

# LendingClub Loan Classification

Module 5 Project  
Andrew Cole



# Overview

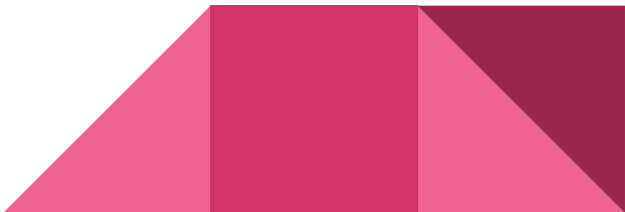
1. Stakeholder Information
2. Methodology
3. Data Exploration & Preprocessing
4. Model Selection
5. Actionable Insights



# Stakeholder Information

- Lending Club is a peer-to-peer lending platform for various types of loans
- LC offers credit at a lower cost than traditional loan sources
- LC provides servicing & screening of loan applicants
- Investors can purchase 'notes' of loans

**Mission Statement:** To develop a machine learning classification model, given applicants' loan application data, to accurately classify whether the applicant will default on the loan.



# Data Exploration & Preprocessing

- LendingClub Statistics - Website
- 24 features, 38646 entries (Misc. Loan Titles; 2007 - 2011)
- Features: Income, Loan Purpose, FICO Avg., Grade, Term, etc.
- Target Variable:
  - Loan Status: **Default** or **Fully Charged Off**
- Percentage of borrowers with loans fully paid: 85.75 %



# Model Selection

- 7 Classification models trained & tested
  - KNN, Random Forest, SVM, Logistic, XGboost etc.
- Evaluation Methods:

Model	Accuracy Score	AUC Score
Random Forest DT	59.21 %	0.61
Random Forest GridSearch	68.90 %	
SVM	63.41	0.62
XGboost		
KNN	72.38 %	0.53
Logistic Regression	65.01%	0.63

# Model Selection

- 7 Classification models trained & tested
  - KNN, Random Forest, SVM, Logistic, XGboost etc.
- Evaluation Methods:

Model	Accuracy Score	AUC Score
Random Forest DT	59.21 %	0.61
Random Forest GridSearch	68.90 %	
SVM	63.41%	0.62
XGboost		
KNN	72.38 %	0.53
Logistic Regression	65.01%	0.63

# Actionable Insights

