SQL Queries For Loan Approval Data

```sql
-- DATA
SELECT *
FROM loan_data;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Loan_ID	Gender	Married	Dependents	Education	Self_Employed	ApplicantIncome	CoapplicantIncome	LoanAmount	Loan_Amount_Term	Credit_History	Property_Area	Loan_Sta
LP001003	Male	Yes	1	Graduate	No	4583	1508	128	360	1	Rural	N
LP001005	Male	Yes	0	Graduate	Yes	3000	0	66	360	1	Urban	Y
LP001006	Male	Yes	0	Not Graduate	No	2583	2358	120	360	1	Urban	Y
LP001008	Male	No	0	Graduate	No	6000	0	141	360	1	Urban	Y
LP001013	Male	Yes	0	Not Graduate	No	2333	1516	95	360	1	Urban	Y
LP001024	Male	Yes	2	Graduate	No	3200	700	70	360	1	Urban	Y
LP001027	Male	Yes	2	Graduate		2500	1840	109	360	1	Urban	Y
LP001029	Male	No	0	Graduate	No	1853	2840	114	360	1	Rural	N
LP001030	Male	Yes	2	Graduate	No	1299	1086	17	120	1	Urban	Y
LP001032	Male	No	0	Graduate	No	4950	0	125	360	1	Urban	Y
LP001036	Female	No	0	Graduate	No	3510	0	76	360	0	Urban	N
LP001038	Male	Yes	0	Not Graduate	No	4887	0	133	360	1	Rural	N
LP001043	Male	Yes	0	Not Graduate	No	7660	0	104	360	0	Urban	N
LP001047	Male	Yes	0	Not Graduate	No	2600	1911	116	360	0	Semiurban	N
LP001050		Yes	2	Not Graduate	No	3365	1917	112	360	0	Rural	N
LP001068	Male	Yes	0	Graduate	No	2799	2253	122	360	1	Semiurban	Y
LP001073	Male	Yes	2	Not Graduate	No	4226	1040	110	360	1	Urban	Y
LP001086	Male	No	0	Not Graduate	No	1442	0	35	360	1	Urban	N
LP001087	Female	No	2	Graduate		3750	2083	120	360	1	Semiurban	Y
LP001095	Male	No	0	Graduate	No	3167	0	74	360	1	Urban	N
LP001097	Male	No	1	Graduate	Yes	4692	0	106	360	1	Rural	N
LP001098	Male	Yes	0	Graduate	No	3500	1667	114	360	1	Semiurban	Y

```sql
-- 1. What is the average loan amount by credit history status?
SELECT Credit_History, AVG(LoanAmount) AS avg_loan_amount
FROM loan_data
GROUP BY Credit_History;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Credit_History	avg_loan_amount
1	103.9809
0	108.2391

```sql
-- 2. How many loans were approved versus rejected?
SELECT Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Loan_Status;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Loan_Status	count
N	91
Y	217

```sql
-- 3. How many loans were approved for male applicants versus female applicants?
SELECT Gender, Loan_Status, COUNT(*) AS count
FROM loan_data
WHERE Gender = 'Male' or Gender = 'Female'
GROUP BY Gender, Loan_Status;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Gender	Loan_Status	count
Male	N	63
Male	Y	170
Female	N	24
Female	Y	46

```sql
-- 4. Is there a difference in loan approval rates between married and unmarried applicants?
SELECT Married, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Married, Loan_Status;

-- 5. How does the education level of the applicant relate to loan approval status?
SELECT Education, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Education, Loan_Status;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Education	Loan_Status	count
Graduate	N	65
Graduate	Y	167
Not Graduate	Y	50
Not Graduate	N	26


```sql
-- 5. How does the education level of the applicant relate to loan approval status?
SELECT Education, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Education, Loan_Status;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Education	Loan_Status	count
Graduate	N	65
Graduate	Y	167
Not Graduate	Y	50
Not Graduate	N	26

```sql
-- 6. Are self-employed individuals more or less likely to get their loans approved?
SELECT Self_Employed, Loan_Status, COUNT(*) AS count
FROM loan_data
WHERE Self_Employed = 'Yes' or Self_Employed = 'No'
GROUP BY Self_Employed, Loan_Status;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Self_Employed	Loan_Status	count
No	N	77
Yes	Y	15
No	Y	187
Yes	N	10

```sql
-- 7. How does the location of the property (urban, semiurban, rural) influence loan approval?
SELECT Property_Area, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Property_Area	Loan_Status	count
Semiurban	Y	102
Urban	Y	64
Rural	Y	51
Rural	N	35
Urban	N	32
Semiurban	N	24

```sql
-- 8. Does the number of dependents impact the likelihood of loan approval?
SELECT Dependents, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Dependents, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Dependents	Loan_Status	count
0	Y	146
0	N	62
2	Y	39
1	Y	32
2	N	15
1	N	14

```sql
-- 9. How does the term of the loan (in months) influence loan approval?
SELECT Loan_Amount_Term, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Loan_Amount_Term, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Loan_Amount_Term	Loan_Status	count
360	Y	190
360	N	73
180	Y	16
180	N	8
480	N	6
120	Y	3
300	N	3
480	Y	3
300	Y	2
84	Y	2
240	Y	1
36	N	1

```sql
-- 10. Is there a correlation between income groups and loan approval status?
SELECT
    CASE
        WHEN ApplicantIncome < ( SELECT avg(ApplicantIncome) FROM loan_data ) THEN 'Below Average'
        ELSE 'Above Average'
    END AS income_group,
    Loan_Status,
    COUNT(*) AS count
FROM loan_data
GROUP BY income_group, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

income_group	Loan_Status	count
Below Average	Y	128
Above Average	Y	89
Below Average	N	50
Above Average	N	41

```sql
-- 11. How does the combined income of applicants and co-applicants relate to loan approval?
SELECT
    CASE
        WHEN ApplicantIncome + CoapplicantIncome < (SELECT AVG(ApplicantIncome) FROM loan_data) THEN 'Below Avg Applicant Income'
        WHEN ApplicantIncome + CoapplicantIncome >= (SELECT AVG(ApplicantIncome) FROM loan_data) THEN 'Above Avg Applicant Income'

    END AS combined_income_group,
    Loan_Status,
    COUNT(*) AS count
FROM loan_data
GROUP BY combined_income_group, Loan_Status;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

combined_income_group	Loan_Status	count
Above Avg Applicant Income	N	66
Below Avg Applicant Income	Y	37
Above Avg Applicant Income	Y	180
Below Avg Applicant Income	N	25

```sql
-- 12. Does the combination of credit history and property area provide additional insights into loan approval?
SELECT Property_Area, Credit_History, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, Credit_History, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Property_Area	Credit_History	Loan_Status	count
Semiurban	1	Y	101
Urban	1	Y	64
Rural	1	Y	49
Rural	1	N	20
Urban	1	N	19
Semiurban	0	N	15
Rural	0	N	15
Urban	0	N	13
Semiurban	1	N	9
Rural	0	Y	2
Semiurban	0	Y	1

```sql
-- 13. How does the combination of education level and employment status affect loan approval?
SELECT Education, Self_Employed, Loan_Status, COUNT(*) AS count
FROM loan_data
WHERE Self_Employed = 'Yes' or Self_Employed = 'No'
GROUP BY Education, Self_Employed, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Education	Self_Employed	Loan_Status	count
Graduate	No	Y	145
Graduate	No	N	54
Not Graduate	No	Y	42
Not Graduate	No	N	23
Graduate	Yes	Y	11
Graduate	Yes	N	8
Not Graduate	Yes	Y	4
Not Graduate	Yes	N	2

```sql
-- 14. Is there a correlation between loan amount and property area in terms of loan approval?
SELECT Property_Area,
    CASE
        WHEN LoanAmount < (SELECT AVG(LoanAmount) FROM loan_data) THEN 'Below Avg'
        WHEN LoanAmount >= (SELECT AVG(LoanAmount) FROM loan_data) THEN 'Above Avg'

    END AS loan_amount_range,
    Loan_Status,
    COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, loan_amount_range, Loan_Status;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Property_Area	loan_amount_range	Loan_Status	count
Rural	Above Avg	N	25
Urban	Below Avg	Y	31
Urban	Above Avg	Y	33
Urban	Below Avg	N	20
Semiurban	Above Avg	N	11
Semiurban	Above Avg	Y	61
Urban	Above Avg	N	12
Semiurban	Below Avg	Y	41
Rural	Below Avg	N	10
Rural	Above Avg	Y	39
Semiurban	Below Avg	N	13
Rural	Below Avg	Y	12

```sql
-- 15. How does the number of dependents and education level jointly influence loan approval?
SELECT Dependents, Education, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Dependents, Education, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Dependents	Education	Loan_Status	count
0	Graduate	Y	116
0	Graduate	N	45
0	Not Graduate	Y	30
2	Graduate	Y	27
1	Graduate	Y	24
0	Not Graduate	N	17
2	Not Graduate	Y	12
1	Graduate	N	10
2	Graduate	N	10
1	Not Graduate	Y	8
2	Not Graduate	N	5
1	Not Graduate	N	4

```sql
-- 16. What is the distribution of credit history among different property areas?
SELECT Property_Area, Credit_History, COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, Credit_History
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Property_Area	Credit_History	count
Semiurban	1	110
Urban	1	83
Rural	1	69
Rural	0	17
Semiurban	0	16
Urban	0	13

```sql
-- 17. Are there any trends in loan approval status based on the loan amount term?
SELECT Loan_Amount_Term, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Loan_Amount_Term, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Loan_Amount_Term	Loan_Status	count
360	Y	190
360	N	73
180	Y	16
180	N	8
480	N	6
120	Y	3
300	N	3
480	Y	3
300	Y	2
84	Y	2
240	Y	1
36	N	1

```sql
-- 19. What is the average loan amount for self-employed applicants compared to non-self-employed ones?
SELECT Self_Employed, AVG(LoanAmount) AS avg_loan_amount
FROM loan_data
WHERE Self_Employed = 'Yes' or Self_Employed = 'No'
GROUP BY Self_Employed;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Self_Employed	avg_loan_amount
No	104.4015
Yes	104.3200

```sql
-- 18. What is the average income of applicants in each property area?
SELECT Property_Area, AVG(ApplicantIncome) AS avg_income
FROM loan_data
GROUP BY Property_Area;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Property_Area	avg_income
Rural	3708.0465
Urban	3443.7083
Semiurban	3441.1190

```sql
-- 20. Are there any significant differences in loan approval rates for different education levels in various property areas?
SELECT
    Education,
    Property_Area,
    Loan_Status,
    COUNT(*) AS count
FROM loan_data
GROUP BY Education, Property_Area, Loan_Status
ORDER BY count DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Education	Property_Area	Loan_Status	count
Graduate	Semiurban	Y	79
Graduate	Urban	Y	51
Graduate	Rural	Y	37
Graduate	Urban	N	24
Not Graduate	Semiurban	Y	23
Graduate	Rural	N	21
Graduate	Semiurban	N	20
Not Graduate	Rural	N	14
Not Graduate	Rural	Y	14
Not Graduate	Urban	Y	13
Not Graduate	Urban	N	8
Not Graduate	Semiurban	N	4

SQL Queries

-- 1. What is the average loan amount by credit history status?

```
SELECT Credit_History, AVG(LoanAmount) AS avg_loan_amount
FROM loan_data
GROUP BY Credit_History;
```

-- 2. How many loans were approved versus rejected?

```
SELECT Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Loan_Status;
```

-- 3. How many loans were approved for male applicants versus female applicants?

```
SELECT Gender, Loan_Status, COUNT(*) AS count
FROM loan_data
WHERE Gender = 'Male' or Gender = 'Female'
GROUP BY Gender, Loan_Status;
```

-- 4. Is there a difference in loan approval rates between married and unmarried applicants?

```
SELECT Married, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Married, Loan_Status;
```

-- 5. How does the education level of the applicant relate to loan approval status?

```
SELECT Education, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Education, Loan_Status;
```

-- 6. Are self-employed individuals more or less likely to get their loans approved?

```
SELECT Self_Employed, Loan_Status, COUNT(*) AS count
FROM loan_data
WHERE Self_Employed = 'Yes' or Self_Employed = 'No'
GROUP BY Self_Employed, Loan_Status;
```

-- 7. How does the location of the property (urban, semiurban, rural) influence loan approval?

```
SELECT Property_Area, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, Loan_Status
ORDER BY count DESC;
```

-- 8. Does the number of dependents impact the likelihood of loan approval?

```
SELECT Dependents, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Dependents, Loan_Status
ORDER BY count DESC;
```

-- 9. How does the term of the loan (in months) influence loan approval?

```
SELECT Loan_Amount_Term, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Loan_Amount_Term, Loan_Status
ORDER BY count DESC;
```

-- 10. Is there a correlation between income groups and loan approval status?

```
SELECT
   CASE
      WHEN ApplicantIncome < ( SELECT avg(ApplicantIncome) FROM loan_data ) THEN 'Below Average'
      ELSE 'Above Average'
   END AS income_group,
   Loan_Status,
   COUNT(*) AS count
FROM loan_data
GROUP BY income_group, Loan_Status
ORDER BY count DESC;
```

-- 11. How does the combined income of applicants and co-applicants relate to loan approval?

```
SELECT
   CASE
      WHEN ApplicantIncome + CoapplicantIncome < (SELECT AVG(ApplicantIncome) FROM loan_data)
THEN 'Below Avg Applicant Income'
      WHEN ApplicantIncome + CoapplicantIncome >= (SELECT AVG(ApplicantIncome) FROM loan_data)
THEN 'Above Avg Applicant Income'

   END AS combined_income_group,
   Loan_Status,
   COUNT(*) AS count
FROM loan_data
GROUP BY combined_income_group, Loan_Status;
```

-- 12. Does the combination of credit history and property area provide additional insights into loan approval?

```
SELECT Property_Area, Credit_History, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, Credit_History, Loan_Status
ORDER BY count DESC;
```

-- 13. How does the combination of education level and employment status affect loan approval?

```
SELECT Education, Self_Employed, Loan_Status, COUNT(*) AS count
FROM loan_data
WHERE Self_Employed = 'Yes' or Self_Employed = 'No'
GROUP BY Education, Self_Employed, Loan_Status
ORDER BY count DESC;
```

-- 14. Is there a correlation between loan amount and property area in terms of loan approval?

```
SELECT Property_Area,
   CASE
      WHEN LoanAmount < (SELECT AVG(LoanAmount) FROM loan_data) THEN 'Below Avg'
      WHEN LoanAmount >= (SELECT AVG(LoanAmount) FROM loan_data) THEN 'Above Avg'

   END AS loan_amount_range,
   Loan_Status,
   COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, loan_amount_range, Loan_Status;
```

-- 15. How does the number of dependents and education level jointly influence loan approval?

```
SELECT Dependents, Education, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Dependents, Education, Loan_Status
ORDER BY count DESC;
```

-- 16. What is the distribution of credit history among different property areas?

```
SELECT Property_Area, Credit_History, COUNT(*) AS count
FROM loan_data
GROUP BY Property_Area, Credit_History
ORDER BY count DESC;
```

-- 17. Are there any trends in loan approval status based on the loan amount term?

```
SELECT Loan_Amount_Term, Loan_Status, COUNT(*) AS count
FROM loan_data
GROUP BY Loan_Amount_Term, Loan_Status
ORDER BY count DESC;
```

-- 18. What is the average income of applicants in each property area?

```
SELECT Property_Area, AVG(ApplicantIncome) AS avg_income
FROM loan_data
GROUP BY Property_Area;
```

-- 19. What is the average loan amount for self-employed applicants compared to non-self-employed ones?

```
SELECT Self_Employed, AVG(LoanAmount) AS avg_loan_amount
FROM loan_data
WHERE Self_Employed = 'Yes' or Self_Employed = 'No'
GROUP BY Self_Employed;
```

-- 20. Are there any significant differences in loan approval rates for different education levels in various property areas?

```
SELECT
   Education,
   Property_Area,
   Loan_Status,
   COUNT(*) AS count
FROM loan_data
GROUP BY Education, Property_Area, Loan_Status
ORDER BY count DESC;
```