

ASH HAMBARDZUMYAN

PREDICTING BUILDING PERMIT APPROVALS IN NYC

INTRODUCTION

WHO WANTS TO KNOW HOW LONG WILL IT TAKE TO GET A BUILDING PERMIT IN THE BIG APPLE?



ARCHITECTS &
ENGINEERS



REAL ESTATE INVESTORS

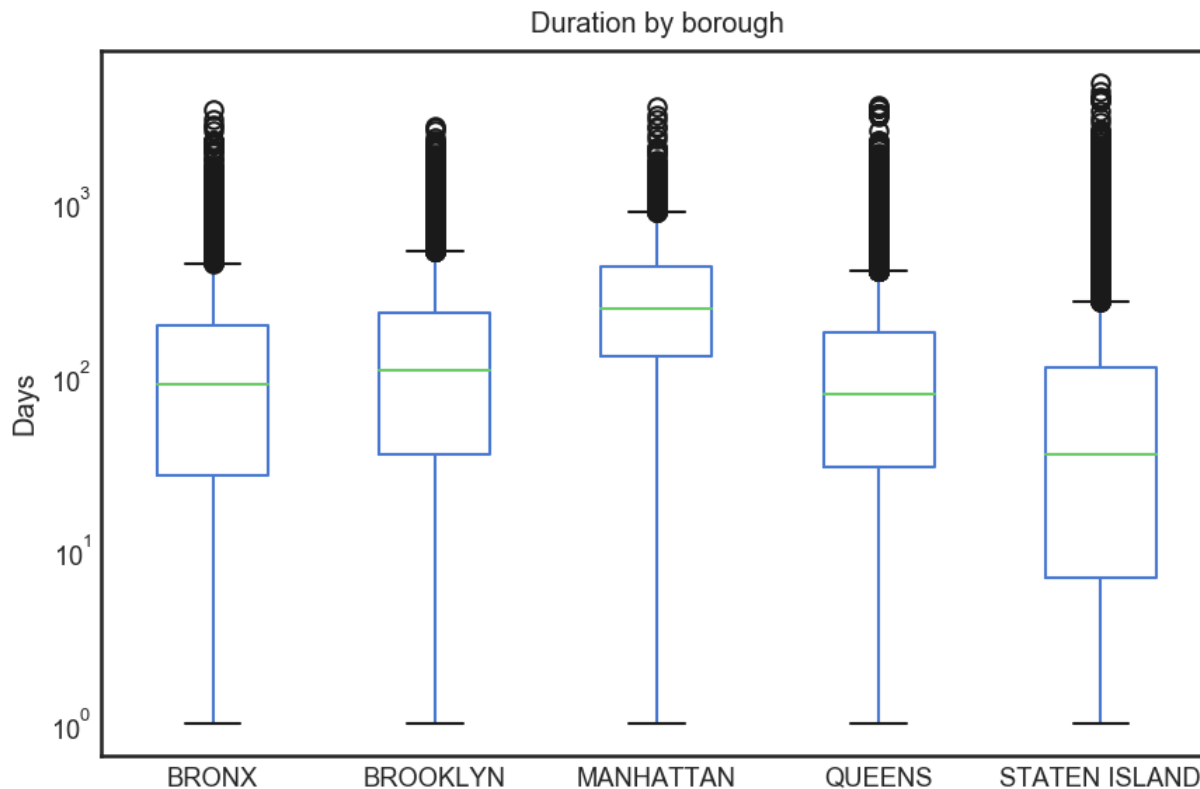


CONSTRUCTION COMPANIES &
CONTRACTORS



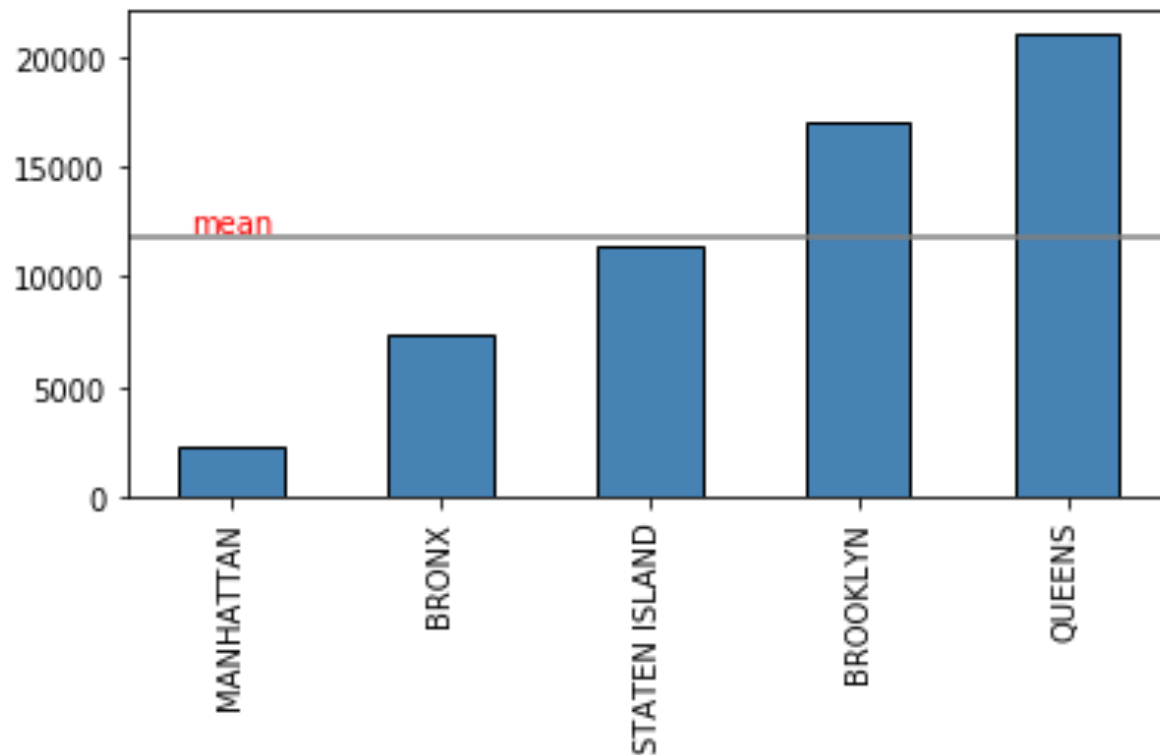
CITY ADMINISTRATION
(DEPARTMENT OF BUILDINGS)

DURATION OF THE BUILDING PERMIT APPLICATION VARIES FROM BOROUGH TO BOROUGH



- ▶ Longest processing times are in MANHATTAN
- ▶ Shortest approvals are in STATEN ISLAND

NUMBER OF BUILDING PERMITS ISSUED

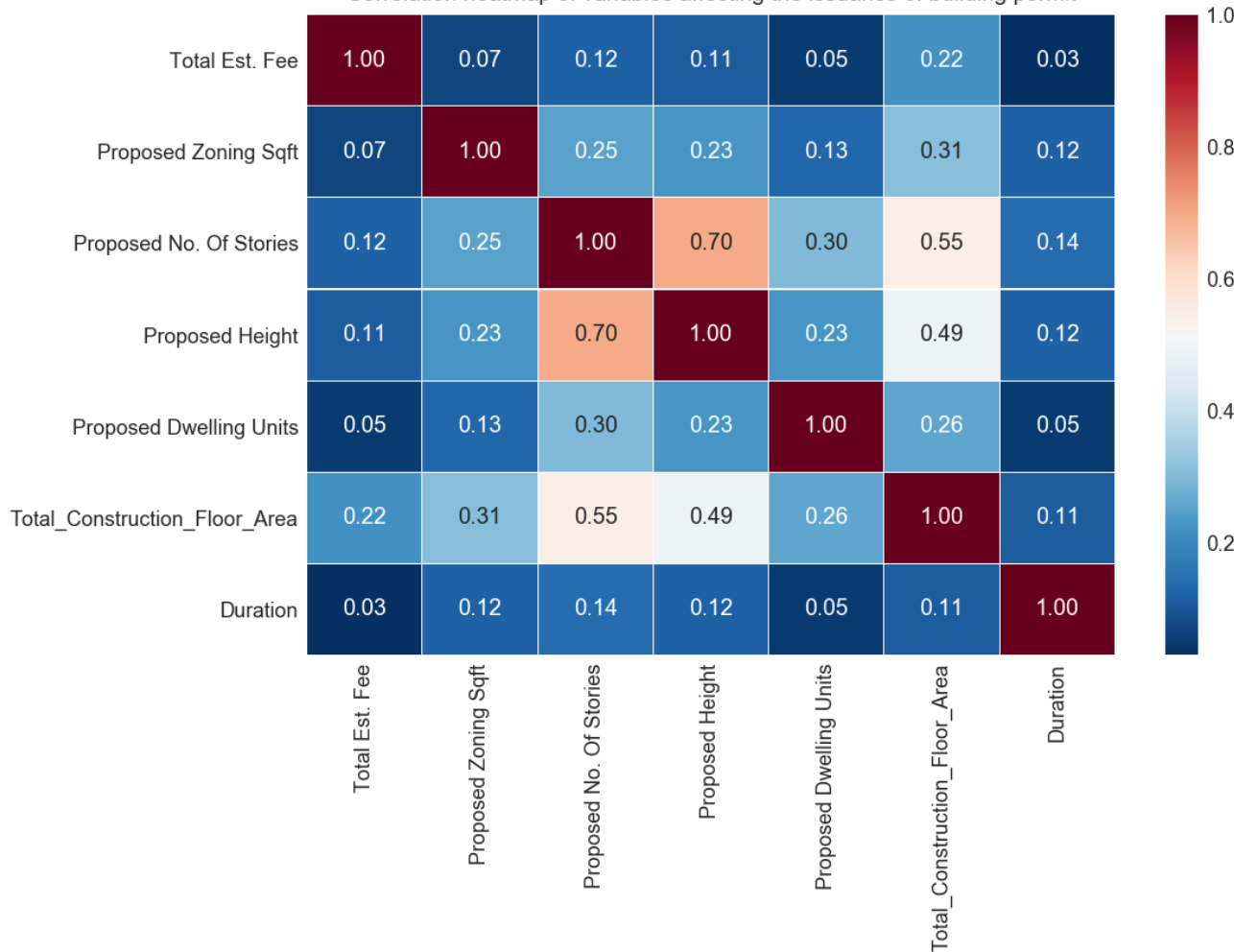


- ▶ Queens has issued the most building permits
- ▶ Manhattan issued the least

Data is from January, 2000 to May, 2019

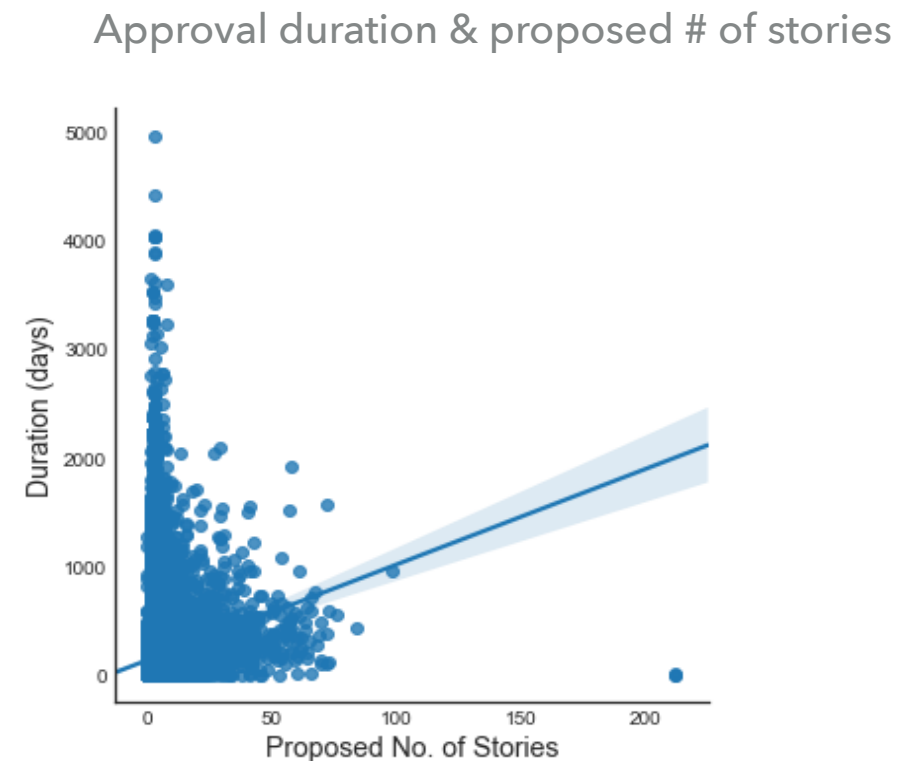
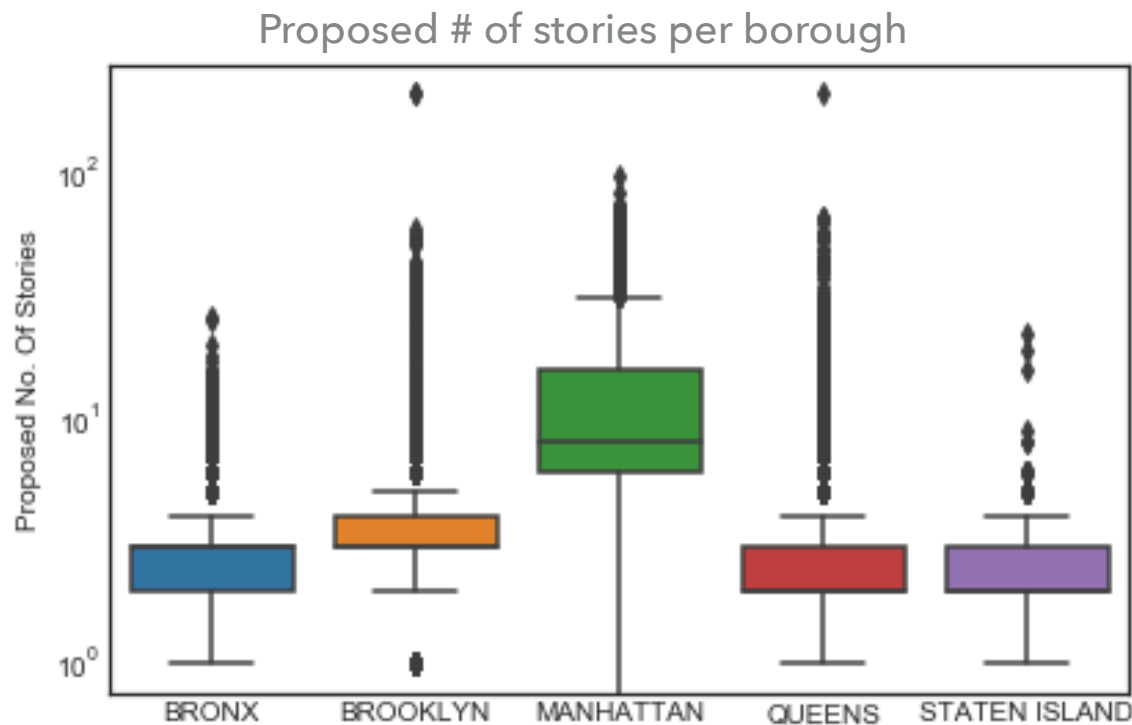
EXPLANATORY VARIABLES AND THEIR CORRELATIONS

Correlation heatmap of variables affecting the issuance of building permit



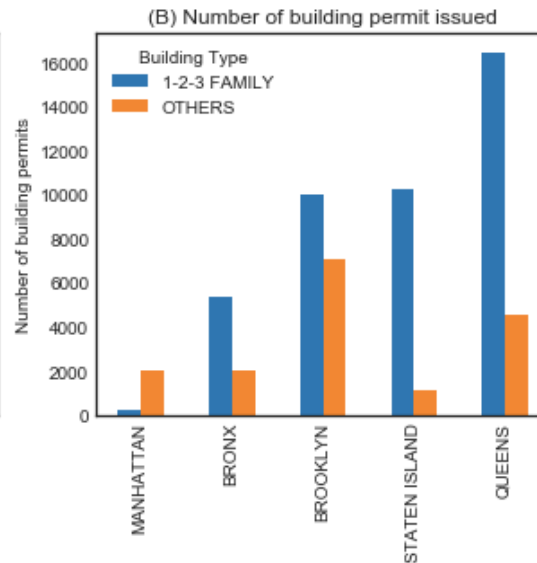
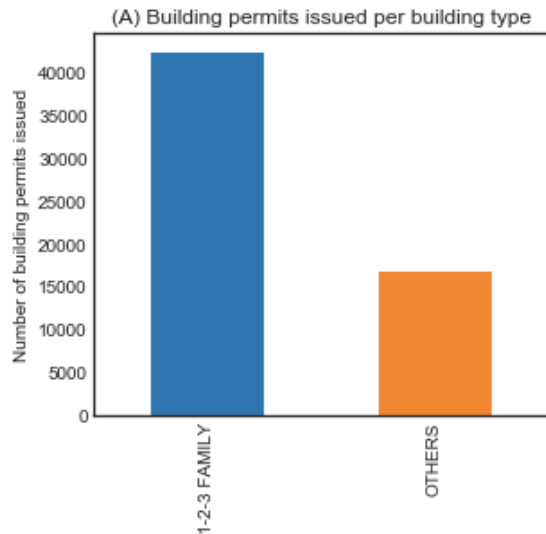
- ▶ No. of stories is strongly correlated with height and total construction area.
- ▶ Those 2 variables were excluded from the regression analysis to avoid bias

MOST SIGNIFICANT EXPLANATORY VARIABLE IS THE NUMBER OF STORIES



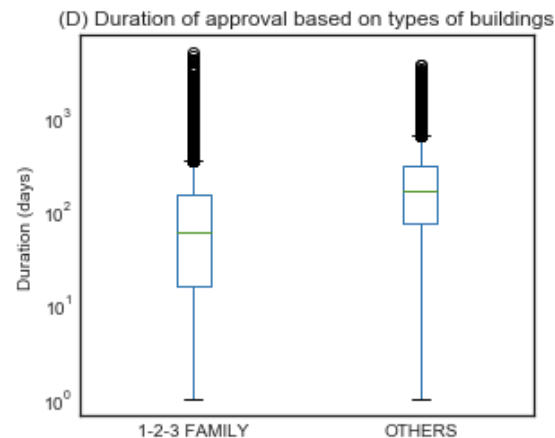
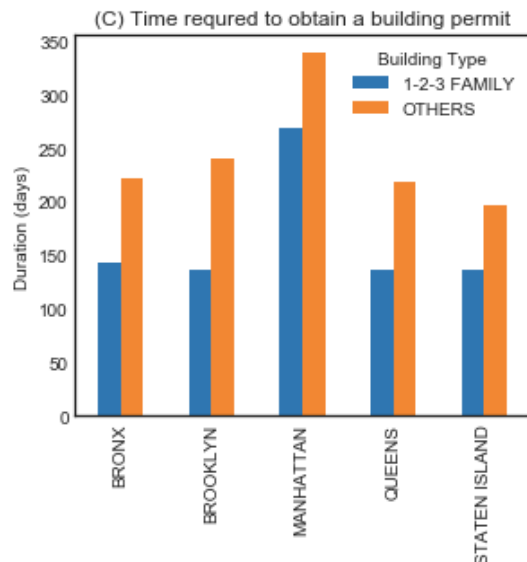
- ▶ The approval process is delayed on average by 9 DAYS per 1 more floor added to the project!
 - ▶ According to the linear regression model

MOST BUILDING PERMITS WHERE ISSUED FOR 1,2,3 FAMILY HOUSES

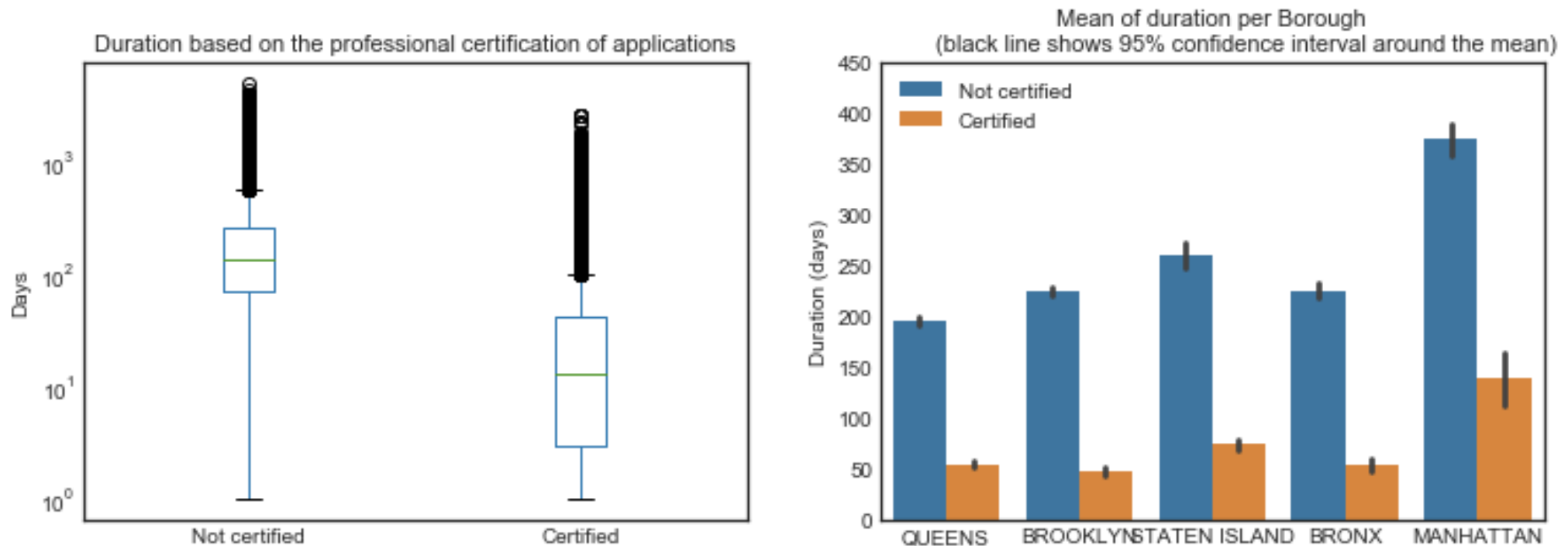


► Manhattan is the only exception where most buildings are of "other" category

► 1,2,3 family houses are approved significantly faster than "other" category buildings

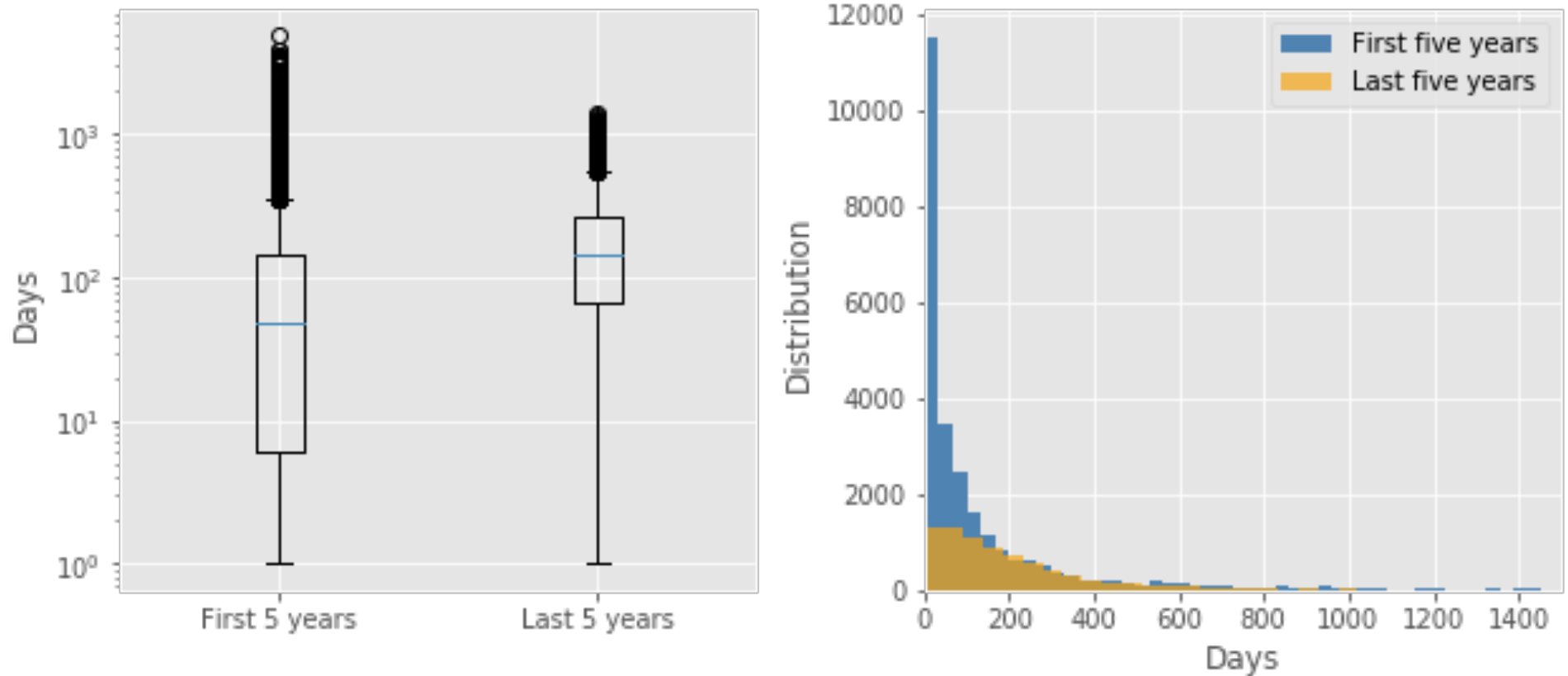


PROFESSIONALLY CERTIFIED APPLICATIONS ARE APPROVED SIGNIFICANTLY FASTER



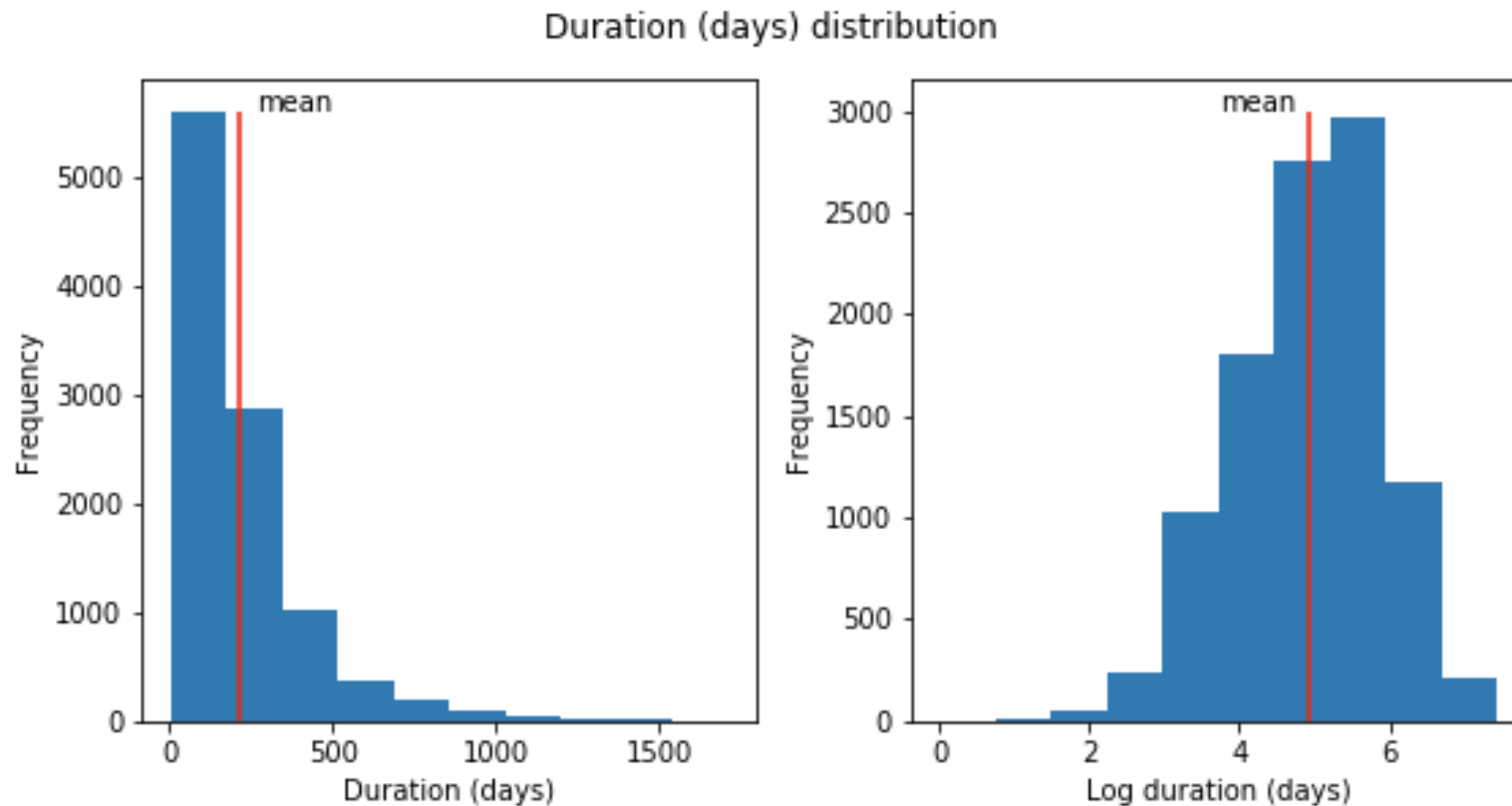
- ▶ Professionally certified application materials are less prone to errors and reduce the back and forth between the issuing agency and the applicant
- ▶ Certified applications are approved on average 62 days vs. 225 days for non-certified applications

THE DURATION WAS SHORTER BUT MORE WIDELY DISTRIBUTED FROM 2000–2005 VS 2015–2019



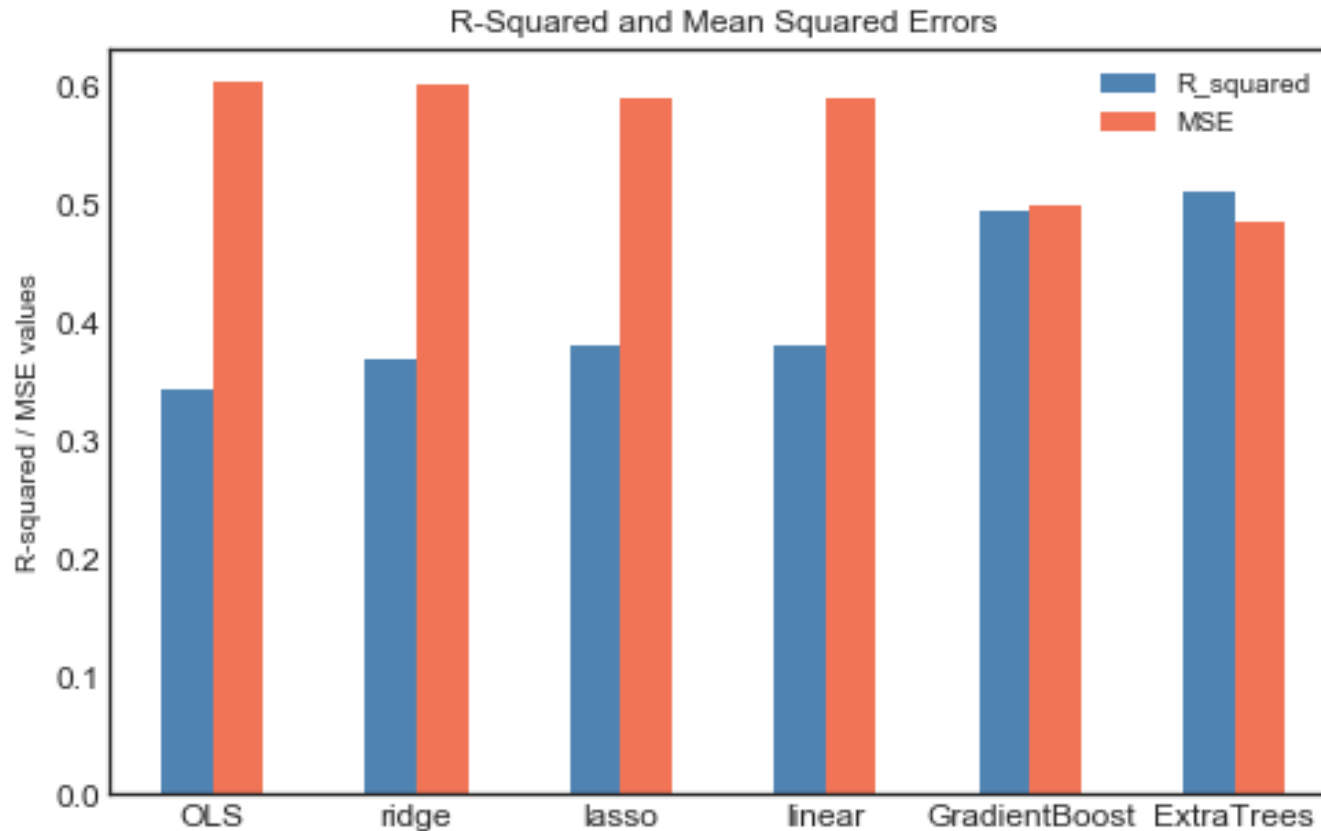
- ▶ The data is NOT normally distributed

DURATION IS SKEWED TO THE RIGHT, LOG DURATION IS MORE NORMALLY DISTRIBUTED



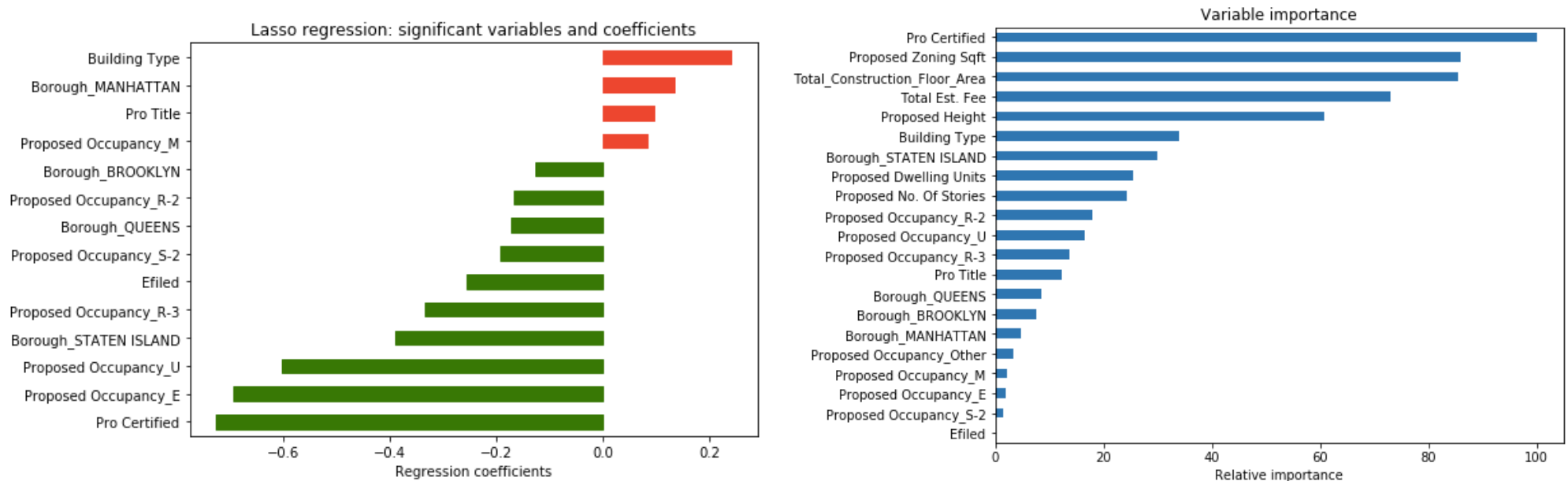
- ▶ We use the log of the duration for the regression and machine-learning analysis

MODELS USED



- ▶ Overall, Gradient Boosting and ExtraTrees models perform the best (low errors, high R-Squared)
- ▶ From linear models, Lasso performs slightly better than the rest

VARIABLE IMPORTANCE: LASSO REGRESSION VS EXTRATREES MODEL



- ▶ Lasso regression and the ExtraTrees models both show the importance of Professionally Certified applications in defining the overall duration of the process