Dataset 1

Gunakan dataset train1.csv untuk menjawab pertanyaan 1-6.

Dataset berikut merupakan data demografi 1556 nasabah bank pada kuarter ke-4 tahun 2017, data ini kemudian dibagi menjadi data train dan test. Variabel target pada data ini adalah not\_paid, variabel biner yang menjadi indikasi suatu loan/ pinjaman lunas (berhasil dilunasi pembayarannya) atau tidak. Suatu pinjaman dikatakan not\_paid (not paid = 1) jika terjadi *default* (gagal bayar), *Charged Off*, atau lewat batas akhir pembayaran (*Grace Period*). Dataset dapat didownload pada tombol download dibawah dan dilakukan pengolahan menggunakan R ataupun python.

Download

Variables Glossary

* initial\_list\_status : indikasi loan termasuk dalam kategori w (whole) atau f(fractional).
* purpose : tujuan peminjaman (loan) terbagi atas 5 kategory yaitu untuk credit\_card, debt\_consolidation, home\_improvement, major\_purchase dan small\_business
* int\_rate : Interest rate (suku bunga) dalam prosentase
* installment : banyaknya installment/ uang bulanan yang wajib dibayarkan peminjam
* annual\_inc : income/ pemasukan tahunan peminjam sesuai yang tertulis saat proses pengajuan pinjaman
* dti : rasio antara pinjaman bulanan yang wajib dibayarkan peminjam dengan gaji/pemasukan peminjam sesuai report
* verification\_status : status verifikasi report pemasukan/gaji peminjam, terbagi atas kategori income verified, not verified, atau source was verified
* grade : grade loan berdasarkan software
* revol\_bal : total kredit dalam revolving balance (pinjaman yang tidak terbayarkan)
* inq\_last\_12m : banyaknya kredit/pinjaman pada akhir bulan 12
* delinq\_2yrs : banyaknya hari telat bayar untuk kriteria 30+ pada histori peminjam selama 2 tahun terakhir
* home\_ownership : kategori kepemilikan rumah peminjam meliputi MORTGAGE, OWN atau RENT
* log\_inc : log dari annual\_inc
* verified : 0 untuk Not\_verified masih dibawah status verifikasi, 1 lainnya
* grdCtoA : 1 untuk grade kredit A, B atau C; 0 untuk grade loan lainnya