**Analytics in Finance:**

The database contains default of credit card payments for 30000 customers in Taiwan. It has 23 variables as explanatory variables including:

1. LIMIT\_BAL: Amount of the given credit (NT dollar): it includes both the individual consumer credit and his/her family (supplementary) credit.
2. 'SEX': Gender (1 = male; 2 = female).
3. Education: (1 = graduate school; 2 = university; 3 = high school; 4 = others).
4. Marital status: (1 = married; 2 = single; 3 = others).
5. Age (year)
6. PAY\_0 -PAY\_6: History of past payment. We tracked the past monthly payment records (from April to September, 2005) as follows: PAY\_0 = the repayment status in September, 2005; PAY\_1 = the repayment status in August, 2005; . . .;PAY\_6 = the repayment status in April, 2005. The measurement scale for the repayment status is: -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.
7. BILL\_AMT1-BILL\_AMT6: Amount of bill statement (NT dollar). BILL\_AMT1 = amount of bill statement in September, 2005; BILL\_AMT2 = amount of bill statement in August, 2005; . . .; BILL\_AMT6 = amount of bill statement in April, 2005.
8. PAY\_AMT1-PAY\_AMT6: Amount of previous payment (NT dollar). PAY\_AMT1 = amount paid in September, 2005; PAY\_AMT2 = amount paid in August, 2005; . . .;PAY\_AMT6 = amount paid in April, 2005.
9. default.payment.next.month(1= default, 0=no-default)

**Questions:**

1. Predict the default rate and assign it to each customer?
2. Which factor(s) is playing significant role in this analysis?
3. Use appropriate plot to derive insight before starting your analysis.
4. Tell a good story about what you have found, even if it is counter intuitive.