This is a collection of 150 individual student records that have seven characteristics that describe the positive and negative points of the academic record and presence of each student. Here is the information:

Student ID – A single number assigned to each student to uniquely identify the student.

First Name & Last Name – The student's whole name.

Marks – The amount of points got by the student in an exam (some entries are absent, and many values are distinctly incorrect, as they are above 100).

Email ID – The student's email address (a few missing entries) and many have invalid syntax.

Attendance (%) – The proportion of classes the student participated in (some missing values).

Status – Shows if the student is "Pass" or "Fail" based on their grades and attendance (some missing entries).

Subject- Shows the subjects students have opted(marks are written for these subjects only) (some missing values).

City- The city where the student lives.

Valid Email (a newly created column in cleaned_data file)- Tells whether the format of email is valid or invalid.

Initial Observations: The dataset is mostly complete but has a few missing values in Marks, Email ID, Attendance, and Status.There's one unrealistic Marks entry (110) that might need correction.The "Status" column suggests a pass/fail system, but the criteria used to decide this aren't provided. The format for Email is written wrong in many rows.