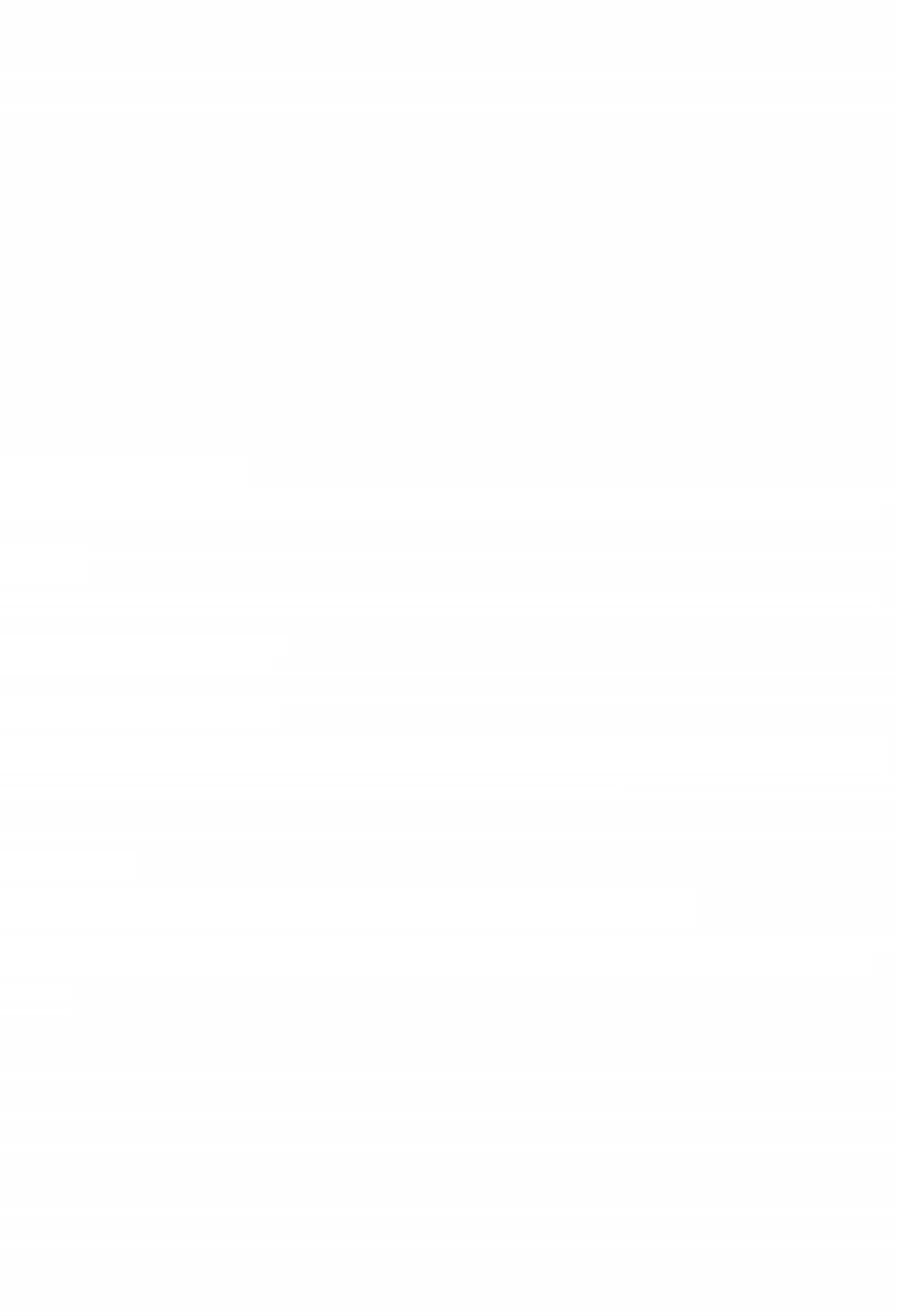
*Who is your customer?*

*A people who was infected or need medical help is our customer.*

*2. JOBS-TO-BE-DONE/PROBLEMS*

*Which jobs-to-be-done (or problems) do you address for your customers? Open to all, to book their appointment using website.*

|  |  |  |
| --- | --- | --- |
| *S. No.* | *Parameter* | *Description* |
| *1.* | *Problem Statement (Problem to besolved)* | *A Patient needs a way to analyze and*  *choose Doctor by their profile because of their experience and comfortable.* |
| *2.* | *Idea / Solution description* | *Instead of assigning Doctor to the patient by the hospital, Patient tend to select the doctor by themselves with*  *experience of Doctor.* |
| *3.* | *Novelty / Uniqueness* | *Instead of displaying Doctor's name anddesignation in their cabin, The Uniqueness is to*  *display Doctor's profile in the*  *Hospital'sDashboard .* |
| *4.* | *Social Impact / Customer Satisfaction* | *Patient can easily analyze and select the respective doctor by their point-of-view.* |
| *5.* | *Business Model (Revenue Model)* | *Here, Salary for 10 doctors reduced to the ratioof 10:7.* |
| *6.* | *Scalability of the Solution* | *The Proposed solution may be inflate by scheduling doctor by Hospital behest for*  *the patient's demand.* |

1. *TRIGGERS*

What triggers customers to act? Fear of fitness.

1. *EMOTIONS: BEFORE / AFTER*

How do customers feel when they face a problem or a job and afterwards?

Before: Patients feel hereafter they don’ t need to continue, lose faith, Feel insecure. After: Patients were satisfied by selecting their desired doctor and got well by quality treatment.

1. *AVAILABLE SOLUTIONS*

Which solutions are available to the customers when they face the problem or need to get the job done?

Pre-appointed schedule for both patient and doctor, Appointing their comfortable doctor with their profile and fee structure.

1. *CUSTOMER CONSTRAINTS*

What constraints prevent your customers from taking action or limit their choices of solutions? Spending time in waiting, fear of bill, availability of doctors or not.

1. *BEHAVIOUR*

What does your customer do to address the problem and get the job done?

Updating them that, Your appointment has x-time left, because the remainder through mails or texting.

1. *CHANNELS of BEHAVIOUR*
   1. *ONLINE*

What kind of actions do customers take online? Extract online channels from #7

Instead of interacting with a doctor in physical mode, Patient can consult the doctor in online mode in an emergency situation.

* 1. *OFFLINE*

What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.



In Spite of consulting the doctor in online mode, a patient can interact with the doctor would be more adequate.

1. *PROBLEM ROOT CAUSE*

What is the real reason that this problem exists? What is the back story behind the need to do this job?

Not having friendly conversation and care with patient. Waiting time for doctor gets patient antagonize.

Reason:

Receptionist may handle up to limited number of patients in a day but a machine don't have that limitation. So patient can register their appointment through the respective website.

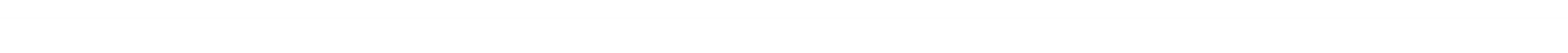
1. *YOUR SOLUTION*

A Patient needs a way to analyze and choose a Doctor by their profile because of their experience and comfort.

4. REQUIREMENT ANALYSIS



*4.1Functional Requirements:*

Following are the functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| *FR No.* | *Functional Requirement*  *(Epic)* | *Sub Requirement (Story / Sub-Task)* |
| *FR-1* | *User Registration* | *Registration through Form*  *Registration through website* |
| *FR-2* | *User Confirmation* | *Confirmation via Email*  *Confirmation via Message* |
| *FR-3* | *Data Cleaning* | *We clean the data because there are many potentialfor data to be duplicated or incorrectly labeled*  *when merging multiple data sources* |
| *FR-4* | *Reliability* | *Users may utilize this dashboard in an*  *effective, efficient, and reliable manner since it is consistent* |

|  |  |  |
| --- | --- | --- |
|  |  | *and reliable for them.* |
| *FR-5* | *Accuracy* | *Dashboard accurately predicts the patient's health*  *risks based on the length of their stay.* |



4.2



*Non-functional Requirements:*

Following are the non-functional requirements of the proposed solution.

|  |  |  |
| --- | --- | --- |
| *FR*  *No.* | *Non-Functional Requirement* | *Description* |
| *NFR-1* | *Usability* | *In order to provide a clear understanding of the patient's Length Of Stay, this dashboard makesuse of data visualization*  *techniques including* |

|  |  |  |
| --- | --- | --- |
|  |  | *charts and graphs.* |
| *NFR-2* | *Security* | *Only users who have the password can accessthe website. High degrees of security are provided through the use of encryption*  *techniques to secure the database.* |
| *NFR-3* | *Reliability* | *Users will find this dashboard to be constant anddependable, assisting them in using it*  *effectively, efficiently, and dependably.* |
| *NFR-4* | *Performance* | *The project must respond quickly to the user's actions or even if the user has to wait the waiting*  *period must be short.* |
| *NFR-5* | *Availability* | *The project is independent of platforms. Onpractically every platform, it functions*  *flawlessly.* |
| *NFR-6* | *Scalability* | *The project enables concurrent usage of the databy several people. Because adding features and improving the website is simple, it is very*  *scalable.* |



*Add user’ s account*

*Health Care System*

*(Dashboard)*

*Monitor Records*

*Add Patient Details*

*Data*

*analyzing*

*IBM*

*Cognit ive*

*Add Doctors*

*Add Bed details*

*User*

*Admin*

*Storing Data’*

*s/Record’*

*s*



*5. PROJECT DESIGN*

5.1



*Data Flow Diagrams*

A Data Flow Diagram (DFD) is a traditional visual representation of the information

flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leavesthe system, what changes the information, and where data is stored.

*Visualize*

* + 1. *Patients Based on age 3. Number of Beds*
    2. *Number of Patients*
  1. *TECHNICAL ARCHITECTURE*

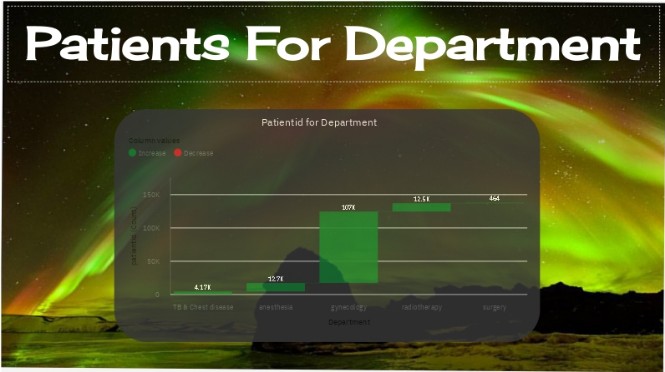


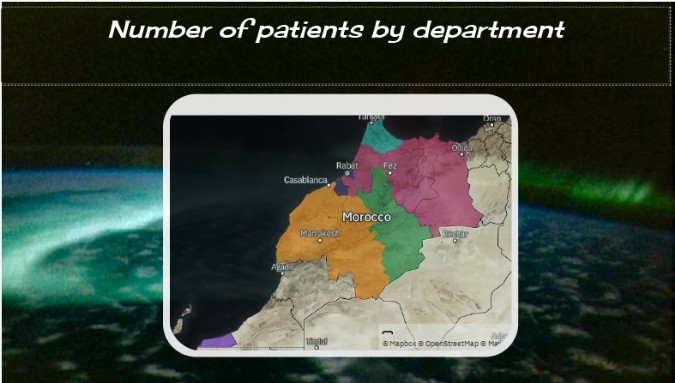
*SOLUTION &*



* 1. *USER STORIES*





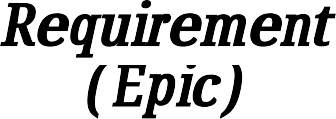


1. *PROJECT PLANNING & SCHEDULING*
   1. *SPRINT PLANNING & ESTIMATION*

|  |  |
| --- | --- |
| *MILESTONES* | *TASKS* |
| *MILESTONE ~~-~~ 1* | *Data Collecting process (Datasets)* |
| *MILESTONE ~~-~~ 2* | *Required Datasets are uploaded on the IBMCognitive Platform.* |
| *MILESTONE ~~-~~ 3* | *Data Exploration and Data Visualization* |
| *MILESTONE ~~-~~ 4* | *To Create a Interactive Dashboard.* |
| *MILESTONE ~~-~~ 5* | *Display the Insights in the Dashboard* |
| *MILESTONE ~~-~~ 6* | *Construct a Standardized Data Set and use theneeded data with the Assistance of a Python Program* |
| *MILESTONE ~~-~~ 7* | *Use of different algorithm with Google Colab toachieve the desired result with more*  *accuracy.* |

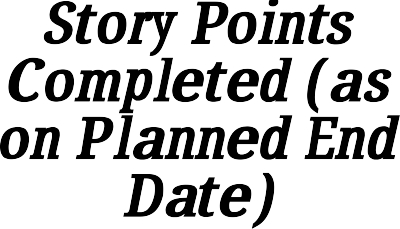
|  |  |
| --- | --- |
| *MILESTONE ~~-~~ 8* | *Making the output simpler to understand andmore efficient.* |
| *MILESTONE ~~-~~ 9* | *Deployed in the Github* |



* 1. *SPRINT DELIEVERY SCHEDULE*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Sprint* | *Functional*  *Requirement (Epic)* | *User*  *Story Number* | *User Story / Task* | *Story* | *Priority* | *Team Members* |
| *Sprint-1* | *Register* | *USN-1* | *Data Collecting process (Datasets) Required Datasets are uploaded on the IBMCognitive*  *Platform* | *10* | *Medium* | *Vijayendren S.L* |
| *Sprint-1* | *Login* | *USN-1* | *Data Exploration and Data Visualization* | *20* | *High* | *Mohamed Zakir Husain. N* |
| *Sprint-* | *Dashboard* | *USN-2* | *To Create a* | *10* | *High* | *Ummul Nasiha. U* |
| *2* |  |  | *Interactive* |  |  |  |
|  |  |  | *Dashboard.* |  |  |  |
|  |  |  | *Display the* |  |  |  |
|  |  |  | *Insights in the* |  |  |  |
|  |  |  | *Dashboard* |  |  |  |
| *Sprint-*  *2* | *Dashboard* | *USN-3* | *Construct a Standardized Data Set and use* | *10* | *High* | *Sumaiya Farveen. N* |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | *theneeded data with the Assistance of a*  *Python Program* |  |  |  |
| *Sprint-3* | *Dashboard* | *USN-4* | *Use of different* | *20* | *High* | *Mohamed Zakir Husain.N* |
|  |  |  | *algorithm with* |  |  |
|  |  |  | *Google Colab to* |  |  |
|  |  |  | *achieve the* |  |  |
|  |  |  | *desired result with*  *more accuracy.* |  |  |
| *Sprint-4* | *Virtualizes* | *USN-5* | *Making the output simpler to understand and more efficient.*  *Deployed in the Github.* | *20* | *High* | *Ummul Nasiha.U* |

Velocity & Burn Chart:



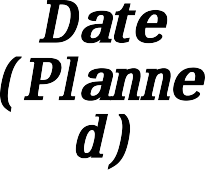
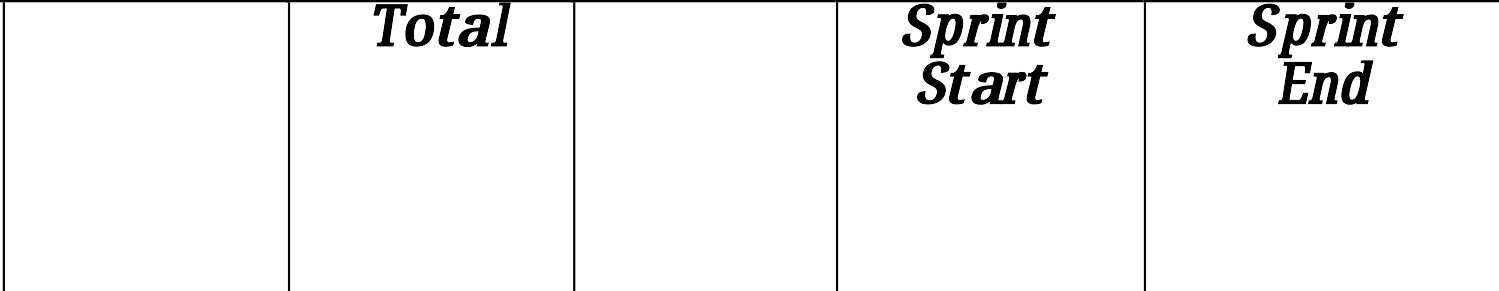
*Project Tracker,*

*down*



*Sprint Release Date (Actual)*

Sprint Total Story



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | | | | | *Story Points Completed (as on Planned End Date)* |  |
| *Sprint-1* | *20* | *6 Days* | *24 Oct*  *2022* | *29 Oct 2022* | *20* | *29 Oct 2022* |
| *Sprint-2* | *20* | *6 Days* | *31 Oct*  *2022* | *05 Nov 2022* | *20* | *05 Nov 2022* |
| *Sprint-3* | *20* | *6 Days* | *07 Nov*  *2022* | *12 Nov 2022* | *20* | *12 Nov 2022* |
| *Sprint-4* | *20* | *6 Days* | *14 Nov*  *2022* | *19 Nov 2022* | *20* | *19 Nov 2022* |

Duration Sprint

Start

Date

Sprint End Date (Planne d)



*Velocity:*

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint).Let’ s calculate the team’ s average velocity (AV) per iteration unit (story points per day)

AV = SPRINT DURATION = 20 = 10

VELOCITY 2

Burn Chart:

*down*

A burn down chart is a graphical representation of work left to do versus time. It is often used inagile software development methodologies such as Scrum. However, burn down charts can beapplied to any project containing measurable progress over time.



Percentage of work

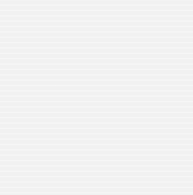
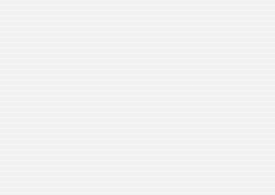
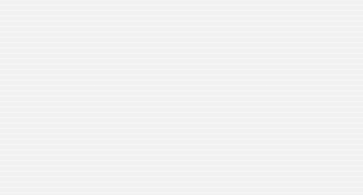
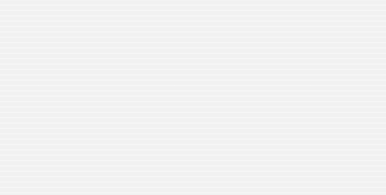
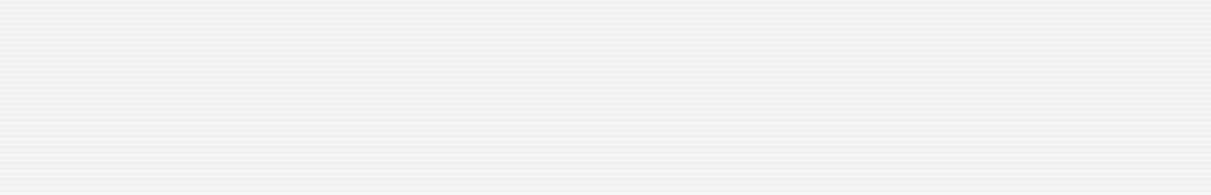


*1*

*1*

1. *TESTING*
   1. *Test Cases*

This report shows the number of test cases that have passed, failed, and untested

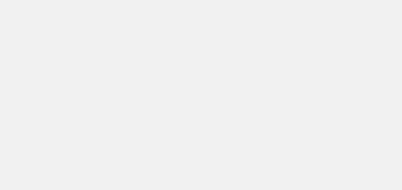
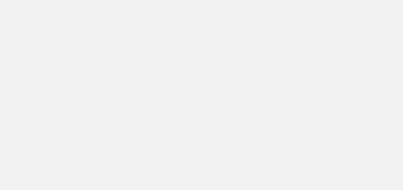
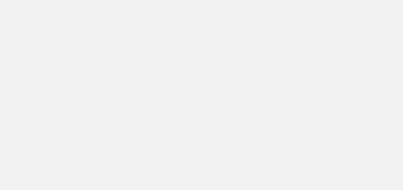
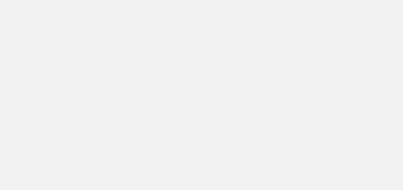
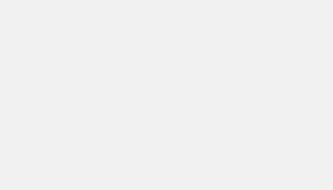
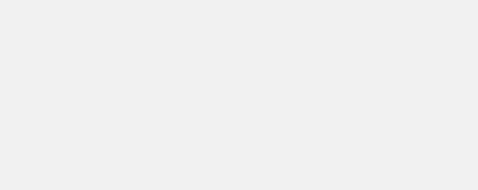


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Section* | *Total Cases* | *Not Tested* | *Fail* | *Pass* |
| *Print Engine* | *0* | *0* | *0* | *0* |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Client Application* | *0* | *0* | *0* | *0* |
| *Security* | *0* | *0* | *0* | *0* |
| *Outsource Shipping* | *0* | *0* | *0* | *0* |
| *Exception Reporting* | *0* | *0* | *0* | *0* |
| *Final Report Output* | *0* | *0* | *0* | *0* |
| *Version Control* | *0* | *0* | *0* | *0* |

* 1. *USER ACCEPTANCE TESTING*

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved.



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Resolution* | *Severity 1* | *Severity 2* | *Severity 3* | *Severity 4* | *Subtotal* |
| *By Design* | *5* | *5* | *3* | *0* | *13* |
| *Duplicate* | *0* | *0* | *0* | *0* | *0* |
| *External* | *7* | *5* | *1* | *0* | *13* |
| *Fixed* | *11* | *8* | *7* | *5* | *31* |
| *Not Reproduced* | *1* | *0* | *0* | *0* | *1* |
| *Skipped* | *0* | *0* | *0* | *0* | *0* |
| *Won't Fix* | *0* | *0* | *0* | *0* | *0* |
| *Totals* | *24* | *18* | *11* | *5* | *58* |

1. *RESULT*

*8.1 PERFORMANCE METRICS*

*Project team shall fill the following information in model performance testing.*

|  |  |  |
| --- | --- | --- |
| *S. No.* | *Parameter* | *Screenshot / Values* |
| *1.* | *Dashboard design* | *17 / 5* |
| *2.* | *Data Responsiveness* | *The final output from IBM Cognos With Watson further*  *converted into PDF or Story file ,So it can be viewed by alldevices.* |
| *3.* | *Amount Data to Rendered (DB2*  *Metrics)* | *0 KB.* |
| *4.* | *Utilization of DataFilters* | *The Utilization of data Filters like Ascending , Descending, Format and so on.* |
| *5.* | *Effective User Story* | *12* |
| *6.* | *Descriptive Reports* | *17 / 5* |

1. *ADVANTAGES &*



*DISADVANTAGES*



*ADVANTAGES:*

*Helps an organization to make a better decision Increase the efficiency of the work*

*The analytics keeps you updated of your customer behavioral changes. Personalization of hospital details.*

*Improving quality of service and health care.*



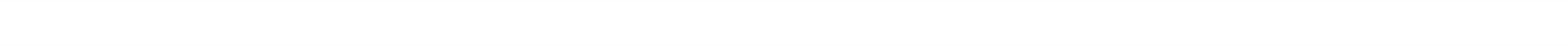
*DISADVANTAGES:*

*Lack of alignment within teams Lack of commitment and patience Low quality of data*

*Privacy Concerns*

*Complexity and Bias*

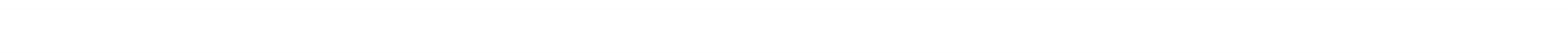
1. *CONCLUSION*



*Data analytics in health care is vital. It helps health care organizations to evaluate*



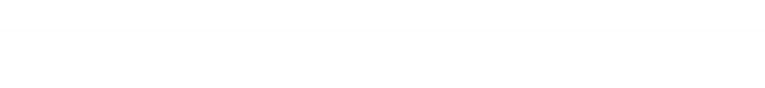
*and develop Number of patients by ward, Age wise patients with department details,*



*Various types of visualizations to analyze the hospital's datasets and hence predict*

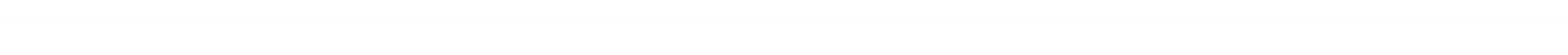


*outbreaks in illness, Data analytics can also lower costs for health care organizations and*

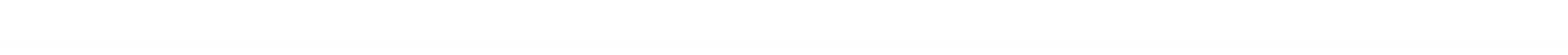


*boost business intelligence.*

1. *FUTURE SCOPE*



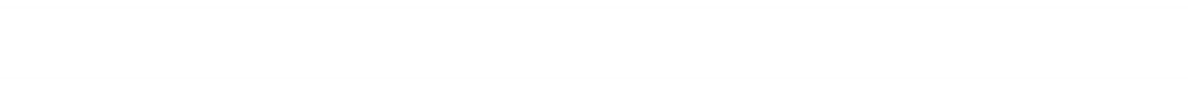
*While every fact of the industry stands to be changed by data analytics in*



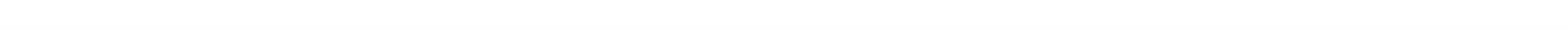
*healthcare, data has significantly improved healthcare in three areas: conducting medical*



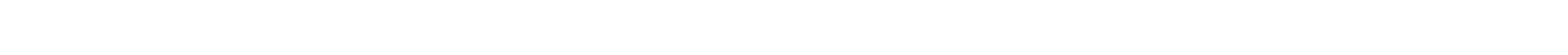
*studies, understanding the cost of medical tests and health insurance, and making*



*preventative recommendations to patients.*



*Hospital Healthcare data analytics helps in analyzing the patient details via*



*hospital that the availability of doctors and number of beds to the patients and hence, it*



*reduce the man power and time of the respective Hospital.*

1. *APPENDIX*

*1.GitHub & Demo*



*Project*

*Link*

*Our GitHub Repository Direct Link*



*https://github.com/IBM-EPBL/IBM-Project-41297-1660640957*

*Project Demonstration Video Direct Link*



*https://*[*www.youtube.com/watch?v=PQs77KJlMA8*](http://www.youtube.com/watch?v=PQs77KJlMA8)

*Webpage to view analyzed visualizations*



*https://mohamedzakirn02.wixsite.com/ibmproject*