School Of Information Science

(A Constituent Institute of Manipal University)



SOIS Students’ Placement Analysis

Sourav K C- 171046024

Anudeep B N- 171046039

Course: ME-Big Data and Data Analytics

Guided by:

Mr.Samarendranath B

Assistant Professor

**Index**

1. Abstract
2. Introduction
3. Specifications
4. Block Diagram
5. Modules
6. Analysis
7. Results
8. Scope for further work
9. Bibliography

**Abstract**

In this project we collected the data from SOIS placement department of the college. We clean the dataset to the requirements and we analysis the data based on branch, Companies, student placements on branch and strength and number of placements that took place in SOIS department. Using python technologies, we analysis and visualise from the given data.

**Introduction**

The age of big data is now coming. But the traditional data analytics may not be able to handle such large quantities of data. The question that arises now is, how to develop a high-performance platform to efficiently analyse the big data and how to develop an appropriate model. To deeply discuss this issue, this project begins with a data collection, data analysis, followed by data visualization.

We gathered the required data set through the placement office of the SOIS department. For the analysis and visualization, we clean the collected dataset into the requirements. We took some of the major questions based on the dataset given and we analysed it. And for the visualization part we plotted some pie and bar graph on placements percentage that took place in SOIS department.

**Specifications**

We are Analysing and visualising the placement data of the SOIS department.

Software Requirements:

1. Python:

Python is a high-level, interpreted, interactive and object-oriented scripting language. Python is designed to be highly readable. It uses English keywords frequently where as other languages use punctuation, and it has fewer syntactical constructions than other languages.

**Python is Interpreted** − Python is processed at runtime by the interpreter. You do not need to compile your program before executing it. This is similar to PERL and PHP.

**Python is Interactive** − you can actually sit at a Python prompt and interact with the interpreter directly to write your programs.

**Python is Object-Oriented** − Python supports Object-Oriented style or technique of programming that encapsulates code within objects.

**Python is a Beginner's Language** − Python is a great language for the beginner-level programmers and supports the development of a wide range of applications from simple text processing to WWW browsers to games.

2. Pandas:

Pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming language.

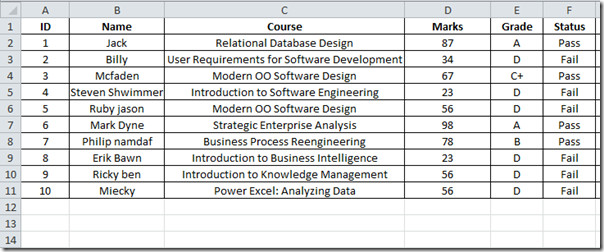
3. Matplotlib (Plotting):

Matplotlib is a Python 2D plotting library which produces publication quality figures in a variety of hardcopy formats and interactive environments across platforms. Matplotlib can be used in Python scripts, the Python and [IPython](http://ipython.org/) shell, the [jupyter](http://jupyter.org/index.html) notebook, web application servers, and four graphical user interface toolkits.

**Block Diagram**

Available Data

Analysing Data



Python



Visualization of Data

Fig: Architecture Flow of the Project

We gathered the required data set through the placement office of the SOIS department. For the analysis and visualization, we clean the collected dataset into the requirements by python script. We took some of the major questions based on the dataset given and we analysed it. And for the visualization part we plotted some pie and bar graph by using python packages like Numpy, Pandas and Matplotlib on placements percentage that took place in SOIS department.

Modules:

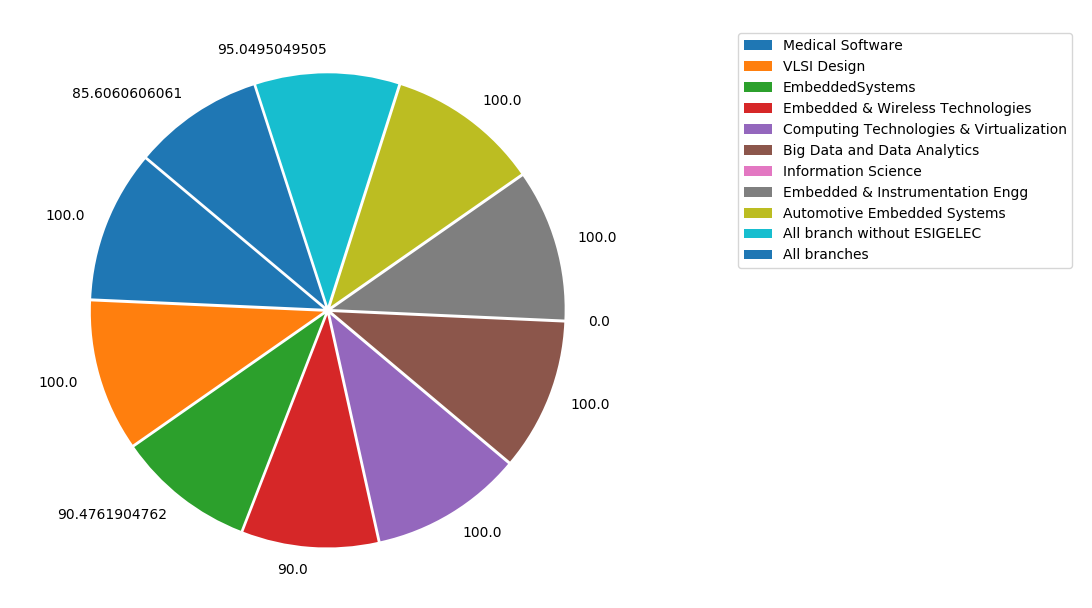


Fig 1: Pie chart for Percentage of Placement at SOIS in the year-2016

Fig 1 represents the Percentage of Placement that took place at SOIS in the year 2016.Here the different colours in the pie chart shows the ME branches of SOIS department , where the percentage of the students’ placed in different companies.

The chart shows that, the Medical Software branch placed is 100%, VLSI design got placed is 100%, Embedded Systems got placed is 90.47%, Embedded and Wireless Technologies got placed is 90%, Computing Technologies and Visualization got placed is 100%, Big Data and Data Analysis got placed is 100%, Information Science got placed is Zero because of Zero students, All branches without ESIGELEC got placed is 95.049% and All branches of SOIS department got placed is 85.606%.

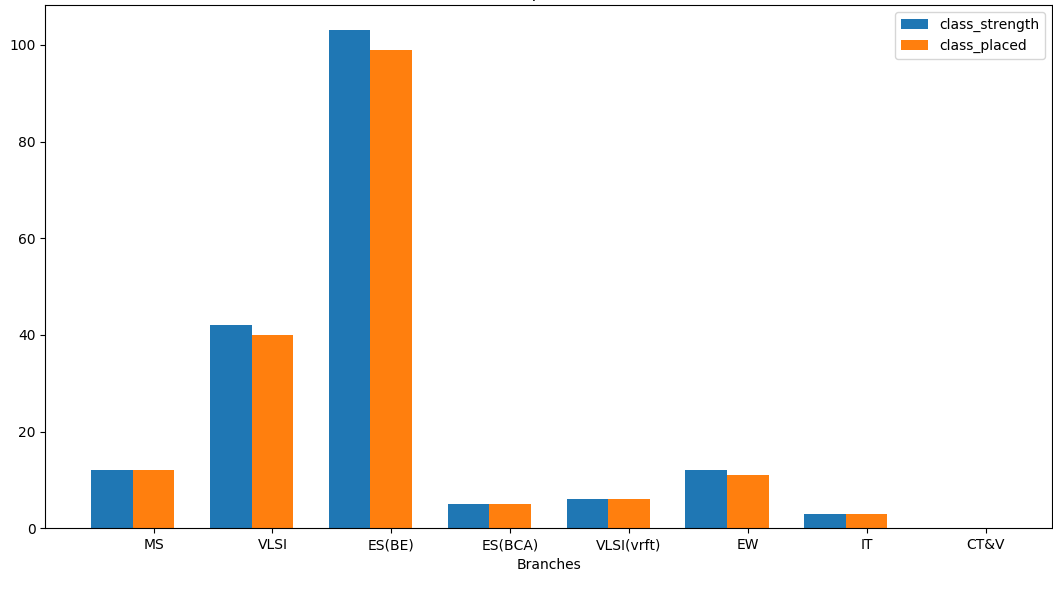


Fig 2: Placement of all the branch\_class in 2013

Fig 2 represents the subplot graphs, where blue bar chart represents the class strength of different branches and orange reprents the students placed with respect to the class strength. Where x-axis represents the different branches of SOIS and y-axis represents the percentage of the class strength and percentage of students placed.

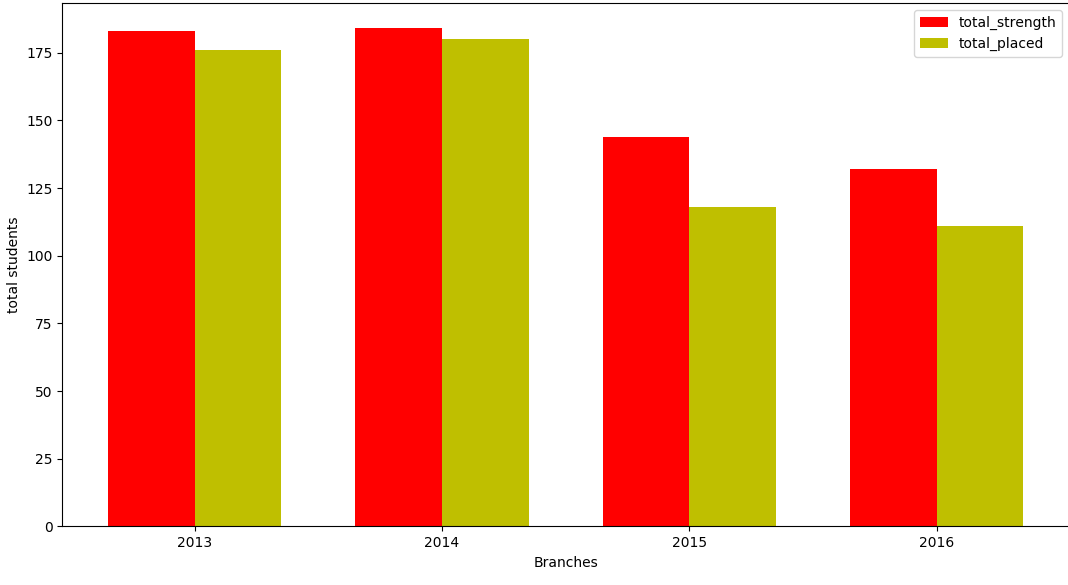


Fig 3: Placements that took in SOIS by Year

Fig 3 represents the subplot graphs, where red bar chart represents the class strength of different branches with respect to year and yellow represents the students placed with respect to the class strength of the particular year. Where x-axis represents the different year and y-axis represents the percentage of the student strength and percentage of students placed on particular year.

Analysis

We *Analysed* the SOIS placement dataset by putting up some of the questions

1. Display all the list of companies of particular\_branch\* in the past five years?

2. List out all the students who got placed in particular\_company\* in VLSI in 2015

3. List out required data of Bigdata-2016

4. List out all the students B.E aggregate, who hired by particular\_company\* in 2012 batch

5. How many number of students are ESIGELEC in the year 2014?

6. What is the percentage of students got placed in Embedded Systems in 2016?

7. What is the percentage of placement in all branches without\* ESIGELEC in 2015?

8. Display all companies that are attending SOIS department on particular\_year\*.

We *Visualized* the SOIS placement dataset by putting up some of the questions

1. Show the Visualization of Percentage of placement of all branches in 2016

2. Show the Visualization for the placement in all the branch class in 2013

3. Display statistics all year of placements took in SOIS department

Results:

We got output for Analysis and Visualization that we had done on placement dataset. On the basis on Analysis and Visualization, we can predict the total placements that took place in SOIS department since past 5 years.

**Contribution of mine:**

All Analysis and visualization part is done by myself.

Scope for further work:

This type of Analysis and Visualization can be used in health care, companies etc.

For the better visualization we can implement D3 (Data Driven Documents), so that it will visualised on Web, the users can access it through the web.

Bibliography

1. The Grammar of Graphics by Leland Wilkinson, Springer
2. Ten Simple Rules for Better Figures by Nicolas P. Rougier, Michael Droettboom, Philip E. Bourne, <https://hal.inria.fr/hal-01063732>
3. A survey of powerful visualization techniques, from the obvious to the obscure by Jeffrey Heer, Michael Bostock, and Vadim Og ievetsky
4. Interactive Data Visualization for the web by Scott Murray.
5. Designing Data Visualizations by Noah Iliinsky and Julie Steele