For level 1, we had to see what Feature 1,2 and 3 represent so to start I plotted a histogram of values as well as a boxplot for all the features to see their individual distribution. I also plotted a heatmap to check correlation with all numerical features.

1) Feature 1 had values belonging between 15 and 22 both inclusive, and had a positive correlation with failures so I concluded that this feature represents age of the students since the range belonged to a typical secondary school student and older students would have higher failures showing why they still are in secondary school.

Before making any immediate conclusion for Feature 2 and 3, I added a new column having sum of G1 and G2 since I thought it might give further insight. I made a pairplot to further visualize the dependance on other features and I plotted boxplots for both these features with all the other columns.

2) I observed that Feature 2 had a direct positive dependance on Fedu (slightly), G1,G2 and G3, a negative dependance on absences, and Feature 3(slightly), and had values between 1 and 4, from which I then came to a conclusion that Feature 2 represents some sort of a GPA of the students (no student has a perfect GPA of 5).
3) I observed that Feature 3 had a negative dependance on nursery, higher and Feature 2(slightly), a positive dependance on go out, Dalc and freetime (slightly) and had values between 1 and 5. From this I concluded that Feature 3 shows students' likelihood of having alcohol with their friends from 5 being most likey to 1 being least likely.

For level 2, first I checked which columns had null values and how many null values each of them had. The number of null values present were at a maximum 10% of the total rows so we can use techniques to fill them since if the entire column was majorly made up of null values it might be somewhat wrong to fill them according to intuition.

1) To fill Fedu, I filled with the mean of all people having that Fjob since Fjob had no null values, for example for a person with Fjob being services with null Fedu I filled it with the mean of all people having Fjob as services since different jobs will have different pay.
2) For famsize I plotted Famsize vs Fedu+Medu and then came to a threshold value of 5 for famsize being Greater than 3 since more educated and more richer parents might be okay with having a larger family.
3) For traveltime I filled it with the mean value for all students since it didn't have any significant correlation with some other column.
4) For filling higher, I had first filled with yes since a very high percentage of students had higher as yes, showing that students going to these schools tend to pursue higher education. But then I saw higher education inclination has an effect on grades so I checked accuracy of my models to then fill higher with yes if G3>10 and no if G3<=10.
5) For freetime, I filled with the same value as go out since these two features had some correlation.
6) For absences, I filled with the mean of the column since there was no significant correlation with some other feature.
7) For G2, I filled with the average of G1 and G3 since all three grades depend on each other.
8) For feature1,2 and 3 I had filled with the median first since the range was pretty small for 2 and 3 and for 1 it was tough to choose any threshold on the failures feature. But then after checking accuracy I made a choice to fill Feature 3 with the same value as Dalc since it had high correlation.

Level 3,

I noted down all the features and tried to plot everything from which I felt might draw an inference from and here are some of my inferences.

EDA inferences

1) Not much difference of grades across the 2 schools but GP 's grades are concentrated at a higher level than MS.
2) Probability of Medu and Fedu being higher is more in an urban address than in a rural one.
3) Bigger families do need support
4) Higher medu, fedu higher grades showing the impact of parents education on their children's academic performance (not much but there).
5) Obvious relationship between failurs and g3 not very linear shows that it is slightly unpredictable (kids having failures can also get good grades and vice versa)
6) Students living farther have lesser grades
7) Students very old still in secondary school showing poor academic performance being the reason they're still in school.
8) School support slightly helps with grades
9) Internet doesn't necessarily help with grades the lower bound is lower upper bound is higher showing it can have effects on both sides.
10) Parents who have done education inspire kids to pursue higher education too.
11) Students who aspire to pursue higher education tend to get higher grades which showing their motivation towards academics.
12) Family relationship has a slight effect on grades, but not so much on health.
13) Students having more freetime tend to spend it with their friends by going out with them.
14) Students who consume more alcohol tend to get lesser grades showing the ill effects that indulging in alcohol has on students' academics.
15) Students belonging to GP are more into extracurricular activities as compared to students in MS.
16) Also Fedu is highly correlated with Medu showing level of education is an important criteria for parents' marriage.

Since we've compared different features and drawn our inferences, now we see what relationships different features have with the student being in a relationship or not which will be later used by the classification techniques to make predictions.

Relationship status vs other features

1) Percentage of students in a romantic relationship is higher in MS than in GP
2) Higher percentage of females are in a romantic relationship than males
3) Students pursuing extracurricular activities have a higher chance of being in a relationship as compared to students who do not. This shows the effect of having a broader personality.
4) Higher vs relationship
5) Freetime or going out has very less effect on being in a relationship. Students having a lot of freetime or those who go out a lot do not necessarily have a higher chance of being in a relationship.
6) Students not in a relationship have lower absences maybe indicating being in a relationship might sway you towards not attending school.

7) Students in a relationship have slightly lower grades showing they might be distracted but not that much.
8) Older students clearly have a higher chance of being in a relationship.

Level 4 and 5,

Next, we split our data into training and testing data keeping 75% for training the model and 25% for testing. (got maximum accuracy using this ratio).

Then we classified our data into being in a relationship or not using 3 classification techniques - Logistic Regression, Decision Trees and Random Forests.

The confusion matrix and accuracy was calculated for all 3 of them which showed that Random Forests works the best for our data.

Next, we get the decision boundaries for all 3 classification models and using SHAP globally, we get to know which feature impacted the output the most in the Random Forests technique.

After this we choose 2 students, one who was predicted to be in a relationship and one who wasn't and then we observe the local feature explanations for both of them revealing us which feature had the major impact on these 2 students specifically.