# MACHINE LEARNING WORKSHEET 3

1. D)
2. D)
3. C)
4. B)
5. D)
6. C)
7. D)
8. A)
9. A)
10. B)
11. A)
12. B)
13. Clustering is useful for exploring data. If there are many cases and no obvious groupings, clustering algorithms can be used to find natural groupings. **Clustering** can also serve as a useful data-preprocessing step to identify homogeneous groups on which to build supervised models.
14. K-means clustering algorithm can be significantly improved by using a better initialization technique, and by repeating (re-starting) the algorithm.When the data has overlapping clusters, k-means can improve the results of the initialization technique.When the data has well separated clusters, the performance of k-means depends completely on the goodness of the initialization.

# SQL WORKSHEET

**Q1. Write SQL query to create table Customers**

CREATETABLEcustomers(

customerNumberint primary key,

customerNamevarchar(30),

contactLastNamevarchar(20),

contactFirstNamevarchar(20),

phoneint,

addressLine1varchar(50),

addressLine2varchar(50),

cityvarchar(10),

statevarchar(10),

postalCodeint,

countryvarchar(10),

salesRepEmployeeNumbervarchar(20),

creditLimitint

);

* 1. Write SQL querytocreate table**Orders.**

CREATE TABLE Orders (  
    orderNumber int NOT NULL,  
    orderdate DATE,  
    requireddateDATE,  
    shippeddate DATE,

Status varchar(20),

Comments varchar(30),  
   customerNumber FOREIGN KEYREFERENCES customers(customerNumber)  
);

* 1. WriteSQL querytoshow allthecolumns datafromthe**Orders**Table.

Select \* from Orders;

* 1. Write SQL querytoshow allthecomments fromthe**Orders**Table.

Select Comments from Orders;

* 1. WriteaSQLqueryto showorderDate andTotalnumberoforders placed on thatdate,from**Orders**table.

Select orderdate ,quantityorder group by orderdate from (join Orders.orderNumber = orderdetails.orderNumber

Where group by ordernumber );

* 1. WriteaSQLquerytoshowemployeNumber,lastName,firstNameofall theemployeesfrom**employees**

table.

Select employeeNumber,lastName, firstName from employees;

* 1. WriteaSQLquerytoshowallorderNumber,customerNameoftheperson whoplacedtherespectiveorder.

Select Orders.orderNumber, customers.customerName From Orders Right Join customers on Orders.customerNumber = customers.customerNumber order by orderNumber;

* 1. Write a SQL query to show name of all the customers in one column and salerepemployee name inanothercolumn.

Select customers.customerName , employees.firstname from customers right join employees on salerepemployeeNumber = employees.employeenumber order by customerName

* 1. Write a SQL query to show Date in one column and total payment amount of the payments made on that datefromthe**payments**table.

Select paymentDate,amount from payments group by paymentdate;

* 1. WriteaSQLqueryto showallthe products productName, MSRP, productDescriptionfromthe**products**

table.

Select productName, MSRP, productDescription from products;

* 1. WriteaSQLquerytoprinttheproductName,productDescription ofthemost orderedproduct.

Select productName, productDescription from products where(join products.productcode = orderdetails.productcode max(quantityorder));

* 1. WriteaSQLquerytoprintthecitynamewheremaximumnumberoforderswereplaced.

Select city from customers (where (max(group by (select oderdetails.quantityorder from join customers.customerNumber= orders.customerNumber where(select orderNumber,quantityorder from orderdetails )))))

* 1. WriteaSQLquerytogetthenameofthestatehavingmaximumnumberofcustomers.

Select state from customers where max(group by state);

* 1. Write a SQL query to print the employee number in one column and Full name of the employee in thesecondcolumn for alltheemployees.

Select employeeNumber, concat(firstName, lastName) as employeeName from employees;

* 1. Write a SQL query to print the orderNumber, customer Name and total amount paid by the customer for thatorder(quantityOrdered×priceEach).

select c.customerName ,o.orderNumber, (SELECT (a.quantityOrdered ×a.priceEach) from orderdetails AS aINNER JOIN order as b wherea.orderNumber=b.orderNumber)as'amount'

from customers as c INNER JOIN orders as o wherec.customerNumber=o.customerNumbergroupbyo.orderNumber;

# STATISTIC WORKSHEET 3

**STATISTICS WORKSHEET 3**

**Q1. b**

**Q2. c**

**Q3. a**

**Q4. a**

**Q5. b**

**Q6. a**

**Q7. b**

**Q8. d**

**Q9. a**

**Q10. In statistics and probability theory, the Bayes theorem (also known as the Bayes rule) is a mathematical formula used to determine the conditional probability of events. Essentially, the Bayes theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event.The theorem is named after English statistician, Thomas Bayes, who discovered the formula in 1763. It is considered the foundation of the special statistical inference approach called the Bayes inference. Besides statistics, the Bayes theorem is also used in various disciplines, with medicine and pharmacology as the most notable examples. In addition, the theorem is commonly employed in different fields of finance. Some of the applications include but are not limited to, modeling the risk of lending money to borrowers or forecasting the probability of the success of an investment.**

**Q11. A Z-score is a numerical measurement that describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviations from the mean. If a Z-score is 0, it indicates that the data point's score is identical to the mean score. A Z-score of 1.0 would indicate a value that is one standard deviation from the mean. Z-scores may be positive or negative, with a positive value indicating the score is above the mean and a negative score indicating it is below the mean.**

**Q12. A t-test is used as a hypothesis testing tool, which allows testing of an assumption applicable to a population. A t-test looks at the t-statistic, the t-distribution values, and the degrees of freedom to determine the statistical significance. To conduct a test with three or more means, one must use an analysis of variance.**

**Q13. In statistics, percentiles are used to understand and interpret data. The n th percentile of a set of data is the value at which n percent of the data is below it. In everyday life, percentiles are used to understand values such as test scores, health indicators, and other measurements**