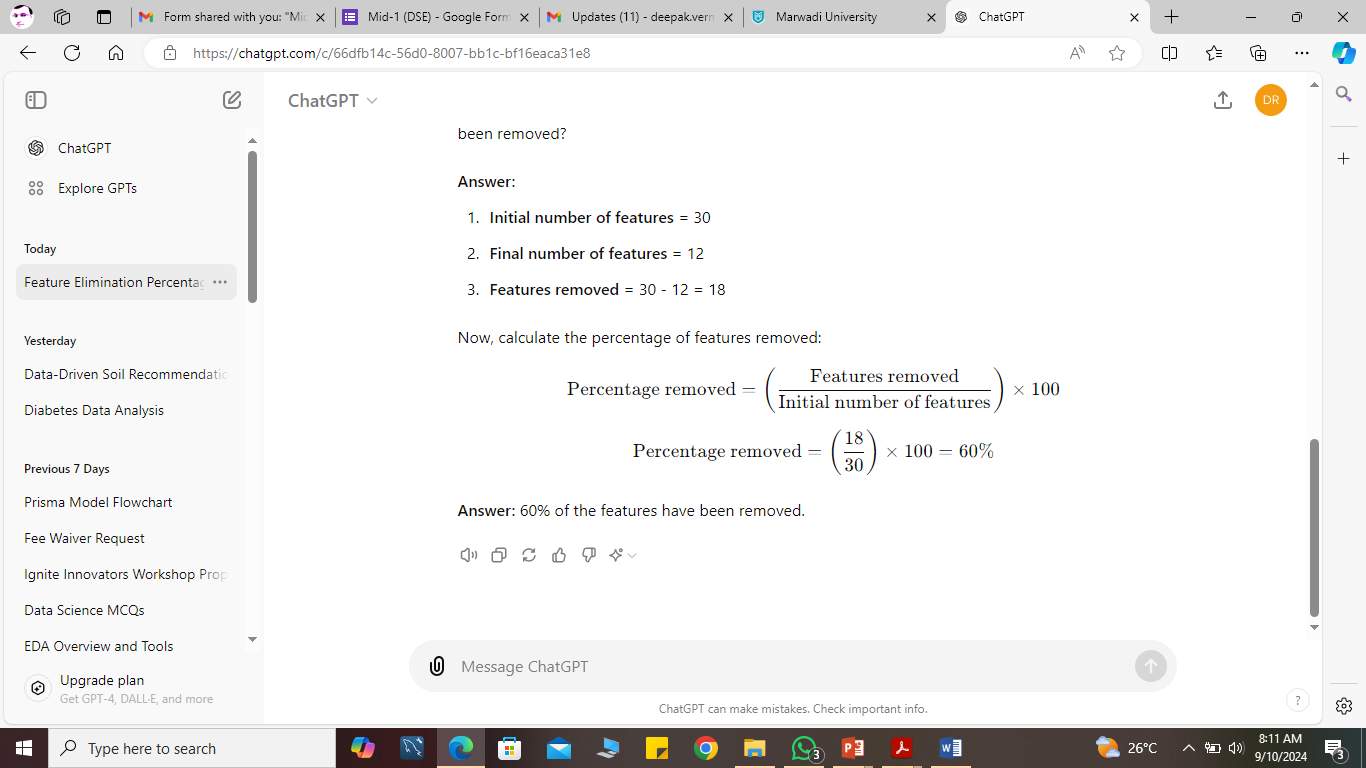
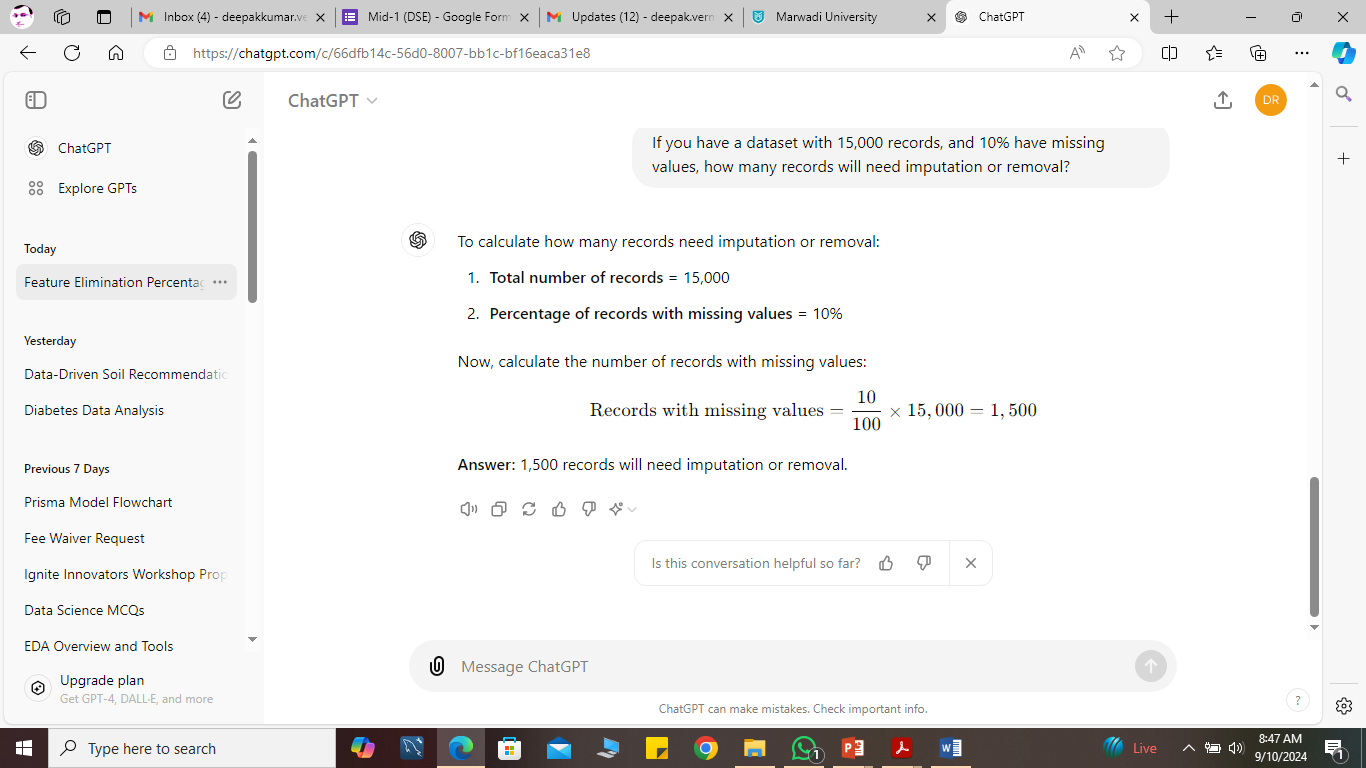
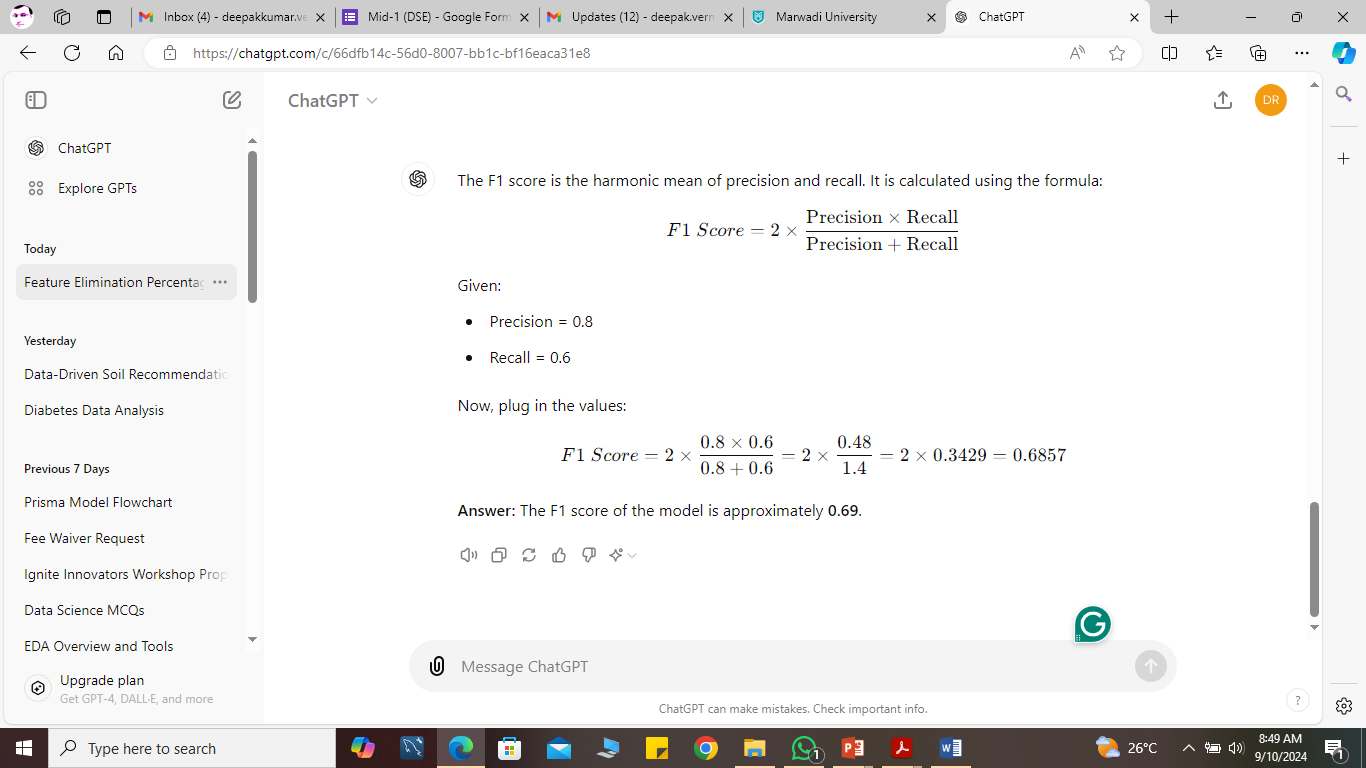
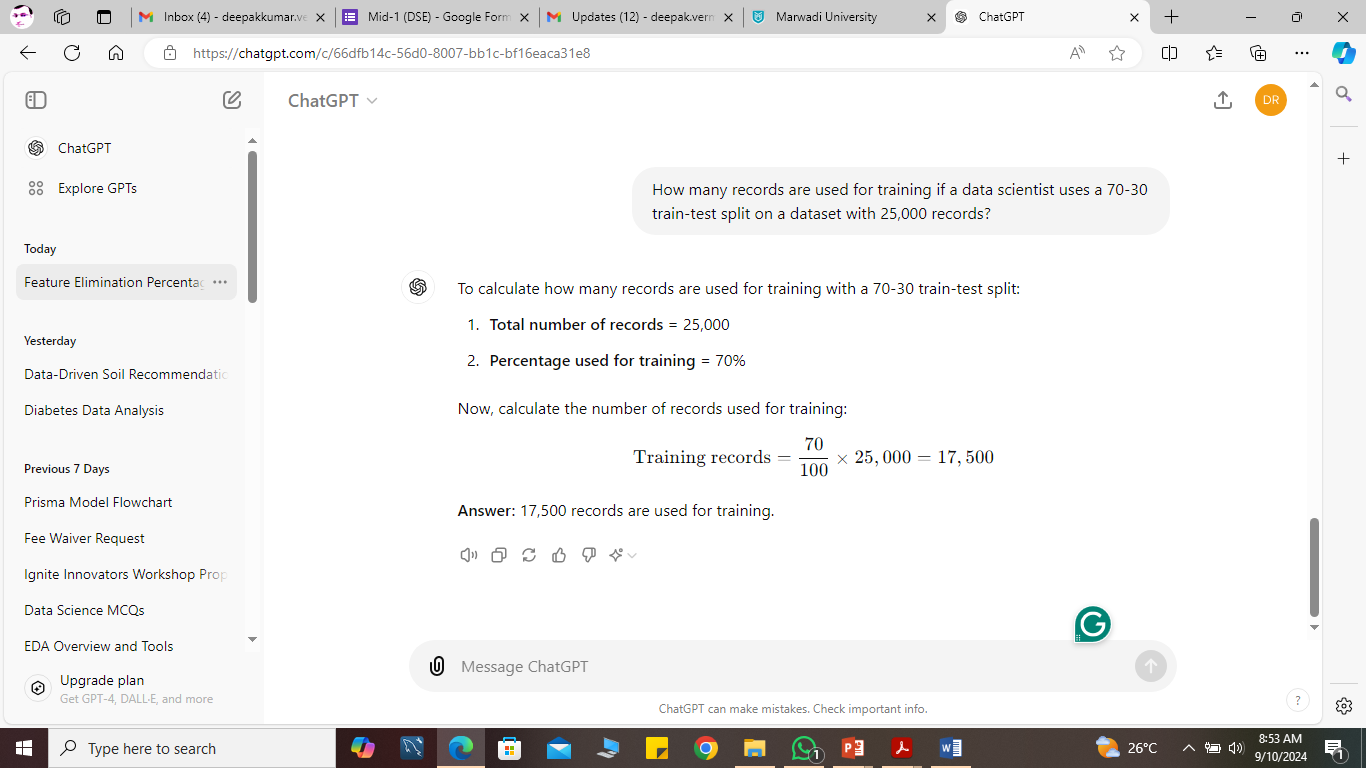
**Question:**

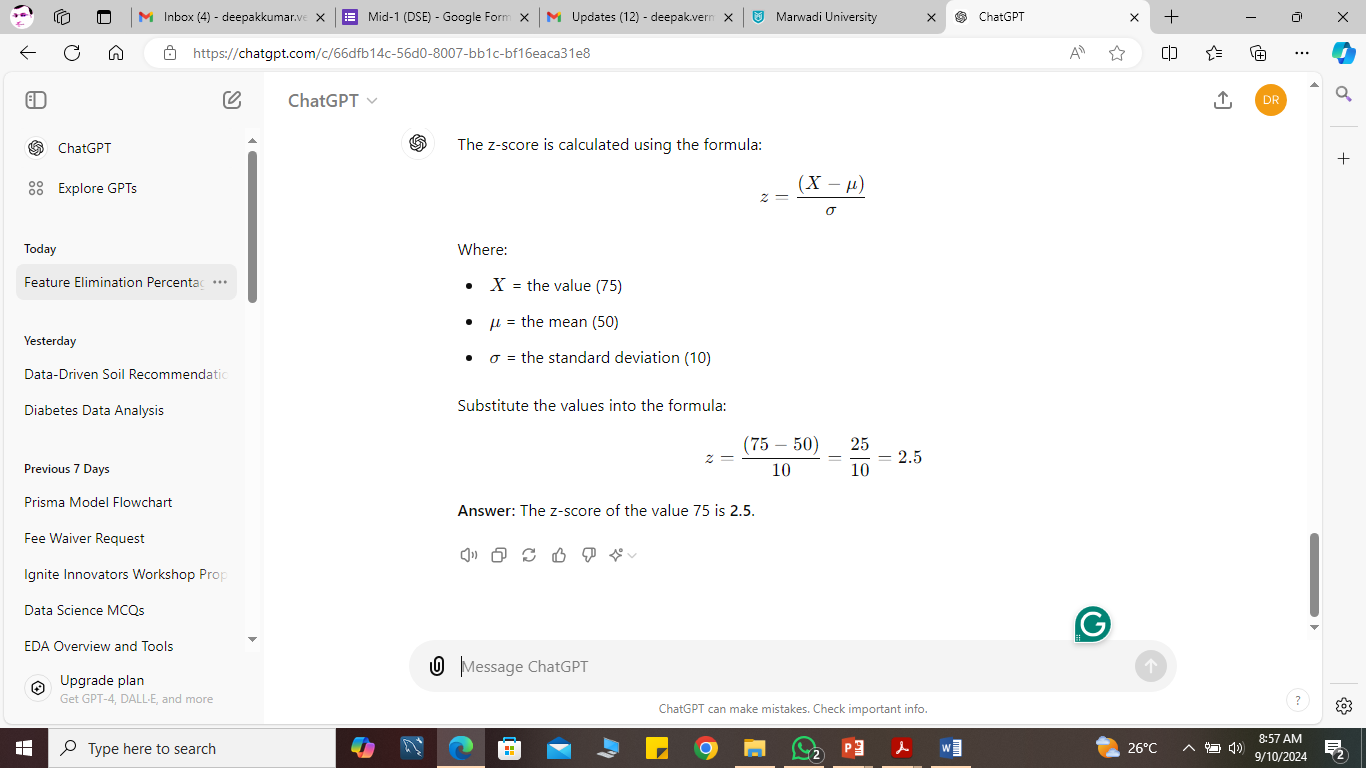
1. A dataset initially has 15,000 records with 30 features. After applying dimensionality reduction techniques, the number of features is reduced to 12. What percentage of the features have been removed?
2. If you have a dataset with 15,000 records, and 10% have missing values, how many records will need imputation or removal?
3. A model has a precision of 0.8 and a recall of 0.6. What is the F1 score of the model?
4. How many records are used for training if a data scientist uses a 70-30 train-test split on a dataset with 25,000 records?
5. You are analyzing a dataset with a mean of 50 and a standard deviation 10. What is the z-score of a value 75 in this dataset?
6. Suppose you perform a principal component analysis (PCA) on a dataset with 100 features and retain 95% of the variance, and the first ten principal components account for 90% of the variance. How many more components would you need to keep to achieve 95% variance?











Given that the first ten principal components account for 90% of the variance, and you need to retain 95% of the variance, you would need to account for an additional 5%.

To determine how many more components are needed, you would need information about the variance contribution of the remaining principal components. However, without specific values, let's assume that additional components beyond the first ten contribute relatively small amounts of variance.

Since 90% of the variance is already explained by the first 10 components, you need to capture an extra 5%. Typically, if each of the following few components contributes a small amount of variance, it is likely that you would need **2 to 3 more components** to reach 95%.

Therefore, to achieve 95% of the variance, you may need 2 to 3 more components beyond the first 10.