SQL

1. CREATE TABLE Customers (
        customerNumber INT,
        customerName VARCAHR(255),
        contactLastName VARCHAR(255),
        contactFirstName VARCHAR(255),
        phone INT,
        addressLine1 VARCHAR(255),
        addressLine2 VARCHAR(255),
        city VARCHAR(255),
        state VARCHAR(255),
        postalcode INT,
        country VARCHAR(255),
        salesRepEmployeeNumber INT,
        creditLimit INT,
        PRIMARY KEY(customerNumber)
);

2. CREATE TABLE Orders (
        orderNumber INT PRIMARY KEY,
        orderDate DATE,
        requiredDate DATE,
        shippedDate DATE,
        status VARCHAR(255),
        comments VARCHAR(255),
        customerNumber INT,
        PRIMARY KEY(orderNumber),
        FOREIGN KEY(customerNumber) REFERENCES Customers(customerNumber)
);

3. SELECT * FROM Orders;

4. SELECT comments FROM Orders;

5. SELECT
         orderDate, COUNT(DISTINCT orderNumber) AS total_order
   FROM
        Orders
   GROUP BY
        orderDate;

6. SELECT employeNumber, lastName, firstName FROM employees;

7. SELECT customerNumber,orderNumber FROM Orders;

8. SELECT customerName,salesRepEmployeeNumber FROM customers;

9. SELECT paymentDate,amount FROM payments;

10. SELECT productName, MSRP, productDescription FROM products;

11. SELECT
       productName,productDescription,COUNT(DISTINCT orderNumber) AS no_of_orders
    FROM
       products
    LEFT JOIN
       orderdetails
    ON
       products.productCode = orderdetails.productCode
    GROUP BY
       productCode
    ORDER BY
       no_of_orders
    LIMIT 1;

12. SELECT
       city,COUNT(DISTINCT orderNumber) AS no_of_orders
    FROM
       customers
    LEFT JOIN
       orders
    ON
       customers.customerNumber = orders.customerNumber
    GROUP BY
       city
    ORDER BY
       no_of_orders DESC
    LIMIT 1;

13. SELECT
       state,COUNT(DISTINCT orderNumber) AS no_of_orders
    FROM
       customers
    LEFT JOIN
       orders
    ON
       customers.customerNumber = orders.customerNumber
    GROUP BY
       state
    ORDER BY
       no_of_orders DESC
    LIMIT 1;

14. SELECT
       employeeNumber, CONCAT(firstName," ",lastName) AS full_name
    FROM
       employees;

15. SELECT

```
      orderNumber,customerName, quantityOrdered*priceEach AS total_price
FROM
      orderdetails;
```

Statistics Question

1. B)
2. C)
3. A)
4. A)
5. C)
6. B)
7. B)
8. D)
9. A)

10. Bayes' theorem thus gives the probability of an event based on new information that is, or may be related, to that event.

11. It's a measure of how many standard deviations below or above the population mean a raw score is.

12. A t-test is a type of inferential statistic used to determine if there is a significant difference between the means of two groups, which may be related in certain features. The t-test is one of many tests used for the purpose of hypothesis testing in statistics.

13. A percentile is a term used in statistics to express how a score compares to other scores in the same set. While there is technically no standard definition of percentile, it's typically communicated as the percentage of values that fall below a particular value in a set of data.

14. A one way ANOVA is used to compare two means from two independent (unrelated) groups using the F-distribution. The null hypothesis for the test is that the two means are equal. ANOVA is abbreviation of Analysis of Variance.

15. ANOVA can help in finding relation between two independent groups by using hypothesis testing.

Machine Learning

1. D)
2. A)
3. C)
4. B)
5. D)
6. C)
7. D)
8. A)
9. A)
10. B)
11. A)
12. B)
13.

- Clustering analysis is broadly used in many applications such as market research, pattern recognition, data analysis, and image processing.

- Clustering can also help marketers discover distinct groups in their customer base. And they can characterize their customer groups based on the purchasing patterns.

- In the field of biology, it can be used to derive plant and animal taxonomies, categorize genes with similar functionalities and gain insight into structures inherent to populations.

- Clustering also helps in identification of areas of similar land use in an earth observation database. It also helps in the identification of groups of houses in a city according to house type, value, and geographic location.

- Clustering also helps in classifying documents on the web for information discovery.

- Clustering is also used in outlier detection applications such as detection of credit card fraud.