1) A: Missing data can be handled in various ways, we can either fill the missing value via the Fillna method by using mean, mode or median method or we can simpljy delete the rows and columns which contains null value by using dropna method.

B: Data types need to be converted to increase accuracy, avoid errors and to enable correct data manipulation and interpretation:

2) T-Test is a statistical test which helps to understand whether there is a significance difference between two groups. There are several types of T-Test: Independent T-test, Paired t-test and One sample t-test.
T-test can be used while comparing two independent groups: whether they are dependent on each other, or it can used while evaluating benchmark, that is comparing the average sales of a product with standard values

Chi-square test is a statistical test that determines whether there is a dependency between two categorical data. It checks whether the frequency of each category in a contingency table is different from the expected contingency.
It can used while testing relationship between two data or it can be used to check whether there is dependency between two data.
To test the relationship between shipping mode and customer segment, the chi2_statistics, p_value and DOF is to be calculated from scipy.stats. If p_value is more than the alpha value which is 0.05 then there is no significant dependency between shipping mode and customer segment else if it is less than 0.05 then there is a significant dependency.

3) Univariate analysis focuses one variable at a time in a dataset and examines its statistical analysis like its distribution, whether there is an outlier present, data visualization and understanding data type

Univariate analysis focuses on “One variable” while Bivariate analysis focuses on the relationship between “more than one variable”.
Example of Univariate analysis: Monthly Sales of a company.
Example of Bivariate analysis: To see the relationship between the profit and sales of a company.

4) A correlation matrix helps to find the correlation coefficient between two variables that is to see how the pairs are linearly related. To interpret the results would be:
1: Perfect Positive Correlation.
-1: negative correlation.
0: No linear correlation

Sales trend can be plotted overtime by using a lineplot from matplotlib or seaborn libraries of python.

5) Top Performing product categories can be identified via sales and profit by using the groupby function with product categories variable, sales and profit and then by sorting them into descending order.

   Seasonal trend can be analysed by using the resampling method. We can resample the sales value wither W- weekly, M-Monthly or Y-yearly wise. To resample we must make the date as index as without the date range resample will not take place and then plot the sales data

6) Grouped statistics is important as it gives a deeper insight of a data via its subsets or categories which might not come to surface when looking into the data as a whole. It helps to understand the relationship of a particular variable with another one via different region, sales, profit, product categories and these can be further grouped into mean, median, sum, quartile, etc.

   We can group regional sales data by using the groupby function between region and sales and then calculating the sales in different region via its sum, mean, median, max, count, standard deviation and variance.