



PROJECT REPORT

Data Analytics  
  
Event data

|  |  |  |  |
| --- | --- | --- | --- |
| **Created By:** | Vikas Yadav | **Approved By:** | Harshada Topale |
| **Created On:** | 01-7-2023 | **Approved On:** | 31-6-2023 |

Page left blank intentionally

**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** | Event Data | | |
| **Project Sponsor** |  | | |
| **Project Manager** | Harshada Topale | | |
| **Start Date** | 1st July 2023 | **Completion Date** | 28 Aug 2023 |

# **SUMMARY**

The given dataset which is student related to event conduct by an organization ,

According to the given data in the event there total 2157 student ,

there average GPA is 8.0282333567 and the average family income of student is about

2,30412.

There is no outlier in the given data set. There is no relation between GPA,Family , income, Expericence with python (month)"

There are many student participated from various college the top 5 college form many of the student participated are

1. ST xavier’s
2. New horizon institute of technology and management
3. B.K.Birla college of arts, science and commerce
4. AP shah institute of technology
5. Thakur institute of Management [TIMSCDR]

No outlier is detected in the data set and the top 5 city from where maximum student belong are

1.Amreli

2.jalor,

3.jodhpur,

4.amer,

5.mumbai this are the city from where maximum student participated in the event

The average GPA of all city is near about 8 but some city like Delhi And Raipur city has higest GPA. Total 657 student who were interest for (IP) Internship Program.

There are total 679 Student were present data and they were graduate by the end of 2024.

Organization promote the event through various channel in that most of student known about this event from Whatsapp which is about 558. And 87 student attended the event related to Data Science . Total 41 college who encourage the student about this event.&

All the visualization were present with code .

# **INTRODUCTION**

## Background

*<Take this section from the problem statement & the discussion points from requirement elicitation meeting>*

## Stakeholders

*<Take this section from the problem statement & the discussion points from requirement elicitation meeting>*

<Type Here>

## Objectives

*<Reference the objectives in the* ***Project Charter*** *and describe what has been achieved. If there were any changes to the project objectives since the* ***Project Charter*** *was approved/submission they should be documented here.>*

# **METHODOLOGY**

These conventions are all about the positions of line breaks, how many characters should go on a line, and everything in between.

## Considerations & Assumption

While solving the problem I personally face challenges in few question like average family income ,relation related to family this because family income column is in object data type and the values which are given is in range so we have to convert the range value in particular value the proper solution

## Approach

The method I use to solve the problem is Label Encoding =to convert the object value into int,

Repalce =use to covert the range value with single value .

to\_numeric,= to convert the object value into float.

this help to me solve the problem

## Activities

Uploading dataset in juoyter ,importing libarary’s , dataset cleaning, filtering ,groupby, visualization ,replace value , converting data type

# **TARGETTED V/S ACHIEVED OUTPUT**

*<Mention the targeted output in the project plan and against it what you have been able to achieve. Also mention the reason of deviation. Please remember the deviations will not impact the evaluation and the results of your internship. This sections will help you to analyse the results and understand the cause for deviation that will be lessons learnt for you and would help you to perform better in the next project>*

# **CONCLUSION**

Successfully solve all the problem which is given in dataset, This given dataset give deep insight of the event conduct by the organization ,who many student attend the event from where they belong there College name and percentage .

From where student get know about the this event ,what is area of interest ?

There qualification and year of graduation.

This will help the organization to make quick decision to take next step on the basis of this data.

# **APPENDICES**

## Appendix A – Title

<Add component table as required>