

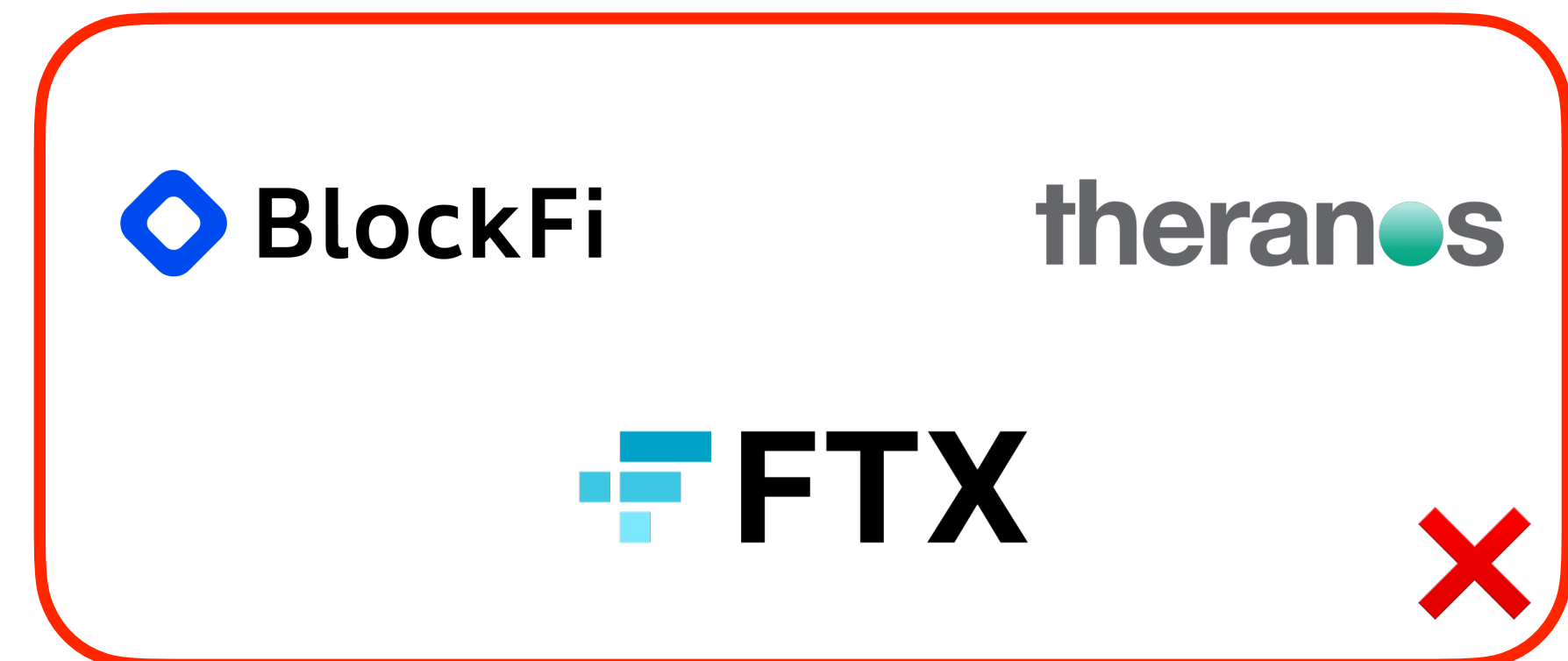
# Project RiskLens

---

## Predicting Startup Success for Sequoia Venture Capital

# Project RiskLens

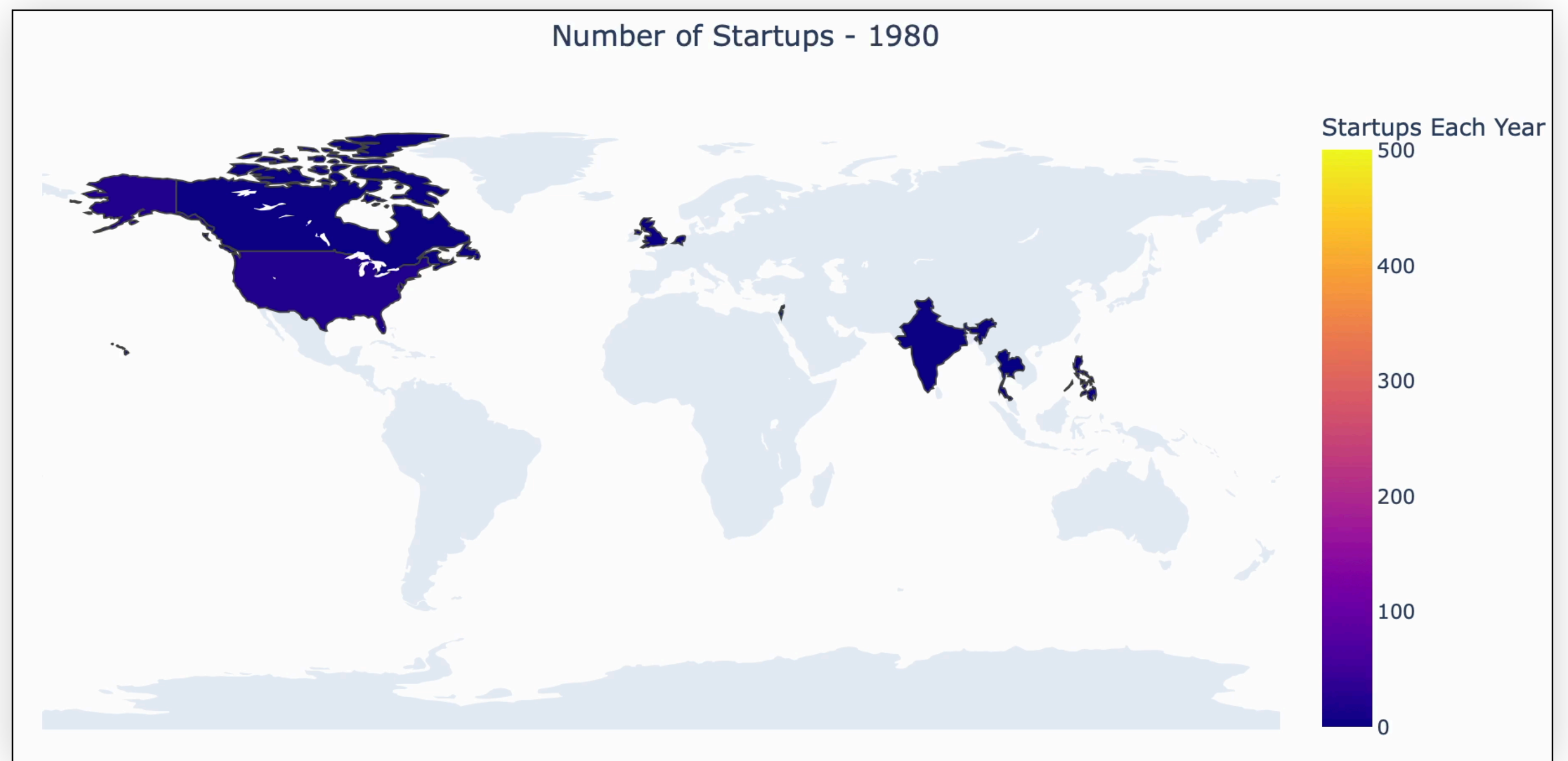
- Venture Capital is high-risk – more than 90% startups fail.
- Even slight improvements in investment decisions can drive multi-million dollar returns.
- Business Goal – predict startup success (IPO/M&A) to complement due diligence and help investment teams prioritize deals.



# Data

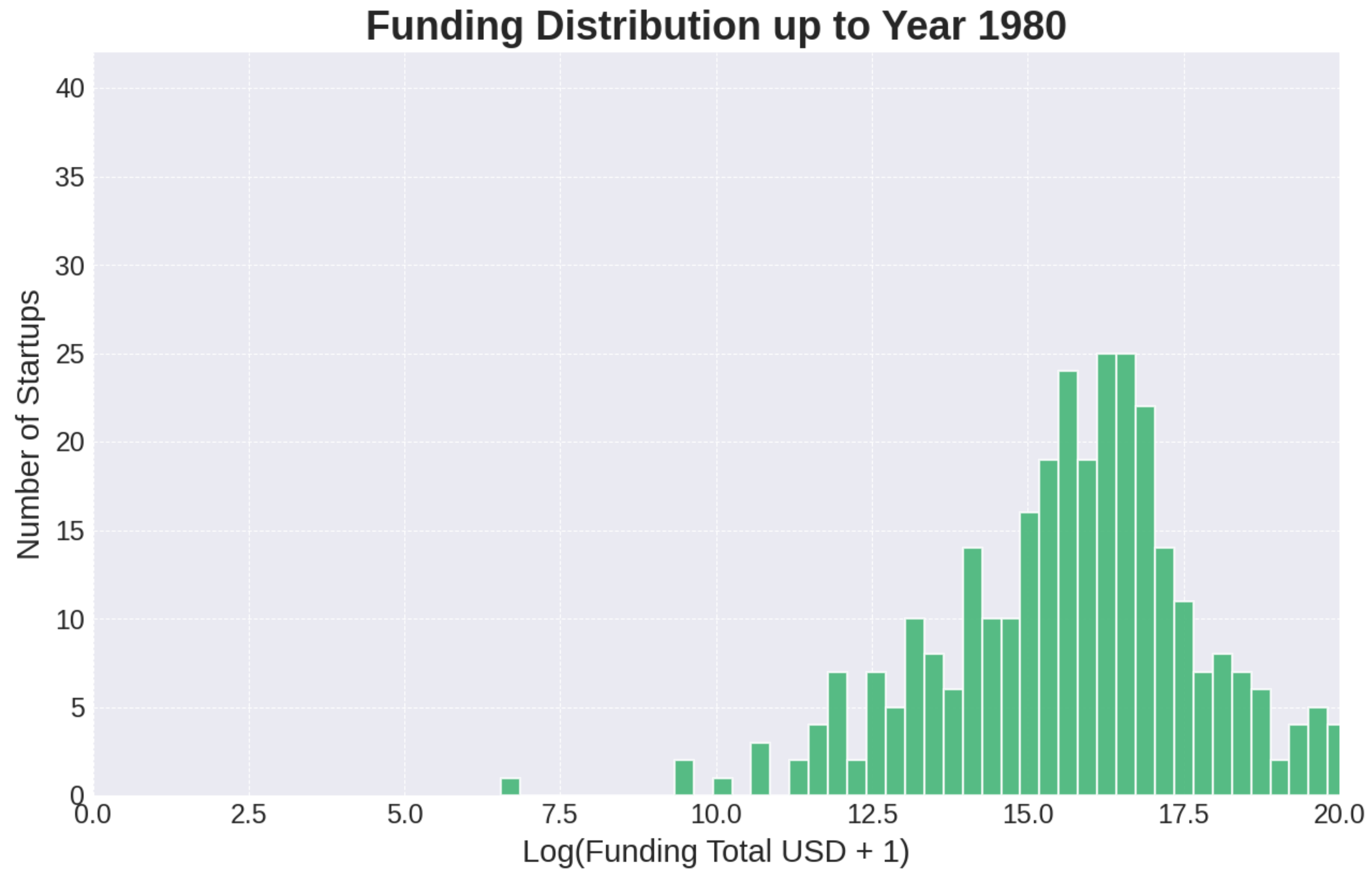
## Crunchbase Dataset

- 54,294 rows with 39 features
- Data about recognized startups founded between 1902 and 2014 across diverse industries and geographies



*Startup Growth Trajectory Across Countries (1980–2014)*

# Data



# Data Transformation

---

## Time Related Features

Founding Year, Diff. Funding Weeks  
First Funding At, Last Funding At

## Funding & Financial Features

Funding Total, Funding Rounds, Seed, Venture,  
Angel, Grant, Private Equity, Round A-H

## Industry & Category

Category List, Market, Industry Group

## Geographic Features

Country, State, Region, City

## Startup Status

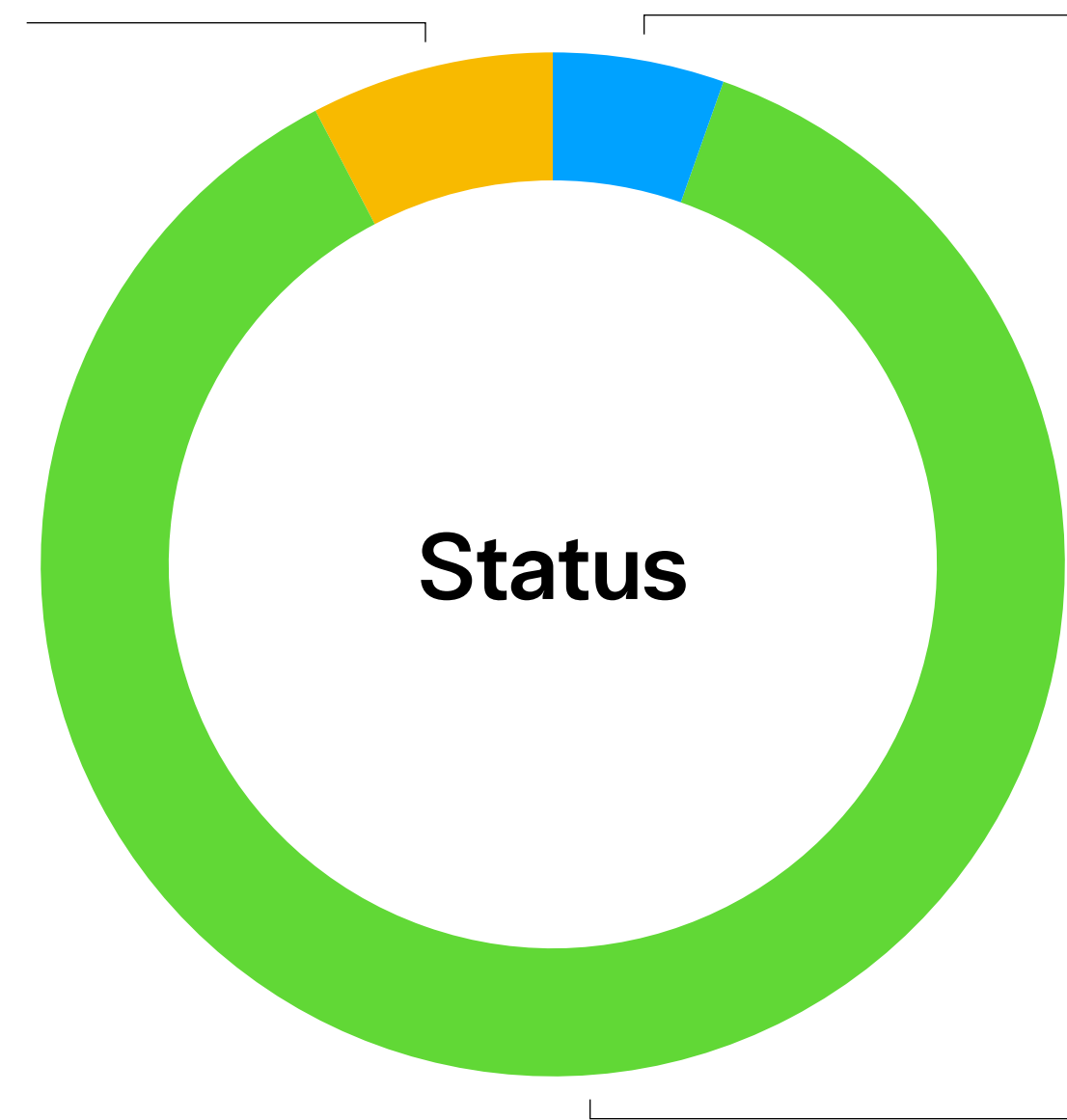
Operating, Acquired, Closed

## Feature Engineering

Filed IPO, Funding Time, Funding Bins

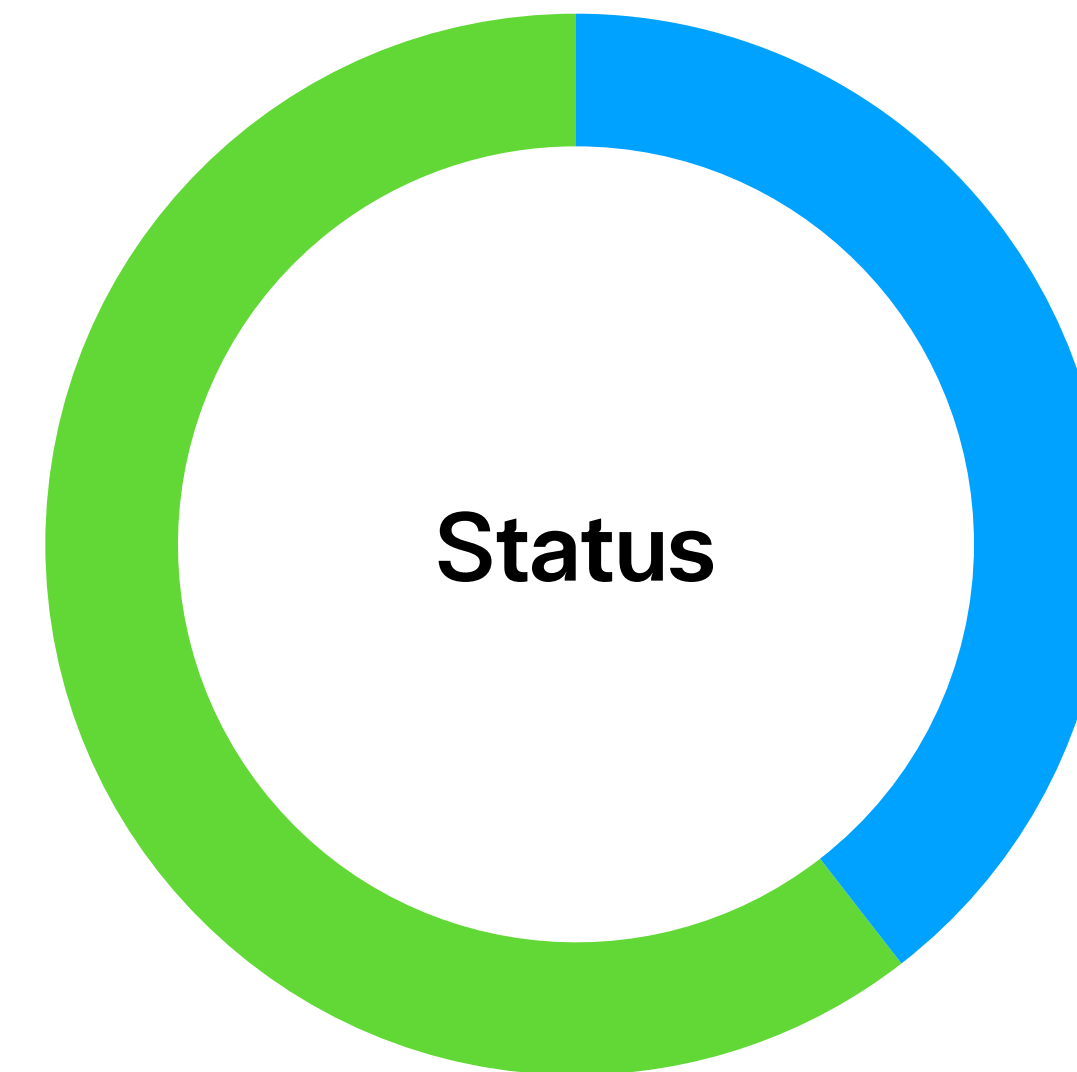
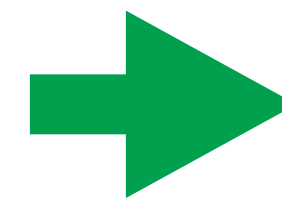
# Data Transformation

---



Highly Imbalanced Dataset

- Closed
- Operating
- Acquired



Balanced Dataset

- Closed
- Successful

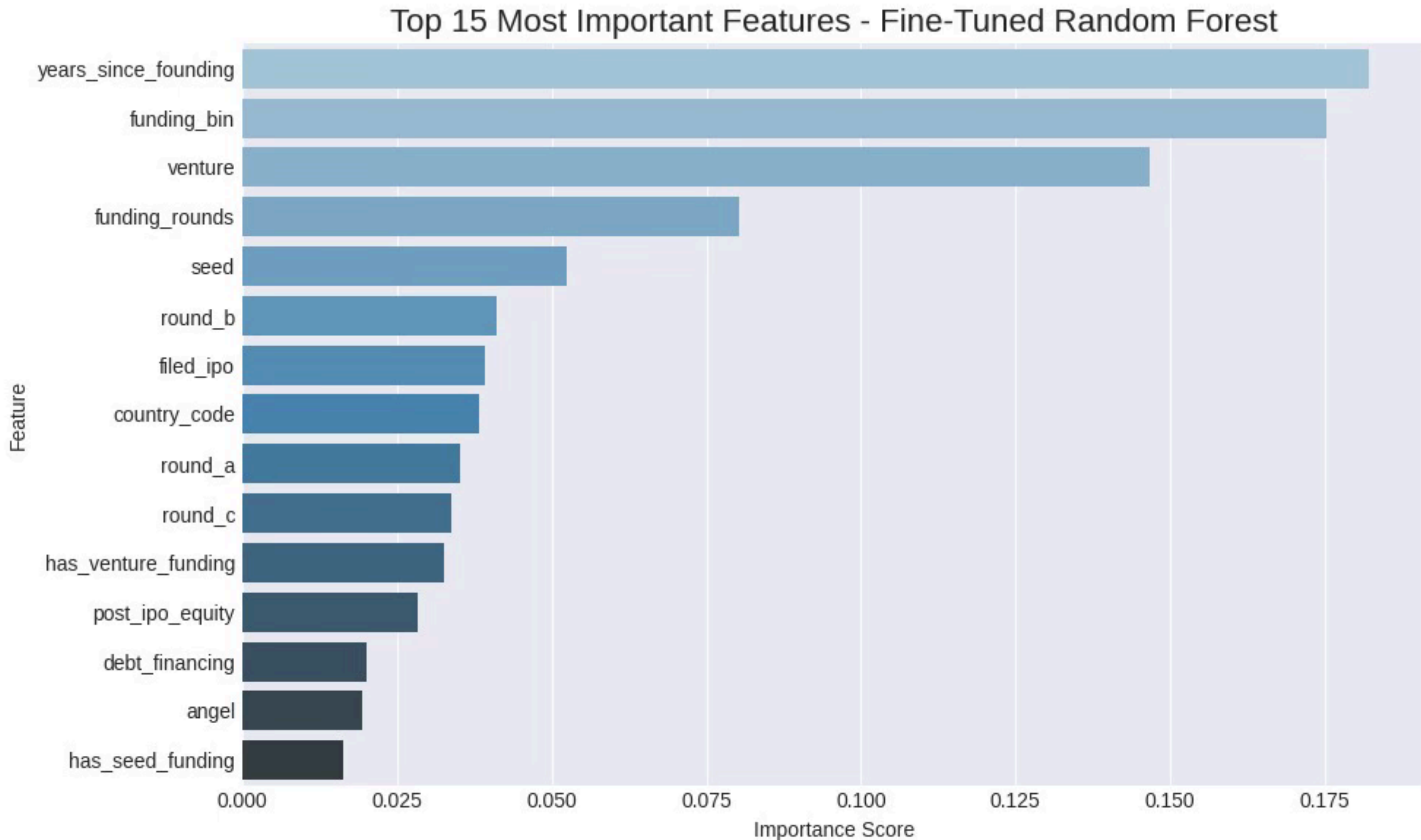
# Modeling and Performance

# Models

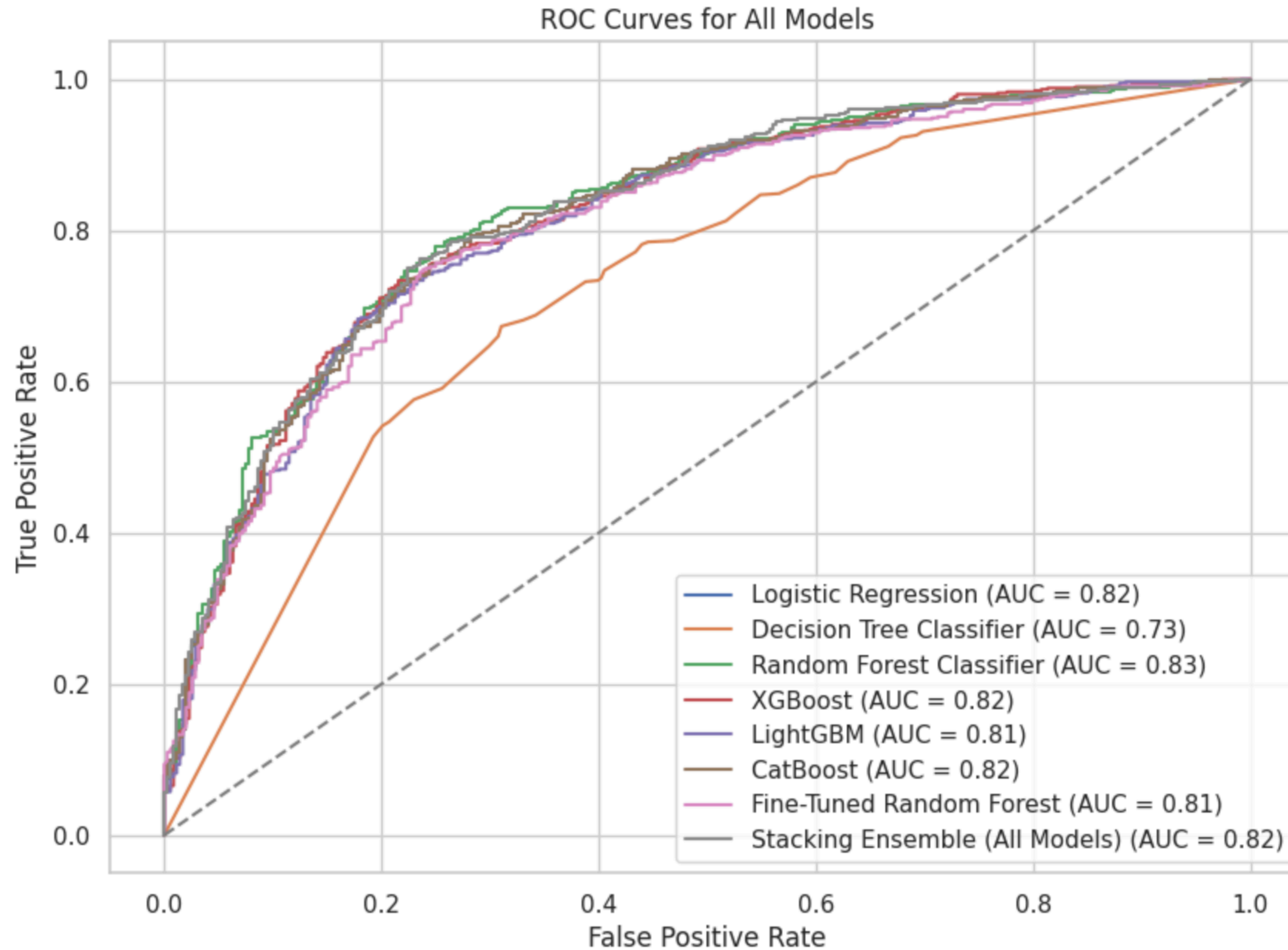
Model	Accuracy	Precision	Recall	F1-Score	AUC
Decison Tree Classifier	0.66	0.76	0.64	0.69	0.70
Logistic Regression	0.72	0.79	0.73	0.76	0.80
Random Forest	0.73	0.75	0.84	0.79	0.80
XGBoost	0.72	0.79	0.73	0.76	0.79
CatBoost	0.72	0.79	0.74	0.76	0.80
LightGBM	0.70	0.78	0.71	0.74	0.79
Stacking Ensemble	0.73	0.79	0.74	0.77	0.80



# Features Importance

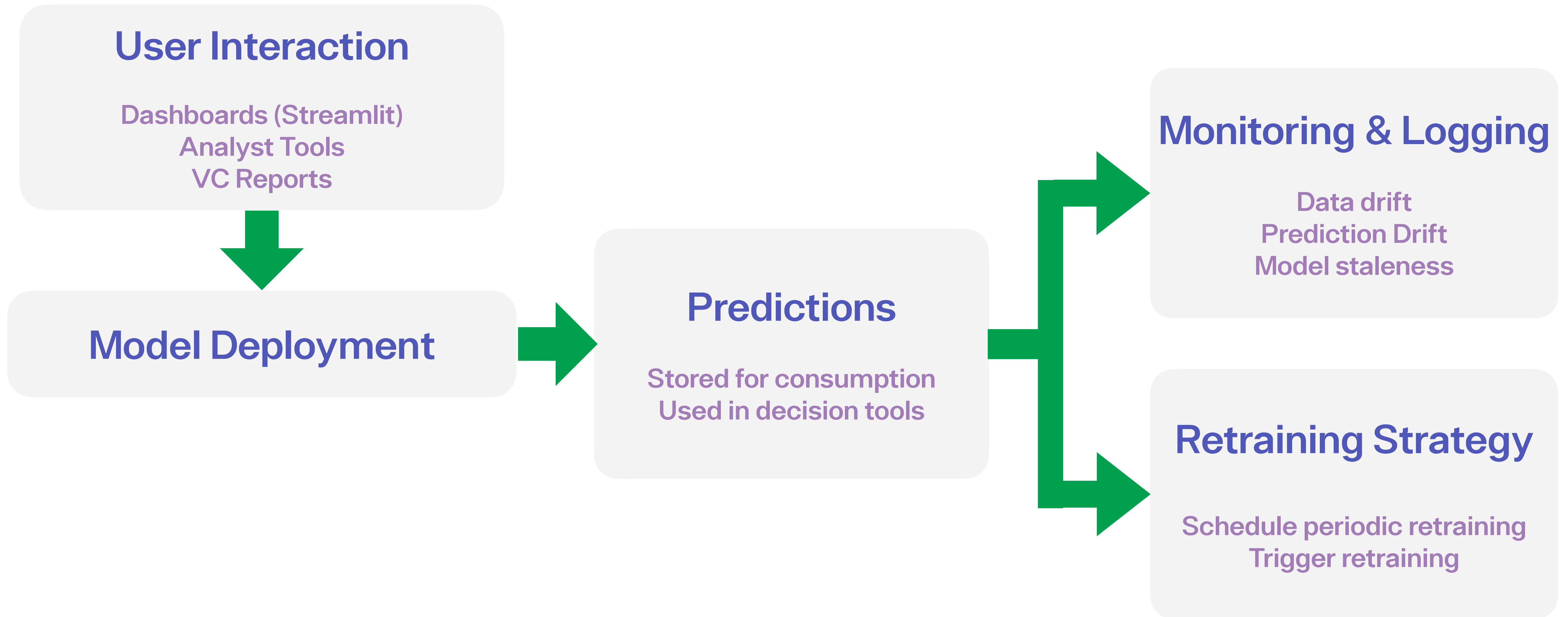


# Model Evaluation



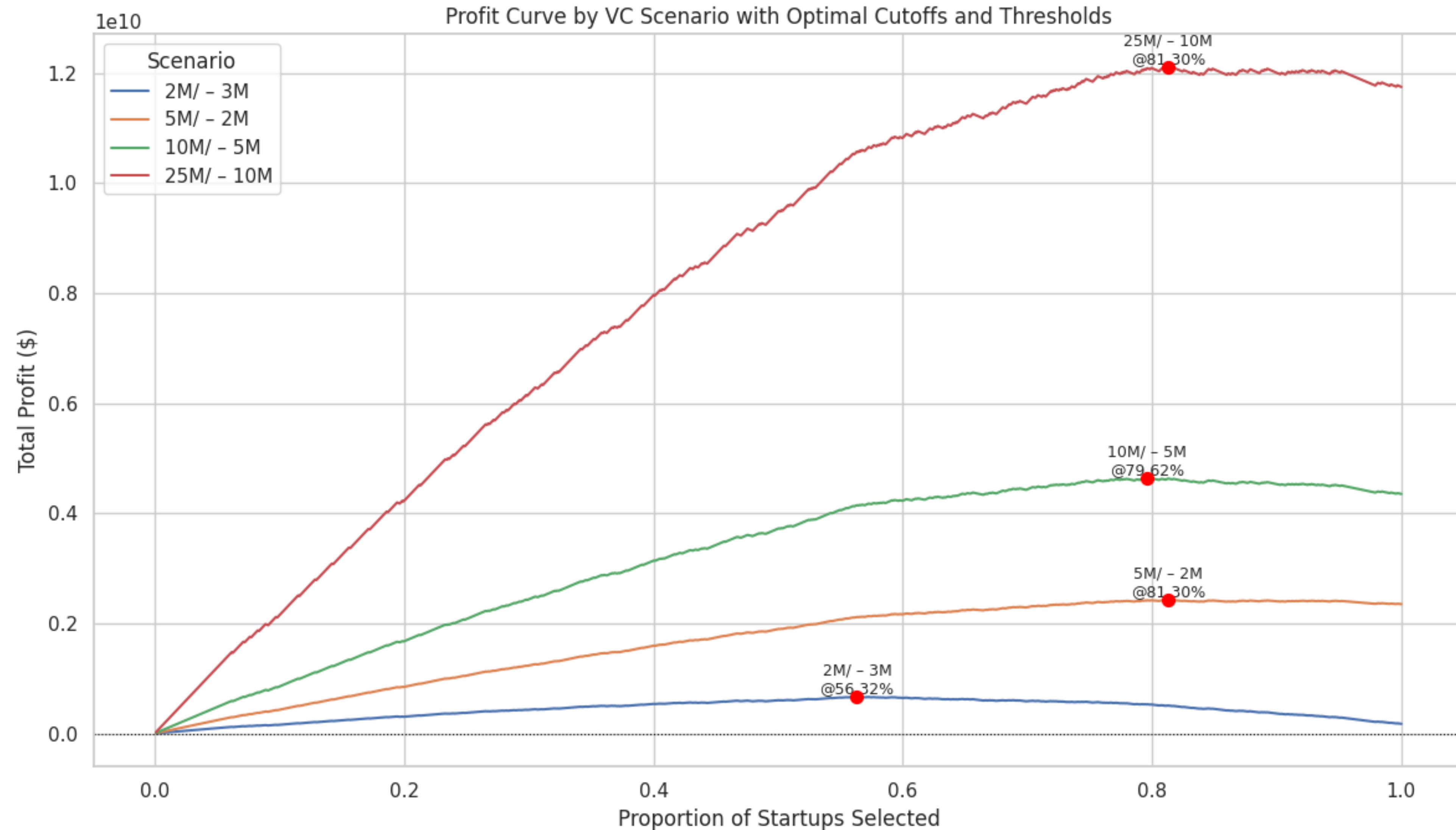
# Model Deployment & Operations

---



# Impact & Future

# Estimated Profit



# Business Impact

---



## Reduced Risk

Smarter, faster investment decisions



## Higher Recall

More likely to catch future unicorns



## Data-driven Screening

Focus analyst time on highest-potential startups

# Further Improvements

---

## External Data Integration

Founder profiles, patents, market trends

## Survival Analysis

Time-to-exit prediction

## Real-Time Monitoring

Update model quarterly to handle market drift

**Thank you for listening!**  
**Open for questions**