Today’s completed work

1. **Dataset:**

   description of columns:

   Student's sex (binary: 'F' - female or 'M' - male)

   Student's age (numeric: from 15 to 22)

   Student's home address type (binary: 'U' - urban or 'R' - rural)

   Family size (binary: 'LE3' - less or equal to 3 or 'GT3' - greater than 3)

   Parent's cohabitation status (binary: 'T' - living together or 'A' - living apart)

   Mother's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education, or 4 - higher education)

   Father's education (numeric: 0 - none, 1 - primary education (4th grade), 2 - 5th to 9th grade, 3 - secondary education, or 4 - higher education)

   Mother's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')

   Father's job (nominal: 'teacher', 'health' care related, civil 'services' (e.g. administrative or police), 'at_home' or 'other')

   Reason to choose this school (nominal: close to 'home', school 'reputation', 'course' preference or 'other')

   Student's guardian (nominal: 'mother', 'father' or 'other')

   Home to school travel time (numeric: 1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)

   Weekly study time (numeric: 1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)

   Number of past class failures (numeric: n if 1<=n<3, else 4)

   Extra educational support (binary: yes or no)

   Family educational support (binary: yes or no)

   Extra-curricular activities (binary: yes or no)

   Attended nursery school (binary: yes or no)

   Wants to take higher education (binary: yes or no)

   Internet access at home (binary: yes or no)

   With a romantic relationship (binary: yes or no)

   Quality of family relationships (numeric: from 1 - very bad to 5 - excellent)

   Free time after school (numeric: from 1 - very low to 5 - very high)

   Going out with friends (numeric: from 1 - very low to 5 - very high)

   Workday alcohol consumption (numeric: from 1 - very low to 5 - very high)

   Weekend alcohol consumption (numeric: from 1 - very low to 5 - very high)

   Current health status (numeric: from 1 - very bad to 5 - very good)

Number of school absences (numeric: from 0 to 93)

2. **Feature extraction:** The relation between different parameters was identified using correlation & count plot.
3. **Analysis:** One of the identified relations was between (internet at home, past failure of students) and its effect on study time.
4. **Algorithms:** Naive Bayes and decision trees were used for its analysis.
5. **Data Visualization:**
   Comparison table between algorithms:

Algorithms	Accuracy score
1. Naive Bayes	75%
2. Decision tree	Between 70% - 87%

Tomorrow's task

1. To carry on the same analysis between (daily alcohol consumption, weekly alcohol consumption, health, family relationship & support) and study time.(Relation between them was determined by correlation)
2. To develop a user interface.