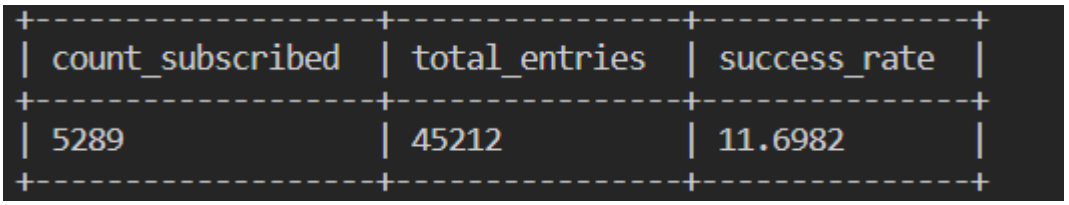
The following questions were answered by data analysis with Hive

1. **Load data and create a Hive table**

```
CREATE TABLE IF NOT EXISTS banking_data (age int, job String, marital
String, education String, default String, balance int, housing String,
loan String, contact String, day int, month String, duration int,
campaign int, pdays int, previous int, poutcome String, status String)
COMMENT 'Banking Details'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ';'
LINES TERMINATED BY '\n'
STORED AS TEXTFILE;

LOAD DATA LOCAL INPATH
'/home/uttam/futurense-datengg-bootcamp/dataset/bankmarketdata.csv'
OVERWRITE INTO TABLE banking_data;
```

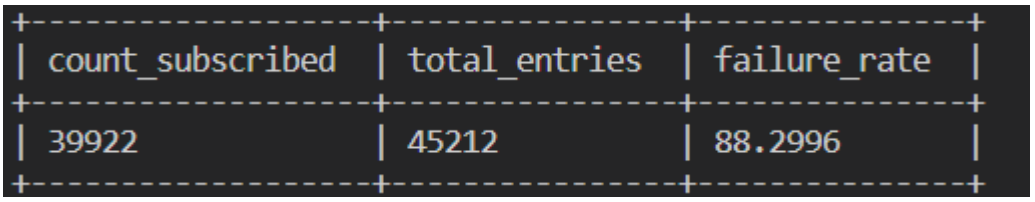
2. **Give marketing success rate. (No. of people subscribed / total no. of entries)**



count_subscribed	total_entries	success_rate
5289	45212	11.6982

```
Select sum(if(status='yes',1,0)) as `count_subscribed`, count(*) as
`total_entries`, round(sum(if(status='yes',1,0))/count(*) * 100, 4) as
`success_rate` from banking_data;
```

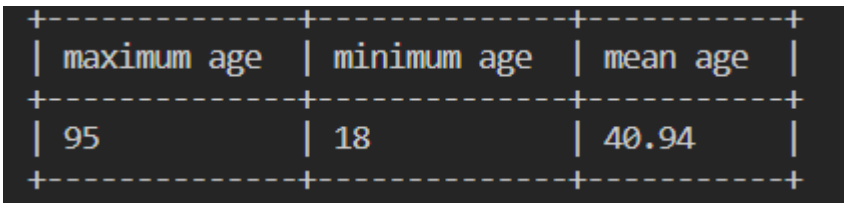
3. **Give marketing failure rate**



count_subscribed	total_entries	failure_rate
39922	45212	88.2996

```
Select sum(if(status='no',1,0)) as `count_subscribed`, count(*) as
`total_entries`, round(sum(if(status='no',1,0))/count(*) * 100, 4) as
`failure_rate` from banking_data;
```

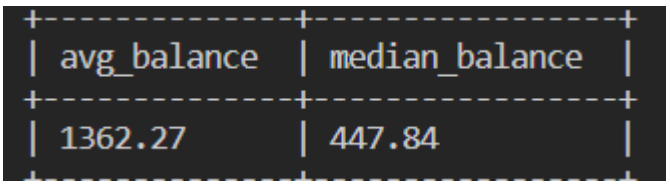
4. **Maximum, Mean, and Minimum age of the average targeted customer**



maximum age	minimum age	mean age
95	18	40.94

```
Select max(age) as `Maximum age`, min(age) as `Minimum age`,
round(avg(age),2) as `Mean age` from banking_data;
```

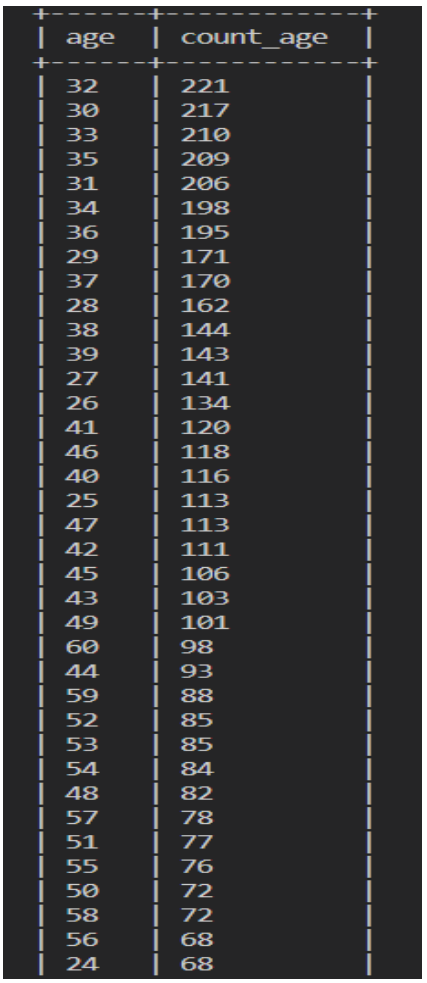
5. **Check the quality of customers by checking the average balance, median balance of customers**



avg_balance	median_balance
1362.27	447.84

```
SELECT round(avg(balance),2) as avg_balance,
round(percentile_approx(balance, 0.5),2) as median_balance FROM
banking_data;
```

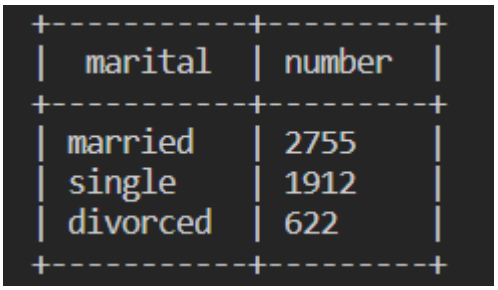
6. **Check if age matters in marketing subscription for deposit**



age	count_age
32	221
30	217
33	210
35	209
31	206
34	198
36	195
29	171
37	170
28	162
38	144
39	143
27	141
26	134
41	120
46	118
40	116
25	113
47	113
42	111
45	106
43	103
49	101
60	98
44	93
59	88
52	85
53	85
54	84
48	82
57	78
51	77
55	76
50	72
58	72
56	68
24	68

```
select age, count(*) as Count_Age from banking_data where status='yes'
group by age order by count_age desc;
```

7. **Check if marital status mattered for subscription to deposit.**

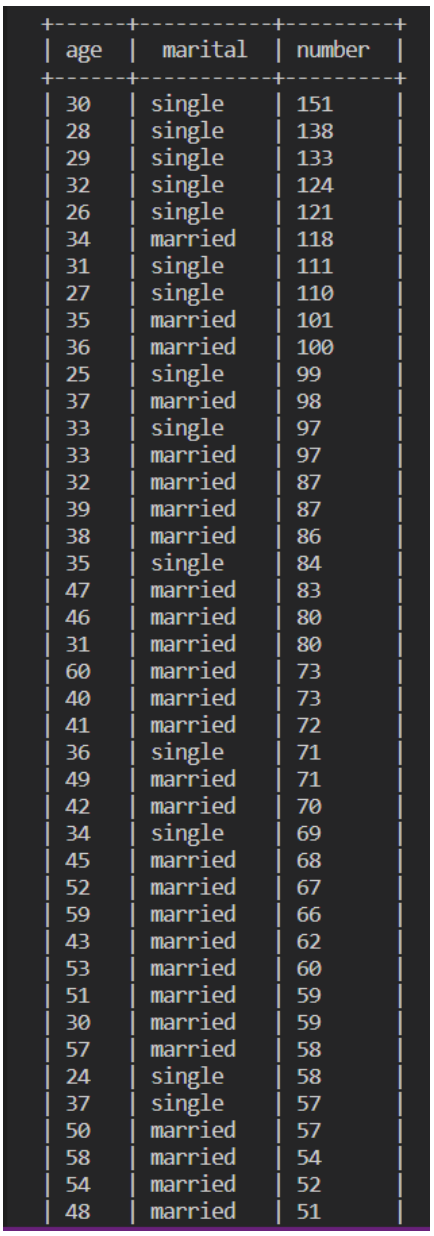


marital	number
married	2755
single	1912
divorced	622

```
select marital, count(*) as number from banking_data where
status='yes' group by marital order by number desc;
```

8. **Check if age and marital status together mattered for subscription to deposit scheme**

```
select age, marital, count(*) as number from banking_data where
status='yes' group by age,marital order by number desc;
```



age	marital	number
30	single	151
28	single	138
29	single	133
32	single	124
26	single	121
34	married	118
31	single	111
27	single	110
35	married	101
36	married	100
25	single	99
37	married	98
33	single	97
33	married	97
32	married	87
39	married	87
38	married	86
35	single	84
47	married	83
46	married	80
31	married	80
60	married	73
40	married	73
41	married	72
36	single	71
49	married	71
42	married	70
34	single	69
45	married	68
52	married	67
59	married	66
43	married	62
53	married	60
51	married	59
30	married	59
57	married	58
24	single	58
37	single	57
50	married	57
58	married	54
54	married	52
48	married	51