Graphical user interface, website

Description automatically generated

A picture containing text, indoor

Description automatically generated

1. Suppose that a data warehouse consists of four dimensions customer, product, salesperson and sales time, and the three measure sales Amt(in rupees), VAT(in rupees) and payment\_type (in rupees). Draw the different classes of schemas that are popularly used for modeling data warehouses and explain it.
2. Use Average-link agglomerative clustering as well as medoid and centroid to cluster the following 5 points: A1(3,11), A2(3,6), A3(9,5), A4(6,9), A5(8,6).
3. Explain the simplistic pattern growth based frequent substructure mining that reduces duplicate graph structure.
4. Briefly explain the advances in processing, mining, and learning complex data.
5. In a study of drug abuse in a local high school, the school board selected 100 eighth graders, 100 sophomores, and 100 seniors randomly from their respective rolls for each grade. Each student was then asked if they used a particular drug frequently, seldom, or never. The data are summarized in the table given below. Is there evidence to suggest that the frequency of drug use is the same across the three different grades? State the null hypothesis, give the test statistic, test criterion, conclusion, and interpretation using the chi-square test.