**TAIWAN DATASET INFORMATION**

**Dataset Information**

This dataset contains information on default payments, demographic factors, credit data, history of payment, and bill statements of credit card clients in Taiwan from April 2005 to September 2005.

**Content**

There are 25 variables:

* **ID**: ID of each client
* **LIMIT\_BAL**: Amount of given credit in NT dollars (includes individual and family/supplementary credit
* **SEX**: Gender (1=male, 2=female)
* **EDUCATION**: (1=graduate school, 2=university, 3=high school, 4=others, 5=unknown, 6=unknown)
* **MARRIAGE**: Marital status (1=married, 2=single, 3=others)
* **AGE**: Age in years
* **PAY\_0**: Repayment status in September, 2005 (-1=pay duly, 1=payment delay for one month, 2=payment delay for two months, … 8=payment delay for eight months, 9=payment delay for nine months and above)
* **PAY\_2**: Repayment status in August, 2005 (scale same as above)
* **PAY\_3**: Repayment status in July, 2005 (scale same as above)
* **PAY\_4**: Repayment status in June, 2005 (scale same as above)
* **PAY\_5**: Repayment status in May, 2005 (scale same as above)
* **PAY\_6**: Repayment status in April, 2005 (scale same as above)
* **BILL\_AMT1**: Amount of bill statement in September, 2005 (NT dollar)
* **BILL\_AMT2**: Amount of bill statement in August, 2005 (NT dollar)
* **BILL\_AMT3**: Amount of bill statement in July, 2005 (NT dollar)
* **BILL\_AMT4**: Amount of bill statement in June, 2005 (NT dollar)
* **BILL\_AMT5**: Amount of bill statement in May, 2005 (NT dollar)
* **BILL\_AMT6**: Amount of bill statement in April, 2005 (NT dollar)
* **PAY\_AMT1**: Amount of previous payment in September, 2005 (NT dollar)
* **PAY\_AMT2**: Amount of previous payment in August, 2005 (NT dollar)
* **PAY\_AMT3**: Amount of previous payment in July, 2005 (NT dollar)
* **PAY\_AMT4**: Amount of previous payment in June, 2005 (NT dollar)
* **PAY\_AMT5**: Amount of previous payment in May, 2005 (NT dollar)
* **PAY\_AMT6**: Amount of previous payment in April, 2005 (NT dollar)
* **default.payment.next.month**: Default payment (1=yes, 0=no)

Distribution of Variables:

Value counts for ID:

ID

1 1

19997 1

20009 1

20008 1

20007 1

..

9996 1

9995 1

9994 1

9993 1

30000 1

Name: count, Length: 30000, dtype: int64

--------------------------------------------------

Value counts for LIMIT\_BAL:

LIMIT\_BAL

50000.0 3365

20000.0 1976

30000.0 1610

80000.0 1567

200000.0 1528

...

730000.0 2

1000000.0 1

327680.0 1

760000.0 1

690000.0 1

Name: count, Length: 81, dtype: int64

--------------------------------------------------

Value counts for SEX:

SEX

2 18112

1 11888

Name: count, dtype: int64

--------------------------------------------------

Value counts for EDUCATION:

EDUCATION

2 14030

1 10585

3 4917

5 280

4 123

6 51

0 14

Name: count, dtype: int64

--------------------------------------------------

Value counts for MARRIAGE:

MARRIAGE

2 15964

1 13659

3 323

0 54

Name: count, dtype: int64

--------------------------------------------------

Value counts for AGE:

AGE

29 1605

27 1477

28 1409

30 1395

26 1256

31 1217

25 1186

34 1162

32 1158

33 1146

24 1127

35 1113

36 1108

37 1041

39 954

38 944

23 931

40 870

41 824

42 794

44 700

43 670

45 617

46 570

22 560

47 501

48 466

49 452

50 411

51 340

53 325

52 304

54 247

55 209

56 178

58 122

57 122

59 83

60 67

21 67

61 56

62 44

63 31

64 31

66 25

65 24

67 16

69 15

70 10

68 5

73 4

72 3

75 3

71 3

79 1

74 1

Name: count, dtype: int64

--------------------------------------------------

Value counts for PAY\_0:

PAY\_0

0 14737

-1 5686

1 3688

-2 2759

2 2667

3 322

4 76

5 26

8 19

6 11

7 9

Name: count, dtype: int64

--------------------------------------------------

Value counts for PAY\_2:

PAY\_2

0 15730

-1 6050

2 3927

-2 3782

3 326

4 99

1 28

5 25

7 20

6 12

8 1

Name: count, dtype: int64

--------------------------------------------------

Value counts for PAY\_3:

PAY\_3

0 15764

-1 5938

-2 4085

2 3819

3 240

4 76

7 27

6 23

5 21

1 4

8 3

Name: count, dtype: int64

--------------------------------------------------

Value counts for PAY\_4:

PAY\_4

0 16455

-1 5687

-2 4348

2 3159

3 180

4 69

7 58

5 35

6 5

1 2

8 2

Name: count, dtype: int64

--------------------------------------------------

Value counts for PAY\_5:

PAY\_5

0 16947

-1 5539

-2 4546

2 2626

3 178

4 84

7 58

5 17

6 4

8 1

Name: count, dtype: int64

--------------------------------------------------

Value counts for PAY\_6:

PAY\_6

0 16286

-1 5740

-2 4895

2 2766

3 184

4 49

7 46

6 19

5 13

8 2

Name: count, dtype: int64

--------------------------------------------------

Value counts for BILL\_AMT1:

BILL\_AMT1

0.0 2008

390.0 244

780.0 76

326.0 72

316.0 63

...

11636.0 1

146034.0 1

20817.0 1

1351.0 1

47929.0 1

Name: count, Length: 22723, dtype: int64

--------------------------------------------------

Value counts for BILL\_AMT2:

BILL\_AMT2

0.0 2506

390.0 231

326.0 75

780.0 75

316.0 72

...

26357.0 1

85195.0 1

6889.0 1

11004.0 1

48905.0 1

Name: count, Length: 22346, dtype: int64

--------------------------------------------------

Value counts for BILL\_AMT3:

BILL\_AMT3

0.0 2870

390.0 275

780.0 74

326.0 63

316.0 62

...

19580.0 1

45129.0 1

227807.0 1

39330.0 1

49764.0 1

Name: count, Length: 22026, dtype: int64

--------------------------------------------------

Value counts for BILL\_AMT4:

BILL\_AMT4

0.0 3195

390.0 246

780.0 101

316.0 68

326.0 62

...

97189.0 1

118839.0 1

23567.0 1

18377.0 1

36535.0 1

Name: count, Length: 21548, dtype: int64

--------------------------------------------------

Value counts for BILL\_AMT5:

BILL\_AMT5

0.0 3506

390.0 235

780.0 94

316.0 79

326.0 62

...

19341.0 1

66726.0 1

80682.0 1

28508.0 1

32428.0 1

Name: count, Length: 21010, dtype: int64

--------------------------------------------------

Value counts for BILL\_AMT6:

BILL\_AMT6

0.0 4020

390.0 207

780.0 86

150.0 78

316.0 77

...

26852.0 1

108660.0 1

480.0 1

15298.0 1

15313.0 1

Name: count, Length: 20604, dtype: int64

--------------------------------------------------

Value counts for PAY\_AMT1:

PAY\_AMT1

0.0 5249

2000.0 1363

3000.0 891

5000.0 698

1500.0 507

...

3391.0 1

7785.0 1

66022.0 1

10121.0 1

85900.0 1

Name: count, Length: 7943, dtype: int64

--------------------------------------------------

Value counts for PAY\_AMT2:

PAY\_AMT2

0.0 5396

2000.0 1290

3000.0 857

5000.0 717

1000.0 594

...

7866.0 1

6206.0 1

10529.0 1

21300.0 1

3526.0 1

Name: count, Length: 7899, dtype: int64

--------------------------------------------------

Value counts for PAY\_AMT3:

PAY\_AMT3

0.0 5968

2000.0 1285

1000.0 1103

3000.0 870

5000.0 721

...

5102.0 1

5368.0 1

28138.0 1

549.0 1

25128.0 1

Name: count, Length: 7518, dtype: int64

--------------------------------------------------

Value counts for PAY\_AMT4:

PAY\_AMT4

0.0 6408

1000.0 1394

2000.0 1214

3000.0 887

5000.0 810

...

18916.0 1

3468.0 1

11476.0 1

4363.0 1

8049.0 1

Name: count, Length: 6937, dtype: int64

--------------------------------------------------

Value counts for PAY\_AMT5:

PAY\_AMT5

0.0 6703

1000.0 1340

2000.0 1323

3000.0 947

5000.0 814

...

9111.0 1

16496.0 1

4819.0 1

10078.0 1

52964.0 1

Name: count, Length: 6897, dtype: int64

--------------------------------------------------

Value counts for PAY\_AMT6:

PAY\_AMT6

0.0 7173

1000.0 1299

2000.0 1295

3000.0 914

5000.0 808

...

7329.0 1

6862.0 1

6525.0 1

11894.0 1

16080.0 1

Name: count, Length: 6939, dtype: int64

--------------------------------------------------

Value counts for default.payment.next.month:

default.payment.next.month

0 23364

1 6636

Name: count, dtype: int64

--------------------------------------------------