

Day -1

Question-1

```
#Roll no: 192224227
age <- c(5, 15, 20, 50, 80, 110)
frequency <- c(200, 450, 300, 1500, 700, 44)
print(median(age))
print(median(frequency))
```

```
[1] 35
[1] 375
```

Question-2

```
#Roll no: 192224227
age <- c(13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70)
print(mean(age))
print(median(age))
mode_age <- names(table(age))[table(age) == max(table(age))]
print(mode_age)
print(range(age))
print(quantile(age, 0.25))
print(quantile(age, 0.75))
```

```
[1] 29.96296
[1] 25
[1] "25" "35"
[1] 13 70
25%
20.5
75%
35
```

Question-3

```
#Roll no: 192224227
data <- c(200, 300, 400, 600, 1000)
min_max_norm <- (data - min(data)) / (max(data) - min(data))
print(min_max_norm)
z_score_norm <- (data - mean(data)) / sd(data)
print(z_score_norm)
```

```
[1] 0.000 0.125 0.250 0.500 1.000
[1] -0.9486833 -0.6324555 -0.3162278  0.3162278  1.5811388
```

Question-4

```
#Roll no: 192224227
data <- c(11, 13, 13, 15, 15, 16, 19, 20, 20, 20, 21, 21, 22, 23, 24, 30, 40, 45, 45, 45, 71, 72, 73, 75)
bins <- 5
bin_indices <- cut(data, bins)
mean_smooth <- tapply(data, bin_indices, mean)
print(mean_smooth)
median_smooth <- tapply(data, bin_indices, median)
print(median_smooth)
min_max_smooth <- tapply(data, bin_indices, function(x) c(min(x), max(x)))
print(min_max_smooth)
```

```
(10.9,23.8] (23.8,36.6] (36.6,49.4] (49.4,62.2] (62.2,75.1]
 17.78571   27.00000   43.75000         NA   72.75000
(10.9,23.8] (23.8,36.6] (36.6,49.4] (49.4,62.2] (62.2,75.1]
  19.5      27.0      45.0         NA      72.5
$` (10.9,23.8] `
[1] 11 23

$` (23.8,36.6] `
[1] 24 30

$` (36.6,49.4] `
[1] 40 45

$` (49.4,62.2] `
NULL

$` (62.2,75.1] `
[1] 71 75
```

Question-5

```
#Roll no: 192224227
age <- c(23, 23, 27, 27, 39, 41, 47, 49, 50, 52, 54, 54, 56, 57, 58, 58, 60, 61)
fat <- c(9.5, 26.5, 7.8, 17.8, 31.4, 25.9, 27.4, 27.2, 31.2, 34.6, 42.5, 28.8, 33.4, 30.2, 34.1, 32.9, 41.2, 35.7)
print(c(mean(age), median(age), sd(age)))
print(c(mean(fat), median(fat), sd(fat)))

[1] 46.44444 51.00000 13.21862
[1] 28.783333 30.700000 9.254395
```

Question-6

```
#Roll no: 192224227
v <- c(23, 23, 27, 27, 39, 41, 47, 49, 50, 52, 54, 54, 56, 57, 58, 58, 60, 61)
# Min-max normalization
min_max <- (35 - min(v)) / (max(v) - min(v))
print(min_max)
# Z-score normalization
z_score <- (35 - mean(v)) / 12.94
print(z_score)
# Decimal scaling
decimal_scaling <- 35 / (10^ceiling(log10(max(abs(v)))))
print(decimal_scaling)

[1] 0.3157895
[1] -0.8844238
[1] 0.35
```

Question-7

```
#Roll no: 192224227
pencils <- c(9, 25, 23, 12, 11, 6, 7, 8, 9, 10)
print(mean(pencils))
print(median(pencils))
mode <- names(table(pencils))[table(pencils) == max(table(pencils))]
print(mode)

[1] 12
[1] 9.5
[1] "9"
```

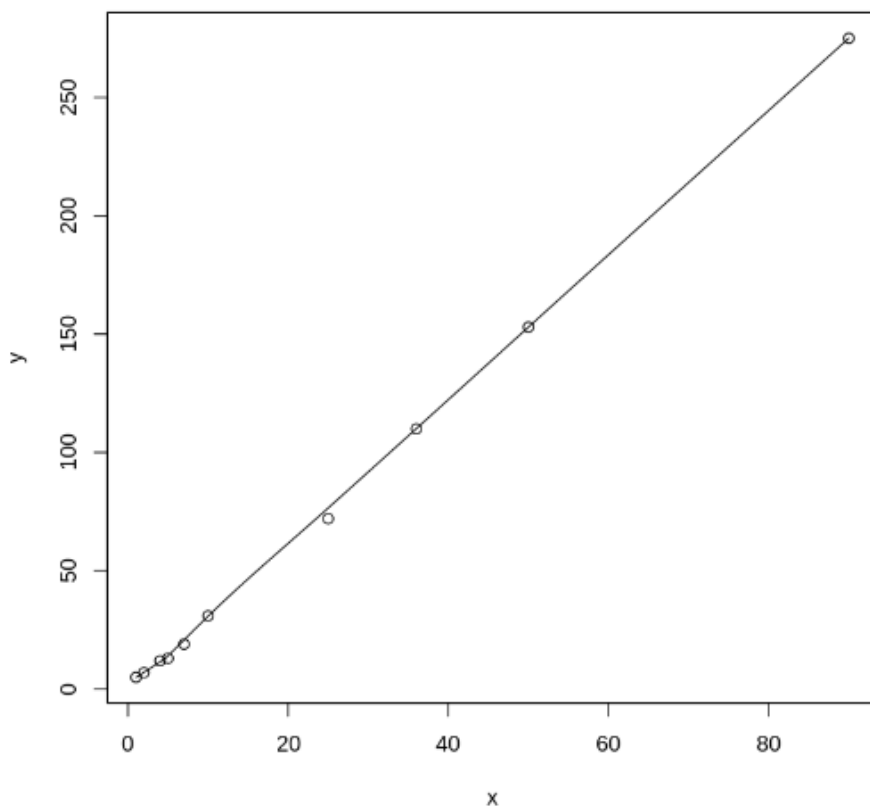
Question-8

```
#Roll no: 192224227
```

```
x <- c(4, 1, 5, 7, 10, 2, 50, 25, 90, 36)
```

```
y <- c(12, 5, 13, 19, 31, 7, 153, 72, 275, 110)
```

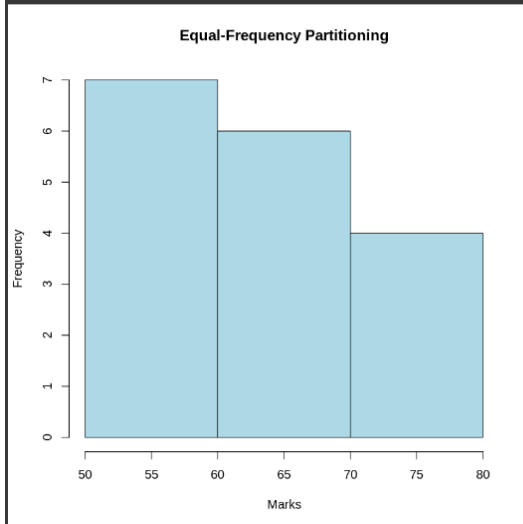
```
scatter.smooth(x, y)
```



Question-9

```
#Roll no: 192224227
marks <- c(55, 60, 71, 63, 55, 65, 50, 55, 58, 59, 61, 63, 65, 67, 71, 72, 75)
num_bins <- 3
bins_eq_frequency <- cut(marks, breaks = num_bins, labels = FALSE)
print(table(bins_eq_frequency))
hist(marks, breaks = num_bins, col = "lightblue", xlab = "Marks", main = "Equal-Frequency Partitioning")
```

```
bins_eq_frequency
1 2 3
5 7 5
```



Question-10

```
#Roll no: 192224227
v <- c(78.3, 81.8, 82, 74.2, 83.4, 84.5, 82.9, 77.5, 80.9, 70.6)
print(IQR(v))
print(sd(v))
```

```
[1] 4.975
[1] 4.445835
```

Question-11

```
#Roll no: 192224227
age <- c(13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70)
print(quantile(age, 0.25))
print(quantile(age, 0.75))
```

25%
20.5
75%
35