

HUIZHE (SUNNY) ZHU



BANK LOAN DEFAULT PREDICTION EXECUTIVE SUMMARY

TAKE-AWAYS AND INSIGHTS WEEK 2

KEY TAKEAWAY

Fraud Analytics using descriptive, predictive and social network techniques

1. fraud types: credit card fraud, insurance fraud, healthcare fraud..
2. fraud detection + fraud prevention
 - expert knowledge matters as much as technical skills
3. big data for fraud detection
 - precision: involve massive data
 - operational efficiency: quick
 - cost efficiency: expert based system is labor intensive & challenging
4. limitation of predictive analysis
 - if fraud type is dramatically different from history
 - encourage to use both supervised & unsupervised

KEY TAKEAWAY

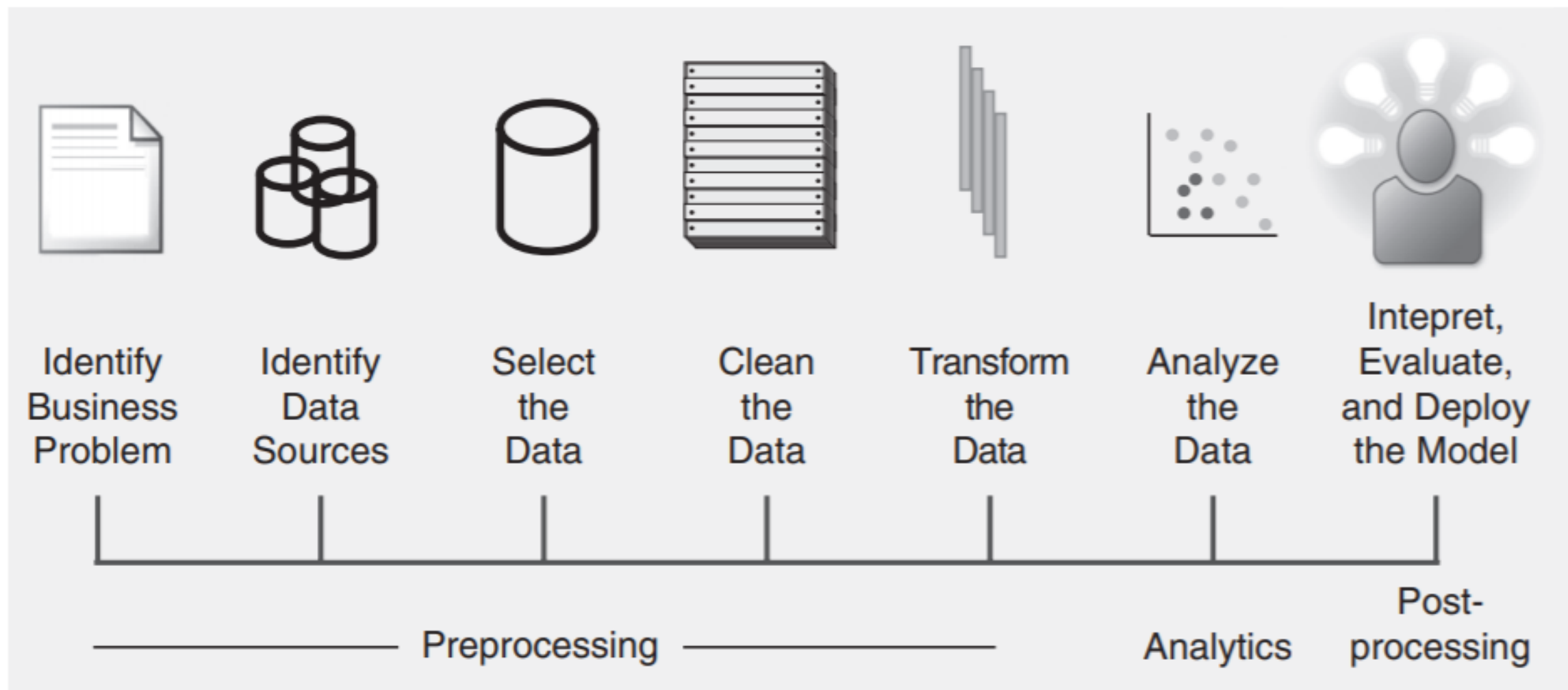
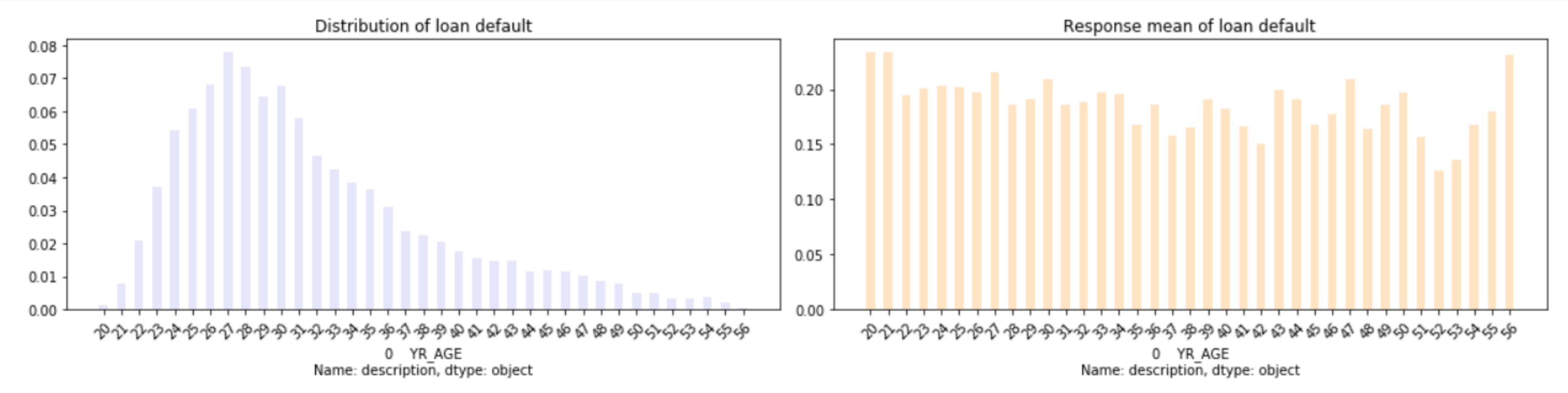
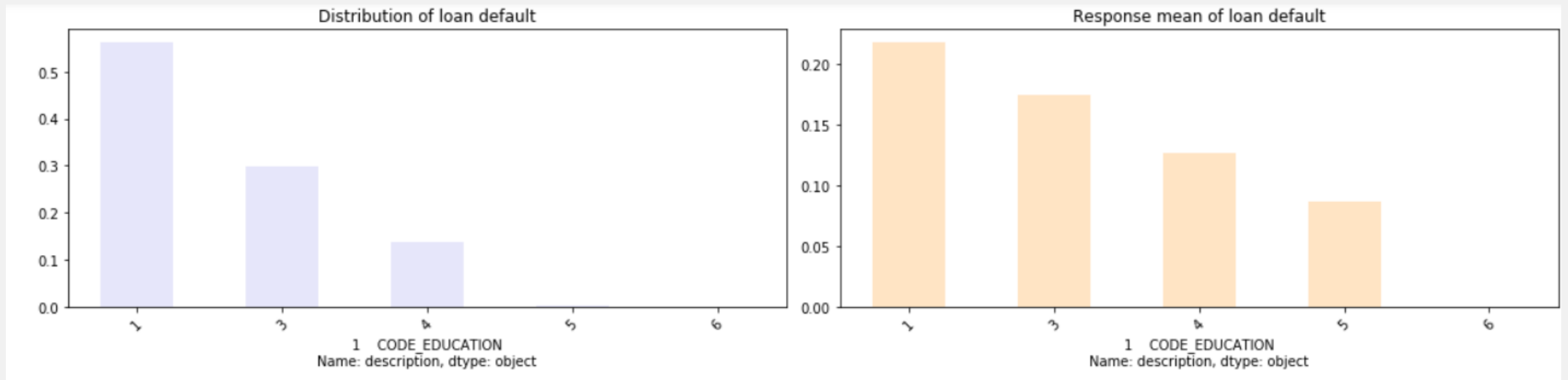


Figure 1.6 The Fraud Analytics Process Model

DEFAULT RISK AND APPLICANT AGE GROUP

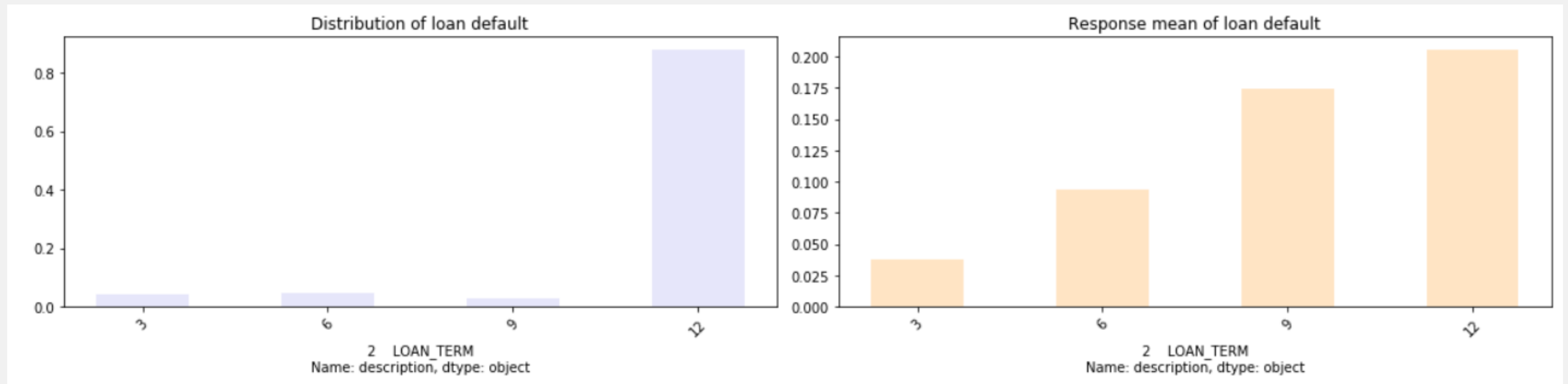


DEFAULT RISK AND APPLICANT EDUCATION LEVEL



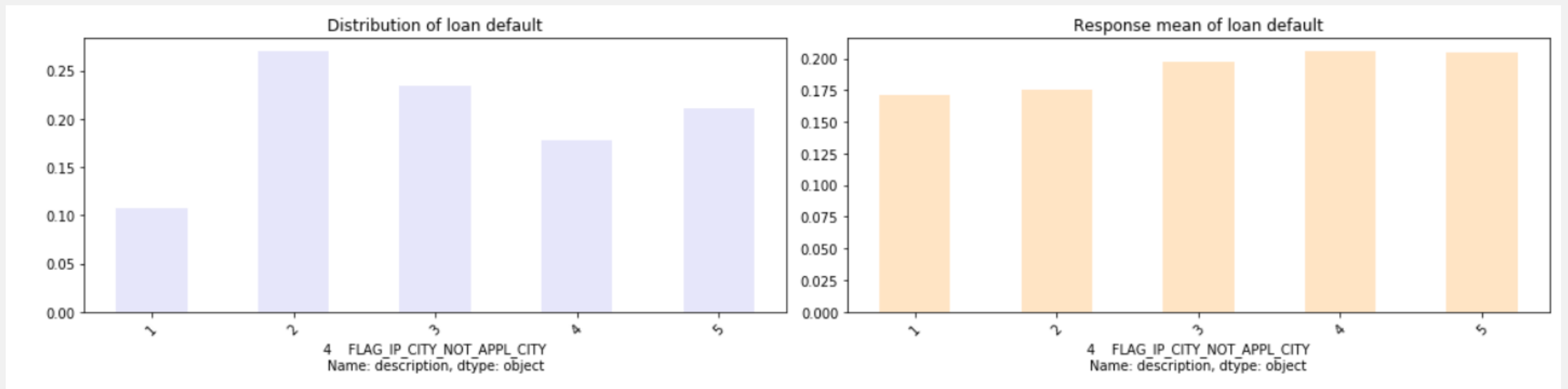
- applicants with less education are more likely to default

DEFAULT RISK AND LENGTH OF LOAN TERM



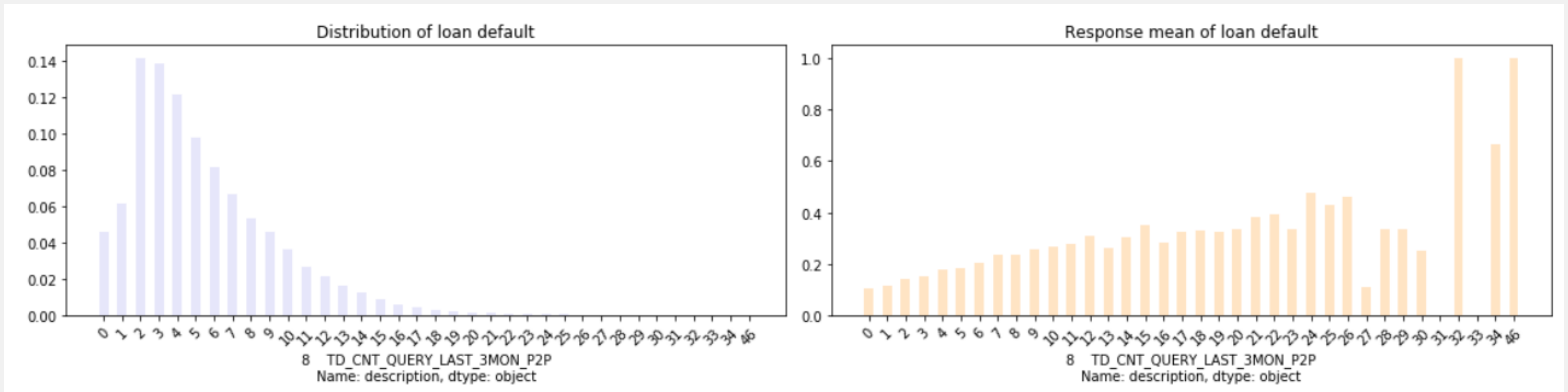
- applicants with longer period of loan are more likely to default

DEFAULT RISK AND LOCATION OF IP



- applicants whose IP is not the same as application IP are more likely to default

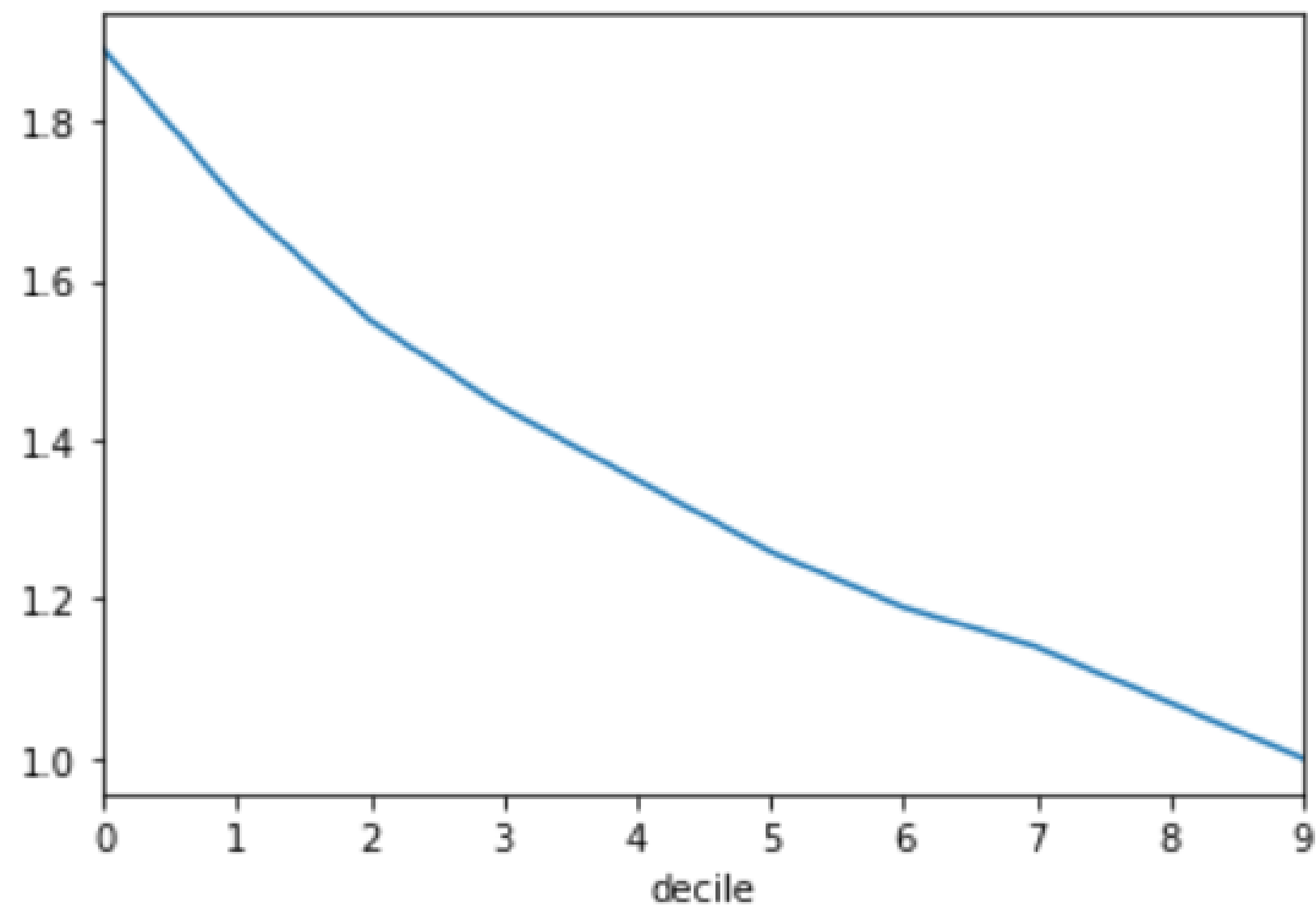
DEFAULT RISK AND QUERY HISTORY ON P2P



- applicants who have continued query from P2P are more likely to default

MODEL EVALUATION

accumulative lift



ROC _auc_score = 0.65

