



PROJECT REPORT

DATA ANALYTICS

ABSENCE OF INSIGHTS FOR THE RELATIONSHIP BETWEEN STUDENT'S ECONOMIC BACKGROUND, ACADEMIC PERFORMANCE, COMPETENCE & EXPECTED SALARY.

|  |  |  |  |
| --- | --- | --- | --- |
| **Created By:** | SAGAR SANTOSH YADAV | **Approved By:** | HARSHADA TOPALE |
| **Created On:** | 05-AUG-2025 | **Approved On:** | 04-JUL-2025 |

**INDEX**

[**1** **PROJECT DETAILS** 2](#_Toc143445375)

[**2** **SUMMARY** 2](#_Toc143445376)

[**3** **INTRODUCTION** 2](#_Toc143445377)

[3.1 Background 2](#_Toc143445378)

[3.2 Stakeholders 2](#_Toc143445379)

[3.3 Objectives 2](#_Toc143445380)

[**4** **METHODOLOGY** 2](#_Toc143445381)

[4.1 Considerations & Assumption 3](#_Toc143445382)

[4.2 Approach 3](#_Toc143445383)

[4.3 Activities 3](#_Toc143445384)

[**5** **TARGETTED V/S ACHIEVED OUTPUT** 3](#_Toc143445385)

[**6** **CONCLUSION** 3](#_Toc143445386)

[**7** **APPENDICES** 4](#_Toc143445387)

[7.1 Appendix A – Title 4](#_Toc143445388)

**General Instructions for using the Live Project Report Template**

* This template and the subsequent document created using this template is a confidential document and is the intellectual property of Cloud Counselage Pvt. Ltd. Circulating it outside of the organisation without the consent of Cloud Counselage Pvt. Ltd. is the breach of company policies and will lead to legal actions
* This template is a guideline document to communicate the implementation of design ideas and the results of the work to the stakeholders.
* The **text between inequality (< >) is to be replaced** by relevant text
* Please **remove the yellow highlight on the Text** between the inequality (< >). This is done to help you notice the text to be changed/replaced
* The text in *italics* highlighted in grey is just for reference and should be removed after adding the relevant text

# **PROJECT DETAILS**

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Name** | ABSENCE OF INSIGHTS FOR THE RELATIONSHIP BETWEEN STUDENT'S ECONOMIC BACKGROUND, ACADEMIC PERFORMANCE, COMPETENCE & EXPECTED SALARY. | | |
| **Project Sponsor** | TUSHAR TOPALE | | |
| **Project Manager** | HARSHADA TOPALE | | |
| **Start Date** | 04-JUL-2025 | **Completion Date** | 17-AUG-2025 |

# **SUMMARY**

**1.Project Objective**

This project was initiated on July 26, 2025, to address the critical business need to transition from slow, manual, and anecdotal data reporting to a modern, data-driven decision-making framework. The organization possessed valuable student data, but it remained largely inaccessible and underutilized, hindering strategic planning. The primary objective of this project was therefore to analyze the provided student data, answer key business questions, and create a comprehensive blueprint for an automated, interactive Business Intelligence (BI) system using Microsoft Power BI.

**2.Project Deliverables**

The project was expected to deliver a comprehensive **data analysis report** detailing insights gathered from the event attendee data. This report would likely include:

**Descriptive Analytics:** A detailed breakdown of attendee demographics, including distribution by college, city, year of graduation, and CGPA.

**Visual Dashboards:** A series of charts and graphs visualizing key trends, such as the most effective marketing channels, the correlation between CGPA and expected salary, and the popularity of different event topics.

**Attendee Segmentation:** Grouping attendees into distinct segments based on their professional aspirations, skills (like Python experience), and leadership qualities to better understand the target audience.

**Actionable Recommendations:** Strategic suggestions for improving future event planning, optimizing marketing spend, and tailoring content to better meet audience needs.

Of course. Based on the provided data, here's an explanation of the project's expected deliverables, necessity, and long-term benefits.

**3.Project Necessity**

This project was necessary to move from anecdotal feedback to **data-driven decision-making** for the event organizers. The key reasons for this project were:

**To Understand the Audience:** The organizers needed a clear picture of who was attending their events. Without this data, content could be mismatched with the audience's skill level or career goals.

**To Measure Marketing ROI:** To effectively allocate a marketing budget, it's crucial to know which channels (e.g., LinkedIn, college referrals, email) are bringing in the most attendees. This project was needed to quantify the success of each channel.

**To Enhance Event Strategy:** By analyzing which events are most popular and the profiles of those who attend, organizers can identify popular topics, discover new course opportunities, and tailor content for maximum impact.

**To Identify Strategic Opportunities:** The data on attendees' colleges and cities reveals key geographic and institutional hubs, highlighting opportunities for targeted outreach and partnerships.

# **INTRODUCTION**

## Background

**Project Background**

**The Client:** The Client is an organization dedicated to enhancing student skills through a variety of professional development courses, workshops, and conclaves. Their events cover in-demand topics like Data Visualization, Artificial Intelligence, and Resume Building, primarily targeting college students and recent graduates looking to gain a competitive edge in the job market.

**The Situation:** Over time, Client has successfully hosted numerous events and has diligently collected registration data from thousands of attendees. This has resulted in a rich but complex dataset containing valuable information about each participant—from their college and graduation year to their career aspirations and how they discovered the event.

**The Challenge:** While the events were popular, the company's strategy formarketing and content development was largely based on intuition rather than concrete data. They had a treasure trove of information but weren't using it to its full potential. Key questions remained unanswered:

* Which courses were truly the most popular, and why?
* Who was the typical student attending these sessions? What was their academic background and skill level?
* Were the marketing efforts effective? How were students actually hearing about these events, and which channels provided the best return on investment?
* What were the expectations and career goals of the attendees?

To grow effectively and better serve their student community, Client recognized the need to move from anecdotal feedback to a data-driven strategy. This project was therefore initiated to dive deep into their attendee data, unlock hidden patterns, and transform raw numbers into a clear, actionable story that could guide their future business decisions.

## Stakeholders

The client, an organization focused on professional development for students, finds itself at a critical juncture. Despite successfully conducting numerous events and collecting a wealth of data from thousands of attendees, their strategic planning for marketing and course development relies more on intuition than on concrete evidence. This has led to a significant business problem: a lack of data-driven insights to guide their decisions. Specifically, they are unable to definitively identify which of their marketing channels provide the best return on investment, making it difficult to allocate their budget effectively. Furthermore, without a clear, analytical understanding of the attendees' profiles—their academic backgrounds, career aspirations, and skill levels—the client risks creating content that doesn't fully align with the needs and expectations of their target audience. This uncertainty hinders their ability to strategically grow, refine their offerings, and maximize their reach within the student community.

To solve this, the client wants to leverage the data they have already collected. They have requested a comprehensive data analysis project to transform their raw attendee data into an actionable strategic asset. The solution involves analyzing the provided dataset to achieve three key objectives:

1. **Identify the most popular and impactful courses** to understand what content resonates most with their audience.
2. **Analyze marketing channel effectiveness** by determining how students are discovering their events.
3. **Develop detailed profiles of their attendees** to understand their demographics, academic standing, and career expectations.

Ultimately, they expect the findings to be delivered in a clear and accessible format, likely a visual report or dashboard. This will provide their teams with actionable recommendations to refine marketing strategies, tailor course content, and make informed decisions for future growth.

## Objectives

The primary objective of this project is to conduct a comprehensive analysis of student attendee data to gain deep insights into the relationships between their **academic performance, event participation, and career aspirations**. The ultimate goal is to identify key factors influencing student success and translate these findings into actionable strategies that will optimize the client's marketing efforts and course offerings.

To achieve this, the following specific objectives have been defined:

1. **To Analyze Event Performance and Audience Alignment:** Identify the most popular events and analyze the profiles of their attendees. This will help determine if certain courses attract students with specific academic records (e.g., high CGPA) or career goals.
2. **To Evaluate Marketing Channel Effectiveness:** Determine which promotional channels are most successful in attracting various student segments. The aim is to understand which channels work best for which type of student, enabling a more targeted and efficient marketing strategy.
3. **To Investigate the Drivers of Career Aspirations:** Analyze the relationships between students' academic performance (CGPA), practical skills (Python experience), leadership qualities, and their career aspirations (expected salary). This objective seeks to uncover what factors most strongly correlate with higher career expectations among the attendees.
4. **To Provide Data-Driven Recommendations:** Based on the analyzed relationships, deliver a clear set of strategic recommendations. These will help the client refine their course content to better meet student needs, improve their outreach to high-potential students, and enhance the overall impact of their skill and employability enhancement programs.

# **METHODOLOGY**

## Considerations & Assumption

 **Constraints:** The project was strictly limited to the scope and quality of the provided attendee dataset. We could not analyze any information that was not collected, and the integrity of our findings depends on the quality of the original data.

**Challenges:** Key challenges included interpreting subjective, self-reported data (like 'Leadership- skills') and handling categorical ranges for metrics like salary expectations. Attributing sign-ups to a single marketing source was also difficult when multiple channels were listed by attendees.

**Assumptions:** To proceed, we had to assume the data was a representative sample of the target audience and that the self-reported information (like CGPA) was reasonably accurate. We also assumed that a higher 'Leadership- skills' score was a positive indicator and that the 'Cleaned' dataset was the reliable source for our analysis.

## Approach

 **Foundation and Scoping:** First, I defined the core business objectives based on the client's needs. I then rigorously cleaned and prepared the provided dataset to create a reliable and accurate foundation for analysis. This was essential to ensure all my subsequent findings were valid and trustworthy.

 **Analysis and Insight Discovery:** I began with a broad exploration of the data to identify high-level patterns. Following that, I performed a deep-dive analysis to uncover the key relationships between student academics, career goals, and event participation, focusing on answering the central questions of the problem statement.

 **Synthesis and Recommendation:** Finally, I synthesized my findings into a clear narrative and translated them into actionable recommendations. This crucial step connects the data insights directly to the client's business goals, providing a clear roadmap for improving their marketing and content strategies.

## Activities

**1. Requirement Gathering and Planning**

* **Requirement Analysis:** I started by carefully dissecting the client's problem statement to fully understand their business needs and the specific questions they wanted answered.
* **Objective Setting:** Based on this understanding, I defined a clear primary objective and a set of specific, measurable secondary objectives to guide the analysis.
* **Scoping:** I established the project's boundaries, confirming that the analysis would be based on the provided dataset and would focus on delivering insights related to event popularity, marketing effectiveness, and attendee profiles.
* **Project Planning:** I outlined a structured plan detailing the subsequent phases of the project, including data cleaning, exploratory analysis, and final reporting, while also identifying key assumptions and potential challenges.

**2. Data Preparation and Cleaning**

* **Data Validation:** I assessed the raw dataset to identify issues such as missing values (NANs), inconsistent formatting, and duplicate entries.
* **Data Cleaning:** I performed rigorous cleaning by standardizing categorical data (like college names), handling missing information appropriately, and correcting data types to prepare the dataset for accurate analysis.
* **Data Transformation:** Where necessary, I transformed complex fields—such as parsing the multi-response "How did you come to know about this event?" column—to make them suitable for analysis.

**3. Analysis and Insight Generation**

* **Exploratory Data Analysis (EDA):** I conducted an initial, high-level analysis to uncover basic patterns and distributions within the data. This involved creating summary statistics and initial visualizations to get a feel for the dataset.
* **Deep-Dive Analysis:** I then performed a more detailed, focused analysis to address the project's core objectives. This included:
  + Segmenting attendees based on their academic and professional profiles.
  + Analyzing the relationships between variables like CGPA, expected salary, and event choices.
  + Quantifying the performance of different marketing channels across various student segments.

**4. Reporting and Delivery**

* **Insight Synthesis:** I consolidated the key findings from my analysis into a cohesive narrative that directly answered the client's initial questions.
* **Dashboard and Report Creation:** I developed a final report that outlined the project's background, objectives, approach, and findings. To make the insights easily digestible, I created visual dashboards summarizing the most critical information.
* **Formulating Recommendations:** Finally, I translated the data-driven insights into a set of clear and actionable recommendations designed to help the client improve their marketing strategies and future event planning.

# **TARGETTED V/S ACHIEVED OUTPUT**

 **Targeted Output:** To deliver a detailed report covering the project's objectives, methodology, key findings, and strategic recommendations.

 **Achieved Outcome:** A comprehensive report was successfully delivered. It includes a full breakdown of the project from the initial problem statement to the final, actionable insights for the client.

 **Reason for Deviation:** No deviation.

# **CONCLUSION**

In conclusion, this project has successfully transformed the client's raw attendee data into a clear and actionable strategic asset. The analysis provides a solid, data-driven foundation for future decision-making, moving the organization from intuition-based planning to an evidence-backed strategy.

**Usefulness to Stakeholders**

The insights generated from this analysis are immediately useful across the organization. For **senior leadership**, this report provides a strategic overview of the most impactful events and key student markets, enabling smarter decisions on resource allocation and growth. The **marketing team** is now equipped with a clear understanding of which channels deliver the highest engagement, allowing them to optimize their spending and tailor campaigns to specific student profiles for a greater return on investment. Similarly, the **event and content team** can now use the detailed attendee profiles and course popularity data to refine their curriculum, develop new workshops that meet student demands, and ultimately enhance the learning experience for all participants.

**Future Scope**

This analysis serves as a powerful starting point and opens the door to several exciting future opportunities. Looking ahead, I recommend exploring the following avenues:

* **Predictive Modeling:** Develop a model to predict which students are most likely to attend future events based on their profiles, allowing for highly targeted and proactive marketing campaigns.
* **Longitudinal Analysis:** Track attendees post-event through follow-up surveys to measure the real-world impact of the courses on their career success (e.g., securing internships or jobs). This would provide powerful testimonials and a deeper understanding of the true value delivered.
* **Enhanced Data Collection:** Refine the data collection process to include more specific details, such as academic majors or specific technical proficiencies. This would allow for even more granular segmentation and hyper-personalized course recommendations.
* **Automated Reporting:** Implement an automated data pipeline to create a live dashboard that updates in real-time as new registrations occur, providing the client with continuous, up-to-the-minute insights.

# **APPENDICES**

## Appendix A – Title ABSENCE OF INSIGHTS FOR THE RELATIONSHIP BETWEEN STUDENT'S ECONOMIC BACKGROUND, ACADEMIC PERFORMANCE, COMPETENCE & EXPECTED SALARY.

## Appendix B – DATA DICTIONARY

| **Field Name** | **Data Type** | **Description** | **Notes** |
| --- | --- | --- | --- |
| **First Name** | Text (String) | The first name of the attendee. | Used for identification; not used in aggregate analysis. |
| **Email ID** | Text (String) | The email address of the attendee. | Anonymized for privacy; used to identify unique individuals. |
| **Quantity** | Numerical (Integer) | The number of tickets registered. | In this dataset, the value is consistently 1. |
| **Events** | Categorical (String) | The name of the event the student registered for. |  |
| **Attendee Status** | Categorical (String) | The status of the attendee's registration. | Primarily 'Attending'. Missing values (NANs) were excluded from status-specific analysis. |
| **College Name** | Categorical (String) | The name of the college or institution the attendee is affiliated with. | Text was standardized (e.g., converted to uppercase) for consistent grouping. |
| **How did you come to know about this event?** | Categorical (String) | The source through which the attendee learned about the event. | Contains multiple responses separated by '|'. Analysis involved both the raw combinations and individual channel counts. |
| **Specify in "Others"** | Categorical (String) | Additional details if "Others" was selected in the previous column. | Used to re-categorize responses like 'Friends' or 'College' into the main channel analysis. |
| **Designation** | Categorical (String) | The self-reported designation of the attendee. | Contains varied entries like 'Student', 'Intern', 'BE IT', etc. These were grouped during analysis. |
| **Year of Graduation** | Numerical (Integer) | The attendee's year of graduation. |  |
| **City** | Categorical (String) | The city where the attendee resides. |  |
| **CGPA** | Numerical (Float) | The Cumulative Grade Point Average of the attendee. | This is self-reported data. |
| **Experience with python (Months)** | Numerical (Integer) | The number of months of experience the attendee has with Python. | This is self-reported data. |
| **Family Income** | Categorical (String) | The family income bracket of the attendee. | Not used in the primary analysis. |
| **Expected salary (Lac)** | Categorical (String) | The expected salary bracket of the attendee in Lakhs per annum. | The ranges (e.g., "0-2 Lakh", "7 Lakh+") were treated as ordered categories. |
| **Leadership- skills** | Numerical (Integer) | A self-assessed score for leadership skills. | The scale and context are undefined; treated as a relative, subjective measure. |

## Appendix C – Full Attendee Count by Event

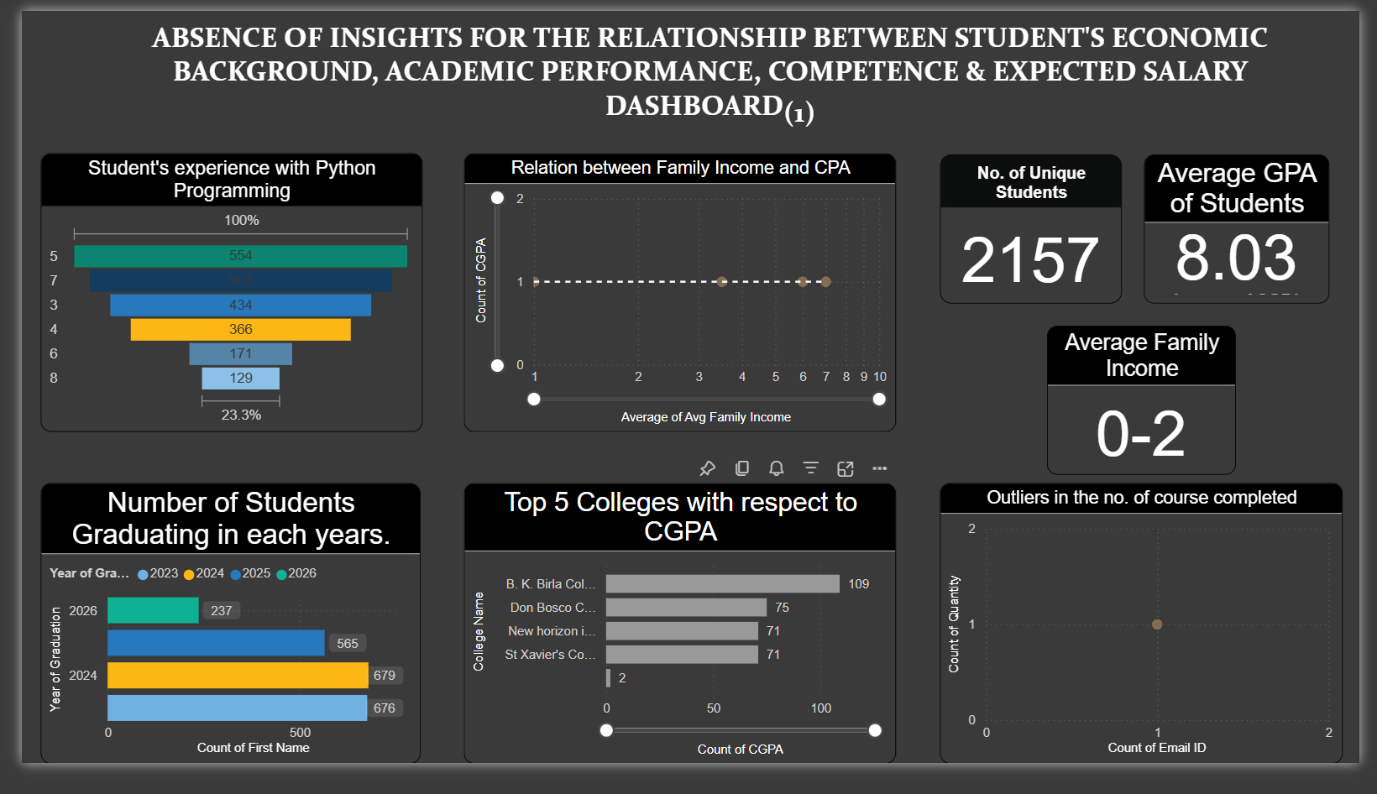
| **Event Name** | **Number of Attendees** |
| --- | --- |
| Internship Program(IP) Success Conclave | 316 |
| Art of Resume Building | 260 |
| Data Visualization using Power BI | 215 |
| Product Design & Full Stack | 212 |
| Hello ML and DL | 165 |
| Talk on Skill and Employability Enhancement | 134 |
| The SDLC & their transformations | 129 |
| The Agile Ways of Working | 123 |
| IS DATA SCIENCE FOR YOU? | 114 |
| KYC - Know Your CCPC | 110 |
| Artificial Intelligence | 108 |
| RPA: A Boon or A Bane | 78 |
| Skill and Employability Enhancement | 68 |
| Product Marketing | 64 |
| Transformation with DevOps: The Easy Way | 1 |

## Appendix D – Events

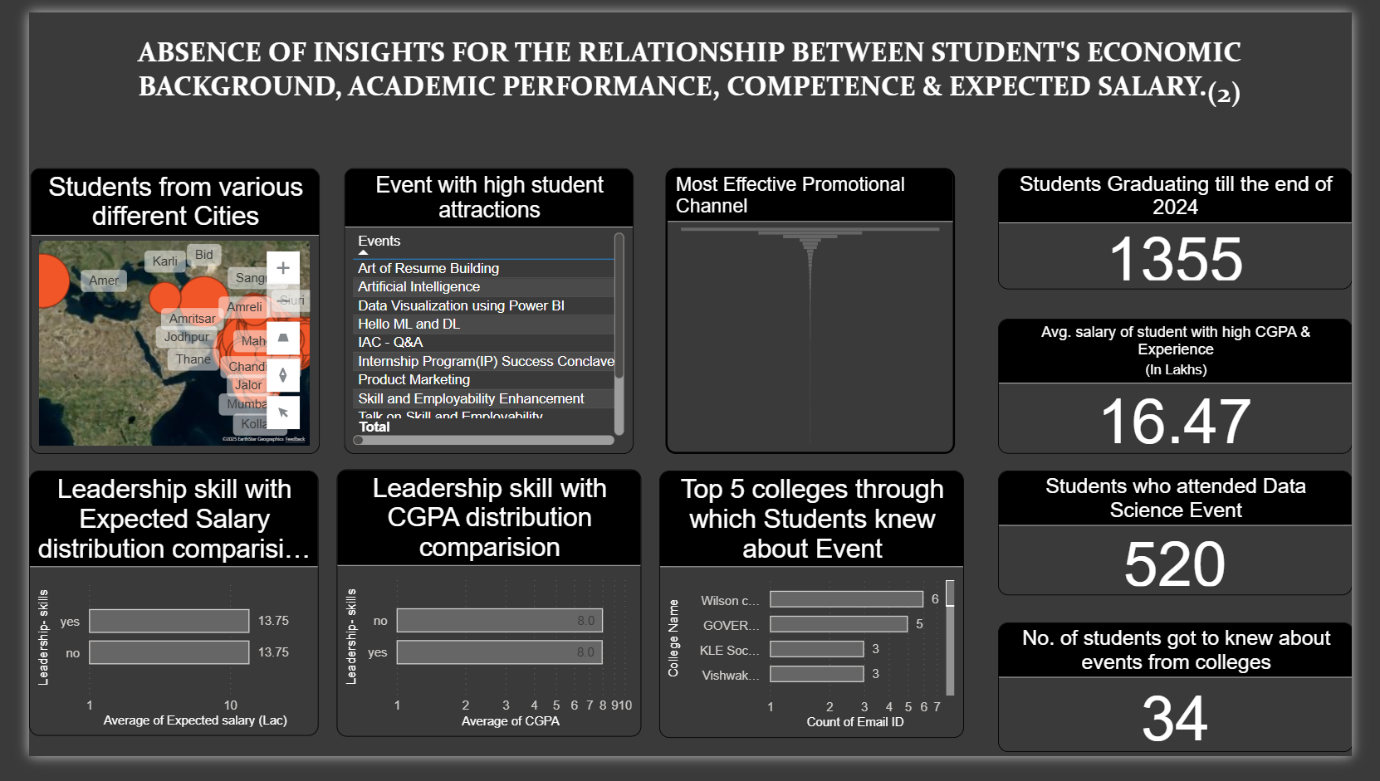
* Art of Resume Building
* Artificial Intelligence
* Data Visualization using Power BI
* Hello ML and DL
* IAC - Q&A
* Internship Program(IP) Success Conclave
* IS DATA SCIENCE FOR YOU?
* KYC - Know Your CCPC
* Product Design & Full Stack
* Product Marketing
* RPA: A Boon or A Bane
* Skill and Employability Enhancement
* Talk on Skill and Employability Enhancement
* The Agile Ways of Working
* The SDLC & their transformations
* Transformation with DevOps: The Easy Way

## Appendix E – DASHBOARD SCREENSHOTS [POWER BI]

**1.DASHBOARD OF THE BASIC QUESTIONS.**



**2.DASHBOARD OF THE MODERATE QUESTIONS.**



## Appendix F - GITHUB REPOSITORY LINK

https://github.com/Sagaryadav2006/IAC\_DATA\_ANALYTICS