

## Research Qualification Test 2018

City University of Hong Kong  
College of Business



### Deadline

Timestamp: 72 hours from you opening the test assignment

### Submission

(1) Report -- A summary report of your methodology and findings. Including both text and graphs. Maximum 4 pages.

(2) Code -- The code you wrote to conduct the analysis. It could be R, Python, Matlab or any other language you use.

### Homework

You work for a lending company that provides small loans to individual borrowers for 3 or 5 years. The company has collected data on the loans granted in the past few years. Your responsibility is to conduct analysis on the data and uncover insights about the loans' performance.

Good loans are the ones where the status is current or fully paid. Bad loans are the ones where the status is default, in grace period or late.

**The dataset in csv format is attached as loan.csv.**

### Data Field Definition

- id Borrower's id
- loan\_amount Amount of loan
- term Term of loan
- int\_rate Interest rate
- installment Monthly payment
- emp\_length Borrower's employment length
- home\_ownership Borrower's home ownership
- annual\_inc Borrower's annual income
- loan\_status Status of the loan
- purpose Purpose of the loan
- addr\_state Borrower's residence state
- dti Borrower's debt-to-income ratio
- delinq\_2yrs The frequency of the borrower's delinquencies in the last 2 years.
- earliest\_cr\_line The date the borrower started his/her first credit line
- mths\_since\_last\_delinq The number of months since the borrower's last delinquency
- open\_acc The number of the borrower's open credit accounts
- revol\_bal The borrower's revolving balance
- total\_acc The number of the borrower's total credit accounts, including currently open & closed
- out\_prncp Outstanding principal of the loan
- total\_pymnt Total payment on the loan
- total\_rec\_prncp Total received principal on the loan
- total\_rec\_int Total received interest on the loan