



Pet adoption speed prediction

By: Simona Tso

The problem



- Millions of stray animals suffer on the streets every day
- Shelters are often overcrowded and lacking resources

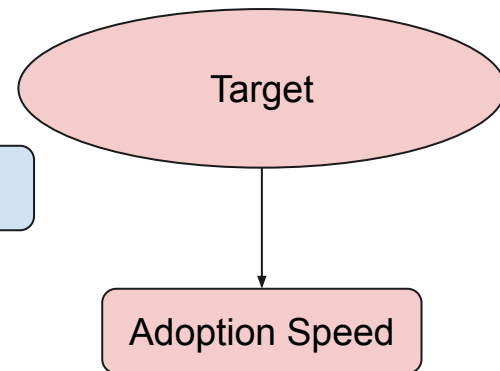
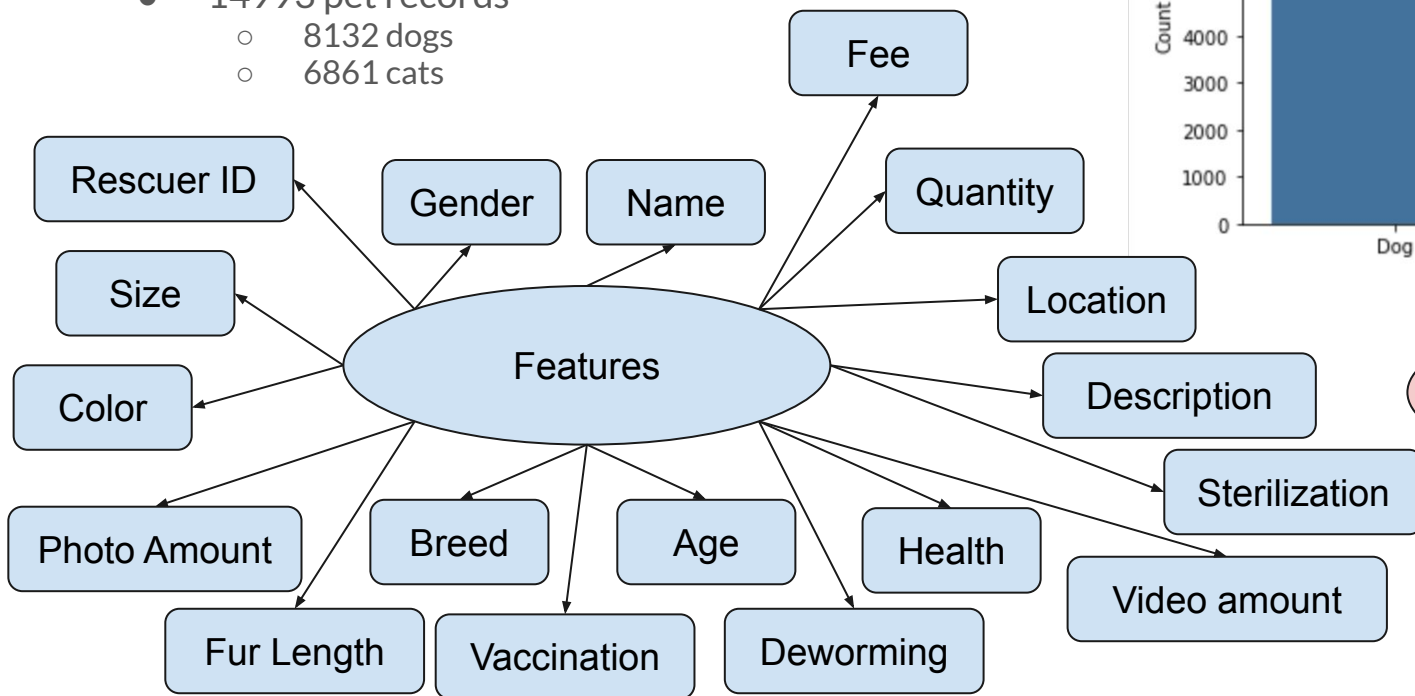
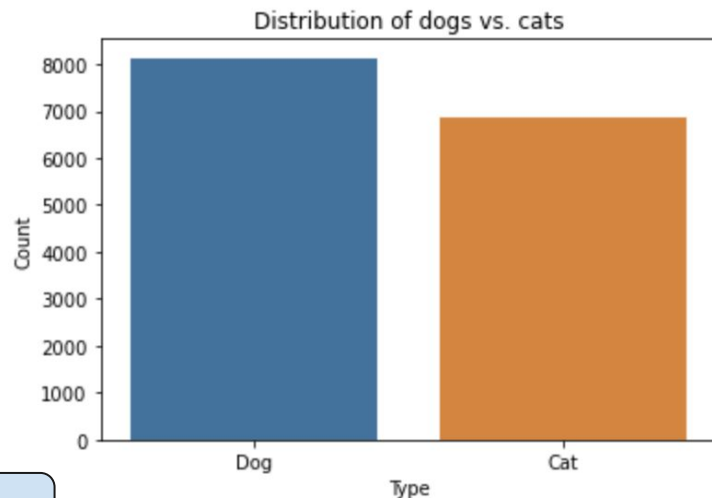
Our solution



- Create a predictive model that can accurately predict whether a pet will be adopted within 100 days
- Identify features that play in important role in pet adoption
- Adopt a data-driven method to increase adoption probability

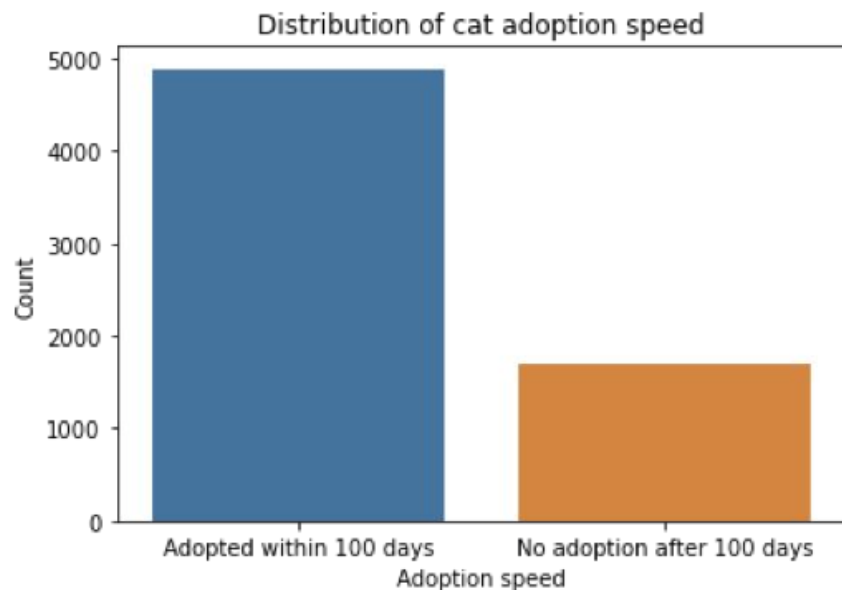
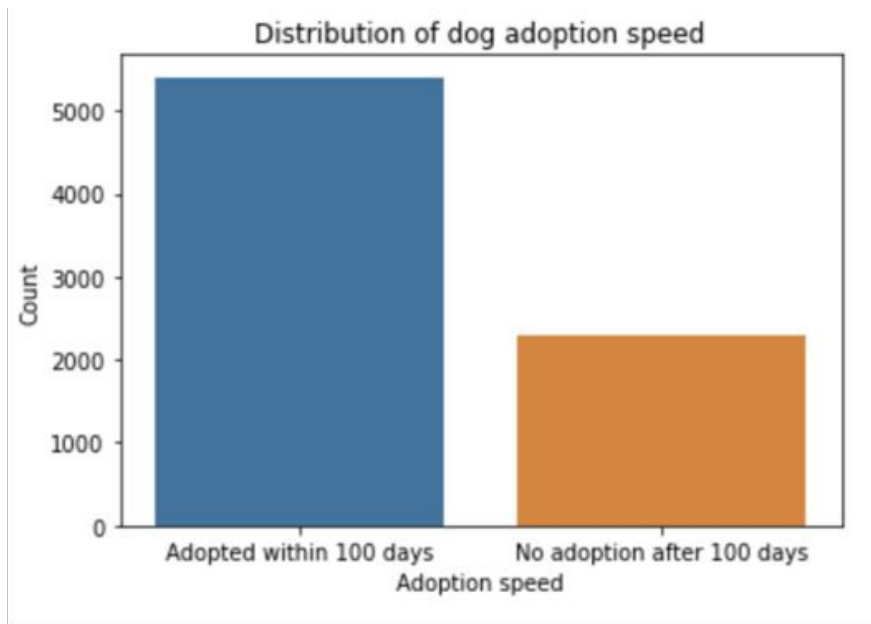
The Data

- Malaysia's leading animal welfare platform: PetFinder.my
- 14993 pet records
 - 8132 dogs
 - 6861 cats



Data Exploration - Target Variable

- 71% of dogs are adopted within 100 days
- 74% of cats are adopted within 100 days



Data Exploration - Adoption Fees

- Dog adoption fees:
 - \$0 - \$3000
- Cat adoption fees:
 - \$0 - 800
- 75% of cats and dogs have no adoption fee

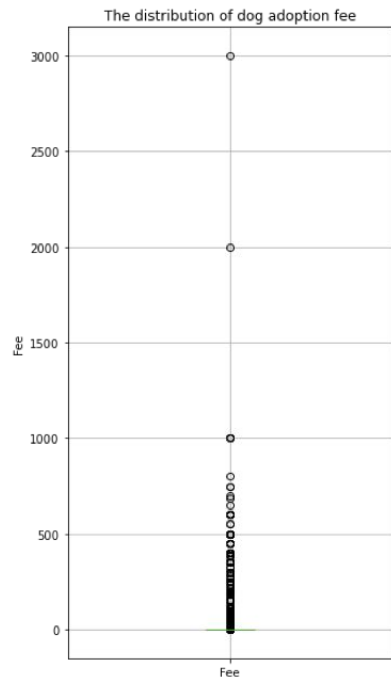


Fig 23.1 Distribution of dog adoption fee

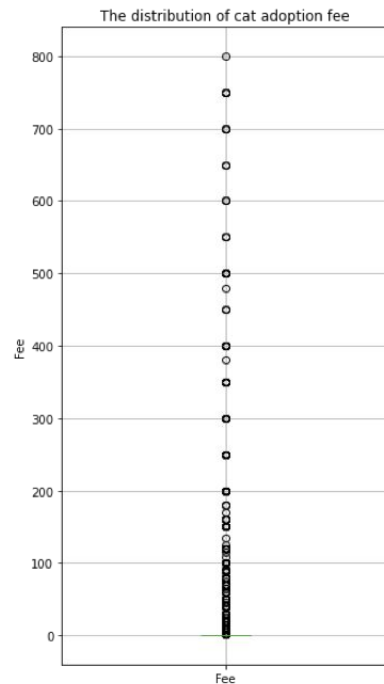
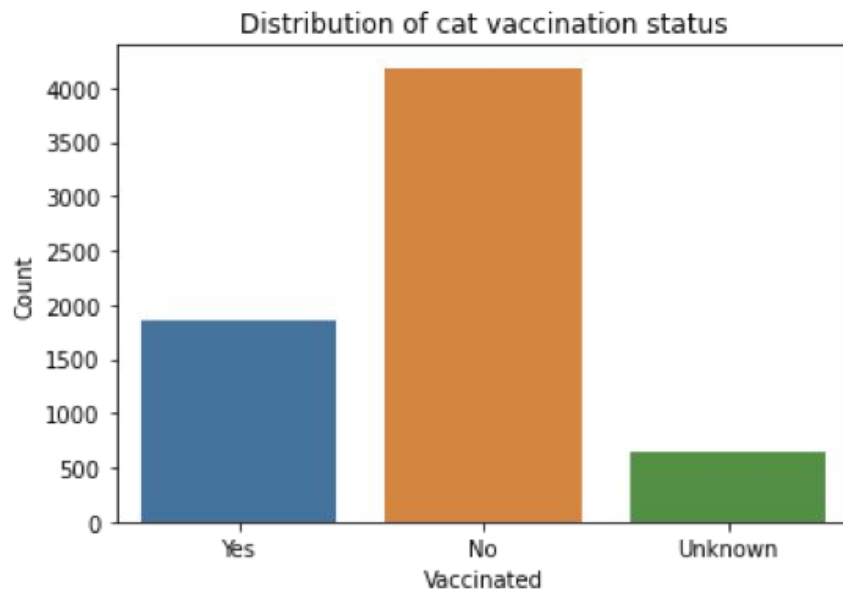
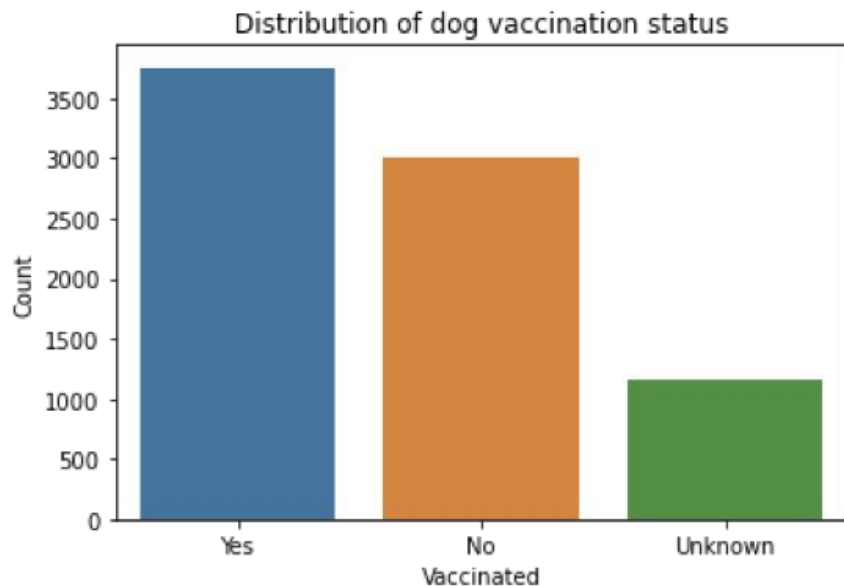


Fig 23.2 Distribution of cat adoption fee

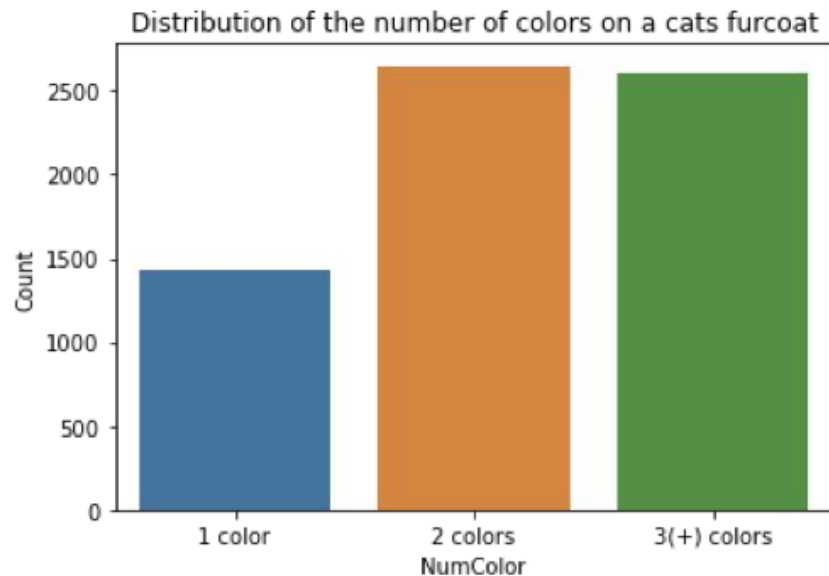
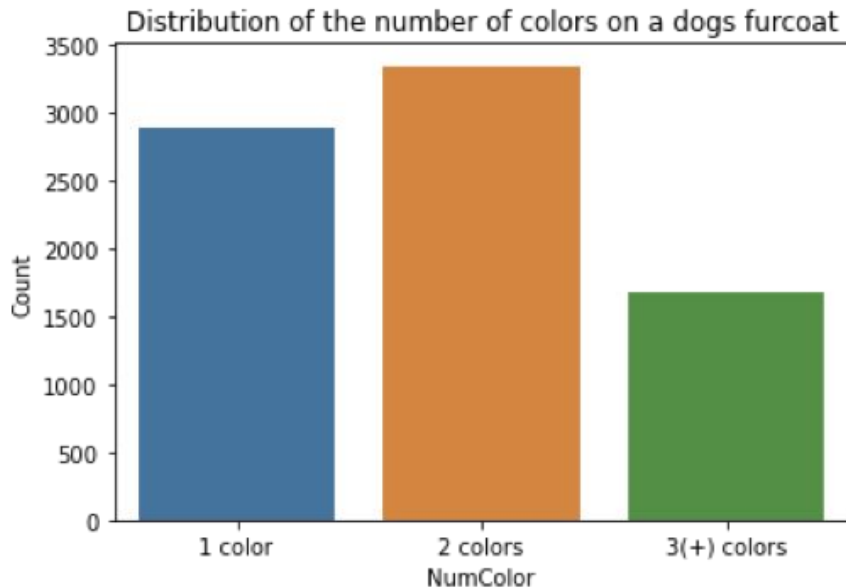
Data Exploration - Vaccination

- Most dogs are vaccinated
- Most cats are not vaccinated



Data Exploration - Number of fur colors

- Most dogs have 1 - 2 colors
- Most cats have 2 - 3+ colors



Model Comparison



Dog adoption dataset

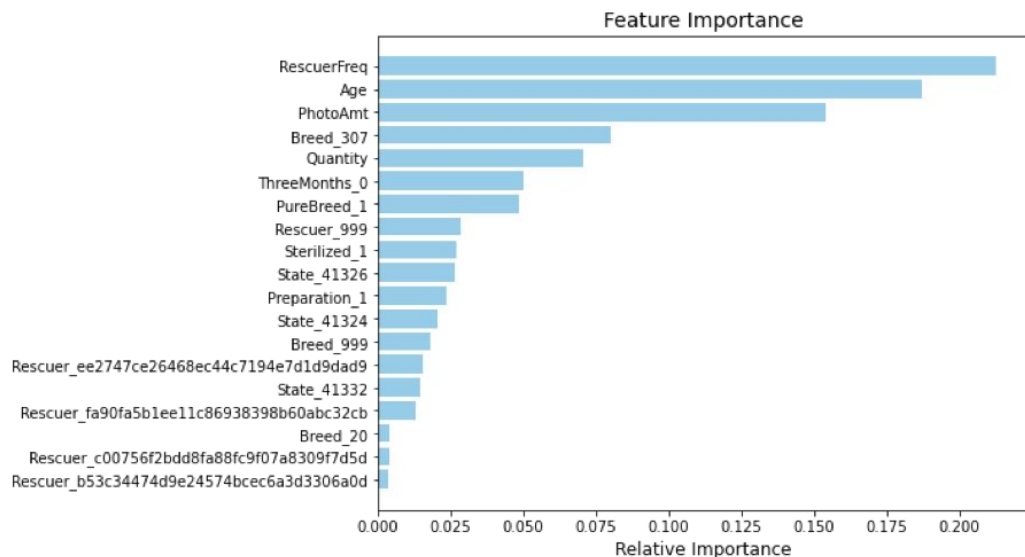
Model	F1-Score	Best Score	ROC AUC	Confusion Matrix
Logistic Regression	0.6176	0.7641	0.797	[[815 254] [146 323]]
Random Forest	0.6553	0.8039	0.8355	[[750 319] [85 384]]
XGBoost	0.5803	0.813	0.8433	[[968 10] [236 233]]

Cat adoption dataset

Model	F1-Score	Best Score	ROC AUC	Confusion Matrix
Logistic Regression	0.522	0.7438	0.7316	[[679 300] [112 225]]
Random Forest	0.5355	0.7761	0.7725	[[646 333] [92 245]]
XGBoost	0.4781	0.7687	0.7717	[[829 150] [184 153]]

For both data sets **Logistic Regression** is the worst and **Random Forest** is the **best**.

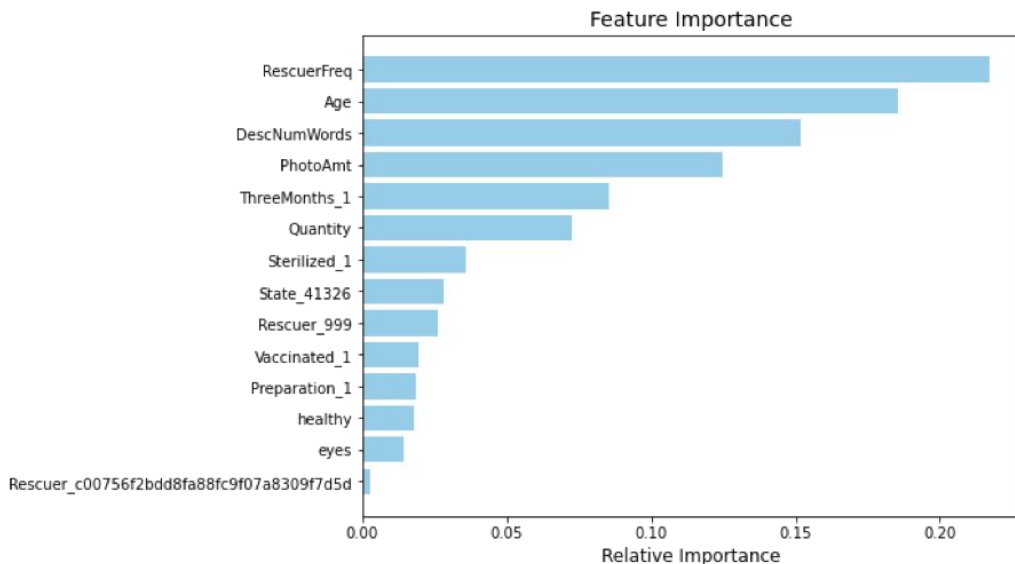
Dog adoption dataset



Most important features:

- Rescuer Frequency
- Age
 - Three Months
- Photo amount
- Pure/ mixed breed
- Quantity
- States
 - Selangor
 - Melaka
 - Negeri Sembilan
- Preparation
 - Sterilization
- Top 10 breeds
- Rescuers:
 - Top 10
 - ee2747ce26468ec44c7194e7d1d9dad9
 - fa90fa5b1ee11c86938398b60abc32cb
 - c00756f2bdd8fa88fc9f07a8309f7d5d
 - b53c34474d9e24574bcec6a3d3306a0d

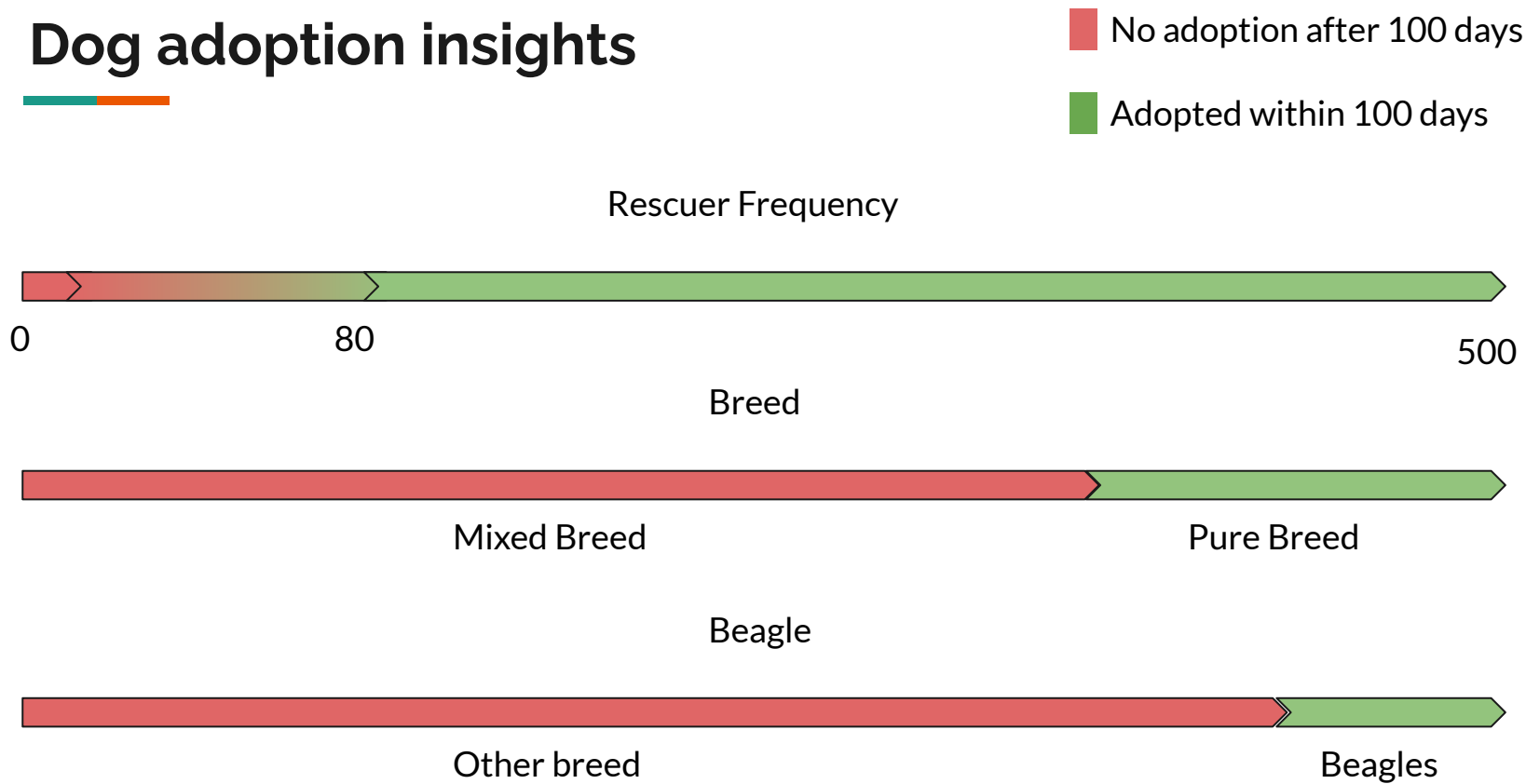
Cat adoption dataset



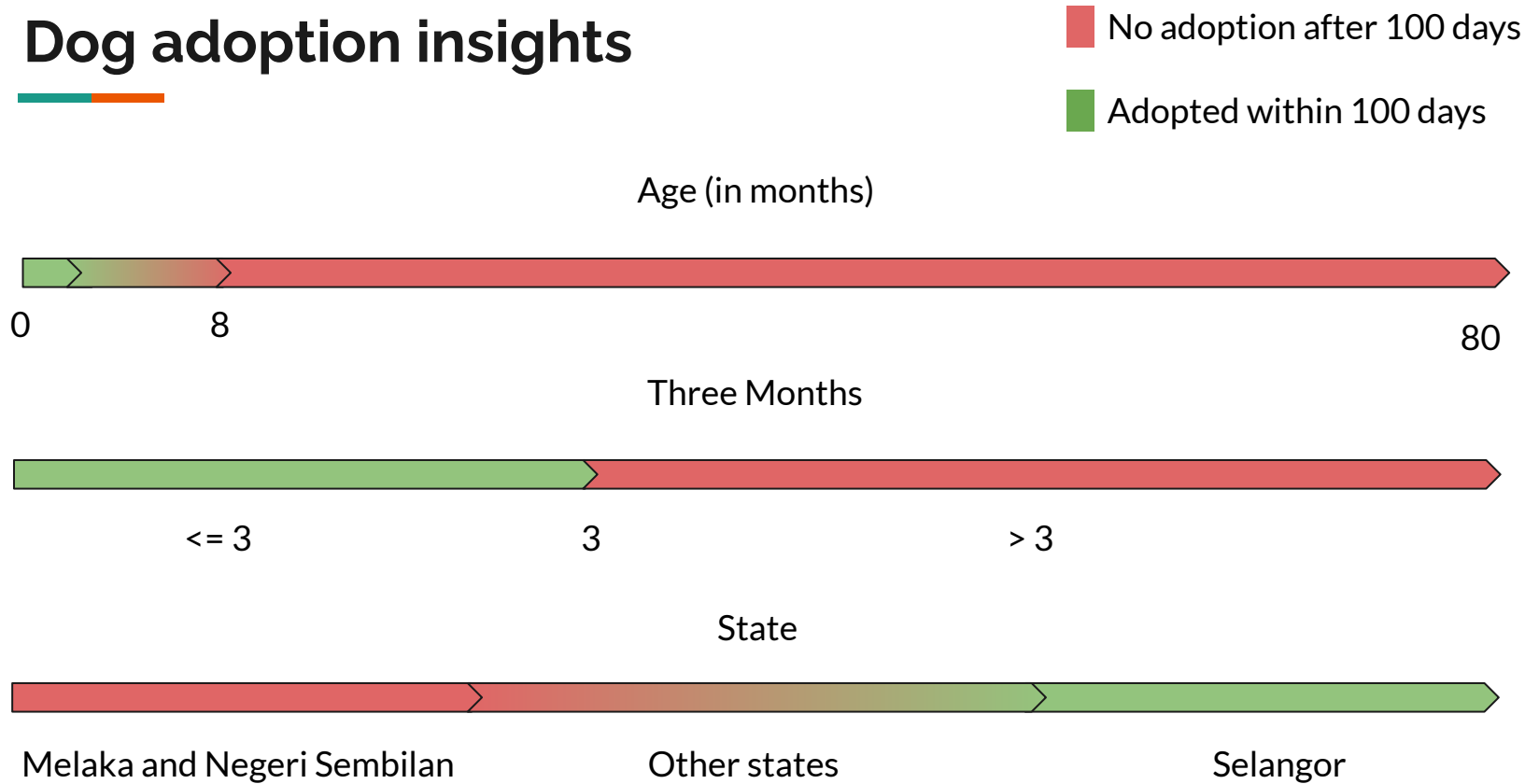
Most important features:

- Rescuer Frequency
- Age
 - Three Months
- Description
 - Length
 - Keywords:
 - 'Healthy'
 - 'Eyes'
- Photo amount
- Quantity
- Selangor state
- Preparation
 - Vaccination
 - Sterilization
- Rescuers:
 - Top 10 rescuers
 - c00756f2bdd8fa88fc9f07a8309f7d5d

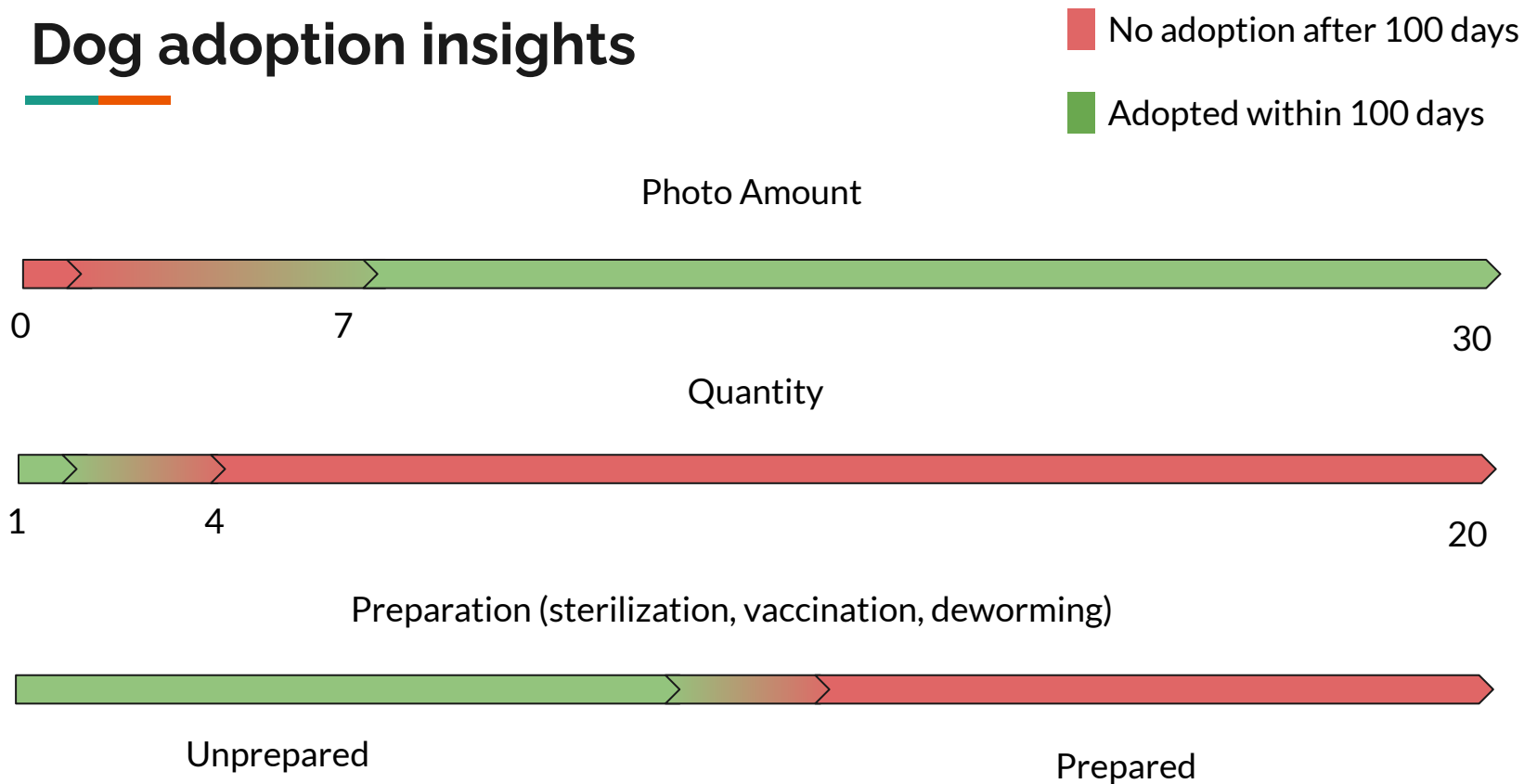
Dog adoption insights



Dog adoption insights



Dog adoption insights



Cat adoption insights

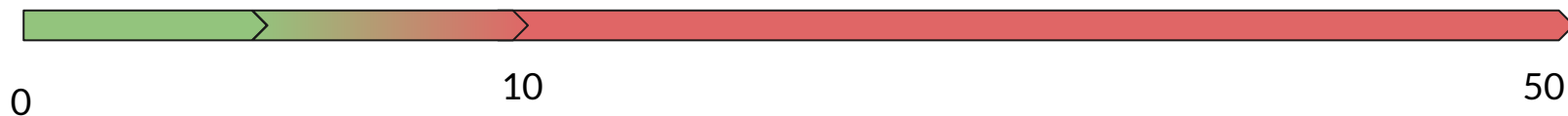
■ No adoption after 100 days

■ Adopted within 100 days

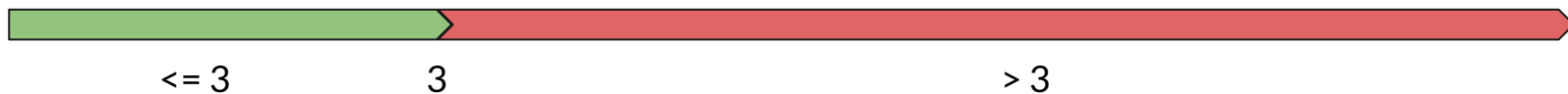
Rescuer Frequency



Age



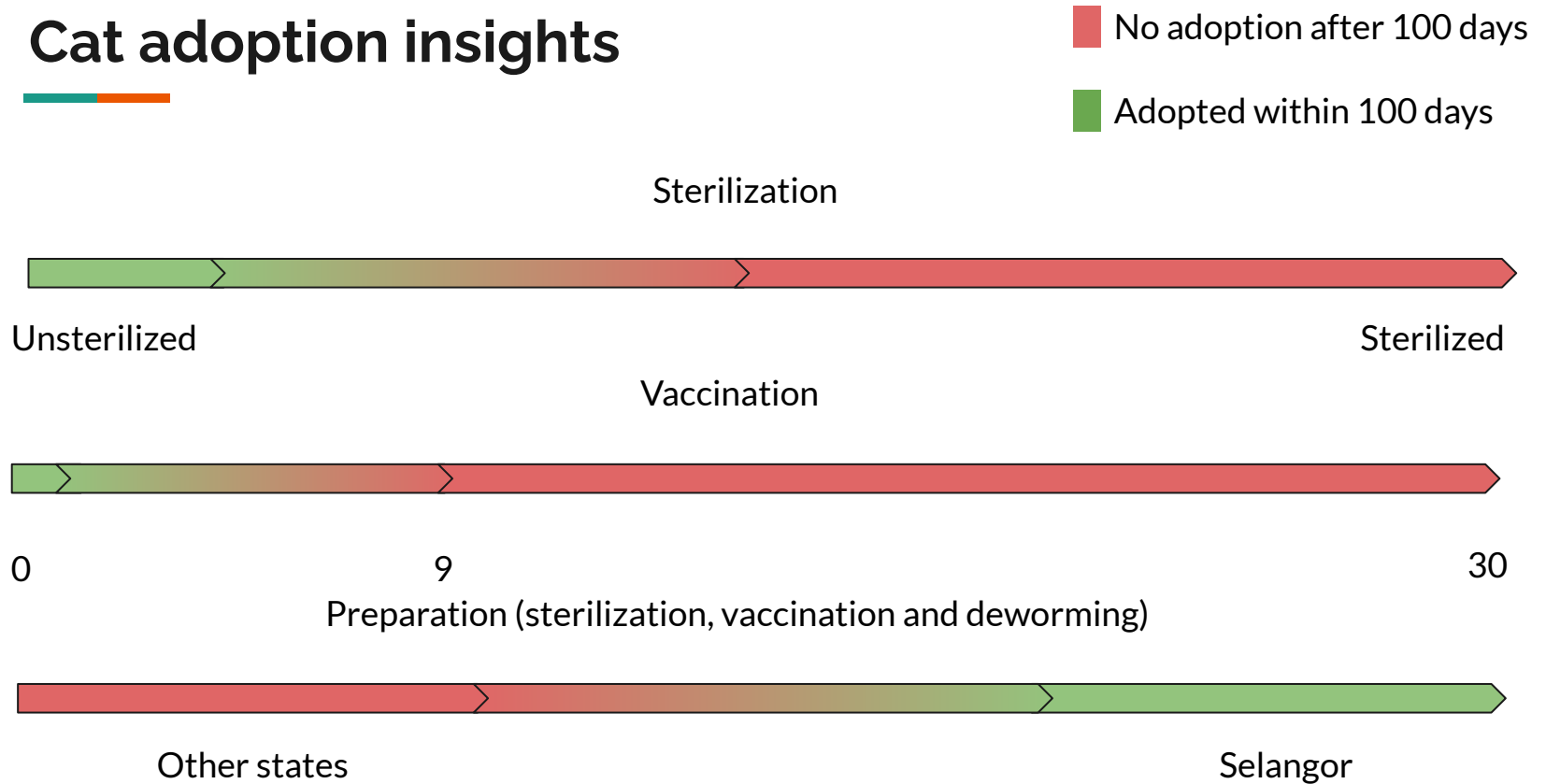
Three Months



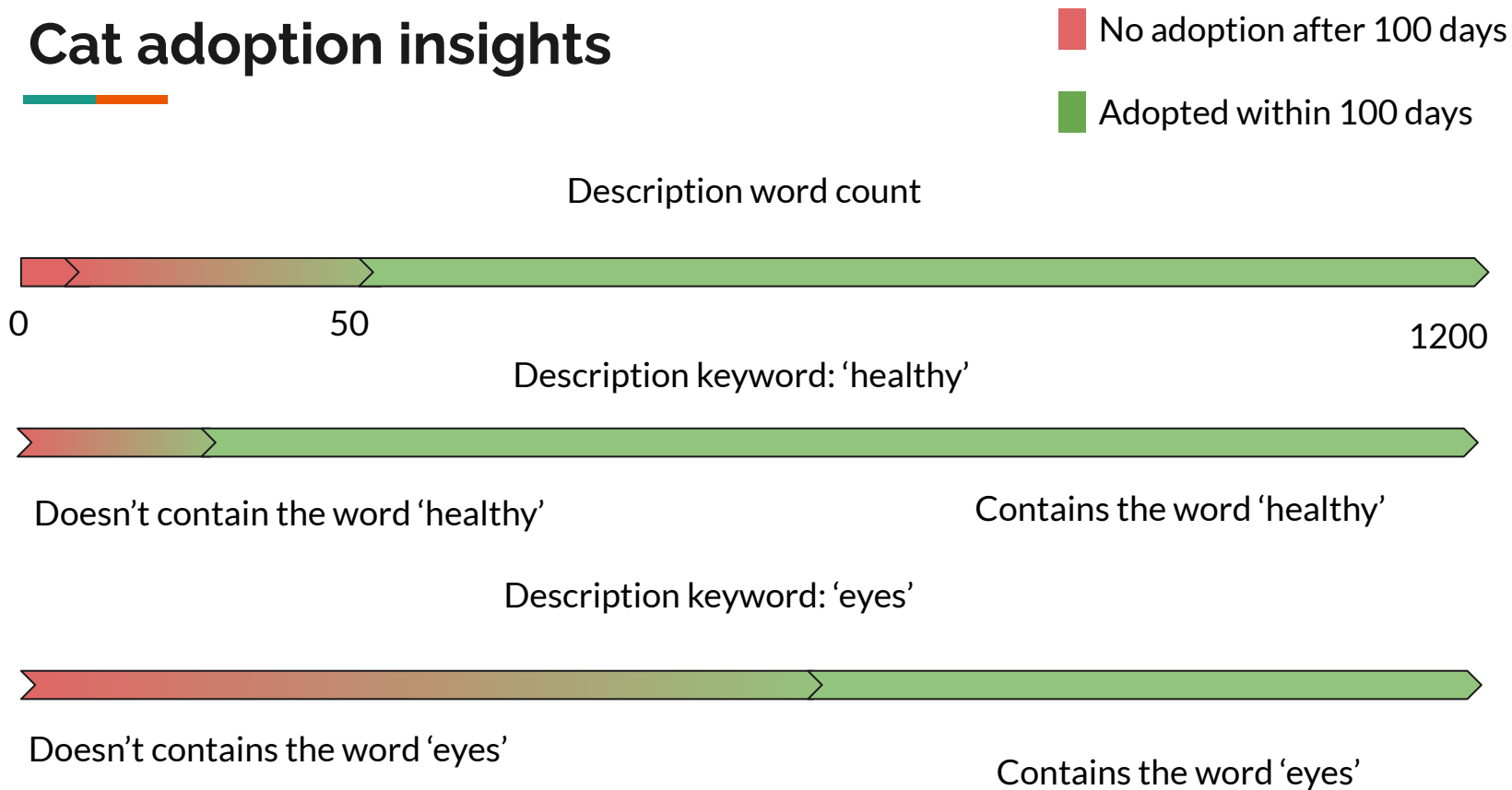
Cat adoption insights



Cat adoption insights



Cat adoption insights



Conclusion - Dog Adoption



- Best model: Random Forest (ROC AUC: 0.8355)
- Important features that increase adoption probability:
 - High rescuer frequency
 - Purebred dogs
 - Young (3 months or younger)
 - Large number of photos (7 - 10)
 - Small quantity
 - Unsterilized
 - Located in Selangor
 - Unprepared (sterilization, vaccination, deworming)
 - Beagles

Conclusion - Cat Adoption



- Best model: Random Forest (ROC AUC: 0.7725)
- Important features that increase adoption probability:
 - High rescuer frequency
 - Young (3 months or younger)
 - Long description (200-400 words)
 - Large number of photos (7 - 10)
 - Small quantity
 - Unsterilized
 - Located in Selangor
 - Vaccinated
 - Including keywords 'eyes' and 'healthy'



Recommendations - for dogs

- Upload 7- 13 photos
- Split up litters (if necessary) to keep quantity at 1
- Study strategies used by rescues in Selangor to understand their success and replicate strategies (if possible)
- Transfer older dogs to Selangor
- Run stray dog awareness programs in Melaka and Negeri Sembilan
- Transfer dogs that are individually rescued to larger rescuers



Recommendations - for cats

- Include a description at least 50 words long
- Include keywords 'eyes' and 'healthy' for high frequency rescuers
- Upload at least 9 - 12 photos
- Split up litters (if necessary) to keep quantity at 1
- Sterilize older cats
- Study strategies used by rescues in Selangor to understand their success and replicate strategies (if possible)
- Transfer cats to Selangor
- Transfer individually rescued cats to larger rescuers



Future Scope

- Use image recognition to include pet photos
- Sentiment analysis on pet profile description
- Additional datasets:
 - GDP
 - Population
 - Pet adoption data sets from other countries