Predicting Pet Adoption

Lusine Kamikyan

Motivation:

- **6.5 million** dogs/cats end up in shelters each year in US
- **1.5 million** of them get **euthanized** each year in US
- **300 million** stray dogs and cats worldwide

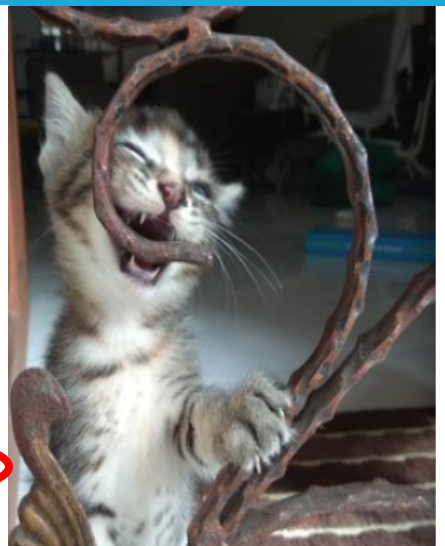
Goal:

- Predict the number of days for pets to be adopted

Data Scraping and Cleaning

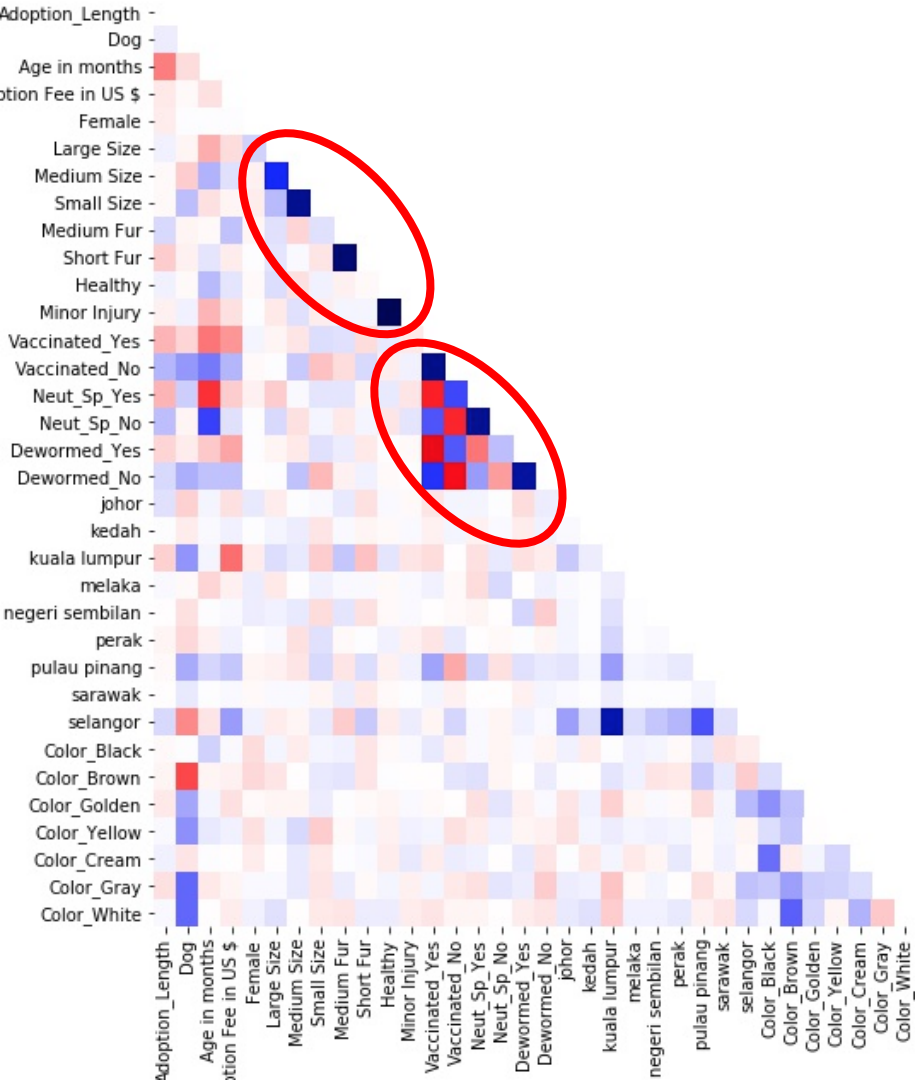
- Source: Petfinder.my
- ~ 1100 pets

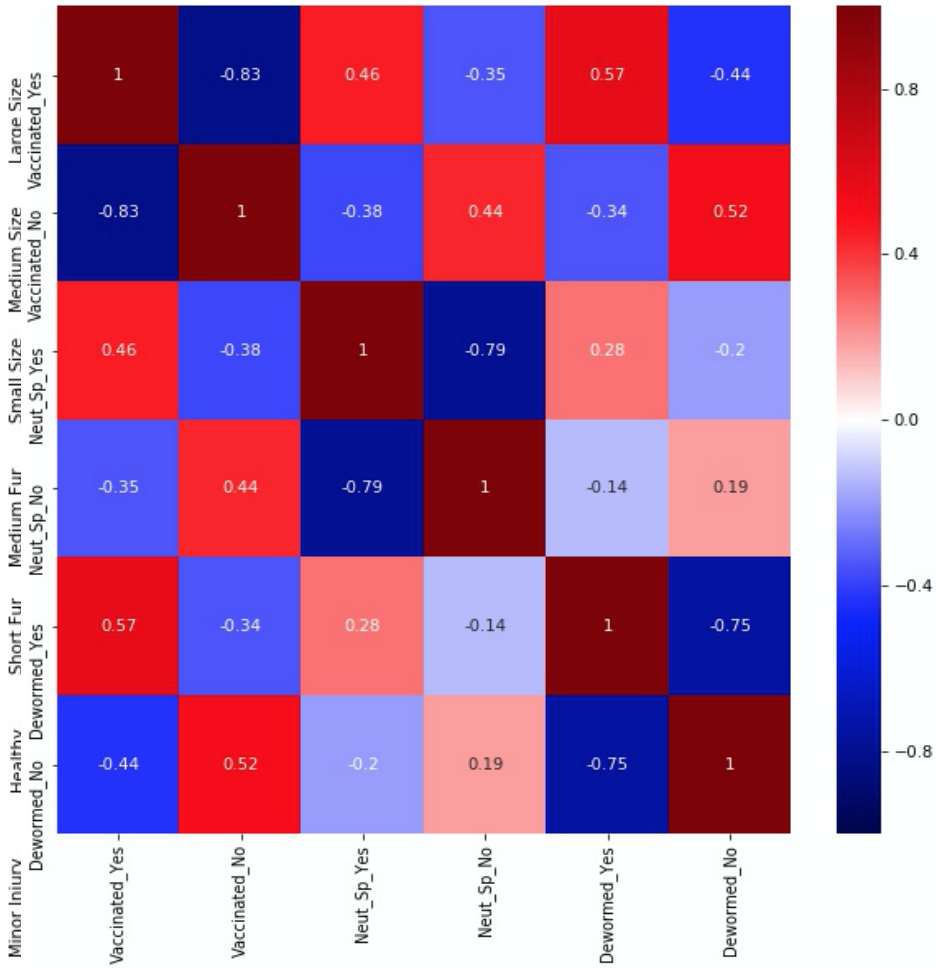
Dog	Labrador Retriever
Profile	Female, 10 Years 9 Months
Vaccinated	Yes
Dewormed	Yes
Spayed	Yes
Condition	Healthy
Body	Large Size, Short Fur
Color	Black
Location	Klang, Selangor
Posted	18 Oct 2016 (Updated 27 Nov 2019)
Adoption Fee	RM 50

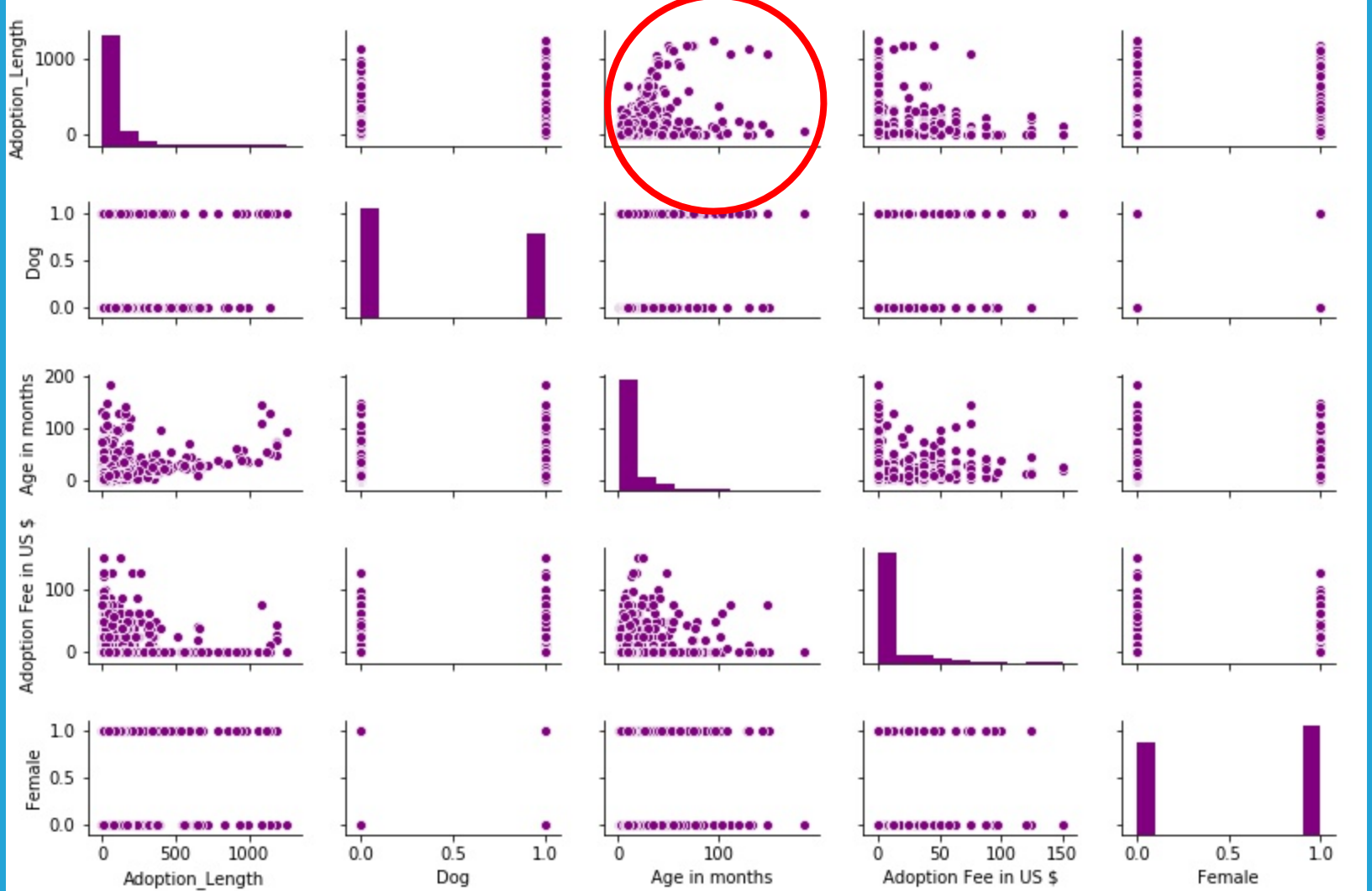


Cat	Domestic Short Hair + Tabby
Profile	Female, 6 Months
Vaccinated	No
Dewormed	Yes
Spayed	No !
Condition	Healthy
Body	Medium Size, Short Fur
Color	Gray
Location	Kuala Lumpur, Wilayah Persekutuan
Posted	3 Aug 2019 (Updated 16 Jan 2020)
Adoption Fee	RM 50

Correlation Heatmap






Adoption_Length
Dog
Age in months
Adoption Fee in US $
Female
0
500
1000
0.0
0.5
1.0
0
100
0
50
100
150
0.0
0.5
1.0

Baseline Model Results

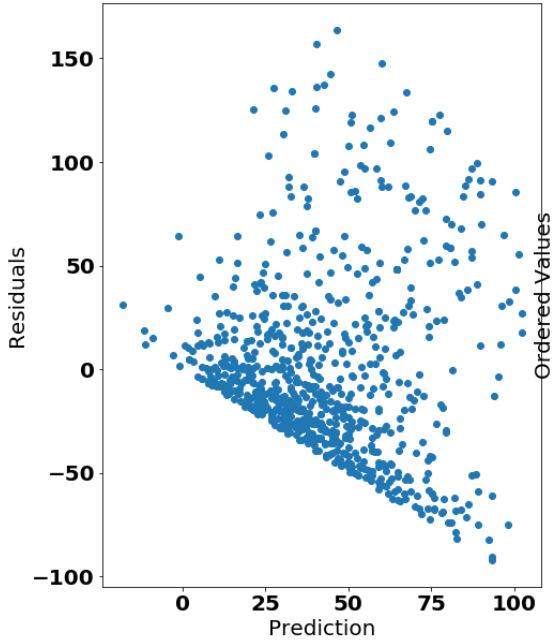
Linear Regression	Linear Regression with Feature Engineering
$R^2 = 0.221$ (Adj $R^2 = 0.185$)	$R^2 = 0.553$ (Adj $R^2 = 0.545$)
RMSE = 42.34	RMSE = 42.7

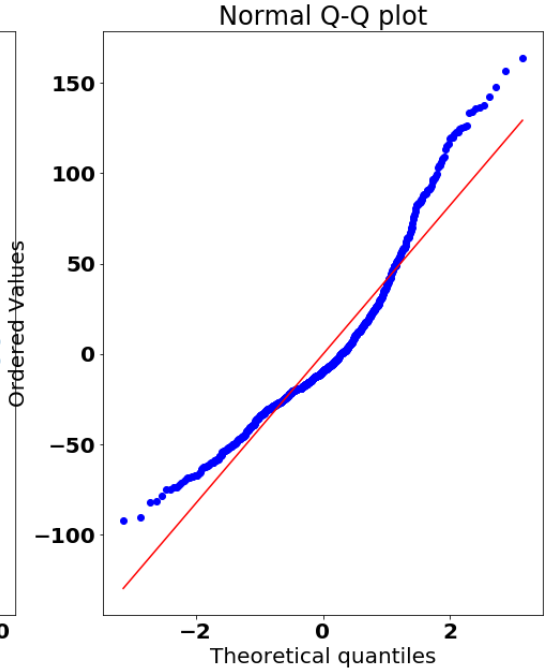
More Models with KFold Cross Validation

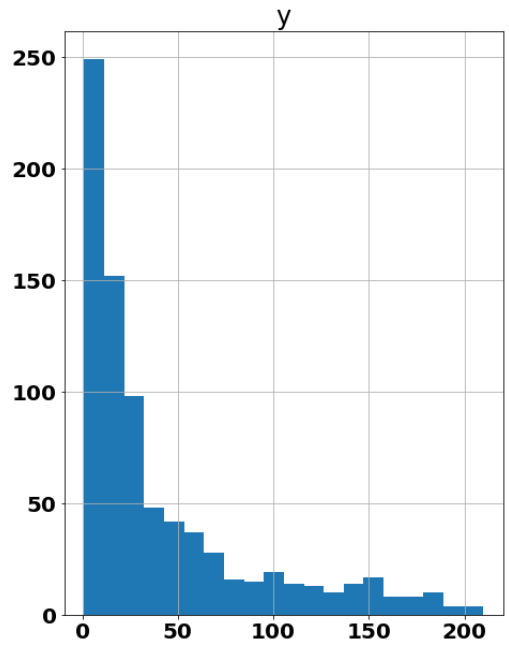
Linear Regression	Ridge Regression	Lasso Regression
$R^2 = 0.173$	$R^2 = 0.205$	$R^2 = 0.207$
RMSE = 43.3	RMSE = 42.7	RMSE = 42.7
		
	$R^2 = 0.165$	
	RMSE = 45.8	

What's wrong?

- Data is very skewed
- Residual plot doesn't look right







Conclusion

- None of the linear models used gave us good results
- Future investigation needed to improve the model

Questions?