

Set 20

1. No of days between two dates

```
# Convert the date strings to Date objects
date1 <- as.Date("2022-01-01")
date2 <- as.Date("2022-02-15")

# Calculate the number of days between the two dates
num_days <- difftime(date2, date1, units = "days")

# Print the result
cat("Number of days between", date1, "and", date2, ":", num_days, "days")
```

2. K means clustering euclidian distance

```
# Define the data points
data <- matrix(c(2,10,2,5,8,4,5,8,7,5,6,4,1,2,4,9), ncol = 2, byrow = TRUE)

# Define the initial centroids
initial_centroids <- matrix(c(2,10,5,8,1,2), ncol = 2, byrow = TRUE)

# Perform k-means clustering
kmeans_result <- kmeans(data, centers = initial_centroids)

# Get the cluster assignments
cluster_assignments <- kmeans_result$cluster

# Get the final centroids
final_centroids <- kmeans_result$centers

# Print the cluster assignments and final centroids
cat("Cluster Assignments:")
print(cluster_assignments)

cat("\nFinal Centroids:")
print(final_centroids)
```

3. Plot the function $f(x)=\sin(x)$

```
# Generate x values in the interval (-3, 3) with a step size of 0.1
x <- seq(-3, 3, 0.1)

# Compute the corresponding y values using the sin function
y <- sin(x)
```

```
# Plot the function
```

```
plot(x, y, type = "l", xlab = "x", ylab = "f(x)", main = "Plot of f(x) = sin(x)")
```

4. Table programs

a)

```
# Create the data frame
```

```
df<-data.frame(
```

```
Item=c("Baby food", "Cereal", "Office supplies", "Fruits", "Office supplies", "Household",  
"Household"),
```

```
OrderPriority = c(1, 2, 3, 1, NA, 3, 3),
```

```
UnitPrice = c(255.28, 205.7, NA, 9.33, 651.21, 668.27, 668.27),
```

```
UnitsSold = c(9925, 2804, 1779, 8102, 5062, 8974, NA),
```

```
stringsAsFactors = FALSE
```

```
)
```

```
# Print the data frame
```

```
df
```

b)

```
mean_order_priority <- mean(df$OrderPriority, na.rm = TRUE)
```

```
df$OrderPriority <- replace(df$OrderPriority, is.na(df$OrderPriority),  
mean_order_priority)
```

c)

```
median_units_sold <- median(df$UnitsSold, na.rm = TRUE)
```

```
df$UnitsSold <- replace(df$UnitsSold, is.na(df$UnitsSold), median_units_sold)
```

d)

```
df <- na.omit(df)
```

e)

```
df$UnitPrice <- unique(df$UnitPrice)
```

f)

```
total_missing <- sum(is.na(df))
```

```
cat("Total number of missing values in df:", total_missing, "\n")
```