A PROJECT REPORT

Submitted by

Akash Gusain (00614202017)

Anuj Mehta (00814202017)

Deepak Pandit (01014202017)

in partial fulfilment for the award of the degree

of

Bachelor of Computer Applications

JAGANNATH INTERNATIONAL MANAGEMENT SCHOOL VASANT KUNJ, NEW DELHI - 110070

Certificate

Certified that the Project Report (BCA-357) entitled “RESULT EXPLORER” done by the following students is completed under my guidance.

Akash Gusain	Anuj Mehta	Deepak Pandit
00614202017	00814202017	01014202017

Signature of the Guide

Date:

Name of the Guide: **Prof. (Dr.) Meenakshi Narula**

Designation: Head Of IT Department

SELF CERTIFICATE

This is to certify that the project report entitled “RESULT EXPLORER” is done by us is an authentic work carried out for the partial fulfilment of the requirements for the award of the degree of BCA under the guidance of **Prof. (Dr.) Meenakshi Narula**. The matter embodied in this project work has not been submitted earlier for award of any degree or diploma to the best of our knowledge and belief.

Akash Gusain	Anuj Mehta	Deepak Pandit
00614202017	00814202017	01014202017

ACKNOWLEDGEMENT

Firstly, we would like to thank the almighty God as nothing in our life is possible without God.

With candour and pleasure we take opportunity to express our sincere thanks and obligation to our esteemed guide **Prof. (Dr.) Meenakshi Narula.** It is because of her able and mature guidance and co-operation without which it would not have been possible for us to complete my project.

We would also like to thank faculty and staff members of IT department who never hesitated and helped us from time to time.

Finally, we gratefully acknowledge the support, encouragement & patience of our family members, without their unconditional support and motivation this project was not possible.

Thank You!

Akash Gusain	Anuj Mehta	Deepak Pandit
00614202017	00814202017	01014202017

Table of Contents

1 INTRODUCTION 1
1.1 PURPOSE 2
1.2 SCOPE 2
1.3 OBJECTIVES 3
1.4 SDLC Methodology 4
2 SYSTEM ANALYSIS STUDY 8
2.1. EXISTING SYSTEM ANALYSIS 8
2.2 SYSTEM REQUIREMENT GATHERING 10
2.3 FEASIBILITY STUDY 11
2.3.1 Technical Feasibility: 11
2.3.2 Economic Feasibility: 12
2.3.3 Operational Feasibility: 12
2.4 SYSTEM REQUIREMENT STUDY 13
2.4.1 System Process Requirement 13
2.5 SOFTWARE REQUIREMENTS/ TECHNOLOGIES 15
3 SYSTEM DESIGN 24
3.1 INTRODUCTION 24
3.2 PHYSICAL DESIGN 24
3.2.1 System Flow Chart 24
3.2.2 Use-Case Diagram 26
3.2.3 DATA FLOW DIAGRAM 27
3.3 DATABASE DESIGN 31
3.3.1 ER DIAGRAM 31
4 SOFTWARE CODING AND IMPLEMENTATION 32
4.1 INTRODUCTION 32
4.1.1 SAMPLE CODE LAYOUT 32
4.2 TESTING 38
4.2.1 Overview And Approach 39
5 OUTPUT FORMS AND REPORTS 41
6 CONCLUSION & FUTURE ENHANCEMENT(S) 49
6.1 CONCLUSION 49
6.2 FUTURE SCOPE 50
BIBLIOGRAPHY 51

1 INTRODUCTION

Guru Gobind Singh Indraprastha University (GGSIPU or IP or IPU), formerly Indraprastha University, is a public state university located in Delhi, India. It was established in 1998 by Government of Delhi, as a teaching-cum-affiliating university. The University is recognised by University Grants Commission (India), under section 12B of the UGC Act. It has been graded 'A' by the NAAC. The medium of instruction is English to meet the international language standards.

Guru Gobind Singh Indraprastha University is a member of the Association of Commonwealth Universities, Association of Indian Universities, Medical Council of India and Distance Education Council. The university was ranked 66 among Indian universities by the National Institutional Ranking Framework (NIRF) in 2019 and 95 in the overall category.

Jagannath International Management School (JIMS) is a private college affiliated with Guru Gobind Singh Indraprastha University (GGSIPU) and located in Vasant Kunj, Delhi. It was established in 2003 by Jagannath Gupta Memorial Education Society. Admission to JIMS is through the Common Entrance Test (CET). The institute offers BCA, BBA and BJMC undergraduate full-time degree programmes. The institute is recognised under section 2(f) by UGC and accredited by the National Assessment and Accreditation Council (NAAC) with an "A" grade.

Technology in today's world has reached to extent that it can be used to do various task in day to day life easily with less effort and time. World today has realised importance of education in one's life which has led to revolution in field of education. Whenever student gets their results in University, GGSIPU uploads results on their website in pdf format in which they are only provided by the marks of the subject which creates a problem for them because they are not provided by any kind of analysis based on any useful perspectives like they don't know how much percentage they have secured in the current semester, what is their college rank or what is the position they have secured in the whole University . every student is lacking these statistics and therefore they cannot measure or compare their performances from the previous semester results therefore they cannot prepare well for the upcoming exams.

In today's scenario colleges including our college(JIMS, Vasant Kunj) needs to analyse their student results manually in which they have to download the result from university website and then manually write excel from the pdf which takes lots of time and effort and then calculating result and other required analysis like college rank and university rank which is a big challenge when work is done manually and also there is always scope for human errors

1.1 PURPOSE

- The purpose of this project is to eliminate the manual process of result analysis carried out by the institutions which consists of downloading the PDF file from the official website of GGSIPU then manually writing each entry in excel and then obtaining the useful statistics which roughly takes 12-14 days, and requires a lot of concentration, with this web application it can be done in no time.
- The next major purpose of this project is to provide the useful statistics like college ranks and university ranks and analysis of the result to the student. This web-based application is capable of giving you ranks on different parameters, overall Percentage and semester wise result of each student

1.2 SCOPE

The Result Explorer will focus on the result that are declared by Guru Gobind Singh Indraprastha University specially for the course of BCA (Bachelor of Computer Applications). This website will provide useful statistics with ranking.

1.3 OBJECTIVES

1. Extracting the PDF in excel or csv format: Firstly, we need to download the PDF of result by the official website of GGSIPU, then extract pdf in form of excel or csv through which we can analyse and export the data in database.

2. Removal of unwanted data from the excel sheet: The second step involves cleaning of the derived excel and removal of unwanted spaces, special characters, unwanted rows and columns and unwanted junk data from the sheet.

3. Exporting the data into the database: The next step is to export the cleaned and desired excel data into database of xampp server (MY SQL) that is been used by our website.

4. Data analysis: Next step involves the applying of different queries and data analysis algorithms to retrieve desired statistics for the students.

5. Developing a web-based application: Final step would be making a web-based application that will act as an interface and can be accessed by its different users i.e. students and faculty members to retrieve results of students.

1.4 SDLC Methodology

SDLC or the Software Development Life Cycle is a process that produces software with the highest quality and lowest cost in the shortest time. SDLC includes a detailed plan for how to develop, alter, maintain, and replace a software system.

SDLC involves several distinct stages, including planning, design, building, testing, and deployment.

How SDLC Works

SDLC works by lowering the cost of software development while simultaneously improving quality and shortening production time. SDLC achieves these apparently divergent goals by following a plan that removes the typical pitfalls to software development projects. That plan starts by evaluating existing systems for deficiencies. Next, it defines the requirements of the new system. It then creates the software through the stages of design, development, testing, and deployment. By anticipating costly mistakes like failing to ask the end user for suggestions, SLDC can eliminate redundant rework and after-the-fact fixes.

Types of SDLC Models

1. Waterfall Model

The Waterfall Model was the first process model to be introduced. It is also referred to as a linear-sequential life cycle model. The waterfall Model illustrates the software development process in a linear sequential flow. This means that any phase in the development process begins only if the previous phase is complete. In this waterfall model, the phases do not overlap.

2. Incremental Model

Incremental Model is a process of software development where requirements are broken down into multiple standalone modules of software development cycle. Incremental development is done in steps from analysis design, implementation, testing/verification, maintenance.

3. Spiral Model

The spiral model combines the idea of iterative development with the systematic, controlled aspects of the waterfall model. This Spiral model is a combination of iterative development process model and sequential linear development model i.e. the waterfall model with a very high emphasis on risk analysis. It allows incremental releases of the product or incremental refinement through each iteration around the spiral.

4. Prototype Model

The Prototyping Model is one of the most popularly used Software Development Life Cycle Models (SDLC models). This model is used when the customers do not know the exact project requirements beforehand. In this model, a prototype of the end product is first developed, tested and refined as per customer feedback repeatedly till a final acceptable prototype is achieved which forms the basis for developing the final product

5. RAD Model

The RAD (Rapid Application Development) model is based on prototyping and iterative development with no specific planning involved. The process of writing the software itself involves the planning required for developing the product. Rapid Application Development focuses on gathering customer requirements through workshops or focus groups, early testing of the prototypes by the customer using iterative concept, reuse of the existing prototypes (components), continuous integration and rapid delivery.

After studying all the models of SDLC and analysing our project we concluded that Incremental model is best suitable for this project. That's why this project has been developed using Incremental Model.

Incremental Model

Incremental Model is a process of software development where requirements divided into multiple standalone modules of the software development cycle. In this model, each module goes through the requirements, design, implementation and testing phases. Every subsequent release of the module adds function to the previous release. The process continues until the complete system achieved.

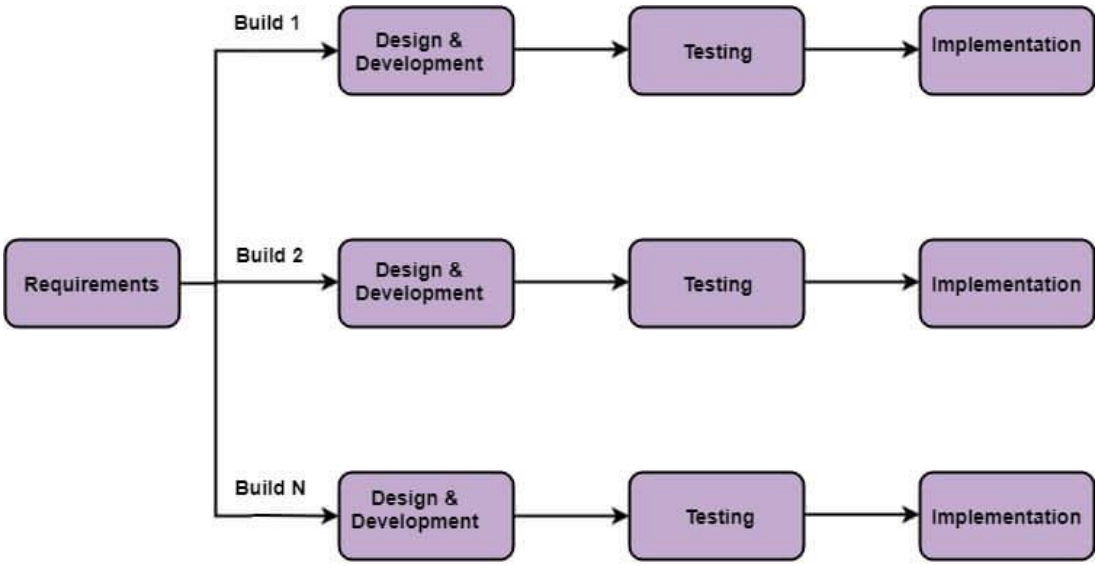


Fig: Incremental Model

Source: https://www.javatpoint.com/software-engineering-incremental-model

When we use the Incremental Model?

- When the requirements are superior.
- A project has a lengthy development schedule.
- When Software team are not very well skilled or trained.
- When the customer demands a quick release of the product.

Advantage of Incremental Model

- Errors are easy to be recognized.
- Easier to test and debug
- More flexible.
- Simple to manage risk because it handled during its iteration.

The various phases of incremental model are as follows:

1. **Requirement analysis:**

   In the first phase of the incremental model, the requirements are analysed for the development of the project. In our project we gathered all the information for analysis by its different end users such as office staffs who types the data from PDF to Excel and by students through one to one interview.

2. **Design & Development:**

In this phase of the Incremental model of SDLC, the design of the system functionality and the development method are finished with success. After the information gathering, we designed DFDs and ERD for our project and then coding part was executed based on the designs.

3. **Testing:**

In the incremental model, the testing phase checks the performance of each existing function as well as additional functionality. In the testing phase, the various methods are used to test the behaviour of each task. After building the project we tested each module separately that is unit testing and then integration testing was performed at the end of the completion of project.

4. **Implementation:**

In implementation phase the finally developed and tested software is deployed at the client site. In our project we have to enhance some of the functionalities for a better experience and then it will be deployed on the client site.

2 SYSTEM ANALYSIS STUDY

2.1. EXISTING SYSTEM ANALYSIS

The existing system that GGSIPU and its associated colleges follows is the university releases the semester results in the form of PDF(Portable Document Format) file on their official website and students have to go their individually to download the result and search their names out of thousands of students. The PDF which the university releases is bulky and not suitable for any kind of analysis by the student or their respective colleges. This increases the workload of college staff as they have to enter the data of PDF file manually in excel file which takes so many days for the extraction of desired results or statistics of the students of their colleges. Also, there is a scope of human error as the work is done by human with complete manual processes

Problems identified with PDF file results

1. **Problem in working with pdf**
   The major problem in working with pdf is as this is an analysis-based project and pdf are not preferable for analysis, so firstly we will have to convert pdf to excel file (.xlsx or .csv) for analysis of data

2. **PDF result files consume a lot of data**
   PDF result file of IP University, Delhi measures anywhere from 75 to 120 MB depending on the course. An average student does not consume the whole PDF file. Thousands of students and hundreds of faculties download them every semester, resulting in wastage of GBs of data

3. **Results are scattered across different PDF**
   The results for each semester are declared in different PDF files, making the

students downloading multiple files. Even after downloading them, it's difficult to manage them, most of the students lose them or delete them, and then redownload them later.

4. **Lack of Statistics**

   The students and the faculties have to manually compute the percentages, CGPAs and it is not feasible to compute rankings without an automated procedure. It makes the task difficult for both faculties and students.

5. **PDF result files are user unfriendly**

Searching result in a big PDF file where almost results of 4,000 in 1000 above pages students are present is not a quick task and in mobiles it is even more unpleasant experience. PDF files are often nonresponsive in mobiles, i.e. they do not adjust to the user's display.

6. **Non-Responding Servers**

Almost every time students and faculties have to wait for the server to respond as there are many users accessing the same PDF file at a time.

2.2 SYSTEM REQUIREMENT GATHERING

It’s difficult to build solution if you don't know the requirements (in spite of the fact that many teams will still try to do it today). The “elicitation” step is where the requirements are first gathered from the client. Many techniques are available for gathering requirements. Each has value in certain circumstances, and in many cases, we need multiple techniques to gain a complete picture from a diverse set of clients and stakeholders. Following are the method that can be followed to gather requirements:

1. ONE-ON-ONE INTERVIEWS:

The most common technique for gathering requirement just sit down with the client and ask them what they need. The discussion should be planned out ahead of time based on the type of requirement you’re looking for. There are many good ways to plan the interview, but generally, you want to ask open ended questions to get the interviewee to start talking and then ask probing questions to uncover requirements.

2. QUESTIONNAIRES:

Questionnaires are much more informal, and they are good tools to gather requirements from stakeholders in remote locations or those who will have only overall requirements. Questionnaires can also be used when you have to gather input from dozens, hundreds, or thousands of people.

For this project we have conducted one on one interview with academic assistants who are one of the users and gathered all the required information

2.2.1 Overview & Analysis of Data Gathered

Based on the requirements that we gathered from different requirement gathering techniques we came to conclusion that the client wants to fully automate the process of getting the data from PDF to Excel which is being done manually at present. The client also wants to develop a website that can be used by students as well as by the faculty members or other academics to retrieve the result just by entering enrolment number. The client also want us to deliver a functionality in which he just has to select the college name and semester and by button click he or she should be able to view the entire rank of the college semester wise also there should be a functionality by which

the client can view the ranking of entire University students semester wise and overall university ranking

2.3 FEASIBILITY STUDY

Feasibility study is used to determine the viability of an idea, such as ensuring a project is legally and technically feasible as well as economically justifiable. It tells us weather a project is worth the investment - in some cases, project may not be doable. It is a preliminary study to investigate the objectives, constraints, resource requirements, cost, benefits and feasibility of a proposed system. the goal of feasibility study is to evaluate alternative methodologies and to propose the most feasible and desirable methodology for development.

The following feasibility studies were conducted for the developed project

- Technical feasibility
- Economic feasibility
- Operational feasibility

2.3.1 Technical Feasibility:

Technical feasibility centres around the existing computer system (Hardware and software etc) and to what extent it supports the proposed edition. This involves financial considerations to accommodate technical enhancements. If the budget is serious constraint, then the project is just not feasible. In this project, all the necessary questions have been taken care to make it technically feasible.

Python has been used for one of the modules of this project for the conversion and extraction of result from pdf to excel which was an open source also all the team members had decent knowledge of the language. The website is designed in PHP with HTML which is very fast and team member were familiar with it as we're studying it as one of our core subject. Also, the tools, operating system and other programming language used in this localization process are compatible with the existing ones.

2.3.2 Economic Feasibility:

Economic analysis is the most frequently used method for evaluating the effectiveness of the candidate system. More commonly known as cost/benefit analysis, the procedures to be determining the benefits and saving that are expected from a candidate and compare them with the cost. If benefits outweigh cost, then the decision is made to design and implement system. A systems financial benefit must exceed the cost of developing that system. i.e. a new system being developed should be a good investment for the organisation. Economic feasibility considers the following:

- The cost to conduct a full system investigation.
- The cost of hardware and software for the class of application.
- The benefits in form of reduced cost or fewer costly error.
- The cost if nothing changes (i.e. the proposed system is not developed).

considering the above stated points, we use open source php development environment and other languages including MYSQL database were also free and no paid software was used for any of the backend functionality also this website is to be integrated with the existing college website and hence there is no additional cost involved for domain name and web hosting thus, we can conclude that the proposed website of **Result Explorer** is economically feasible.

2.3.3 Operational Feasibility:

The operational feasibility is also known as functional feasibility. In this, we study how the project plan satisfy the requirements identified in the requirement analysis face of the system development. it is the Measure of how well a proposed system solve the problem and take advantage of the opportunities identify during scope definition and how it satisfies the requirements identified in the requirement analysis phase of the system development.

As the project has been divided in two separated modules one for conversion of pdf to excel and the other one is for viewing the result analysis which makes it feasible to operate.

2.4 SYSTEM REQUIREMENT STUDY

Requirement analysis/study is done in order to understand the problem the software system is to solve. The problem could be automating an existing manual process, developing a new automated system or a combination of the two. The emphasis requirement analysis is on identifying what is needed from the system, not how the system will achieve it goals. There are at least 2 parties involved in the software development -a client and a developer. The developer has to develop a system to satisfy the client's needs. The developer does not understand the client's problem domain, and the client does not understand the issues involved in the software system. This causes a communication gap, which has to be adequately bridged during the requirement analysis. For this project, the following requirements were analysed

- Functional requirements
- Non-Functional requirements

2.4.1 System Process Requirement

System process requirements are all the requirements at the system level that describe the function which the system as a whole should fulfil to satisfy the stakeholders needs and requirements, And is expressed in an appropriate combination of textual statement, views and non-functional requirement; the later expressing the levels of safety, security, reliability etc. that will be necessary. system requirements play major role in system engineering as they:

- form the basis of system architecture and design activity
- form the basis of system integration and verification activities.
- access reference for validation and stakeholder's acceptance.
- provider means of communication between the various technical staff that interact through the project.

Functional Requirements

Functional requirement defines the functionality of a system or one of its subsystems. It also depends upon the type of software, expected users and the type of system where the software is used. Functional user requirements maybe high-level statements of what

the system should do but functional system requirement should also describe clearly about the system service in detail.

Non-Functional Requirements

A Non- functional requirement is requirement that specifies criteria that can be used to judge the operation of a system, Rather than specific behaviour. They are contrasted with functional requirement that define specific behaviour or functions. The plan for implementing functional requirement is detailed in the system design. The plan for implementing non-functional requirement is detailed in the system architecture because they are usually architecturally significant requirements.

NON- FUNCTIONAL REQUIREMENTS OF THIS PROJECT ARE:

- **Accessibility:** The interface will be accessible to all the users.
- **Efficiency:** The interfaces officiant as it displays to the point details maintaining the efficiency of the system not to be involved in useless things
- **Effectiveness:** the interfaces effective because it provides the basic information of the student and also some additional information
- **Extensibility:** the interface can be extended to add more functionality
- **Interoperability:** Most of the popular browsers can be used to open this website on different computers.

Hardware Requirement

- OS: Windows 8.1/10 - 64-Bit
- CPU: Intel Core i3-2100 @ 3.1GHz or AMD Phenom II X4 965 @ 3.4 GHz
- RAM: 2GB or Higher
- Hard Drive Space Required: 50.0 GB

2.5 SOFTWARE REQUIREMENTS/ TECHNOLOGIES

Jupyter Notebook

jupyter Notebook (formerly Python Notebooks) is a web-based interactive computational environment for creating Jupyter notebook documents. The "notebook" term can colloquially make reference to many different entities, mainly the Jupyter web application, Jupyter Python web server, or Jupyter document format depending on context. A Jupyter Notebook document is a JSON document, following a versioned schema, and containing an ordered list of input/output cells which can contain code, text (using Markdown), mathematics, plots and rich media, usually ending with the ". ipynb" extension.

A Jupyter Notebook can be converted to a number of open standard output formats HTML, presentation slides, LaTeX, PDF, ReStructuredText, Markdown and Python.

MICROSOFT EXCEL

Microsoft Excel is a spreadsheet developed by Microsoft for Windows, macOS, Android and iOS. It features calculation, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications. It has been a very widely applied spreadsheet for these platforms, especially since version 5 in 1993, and it has replaced Lotus 1-2-3 as the industry standard for spreadsheets. Excel forms part of the Microsoft Office suite of software.

Microsoft Excel has the basic features of all spreadsheets, using a grid of cells arranged in numbered rows and letter-named columns to organize data manipulations like arithmetic operations. It has a battery of supplied functions to answer statistical, engineering and financial needs. In addition, it can display data as line graphs, histograms and charts, and with a very limited three-dimensional graphical display. It allows sectioning of data to view its dependencies on various factors for different perspectives (using pivot tables and the scenario manager). It has a programming aspect, Visual Basic for Applications, allowing the user to employ a wide variety of numerical methods, for example, for solving differential equations of mathematical physics, and then reporting the results back to the spreadsheet.

Web Browser (Google Chrome)

As the developed project is a web-based application, therefore it requires a web browser. Any web browser can be used to access this website, but it has been optimized for Google Chrome. Google Chrome is a cross-platform web browser developed by Google. It was first released in 2008 for Microsoft Windows, and was later ported to Linux, macOS, iOS, and Android. The browser is also the main component of Chrome OS, where it serves as the platform for web apps.

Most of Chrome's source code comes from Google's open-source Chromium project, but Chrome is licensed as proprietary freeware. Web Kit was the original rendering engine, but Google eventually forked it to create the Blink engine; all Chrome variants except iOS now use Blink.

As of July 2019, Stat Counter estimates that Chrome has a 71% worldwide browser market share on traditional PCs and 63.34% share across all platforms. Because of this success, Google has expanded the "Chrome" brand name to other products: Chrome OS, Chromecast, Chromebook, Chromebit, Chromebox, and Chromebase.

Xampp

Xampp is a free and open-source cross-platform web server solution stack package developed by Apache Friends, consisting mainly of the Apache HTTP Server, MariaDB database, and interpreters for scripts written in the PHP and Perl programming languages. Since most actual web server deployments use the same components as XAMPP, it makes transitioning from a local test server to a live server possible.

XAMPP's ease of deployment means a WAMP or LAMP stack can be installed quickly and simply on an operating system by a developer. With the advantage of common add-in applications such as WordPress and Joomla! can also be installed with similar ease using Bitnami.

Visual Basic for Applications (VBA)

Visual Basic for Applications (VBA) is an implementation of Microsoft's event-driven programming language Visual Basic 6, which was declared legacy in 2008, and its associated integrated development environment (IDE). Although pre-.NET Visual Basic is no longer supported or updated by Microsoft, the VBA programming language was upgraded in 2010 with the introduction of Visual Basic for Applications 7 in Microsoft Office applications.

Visual Basic for Applications enables building user-defined functions (UDFs), automating processes and accessing Windows API and other low-level functionality through dynamic-link libraries (DLLs). It supersedes and expands on the abilities of earlier application-specific macro programming languages such as Word's Word BASIC. It can be used to control many aspects of the host application, including manipulating user interface features, such as menus and toolbars, and working with custom user forms or dialog boxes.

Python

Python is an interpreted, high-level, general-purpose programming language. Created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant whitespace. Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects.

Python is dynamically typed and garbage collected. It supports multiple programming paradigms, including procedural, object-oriented, and functional programming. Python is often described as a "batteries included" language due to its comprehensive standard library.

Python was conceived in the late 1980s as a successor to the ABC language. Python 2.0, released in 2000, introduced features like list comprehensions and a garbage collection system capable of collecting reference cycles.

Hyper Text Markup Language

Hypertext Markup Language (HTML) is the standard markup language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets (CSS) and scripting languages such as JavaScript.

Web browsers receive HTML documents from a web server or from local storage and render the documents into multimedia web pages. HTML describes the structure of a web page semantically and originally included cues for the appearance of the document.

HTML elements are the building blocks of HTML pages. With HTML constructs, images and other objects such as interactive forms may be embedded into the rendered page. HTML provides a means to create structured documents by denoting structural semantics for text such as headings, paragraphs, lists, links, quotes and other items. HTML elements are delineated by tags, written using angle brackets. Tags such as <img /> and <input /> directly introduce content into the page. Other tags such as <p> surround and provide information about document text and may include other tags as sub-elements. Browsers do not display the HTML tags, but use them to interpret the content of the page**.**

PHP

PHP: Hypertext Preprocessor (or simply PHP) is a general-purpose programming language originally designed for web development. It was originally created by Rasmus Lerdorf in 1994; the PHP reference implementation is now produced by The PHP Group. PHP originally stood for Personal Home Page, but it now stands for the recursive initialism PHP: Hypertext Preprocessor.

PHP code may be executed with a command line interface (CLI), embedded into HTML code, or used in combination with various web template systems, web content management systems, and web frameworks. PHP code is usually processed by a PHP interpreter implemented as a module in a web server or as a Common Gateway Interface (CGI) executable. The web server outputs the results of the interpreted and executed PHP code, which may be any type of data, such as generated HTML code or binary image data. PHP can be used for many programming tasks outside of the web context, such as standalone graphical applications and robotic drone control.

The standard PHP interpreter, powered by the Zend Engine, is free software released under the PHP License. PHP has been widely ported and can be deployed on most web servers on almost every operating system and platform, free of charge.

MY SQL

MY SQL is an open-source relational database management system (RDBMS).Its name is a combination of "My", the name of co-founder Michael Widenius's daughter and "SQL", the abbreviation for Structured Query Language.

MySQL is free and open-source software under the terms of the GNU General Public License, and is also available under a variety of proprietary licenses. MySQL was owned and sponsored by the Swedish company MySQL AB, which was bought by Sun Microsystems (now Oracle Corporation). In 2010, when Oracle acquired Sun, Widenius forked the open-source MySQL project to create MariaDB.

MySQL is a component of the LAMP web application software stack (and others), which is an acronym for Linux, Apache, MySQL, Perl/PHP/Python. MySQL is used by many database-driven web applications, including Drupal, Joomla, phpBB, and WordPress. MySQL is also used by many popular websites, including Facebook, Flickr, MediaWiki, Twitter, and YouTube.

3 SYSTEM DESIGN

3.1 INTRODUCTION

Systems design is the process of defining the architecture, modules, interfaces, and data for a system to satisfy specified requirements. System design involves first logical design and then physical construction of the system. The logical design describes the structure and characteristics of features, such as output, input, files, database and procedures. The physical construction produces actual program software, files and working system.

OBJECTIVE OF DESIGN

Since the simplification of a program should be as free as possible aspects imposed by "how" The program will work, it is seldom document from which coding can directly be done.

So, design the gap between specification and coding taking the specification, decide how the program will be organised, and the method it will use, insufficient details has to be direct code table.

If the specification calls for larger a complex program, then the design is quite likely to work down through a number of level. At each level breaking the implementation problem into a combination of smaller and simple problems. Filling a large gap will involve no stepping stone! The wider the gap large number of stepping stones is required.

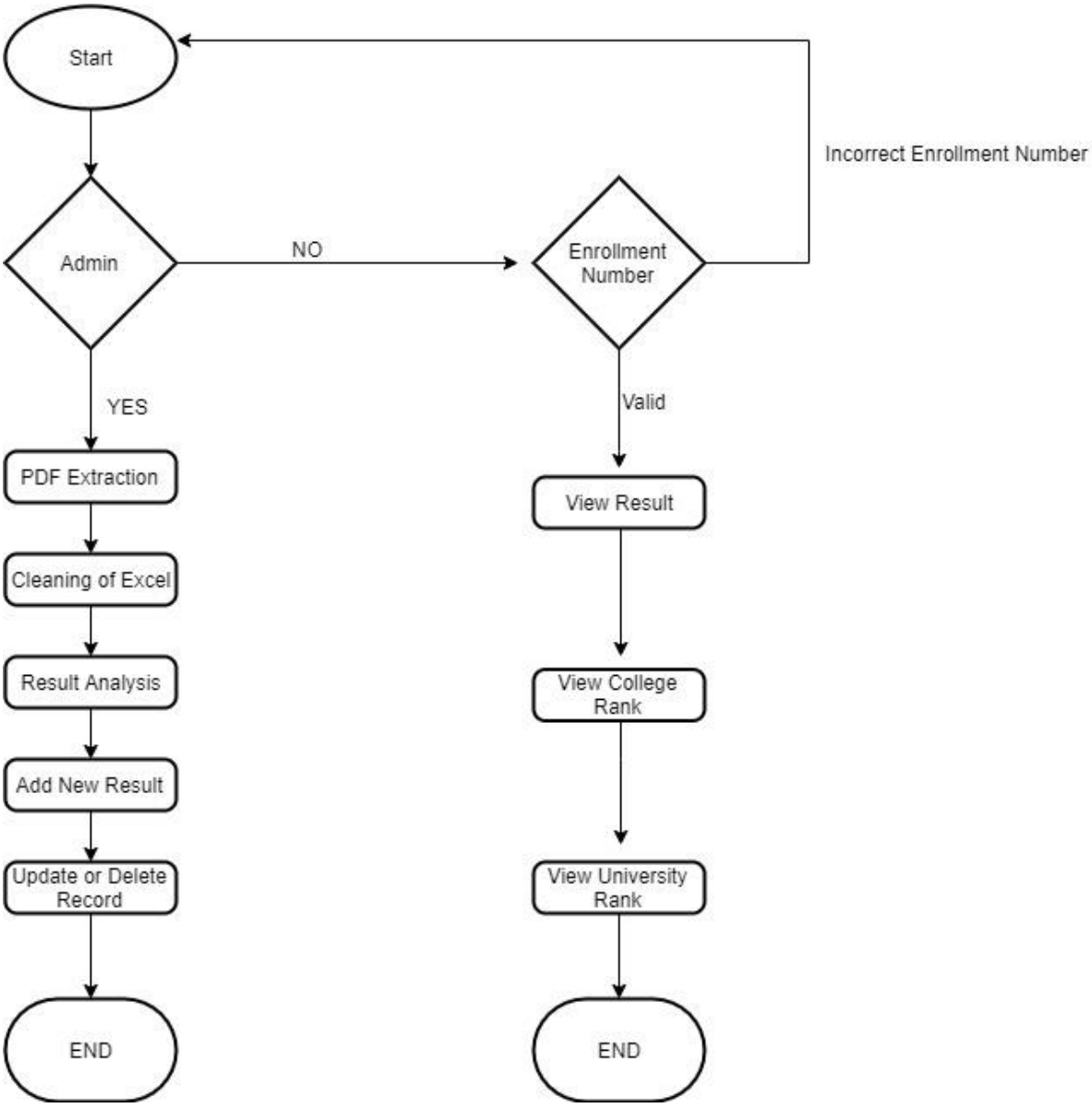
3.2 PHYSICAL DESIGN

3.2.1 System Flow Chart

System flow charts are a way of displaying how data flows in a system and how decisions are made to control events. To illustrate this, symbols are used. They are connected together to show what happens to data when it goes.

Given below is the flowchart of the system

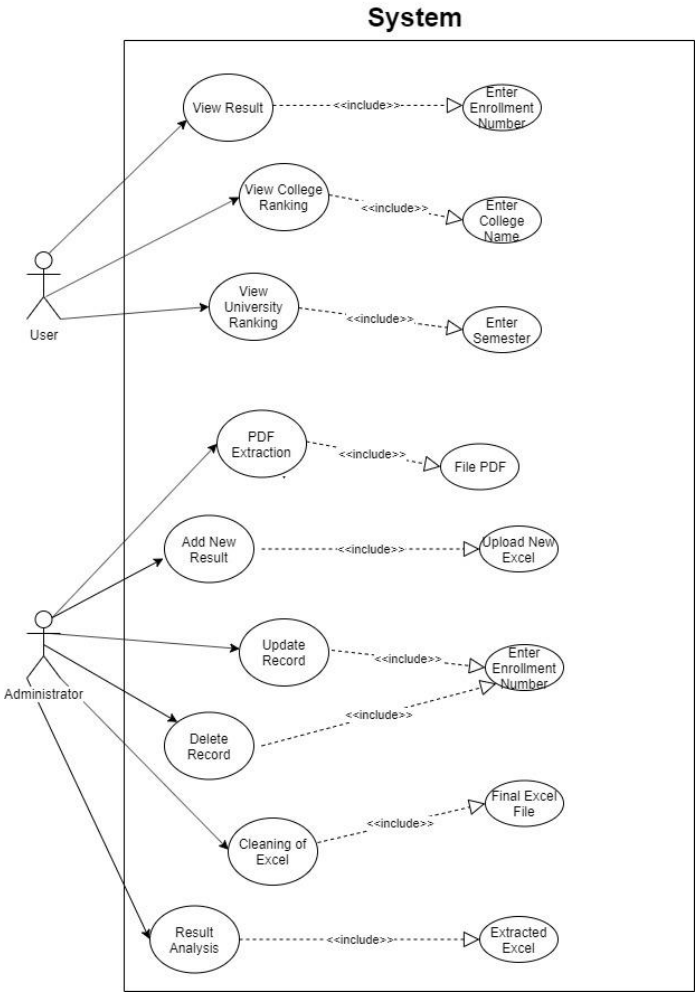
Flowchart of the system



3.2.2 Use-Case Diagram

A use case is a methodology used in system analysis to identify, clarify and organise system requirements. The use cases made up of a set of possible sequences of interaction between system and user in a particular environment and related to a particular goal. Stop it consist of a group of elements (for example, classes and interfaces) that can be used together in a way that will have an effect larger than the sum of the separate elements combined. The use case should contain all the system activities that have significance to the user. Use case can be thought of as a collection of possible scenarios related to a particular goal, indeed, the use case and goals are sometime considered to be synonymous.

Use Case Diagram of the Web Based Application

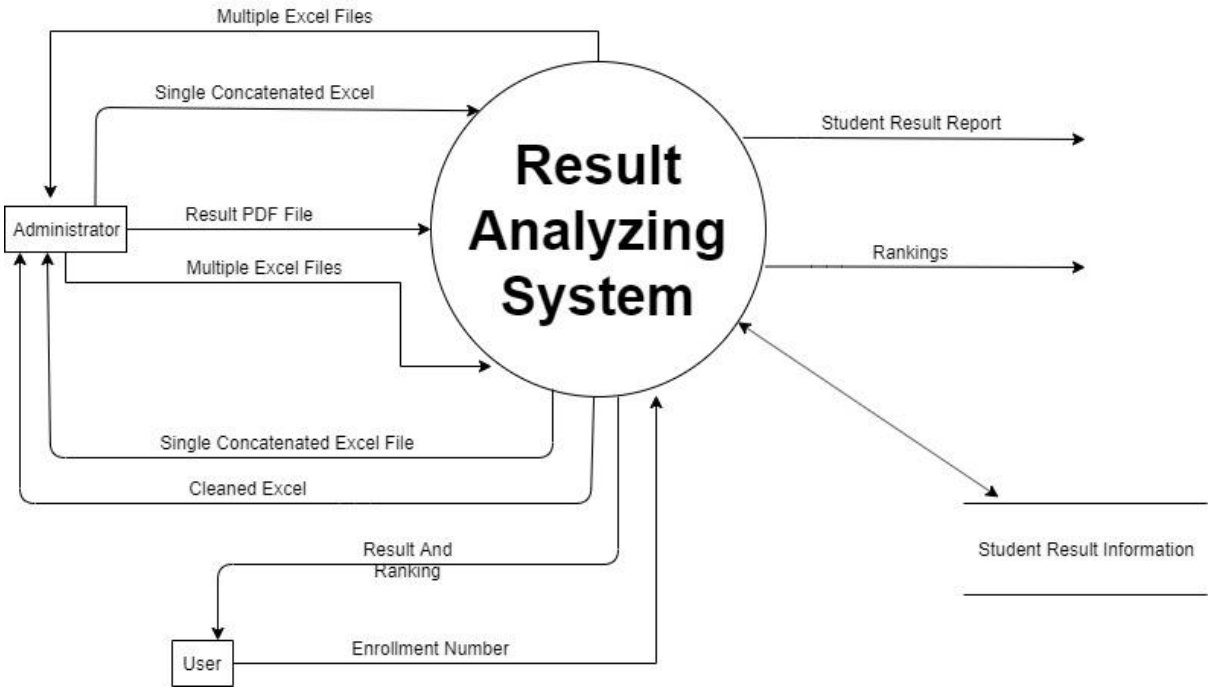


3.2.3 DATA FLOW DIAGRAM

A data flow diagram (DFD) Maps out the flow of information for any process or system. It uses defined symbols like rectangles, circles and arrows plus short text labels, to show data inputs, outputs storage points and the roads between destinations. Data flow charts can range from simple, even hand drawn process overviews 2 in depth multilevel DFDs that dig progressively how data is handled. They can be used to analyse an existing system or model a new one. Like all the best diagrams and charts, a DFD can often visually "say" things that would be hard to explain in words, and they work both for technical non-technical audiences, from developer to CEO. That's why DFDs remain so popular after all these years. While they work well for data flow software and system, they are less applicable nowadays to visualizing Interactive, real time database oriented software systems or system.

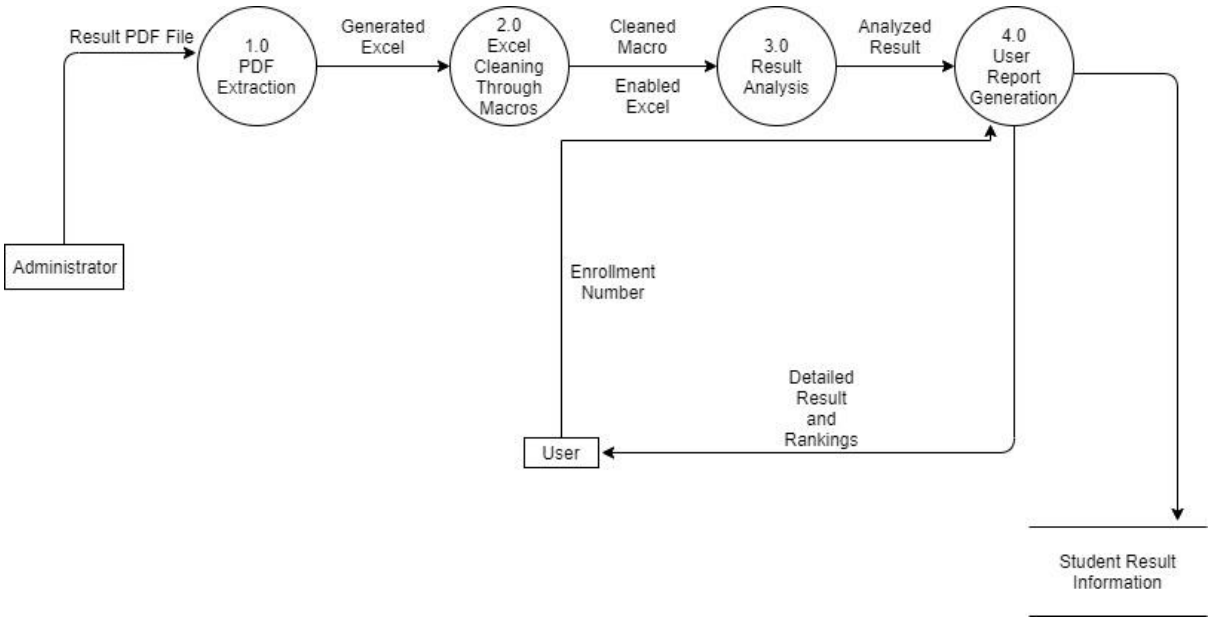
Context Level DFD

Following figure represents the context level diagram of the whole system and its functions from conversion of PDF to excel to final result retrieval.



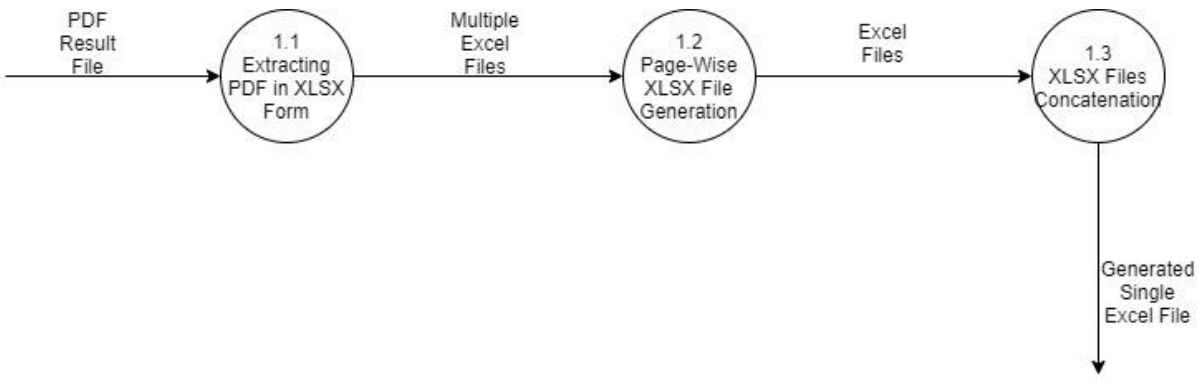
Level-1 DFD

Given figure represents the level one DFDs which includes more detailed processes that were shown in context level diagram.



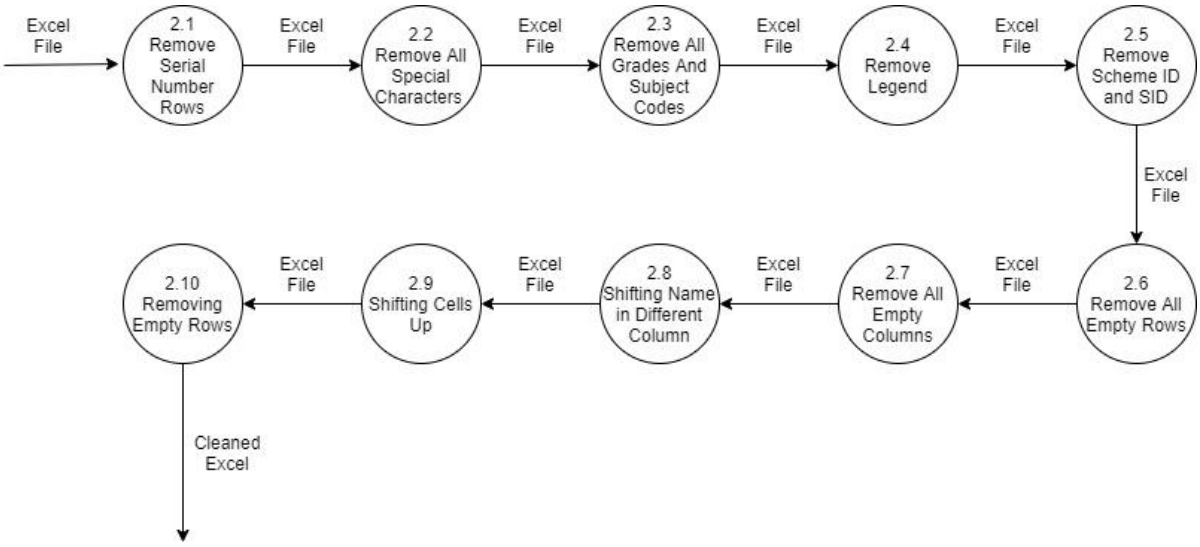
Level-2 DFD Of PDF Extraction

Following figure shows the process of conversion of PDF file that was uploaded by GGSIPU into desired excel file.



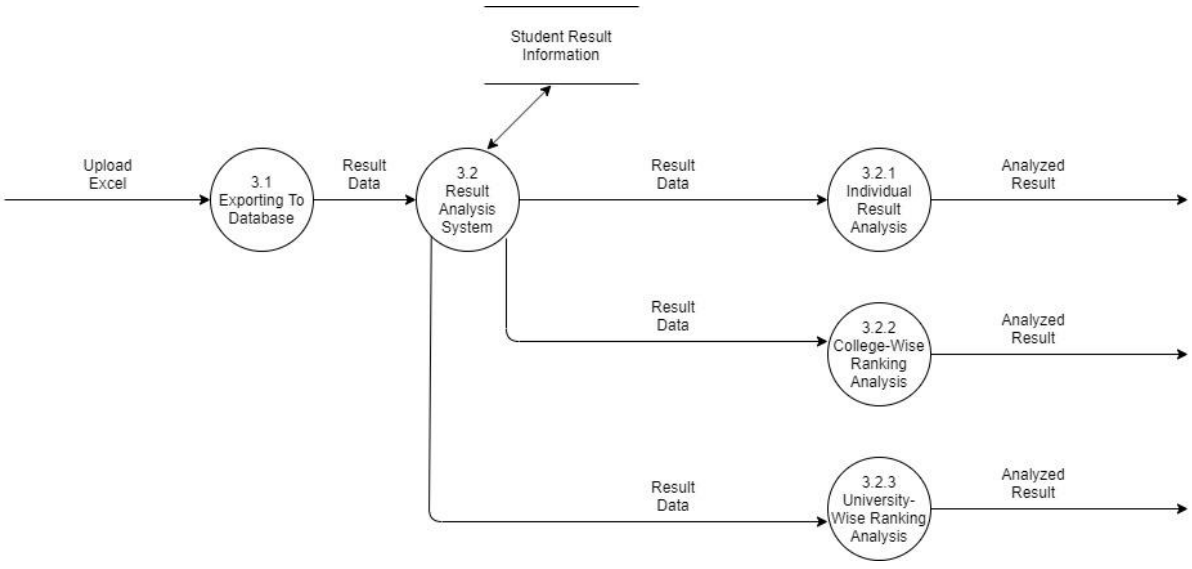
Level-2 DFD Of Excel Cleaning

This figure shows the next process in which the cleaning process of extracted excel Is shown.



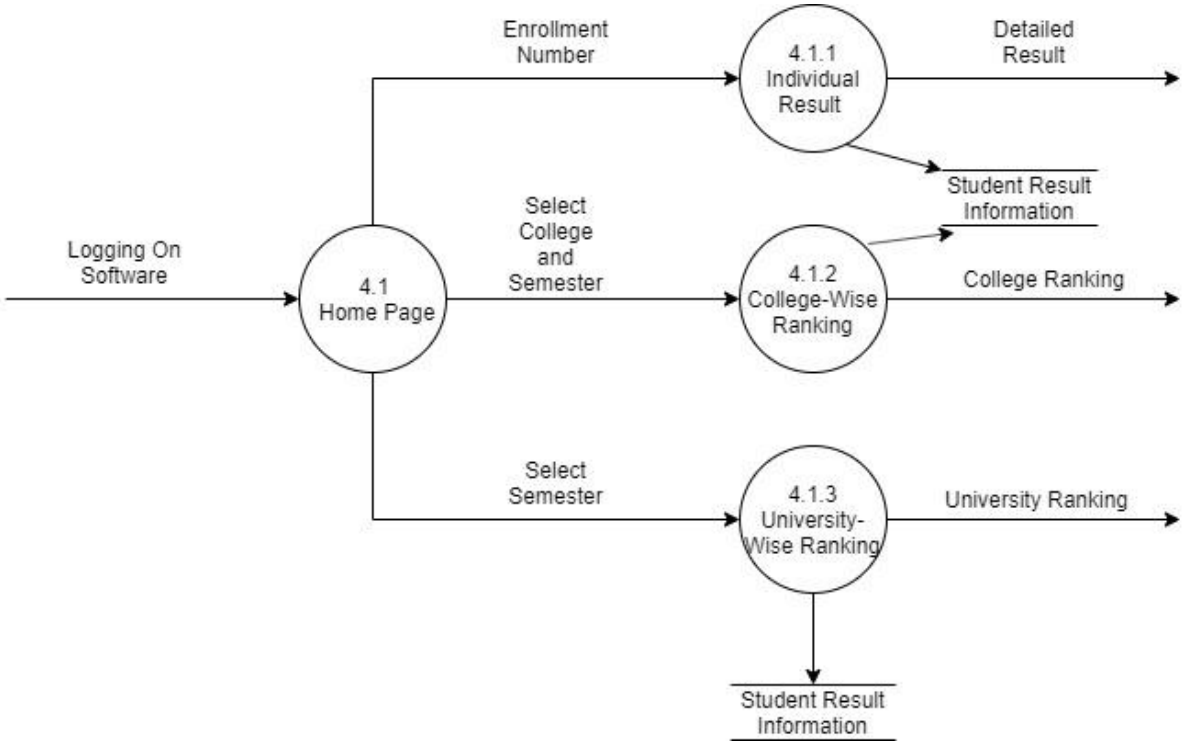
Level-2 DFD Of Result Analysis

This figure shows the detailed process of result analysis



Level-2 DFD Of User Report Generation

This diagram represents the frontend or the web-based application with different features.

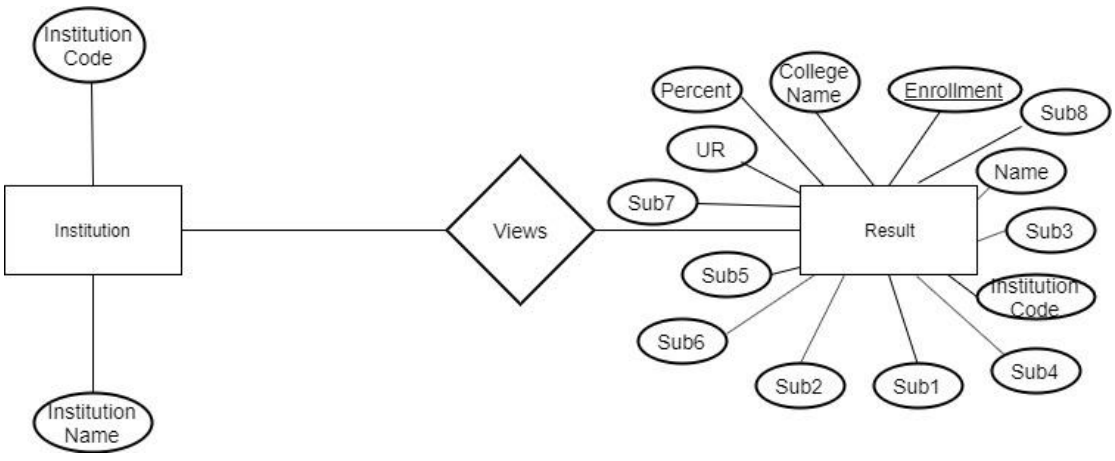


3.3 DATABASE DESIGN

Database design is the organization of data according to the requirements. The designer determines what data must be stored and how the data elements interrelate. With this information, they can begin to fit the data to the database model. Database management system manages the data accordingly. Database design involves classifying data and identifying interrelationships.

3.3.1 ER DIAGRAM

Entity relationship diagram displays the relationships of entity set stored in a database. In other words, we can say that ER diagrams help you to explain the logical structure of databases. At first look, an ER diagram looks very similar to the flowchart. However, ER Diagram includes many specialized symbols, and its meanings make this model unique



Above diagram represents the ER Diagram of the developed system

4 SOFTWARE CODING AND IMPLEMENTATION

4.1 INTRODUCTION

4.1.1 SAMPLE CODE LAYOUT

Code for individual result

```
   ?>
<?php
if(isset($_POST['search']))
{
$id=$_POST['id'];
if(empty($_POST['id']))
{
echo"<script>alert('helllo');</script>";
header("Location:first.php");
}
else{
    $sql = "SELECT * FROM sem3 WHERE Enrollment='$id' ";
 $result = $conn->query($sql) or die($conn->error);

while($row = $result->fetch_assoc())
{
?>
            <div class="panel ">    <div class="heading" style="background-color:#1
                  <span class="title">SEMESTER - 3</span>
               </div><div class="content"><div class="block-shadow" style="overflow
    <tr>
   <th><b>Sr. No.</b></td>
    <th><b>SUBJECT</b></td>
    <th><b>MARKS</b></td>
    </tr><tr ><td>
    1.</td><td>Mathematics  III</td><td><?php echo $row['Mathematics_3']; ?></td>
       </tr><tr ><td>
       2.</td><td>Computer Architecture </td><td><?php echo $row['CA']; ?></td>
       </tr><tr ><td>
       3.</td><td> Front End Design Tool
```

Code For College Ranking

```
<tr>
<th colspan="4"><h2 align="center">COLLEGE RANKING</h2></th>
</tr>
 <tr>
<th>Name</th>
<th>Enrollment</th>
<th>Percentage</th>
<th>RANKING</th>
</tr>
<?php

$servername = "localhost";
$username = "username";
$password = "password";
$dbname = "college";


$conn = new mysqli('localhost', 'root', "", 'college');

if(isset($_POST['search']))
{
$id=$_POST['List1'];
$id1=$_POST['li'];

$sql = "SELECT Name,Enrollment, Percent ,CR FROM $id1 where IC='$id' order by CR ";
 $result = $conn->query($sql) or die($conn->error);

while($row = $result->fetch_assoc())
{
?>
<tr>
<td align="center"><?php echo $row['Name']; ?></td>
<td align="center"><?php echo $row['Enrollment']; ?></td>
<td align="center"><?php echo $row['Percent']; ?></td>
<td align="center"><?php echo $row['CR']; ?></td>
</tr>

<?php
}
```

Code For University Ranking

```
<th colspan="5"><h2 align="center">UNIVERSITY RANKING</h2></th>
</tr>
 <tr>
<th>Name</th>
<th>Enrollment</th>
<th>College Code</th>
<th>Percentage</th>
<th>RANKING</th>
</tr>
<?php

$servername = "localhost";
$username = "username";
$password = "password";
$dbname = "college";


$conn = new mysqli('localhost', 'root', "", 'college');

if(isset($_POST['search']))
{
$id=$_POST['List1'];
$sql = "SELECT Name,Enrollment,IC,Percent ,UR FROM $id ";
 $result = $conn->query($sql) or die($conn->error);

while($row = $result->fetch_assoc())
{
?>
<tr>
<td align="center"><?php echo $row['Name']; ?></td>
<td align="center"><?php echo $row['Enrollment']; ?></td>
<td align="center"><?php echo $row['IC']; ?></td>
<td align="center"><?php echo $row['Percent']; ?></td>
<td align="center"><?php echo $row['UR']; ?></td>
</tr>

<?php
}
```

Code For Aggregate Percentage

```
<tr>
<th colspan="3"><h2 align="center">COLLEGE RANKING</h2></th>
</tr>
 <tr>
<th>Name</th>
<th>Enrollment</th>
<th>Percentage</th>

</tr>
<?php

$servername = "localhost";
$username = "username";
$password = "password";
$dbname = "college";


$conn = new mysqli('localhost', 'root', "", 'college');

if(isset($_POST['search']))
{
$id=$_POST['List1'];


$sql = "SELECT sem1.Enrollment,sem1.Name,(sem1.Percent+sem2.Percent+sem3.Percent+sem4.Percent)/4 as Percent FROM sem1,sem2,sem3,sem4 WHERE sem1.Enrollment=sem2.Enrollment
 AND sem1.Enrollment=sem3.Enrollment AND sem1.Enrollment=sem4.Enrollment AND sem1.IC='$id' ";

 $result = $conn->query($sql) or die($conn->error);

while($row = $result->fetch_assoc())
{
?>
<tr>
<td align="center"><?php echo $row['Name']; ?></td>
<td align="center"><?php echo $row['Enrollment']; ?></td>
<td align="center"><?php echo $row['Percent']; ?></td>
```

Sample Code For Conversion Of PDF To Excel

```
In [6]: table.to_excel("Result4.xlsx") #to export single table.

In [22]: tables.export('Try.csv',f='csv',compress=True) #to export all the tables in different sheets.

In [4]: tables.export('Try1.html',f='html',compress=True)

In [9]: tables.export('Try3.xlsx',f='excel',compress=True)

In [5]: tables.export('Try6.csv',f='csv',compress=True) #to export all the tables in different sheets.(flavour=stream)

In [3]: tables.export('result101',f='sqlite')

In [14]: camelot.plot(tables[1],kind='text')
         plt.show()
```

Sample Code For Concatenation of Excel Sheets

In [7]:
```
import pandas as pd
import os
import glob
import csv
# import numpy as np
```

In [15]:
```
path=r'C:\Users\TheBOT\Python Workspace\PDF Extraction\Test to combine'

filenames = glob.glob(path + "/*.csv")
print(filenames)

# Dataframe Initialization
concat_all_sheets_all_files = pd.DataFrame()

for file in filenames:
    df = pd.read_csv(file)
    testlist=[df]
#     print(testlist)
    concat_all_sheets_single_file = pd.concat(testlist,sort=False)
    concat_all_sheets_all_files=concat_all_sheets_all_files.append(concat_all_sheets_single_file)
```

Sample Code For Cleaning Of Excel

```
Sub remove_empty_lines()
Dim start_time, arr(), rng, row_num As Long, column_num As Long, I, j
    Application.ScreenUpdating = False
    row_num = ActiveSheet.UsedRange.Rows.Count
    column_num = ActiveSheet.UsedRange.Columns.Count
    ReDim arr(1 To row_num)
    rng = ActiveSheet.UsedRange
    start_time = Timer
    For I = 1 To row_num
       For j = 1 To column_num
        If rng(I, j) <> "" Then arr(I) = I
     Next j, I
    With Cells(ActiveSheet.UsedRange.Row, ActiveSheet.UsedRange.Column + column_num).Resize(row_num, 1)
    .Value = WorksheetFunction.Transpose(arr)
    ActiveSheet.UsedRange.Sort Key1:=Cells(ActiveSheet.UsedRange.Row, ActiveSheet.UsedRange.Column + column_num)
    .Value = ""
   End With
    Application.ScreenUpdating = True
    MsgBox Format(Timer - start_time, "0.00s")
End Sub
```

4.2TESTING

Software testing is a process, to evaluate the functionality of a software application with an intent to find whether the developed software met the specified requirements or not and to identify the defects to ensure that the product is defect free in order to produce the quality product.

Testing Approaches:

There are three types of software testing approaches.

1. White Box Testing
2. Black Box Testing
3. Grey Box Testing

White Box Testing: It is also called as Glass Box, Clear Box, Structural Testing. White Box Testing is based on applications internal code structure. In white-box testing, an internal perspective of the system, as well as programming skills, are used to design test cases. This testing is usually done at the unit level.

Black Box Testing: It is also called as Behavioural/Specification-Based/Input-Output Testing. Black Box Testing is a software testing method in which testers evaluate the functionality of the software under test without looking at the internal code structure.

Grey Box Testing: Grey box is the combination of both White Box and Black Box Testing. The tester who works on this type of testing needs to have access to design documents. This helps to create better test cases in this process.

Types of Black Box Testing:

1. Functionality Testing
2. Non-functionality Testing

Functional testing:

In simple words, what the system actually does is functional testing. To verify that each function of the software application behaves as specified in the requirement document. Testing all the functionalities by providing appropriate input to verify whether the actual output is matching the expected output or not. It falls within the scope of black box testing and the testers need not concern about the source code of the application.

Non-functional testing:

In simple words, how well the system performs is non-functionality testing. Non-functional testing refers to various aspects of the software such as performance, load, stress, scalability, security, compatibility etc., Main focus is to improve the user experience on how fast the system responds to a request.

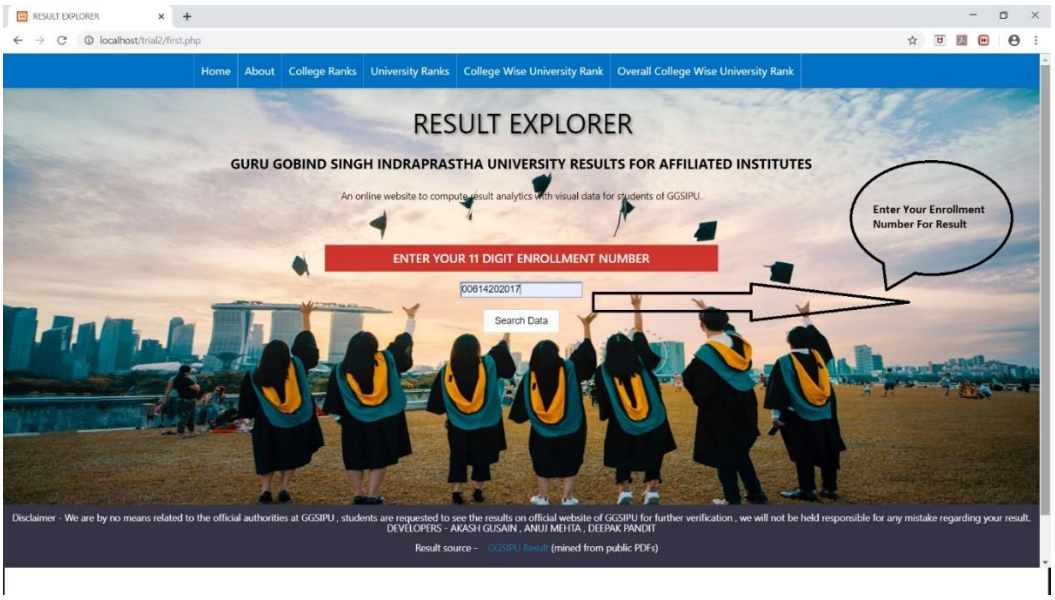
4.2.1 Overview And Approach

Out of several testing approaches we did functional testing in all the functionality of website.

The results were satisfactory, website responded to every functionality the way it should do and hence fulfils the purpose and satisfies every test case.

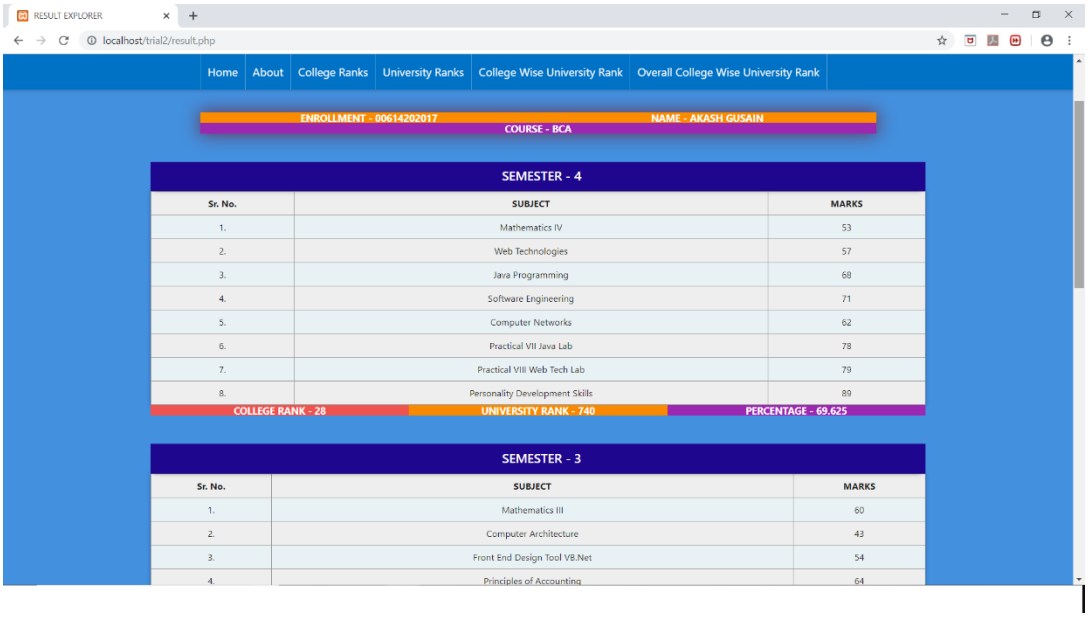
Test Sample is Given Below

Enter Your Enrolment Number To Get Result



Student's Individual Result

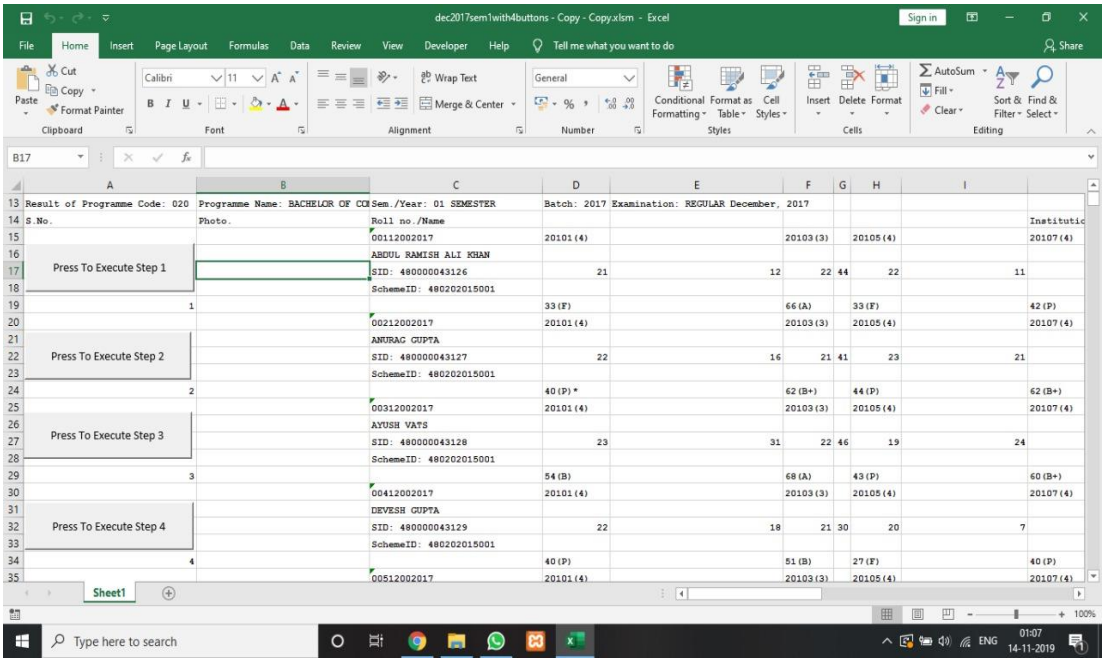
Subject wise marks were verified from the original PDF and testing was successful.



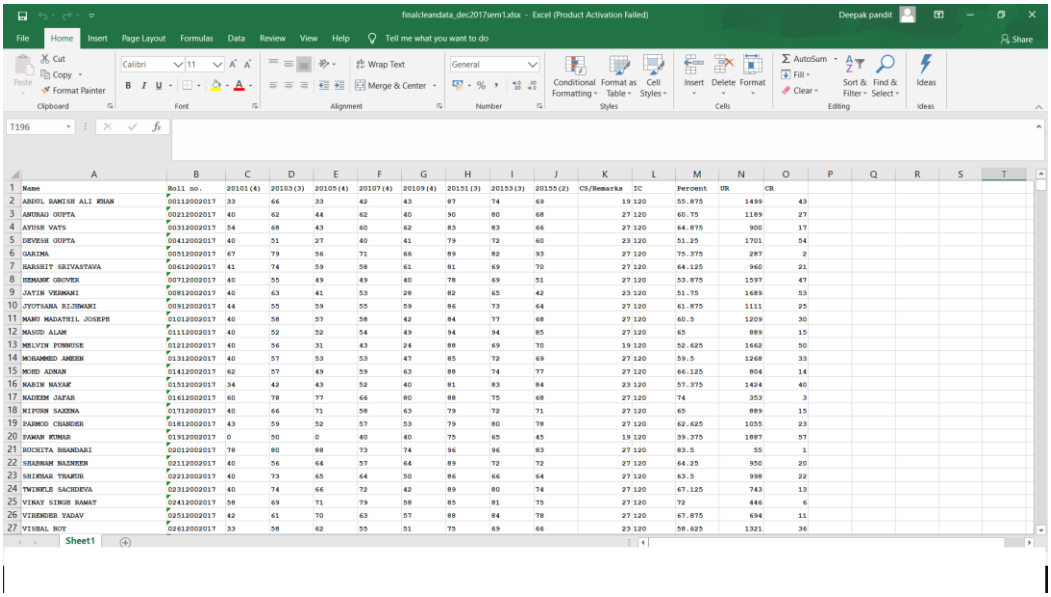
5 OUTPUT FORMS AND REPORTS

The input and output forms of website are displayed below

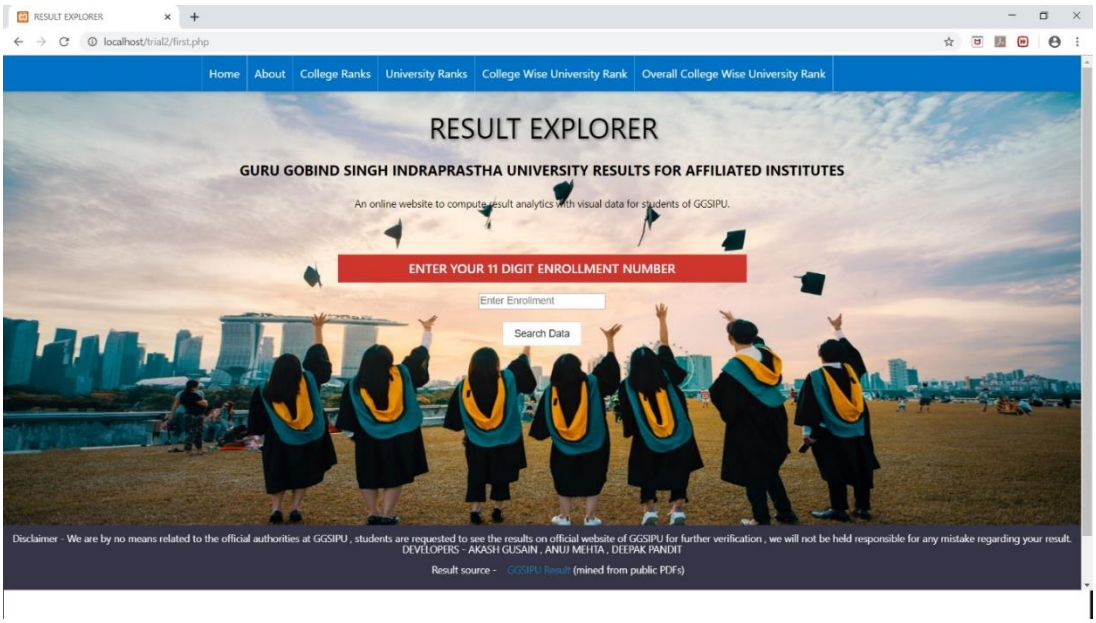
Uncleaned Excel With Macros



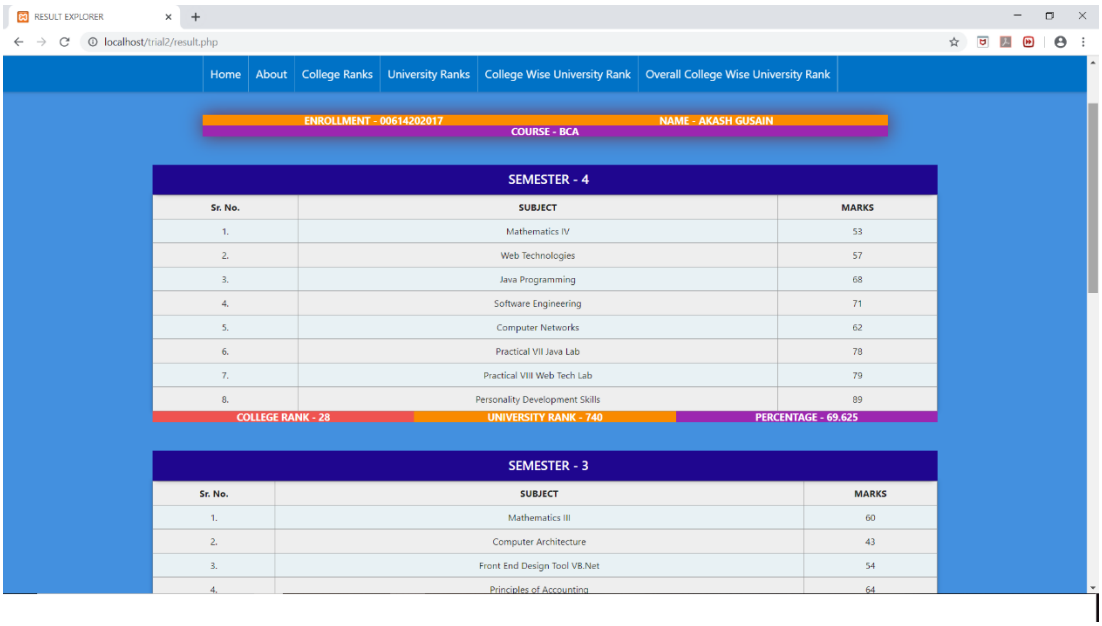
After performing all the cleaning, the desired excel looks in this format



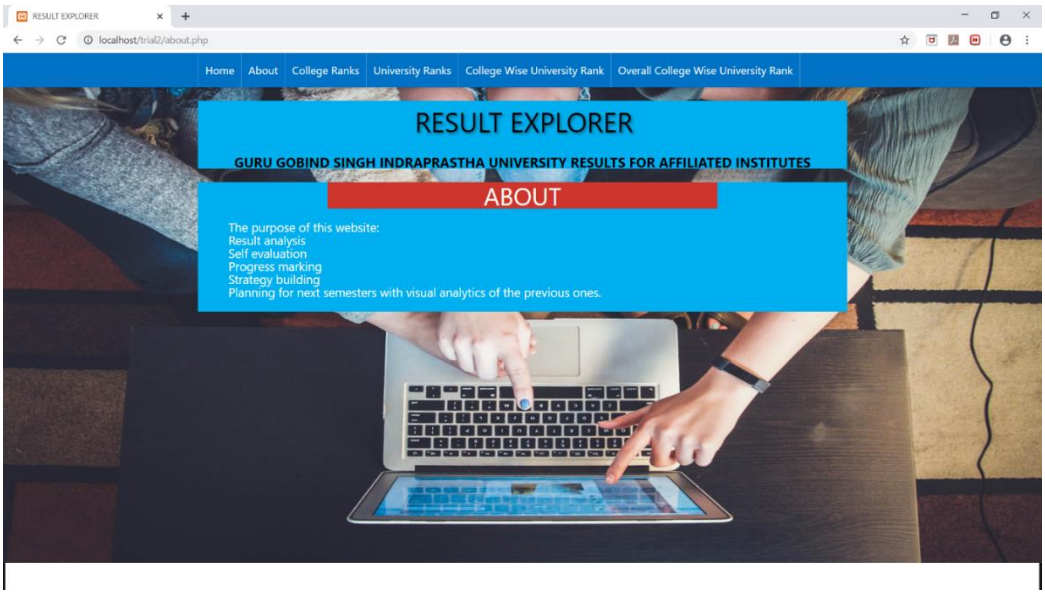
Front Page of The Website



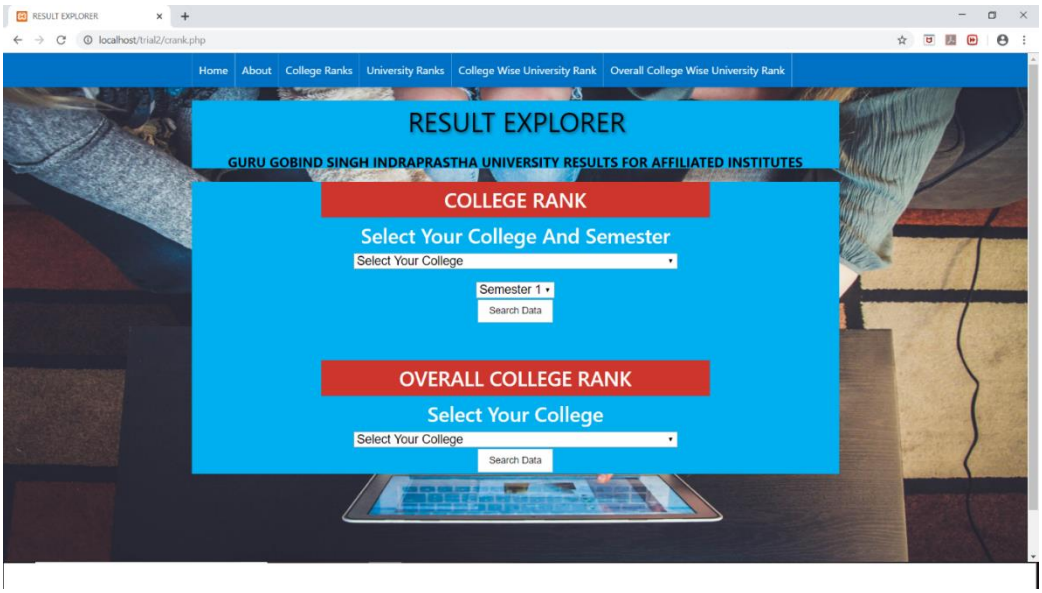
Semester Wise Result of Individual Student



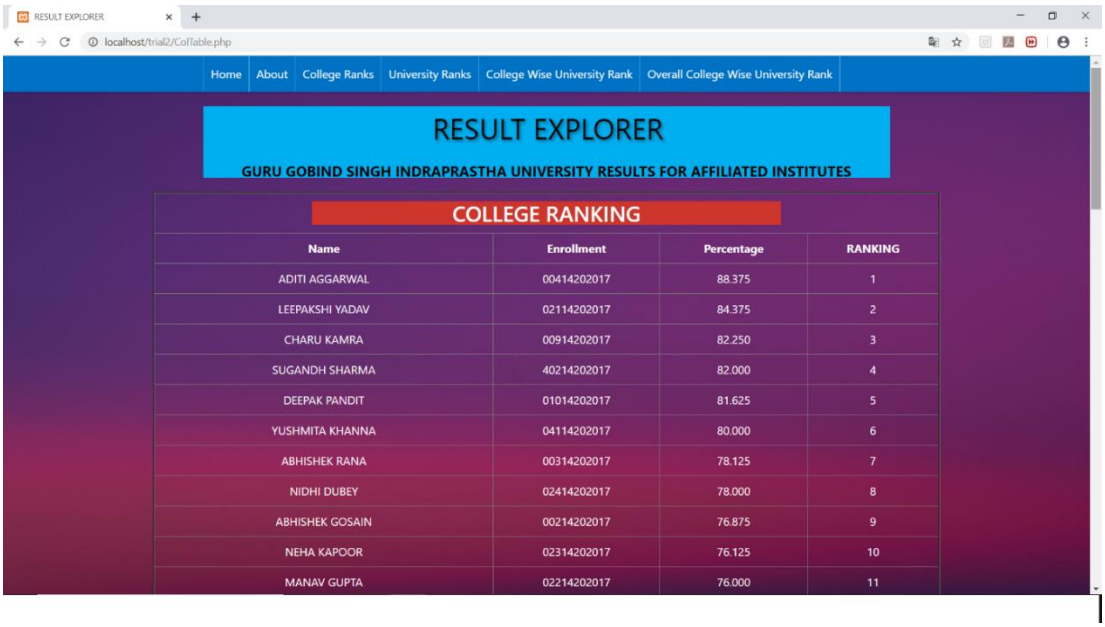
About Page Of The Web Based Application



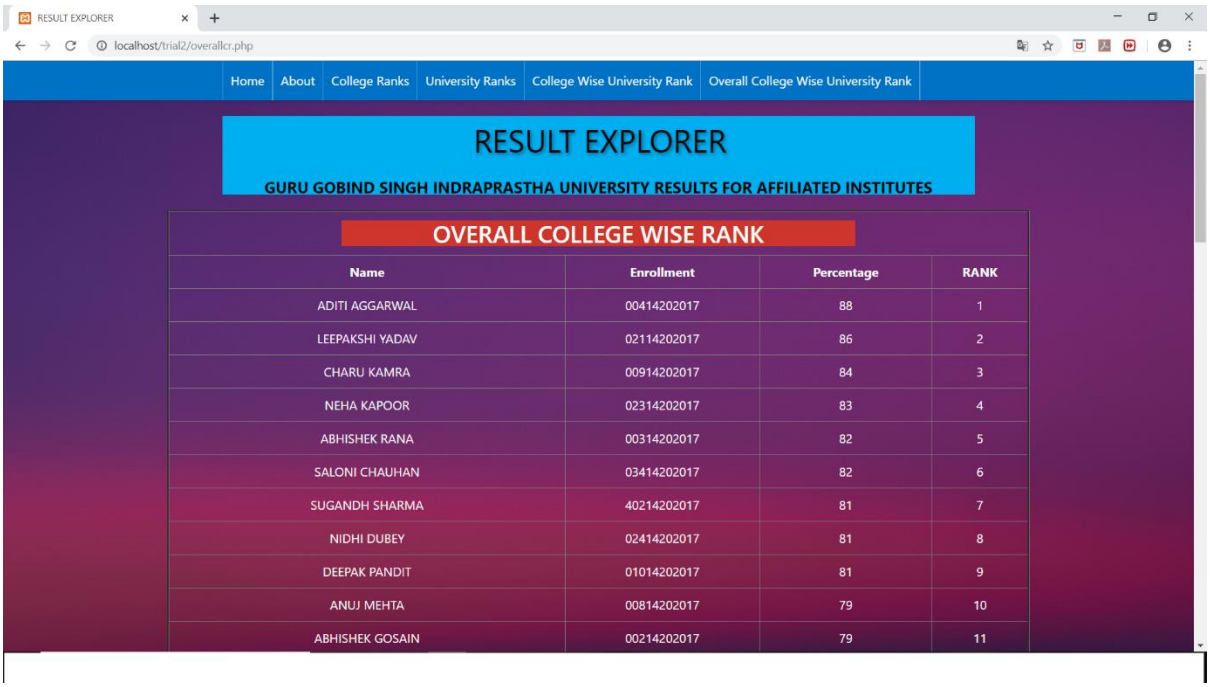
College Ranks



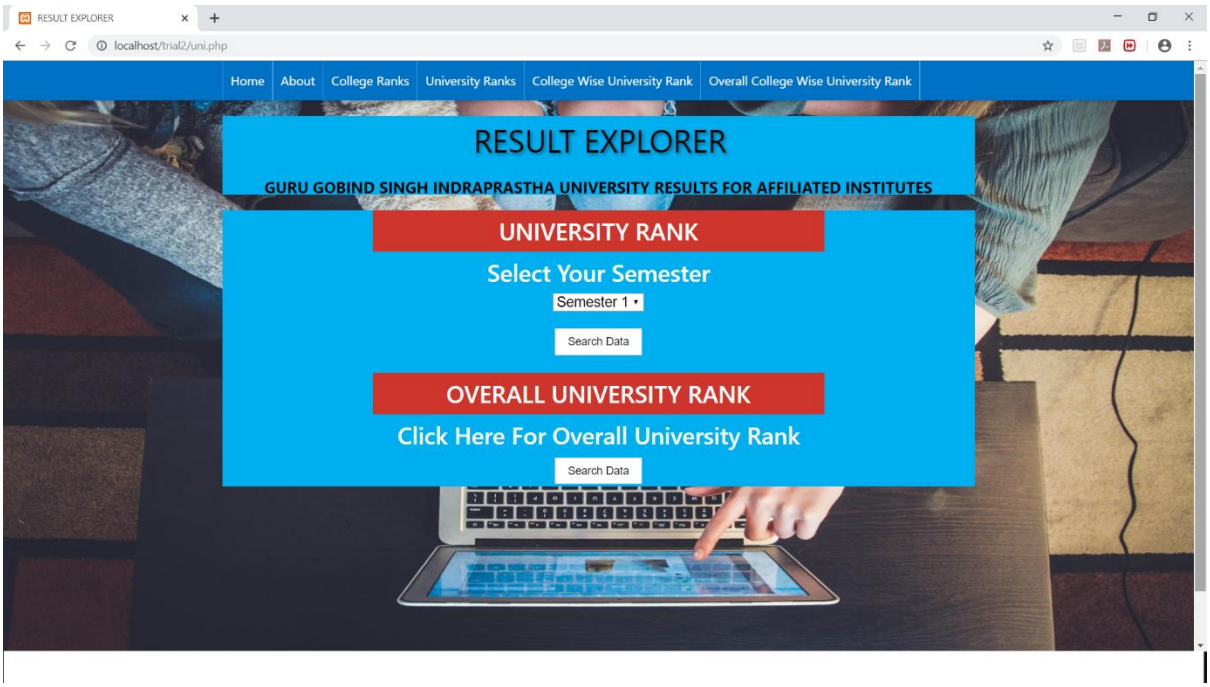
Semester Wise College Rank



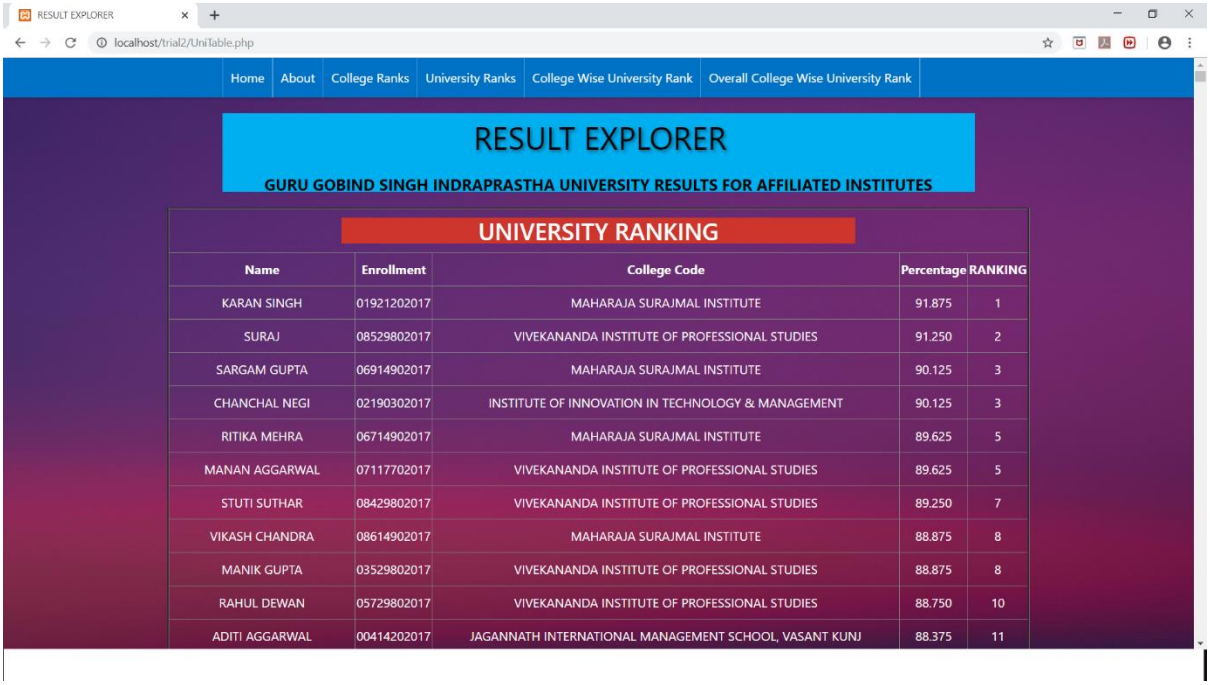
Overall College Rank



University Ranks

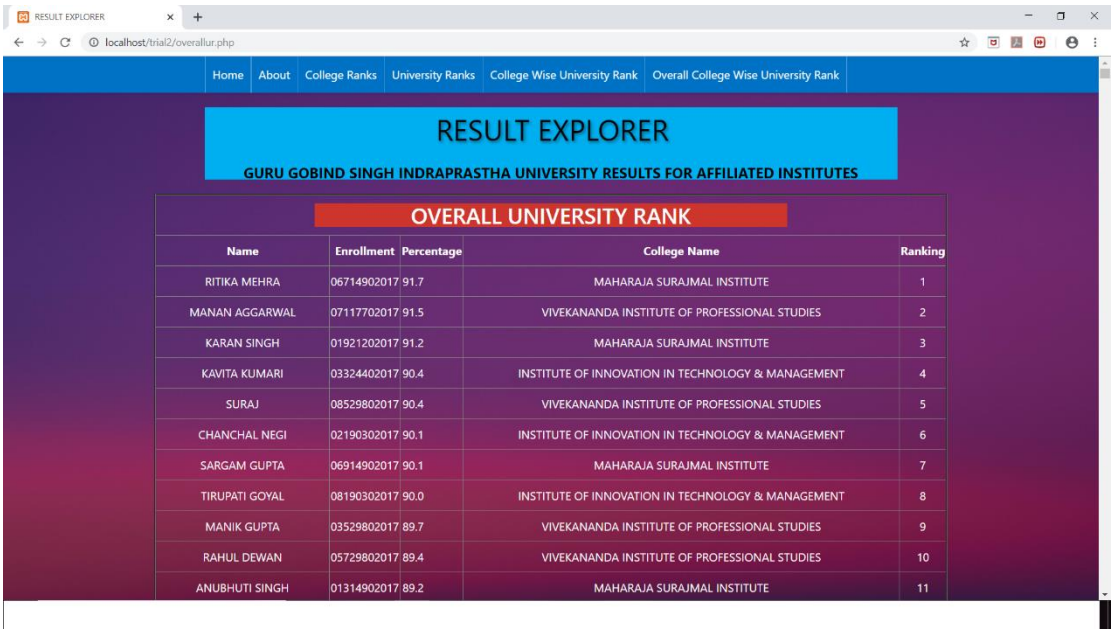


Semester Wise University Rank



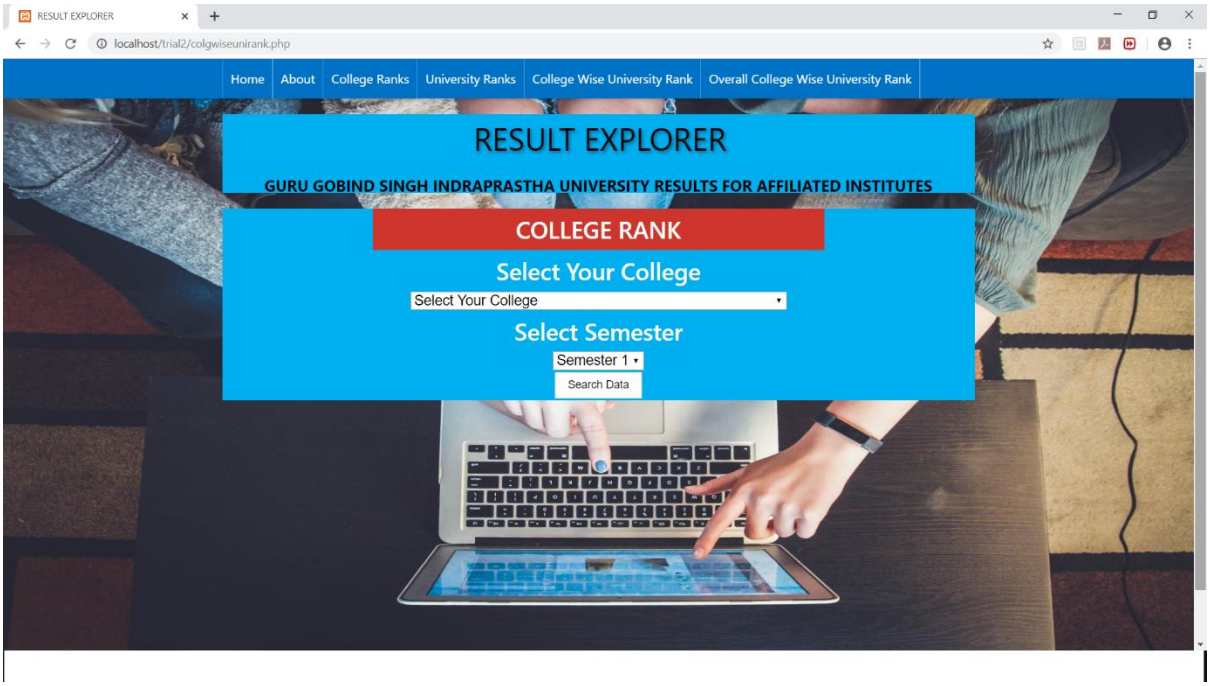
Name	Enrollment	College Code	Percentage	RANKING
KARAN SINGH	01921202017	MAHARAJA SURAJMAL INSTITUTE	91.875	1
SURAJ	08529802017	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	91.250	2
SARGAM GUPTA	06914902017	MAHARAJA SURAJMAL INSTITUTE	90.125	3
CHANCHAL NEGI	02190302017	INSTITUTE OF INNOVATION IN TECHNOLOGY & MANAGEMENT	90.125	3
RITIKA MEHRA	06714902017	MAHARAJA SURAJMAL INSTITUTE	89.625	5
MANAN AGGARWAL	07117702017	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	89.625	5
STUTI SUTHAR	08429802017	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	89.250	7
VIKASH CHANDRA	08614902017	MAHARAJA SURAJMAL INSTITUTE	88.875	8
MANIK GUPTA	03529802017	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	88.875	8
RAHUL DEWAN	05729802017	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	88.750	10
ADITI AGGARWAL	00414202017	JAGANNATH INTERNATIONAL MANAGEMENT SCHOOL, VASANT KUNJ	88.375	11

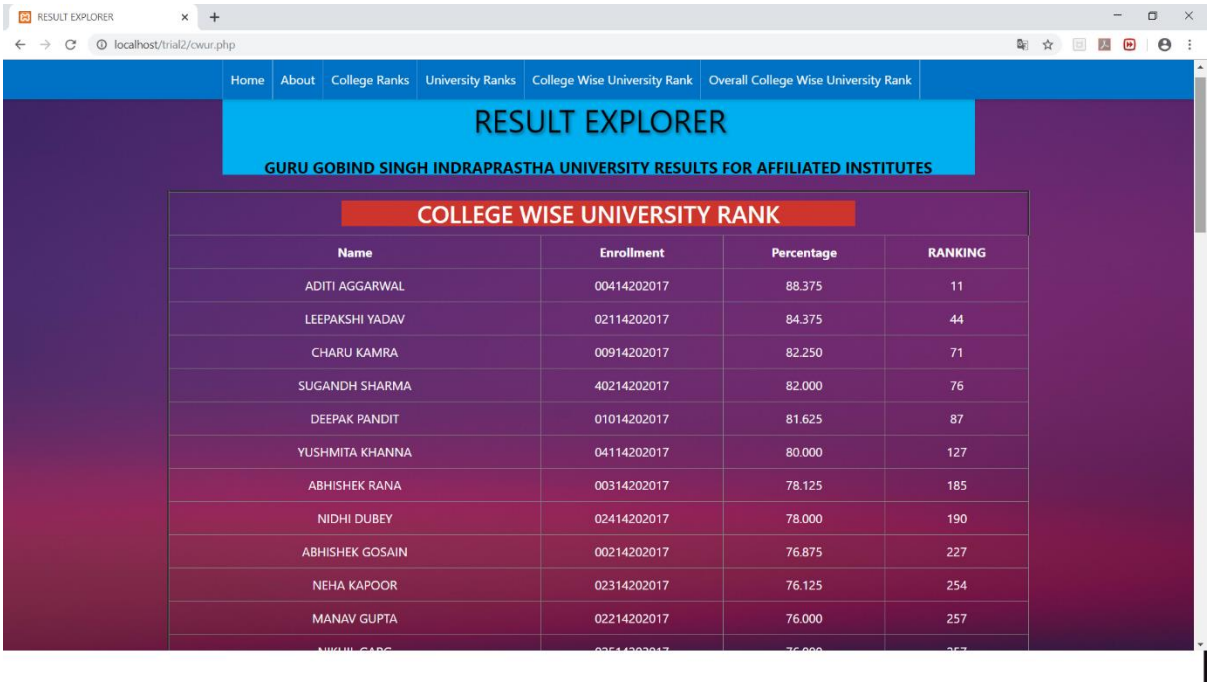
Overall University Rank



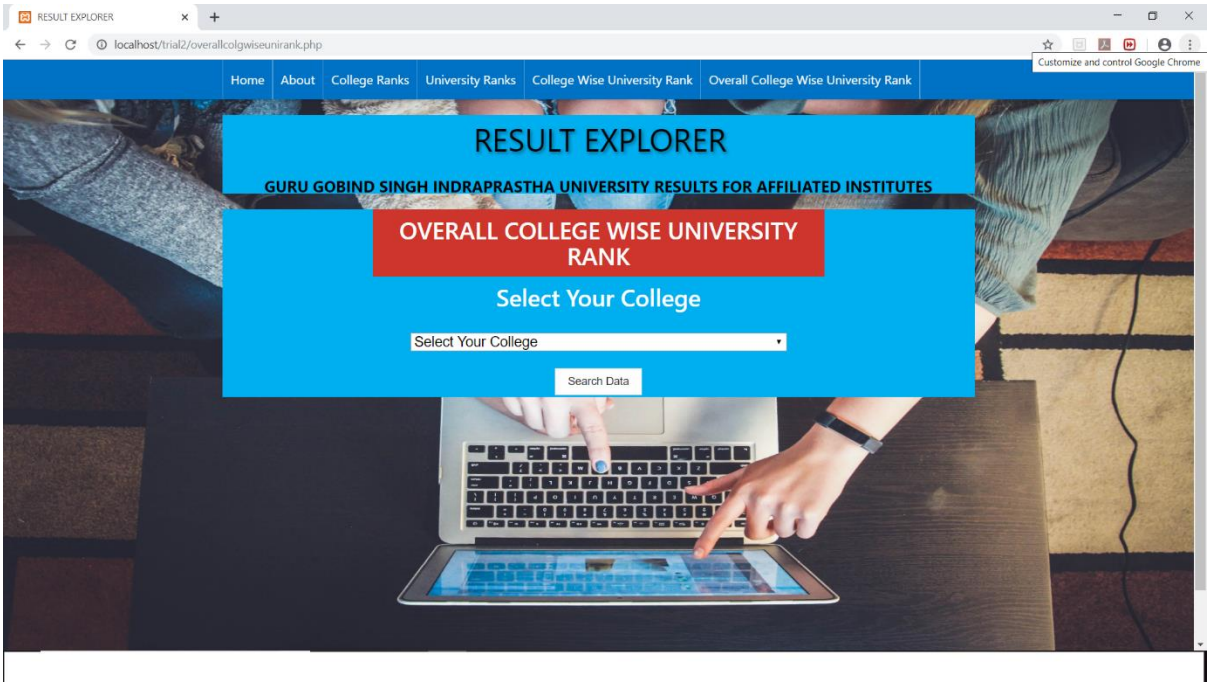
Name	Enrollment	Percentage	College Name	Ranking
RITIKA MEHRA	06714902017	91.7	MAHARAJA SURAJMAL INSTITUTE	1
MANAN AGGARWAL	07117702017	91.5	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	2
KARAN SINGH	01921202017	91.2	MAHARAJA SURAJMAL INSTITUTE	3
KAVITA KUMARI	03324402017	90.4	INSTITUTE OF INNOVATION IN TECHNOLOGY & MANAGEMENT	4
SURAJ	08529802017	90.4	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	5
CHANCHAL NEGI	02190302017	90.1	INSTITUTE OF INNOVATION IN TECHNOLOGY & MANAGEMENT	6
SARGAM GUPTA	06914902017	90.1	MAHARAJA SURAJMAL INSTITUTE	7
TIRUPATI GOYAL	08190302017	90.0	INSTITUTE OF INNOVATION IN TECHNOLOGY & MANAGEMENT	8
MANIK GUPTA	03529802017	89.7	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	9
RAHUL DEWAN	05729802017	89.4	VIVEKANANDA INSTITUTE OF PROFESSIONAL STUDIES	10
ANUBHUTI SINGH	01314902017	89.2	MAHARAJA SURAJMAL INSTITUTE	11

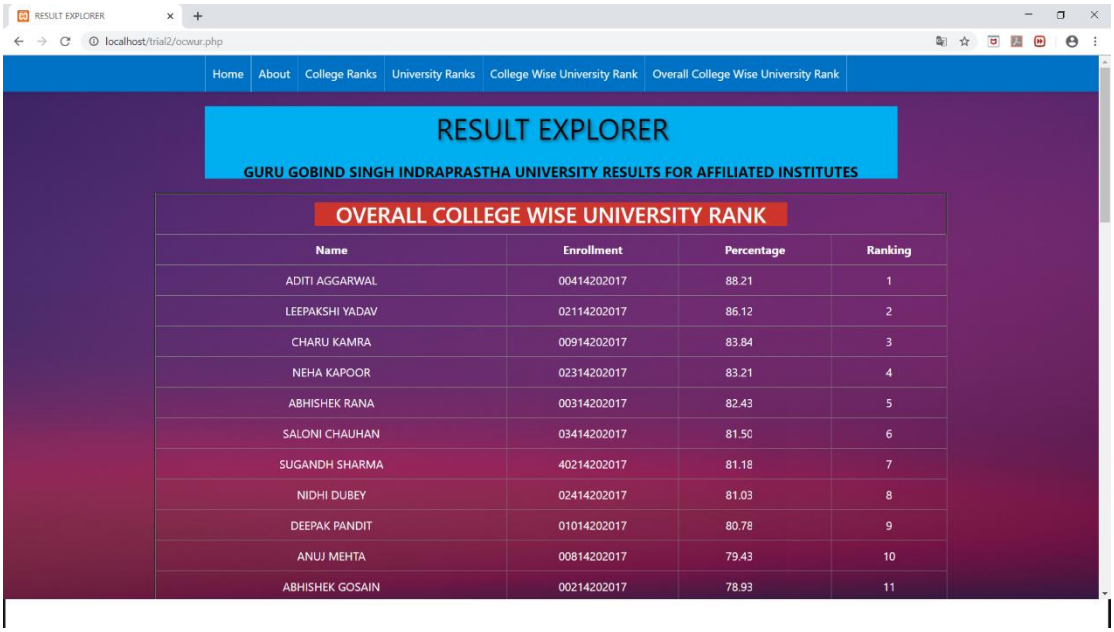
College wise University Rank





Overall College Wise University Rank




RESULT EXPLORER
localhost/trial2/ocwur.php
Home
About
College Ranks
University Ranks
College Wise University Rank
Overall College Wise University Rank
RESULT EXPLORER
GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY RESULTS FOR AFFILIATED INSTITUTES
OVERALL COLLEGE WISE UNIVERSITY RANK
Name
Enrollment
Percentage
Ranking
ADITI AGGARWAL 00414202017 88.21 1
LEEPAKSHI YADAV 02114202017 86.12 2
CHARU KAMRA 00914202017 83.84 3
NEHA KAPOOR 02314202017 83.21 4
ABHISHEK RANA 00314202017 82.43 5
SALONI CHAUHAN 03414202017 81.50 6
SUGANDH SHARMA 40214202017 81.18 7
NIDHI DUBEY 02414202017 81.03 8
DEEPAK PANDIT 01014202017 80.78 9
ANUJ MEHTA 00814202017 79.43 10
ABHISHEK GOSAIN 00214202017 78.93 11

6 CONCLUSION & FUTURE ENHANCEMENT(S)

6.1 CONCLUSION

In this project a website is developed, and all the functionalities were achieved successfully as per the requirement of the user.

Following are the achieved result:

- Conversion of PDF format of result into useful cleaned excel that can be used for analysis.
- Importing the excel sheet of data into our MYSQL Database
- Developed a Web based Application for user with the following functionalities:
    - Semester wise ranking is provided of different colleges to the user.
    - User can also view the semester wise rank of all the students studying in the university
    - In order to compare overall performance, user can also get their aggregate percentages of all the semesters.

6.2 FUTURE SCOPE

Like any other project, this project can also have some scope for better implementation, we can enhance its functionalities like:

- This project can be further scaled for different batches of different courses that are offered by GGSIPU like BBA, BAJMC etc.
- For a better experience and to target more number of audience the functionality of this website can be implemented in an android application.

BIBLIOGRAPHY

- https://www.w3schools.com/php/
- https://stackoverflow.com/questions/20407699/php-connect-to-mysql-database
- https://www.geeksforgeeks.org/php
- https://www.wikipedia.org/
- https://www.youtube.com/
- https://www.javatpoint.com/