We can use the Gale – Shapley algorithm for this problem.

We have a total of 300 projects to be distributed among 1200 students. So, a stable distribution is not possible without a project being given to multiple students or to teams of students.

For the sake of simplicity, let us assume a project is being given to 4 individual students. Each project must have some prerequisites for the students. This can be analysed using the multiple tests taken. These tests will help in defining the order of preference of a particular project towards multiple students.

The preference order of the students for the projects can be based solely on the interests of the students.

For simplification, we should first start with a smaller data set, i.e. with 3 Projects and 6 students. Such that each Project is assigned to 2 students.

Let us denote the Projects as X, Y and Z. And the students as A, B, C, D, E and F.

Let us also define an arbitrary preference order for both the students and the Projects.

|  |  |  |  |
| --- | --- | --- | --- |
|  | X | Y | Z |
| A | 4 | 2 | 6 |
| B | 2 | 5 | 3 |
| C | 5 | 1 | 1 |
| D | 3 | 4 | 5 |
| E | 6 | 3 | 2 |
| F | 1 | 6 | 4 |

Preference of Projects towards students

|  |  |  |  |
| --- | --- | --- | --- |
|  | X | Y | Z |
| A | 1 | 3 | 2 |
| B | 2 | 1 | 3 |
| C | 1 | 2 | 3 |
| D | 1 | 2 | 3 |
| E | 3 | 1 | 2 |
| F | 2 | 3 | 1 |

Preference of Students towards Projects

Now, to solve this problem we should stimulate the Gale – Shapley Algorithm on this dataset.

**CASE I –**

* Students apply for projects. Each student applies for a project which is his/her first preference.
* A Project only accepts the top two candidates according to its preference of students and rejects the rest.
* The rejected students reapply to the next project according to their preference.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | Y | Z |  | X | Y | Z |  | X | Y | Z |  | X | Y | Z |
| A  C  D | B  E | F |  | A  D | B  E  C | F |  | A  D  B | C  E | F |  | B  D | C  E | F  A |

Analysing the results, we note that in this case, A got his/her 2nd choice, B got his/her 2nd choice, C got his/her 2nd choice, D got his/her 1st choice, E got his/her 1st choice and F got his/her 1st choice. Conclusively when the students apply for the projects, the matching/allotment is in favour of the students.

**CASE II –**

* Projects invite the top two students according to its preference.
* A student in case of multiple invitations only accepts the invitation of the project which has a higher preference according to him/her.
* The rejected projects now invite the next student according to its preference

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A | Y |  | A | Y |  | A | Y |  | A | Y |
| B | X |  | B | X Z |  | B | X |  | B | X |
| C | Y Z |  | C | Y |  | C | Y |  | C | Y |
| D |  |  | D |  |  | D |  |  | D | X |
| E | Z |  | E | Z |  | E | Z |  | E | Z |
| F | X |  | F | X |  | F | X Z |  | F | Z |

Analysing the distribution – X got B and D which are its 2nd and 3rd choices, Y got A and C which are its 1st and 2nd choices and Z got E and F which are its 2nd and 4th choices. In this case, the distribution is in favour of the preference of the projects.

Similar to this small dataset of 3 Projects X, Y and Z and 6 students, A, B, C, D, E and F, we can very easily expand this dataset to the actual data of 300 Projects and 1200 students.