Predicting Pet Adoption Speeds

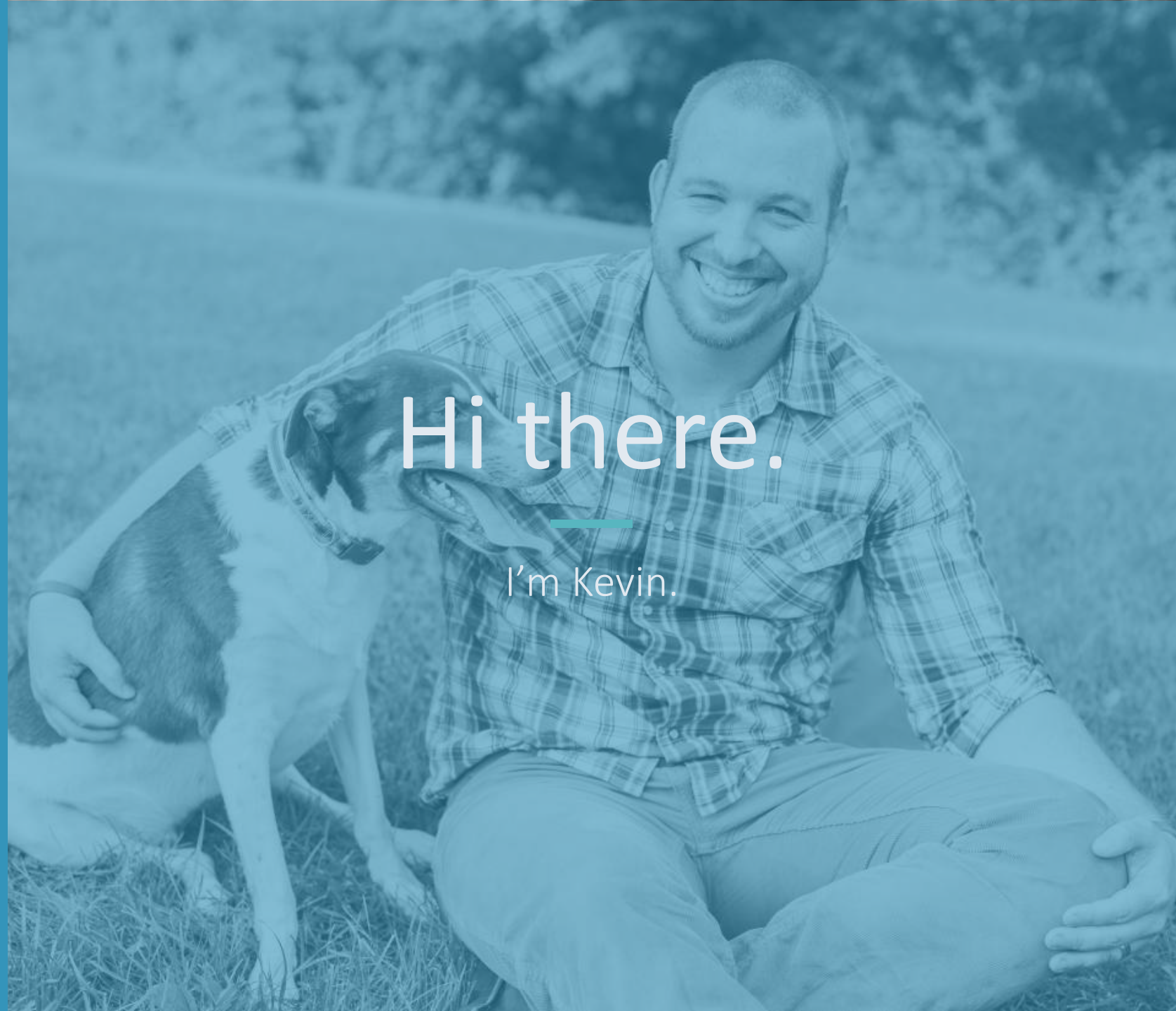
How Cute is that Doggy in the Window

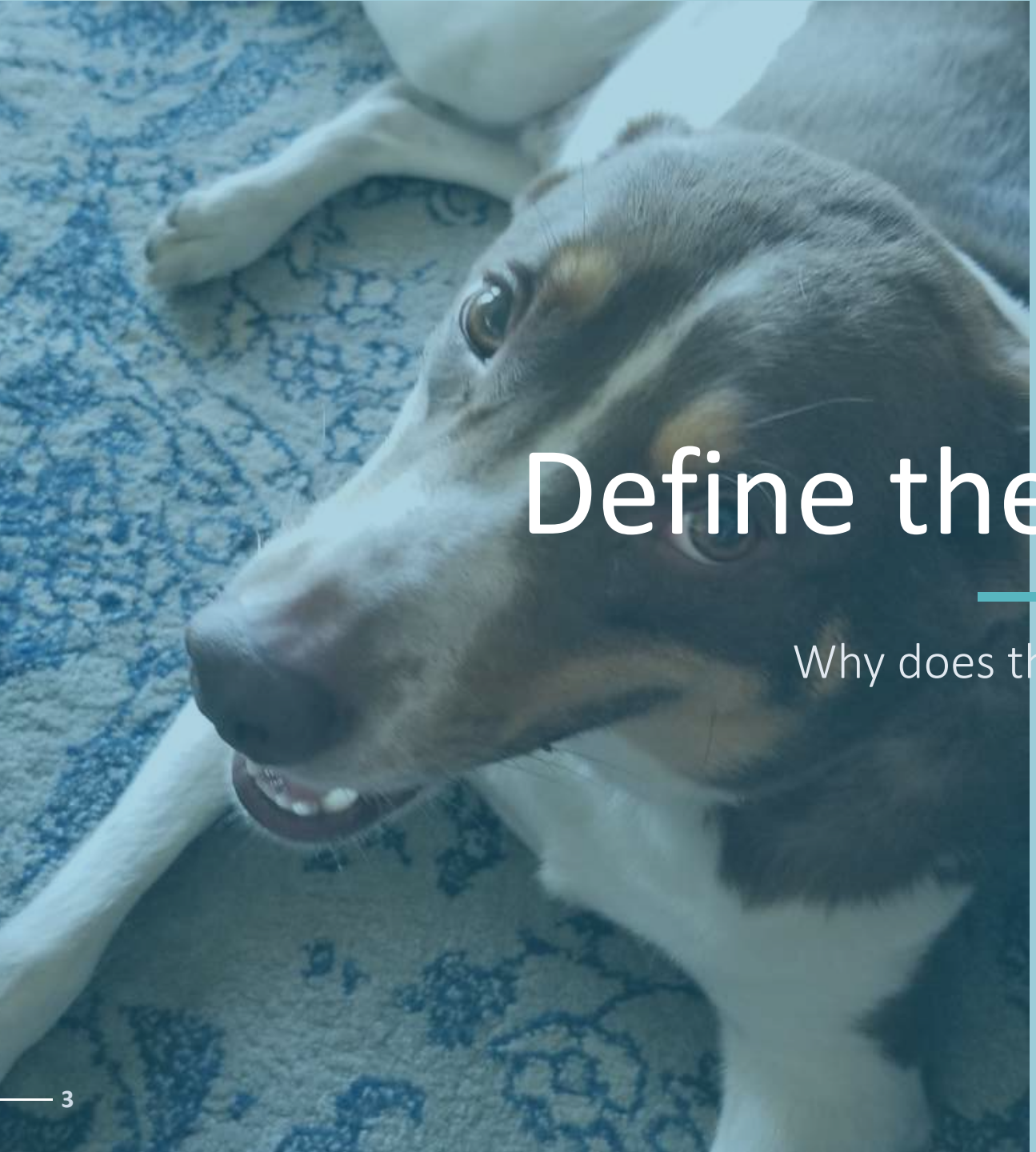
Kevin Roesch
Data Scientist | General Assembly
kevincroesch@gmail.com

GENERAL ASSEMBLY

Agenda

- Define the problem
- Obtain the data
- Explore the data
- Model the data
- Evaluate the model
- Respond to the problem





Define the Problem

Why does this matter?

GENERAL ASSEMBLY

Define the Problem

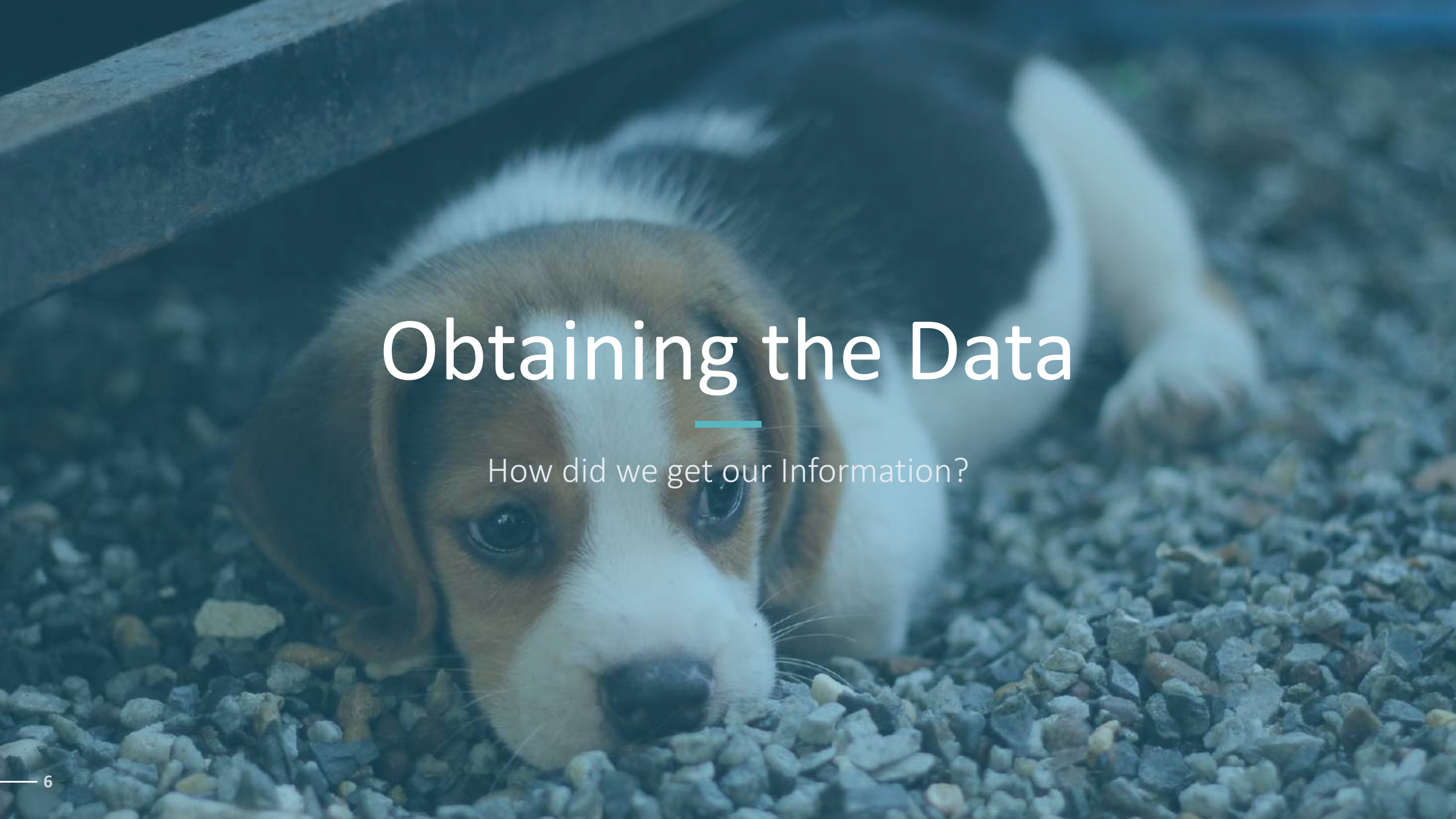
Create a model predicting the adoptability of pets – specifically, how quickly is a pet adopted

- Improve pet profiles' appeal
- Reduce animal suffering and euthanization
- Build more happy families

GENERAL ASSEMBLY

Define the Problem

“Based on an animals profile features, are we able to predict it’s adoptability?”



Obtaining the Data

How did we get our Information?

GENERAL ASSEMBLY

PetFinder.my Adoption Prediction

GENERAL ASSEMBLY

Obtain the Data

Train

Test

Breed

Color

Obtain the Data

Train

TRAIN is the data used to build the model. This data includes actual adoption speeds of the animals from the shelter

Test

TEST is the data used to validate the model. This data is the structured the same as the Train but without the adoption speeds we are attempting to predict.

Obtain the Data

Breed

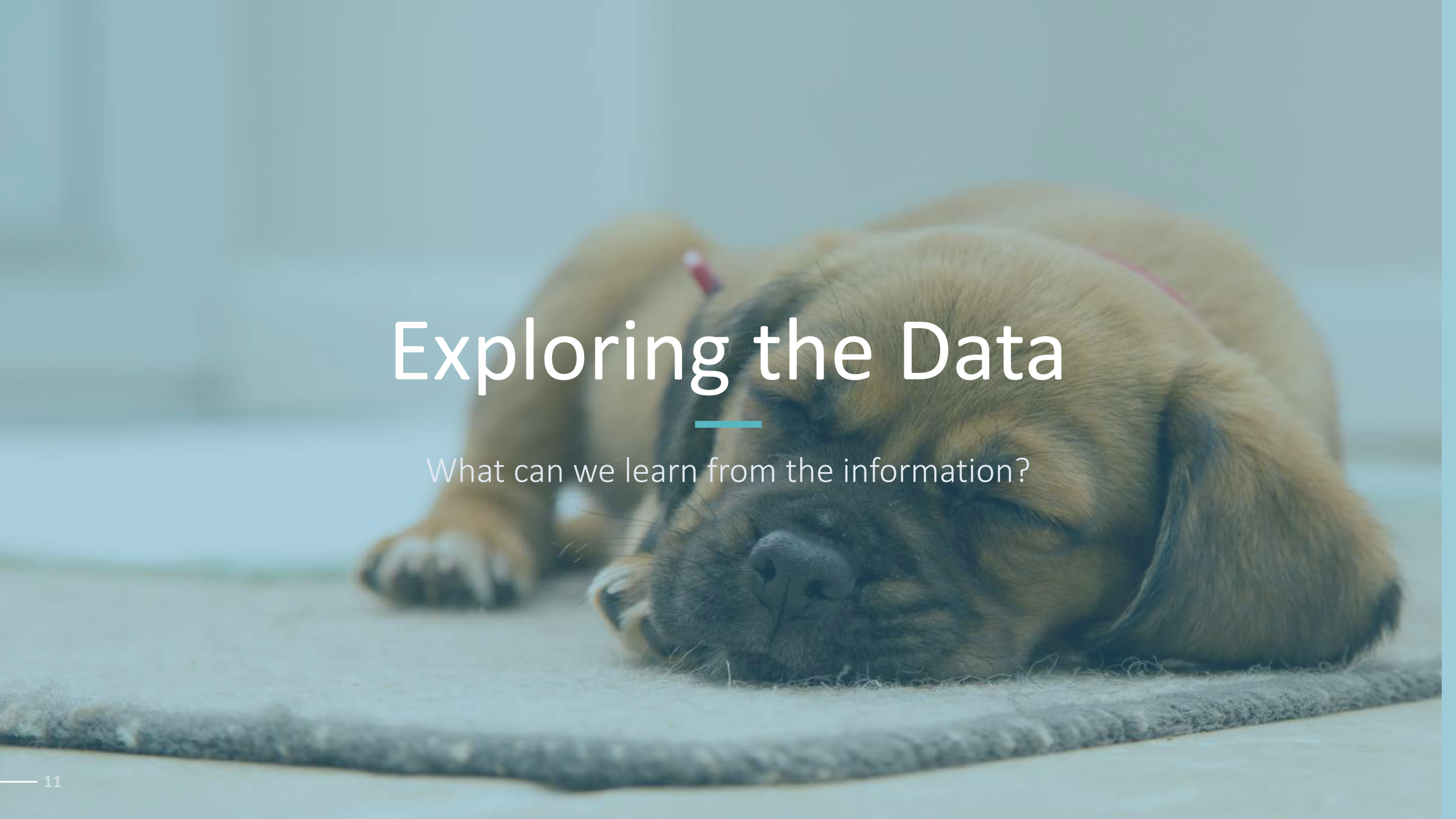
Breed is the data map used to identify the breed in the train/test data sets.

Contains Type, and Breed name for each Breed ID. Type 1 is dog, 2 is cat

Color

Color is the data map used to identify the color name in the train/test data sets.

Contains ColorName for each ColorID



Exploring the Data

What can we learn from the information?

Explore the Data

Train

The training data contained 14,993 total pets and 25 features for each pet

Test

The testing data contained 3,948 total pets and 24 features for each pet

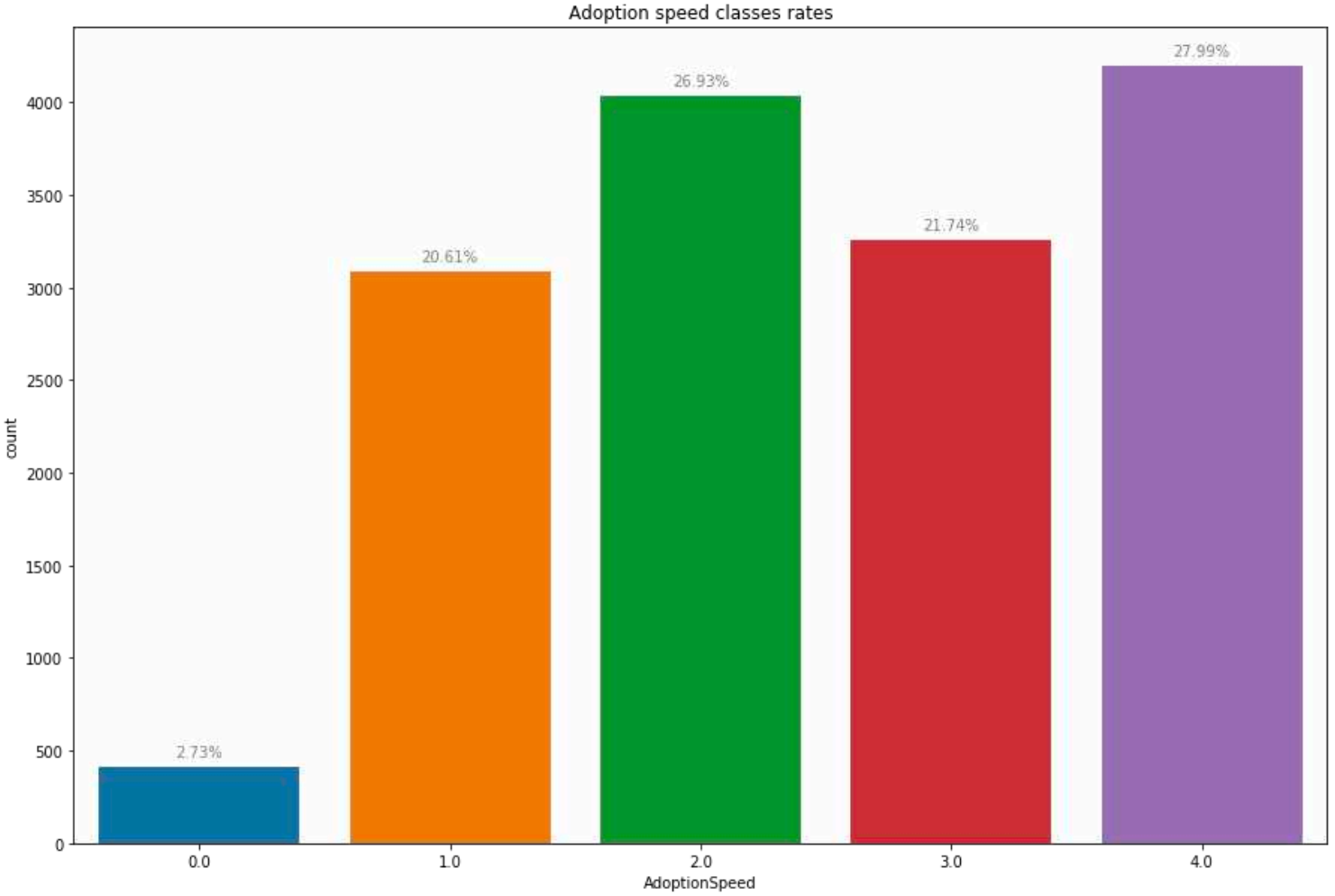
Breed

The breed data contained 307 total breed descriptions for the pets

Color

The color data contained seven color options for each pet

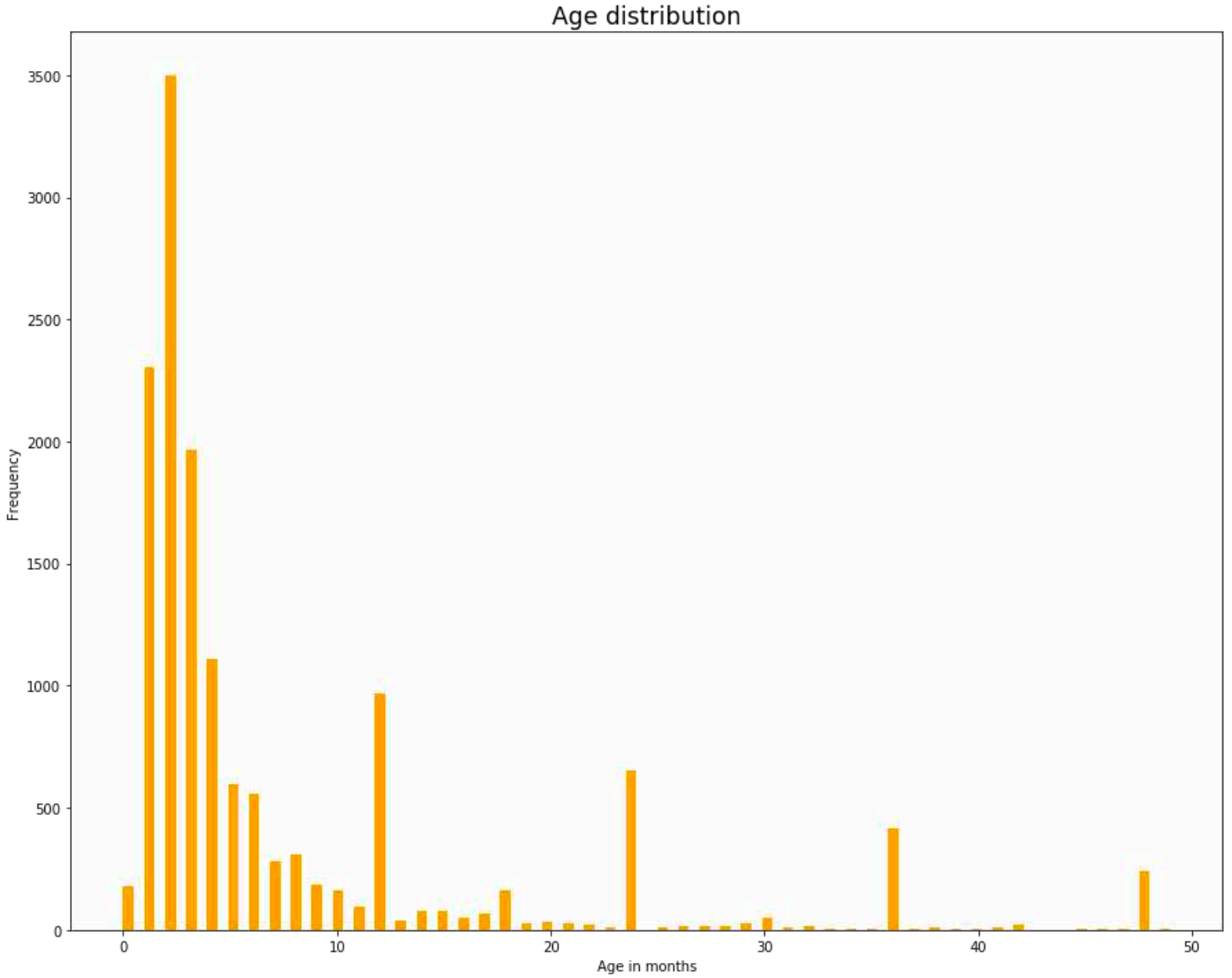
Explore the Data



- **0** Adopted on the same day as it was listed
- **1** Adopted within the first week it was listed
- **2** Adopted within the first month it was listed
- **3** Adopted within 90 days of being listed
- **4** No adoption after 100 days of being listed

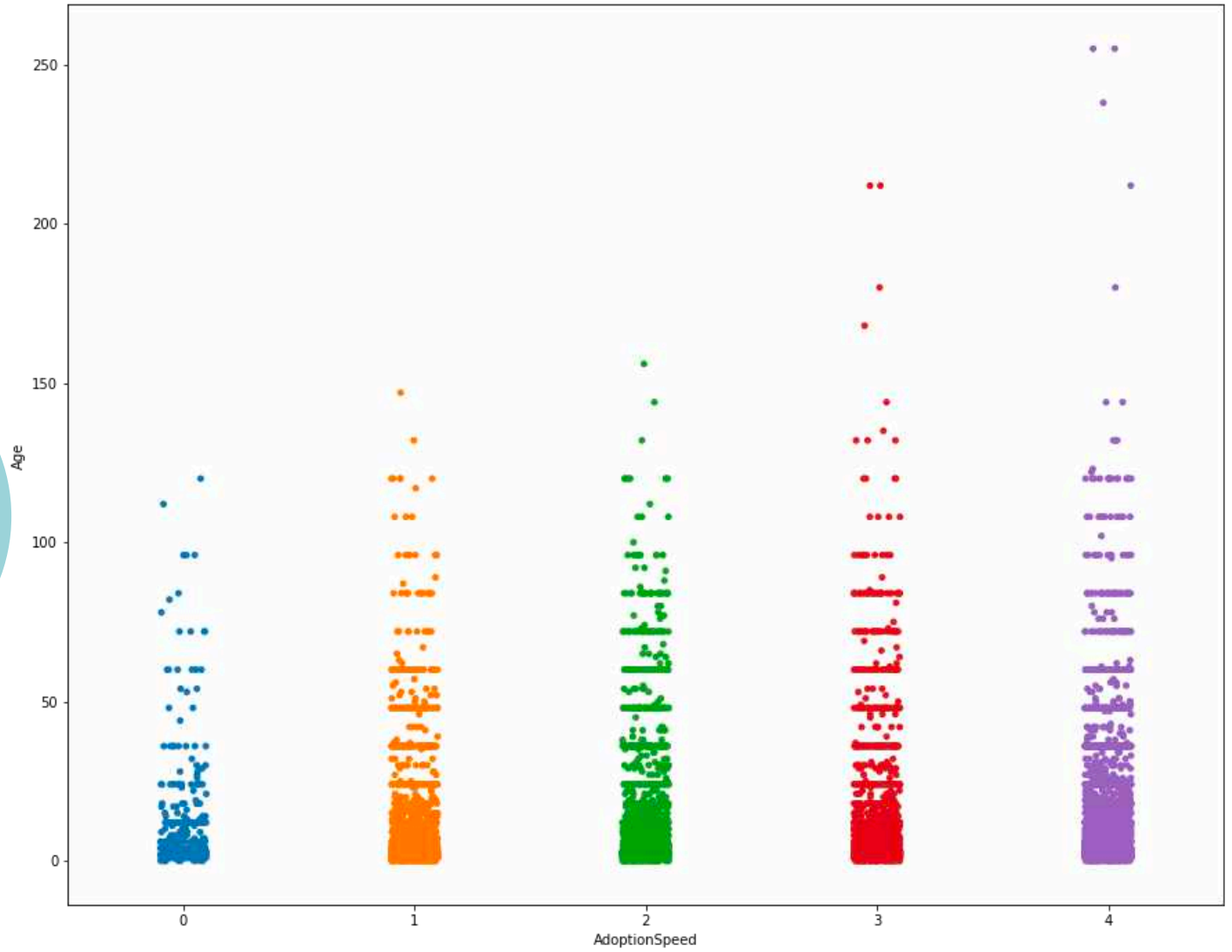
Explore the Data

Age



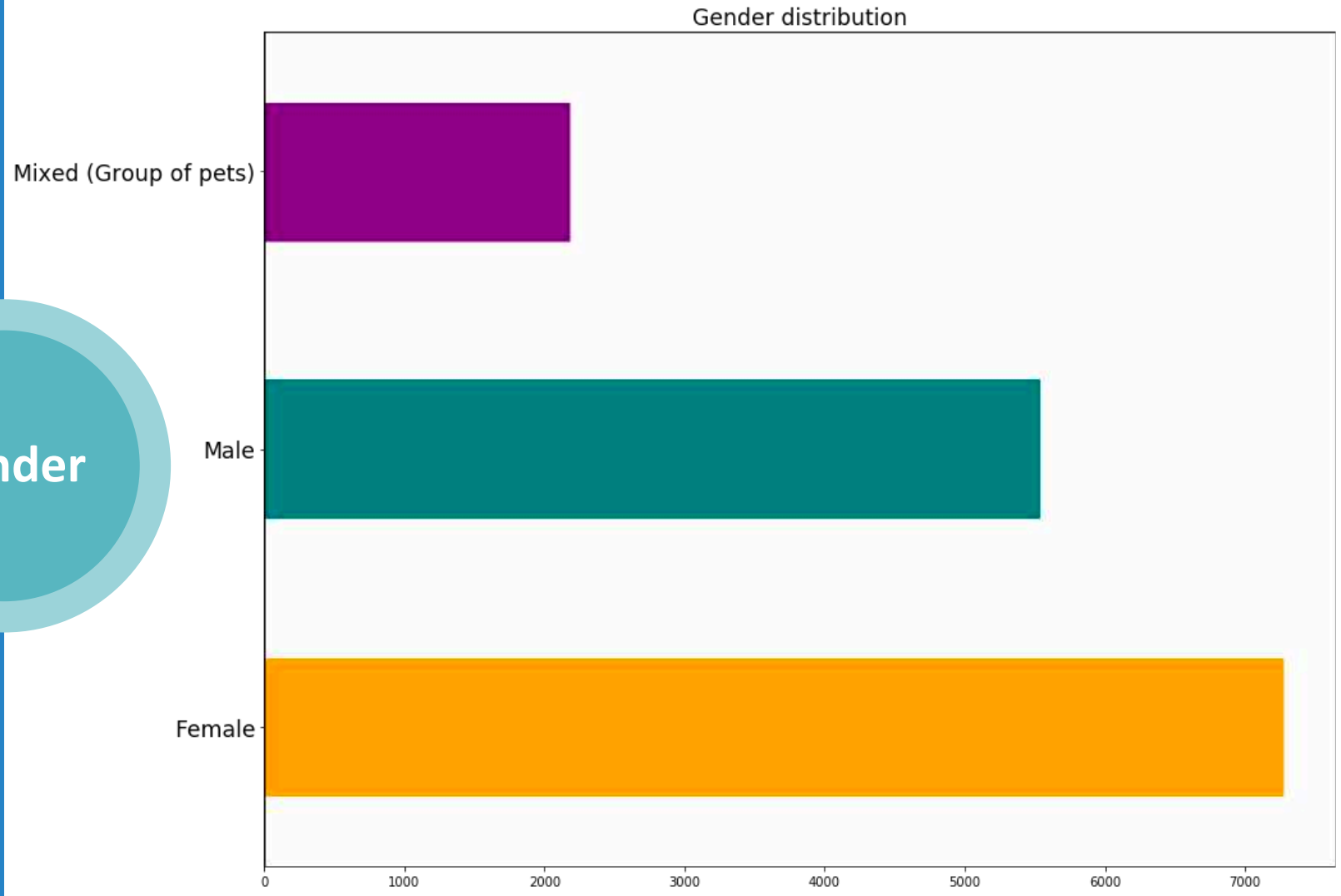
Explore the Data

Age



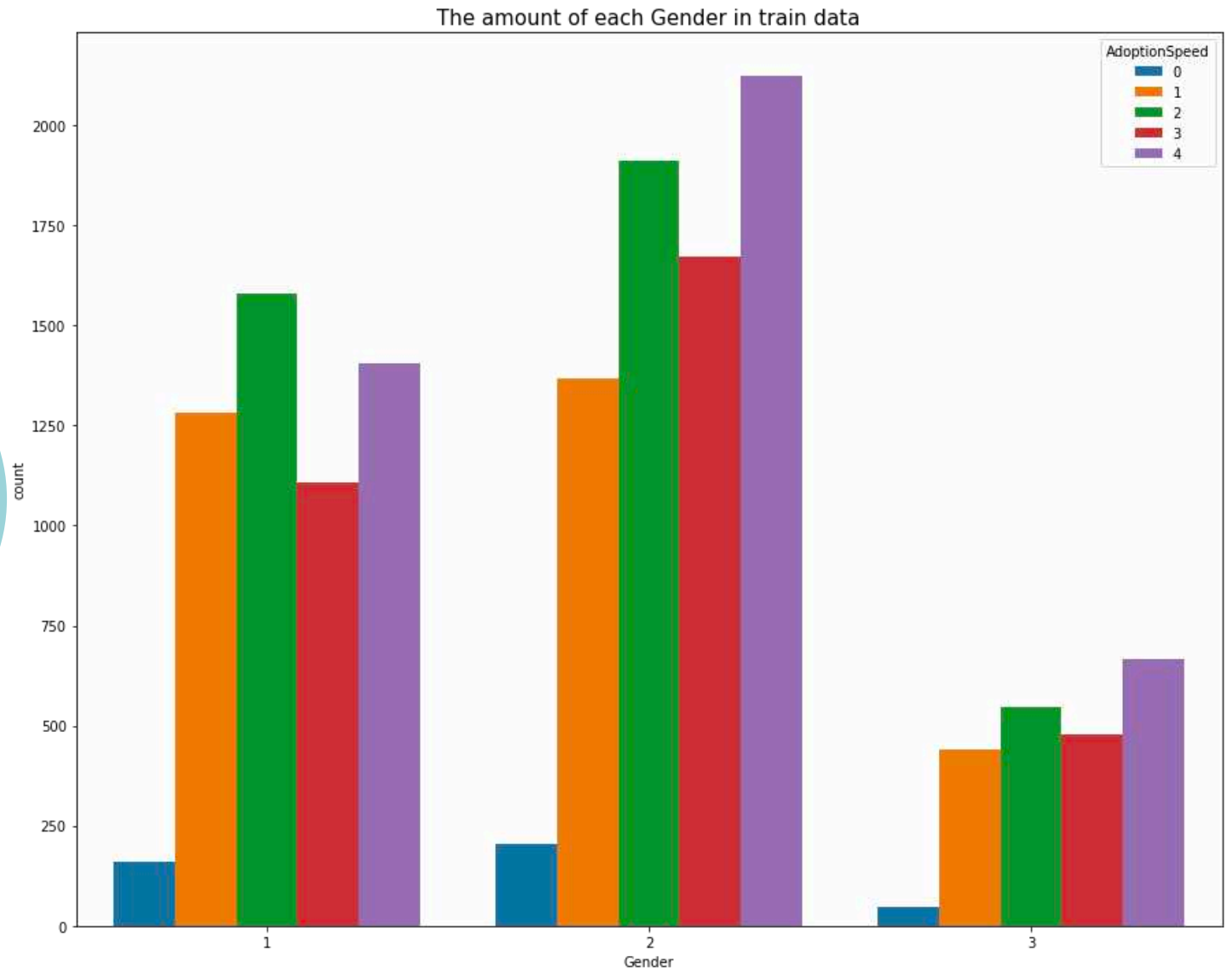
Explore the Data

Gender



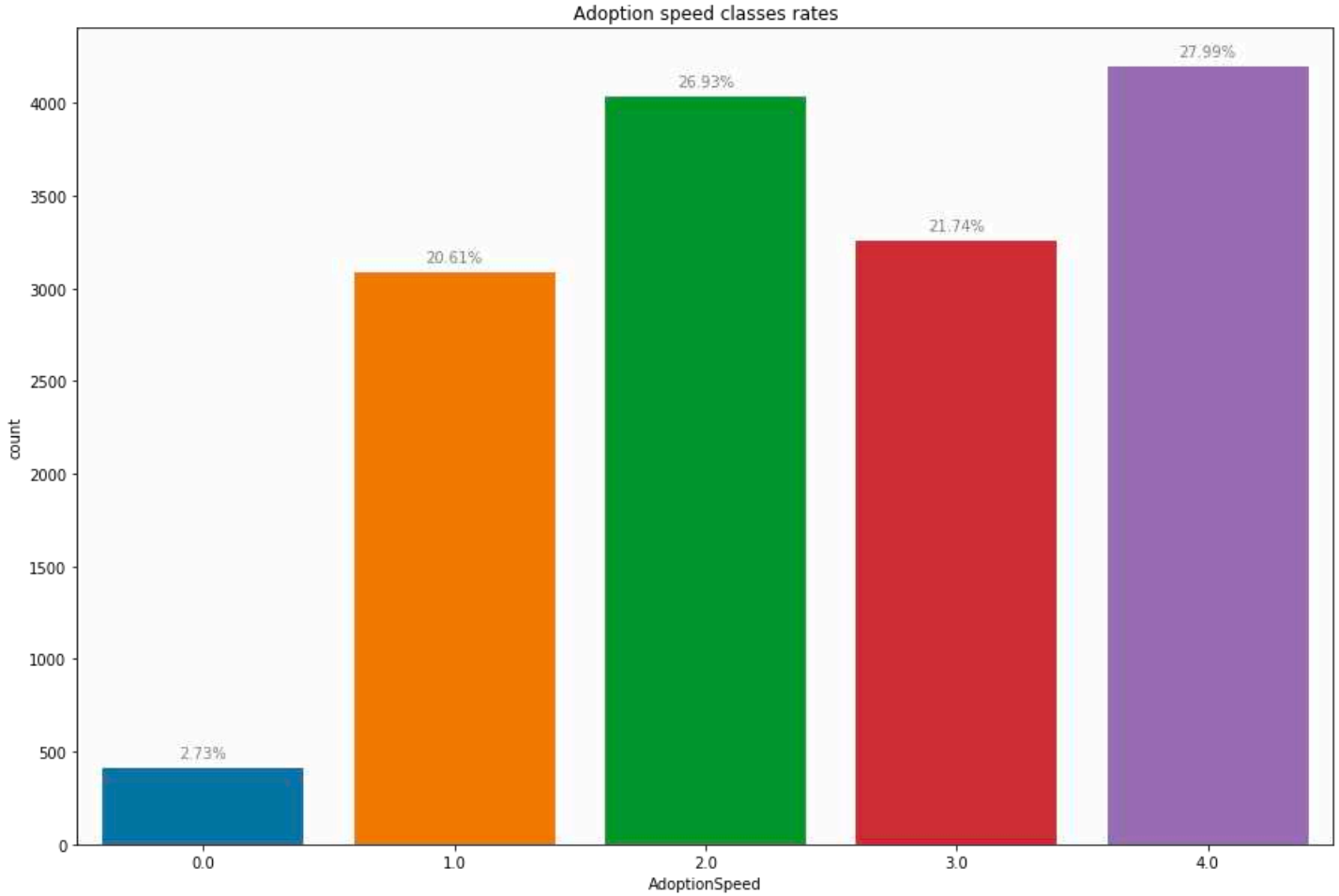
Explore the Data

Gender



Explore the Data

Gender

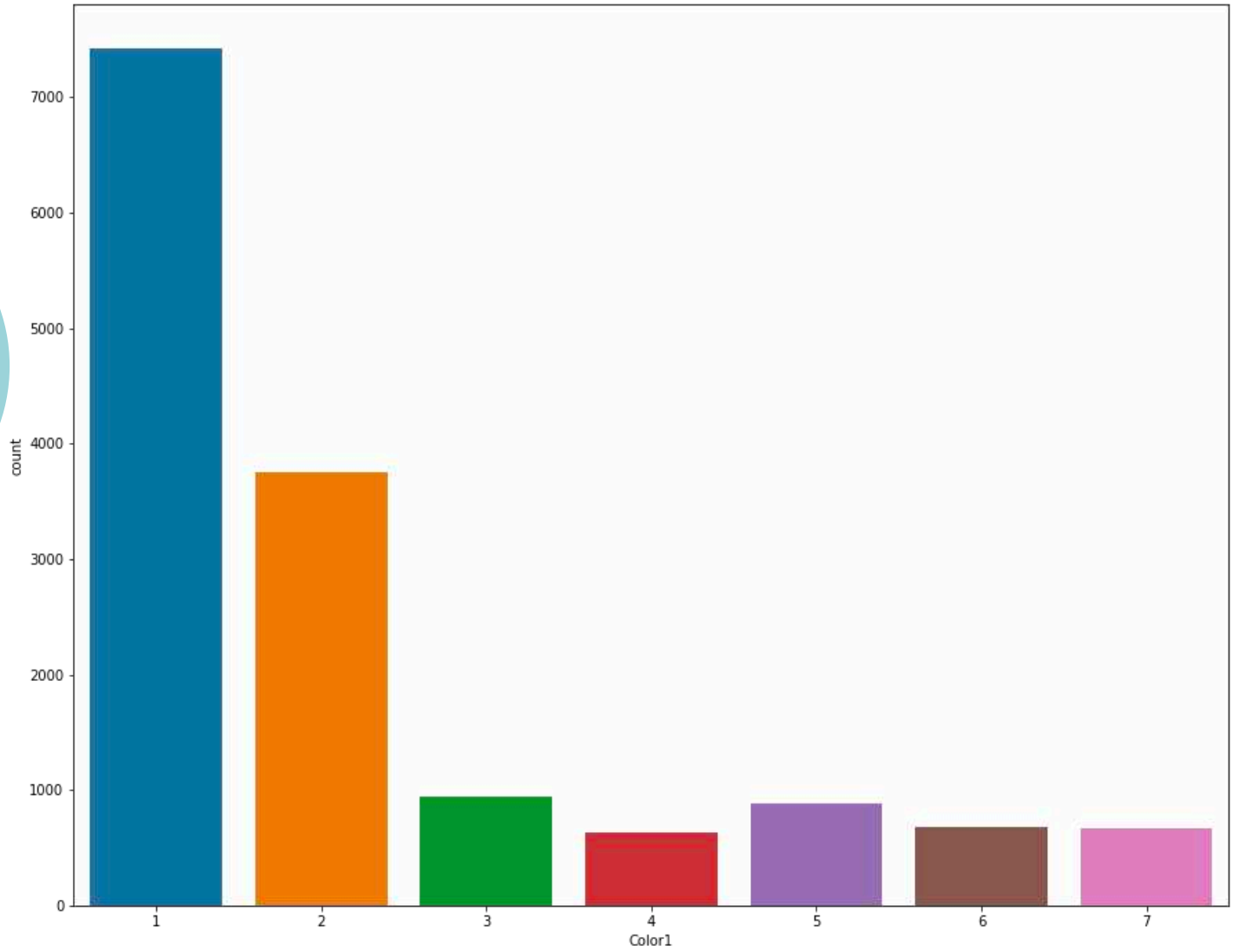


Explore the Data

Color

Color ID:

1 = Black | 2 = Brown | 3 = Golden| 4 = Yellow | 5 = Cream | 6 = Gray | 7 = White

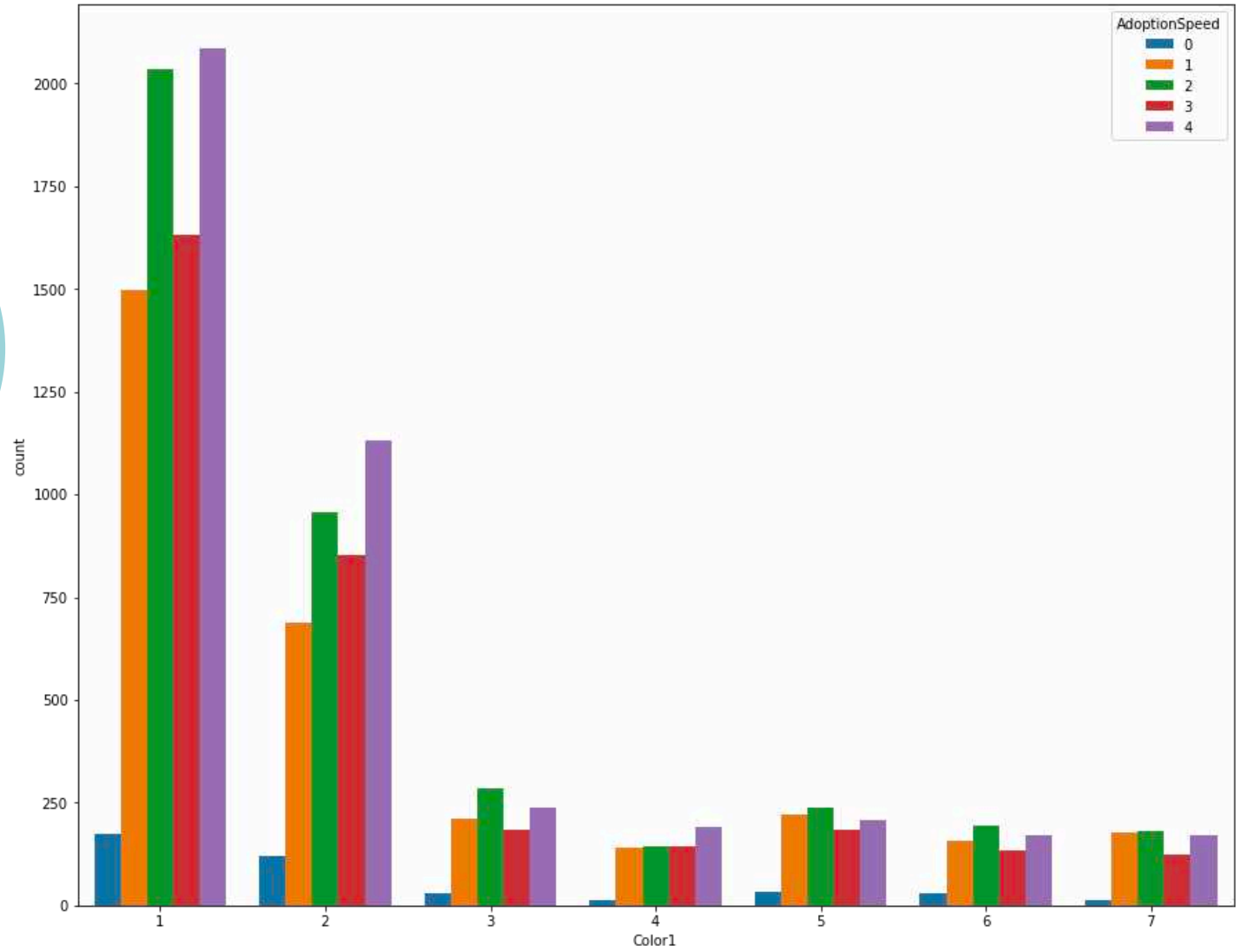


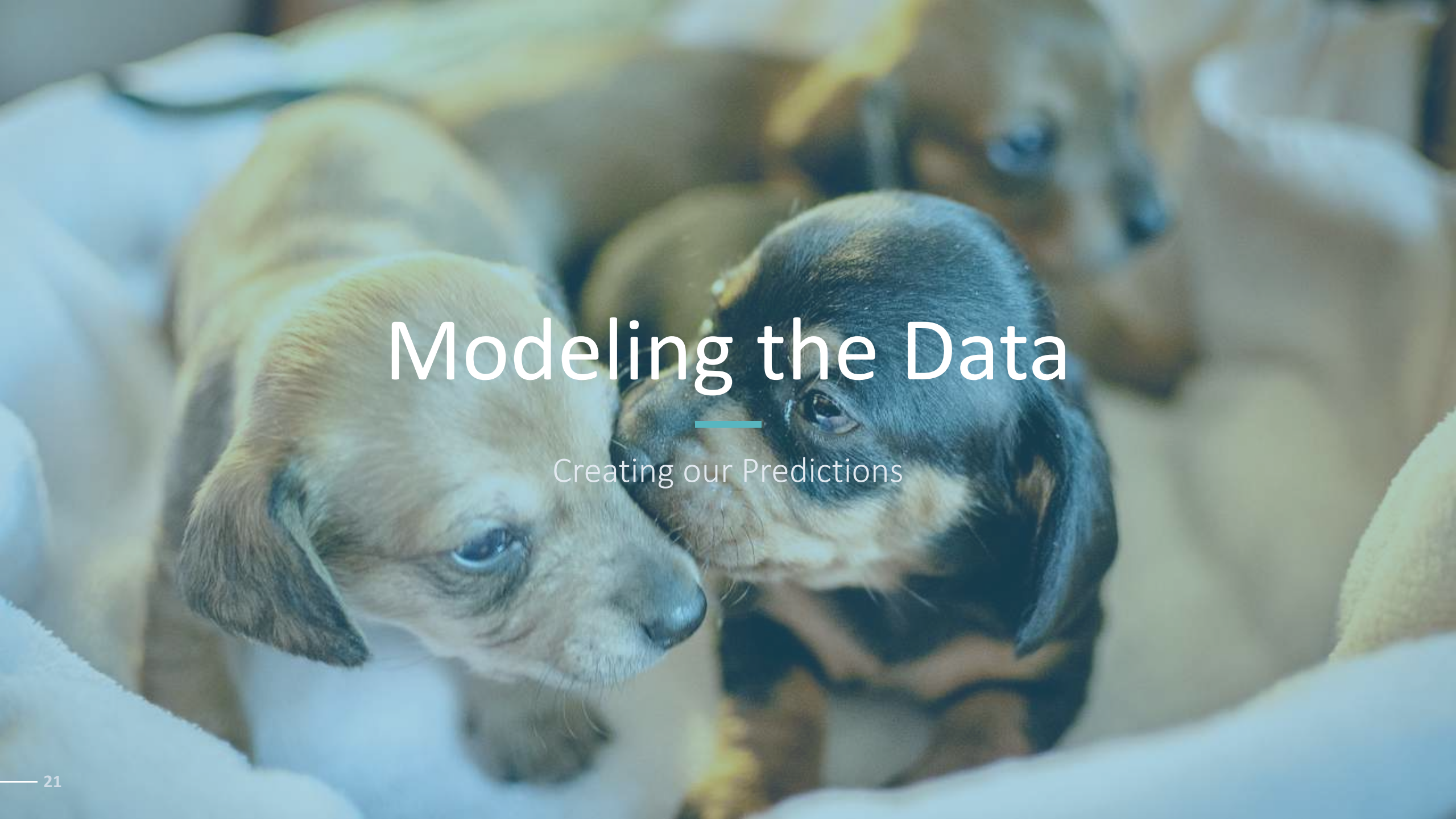
Explore the Data

Color

Color ID:

1 = Black | 2 = Brown | 3 = Golden| 4 = Yellow | 5 = Cream | 6 = Gray | 7 = White



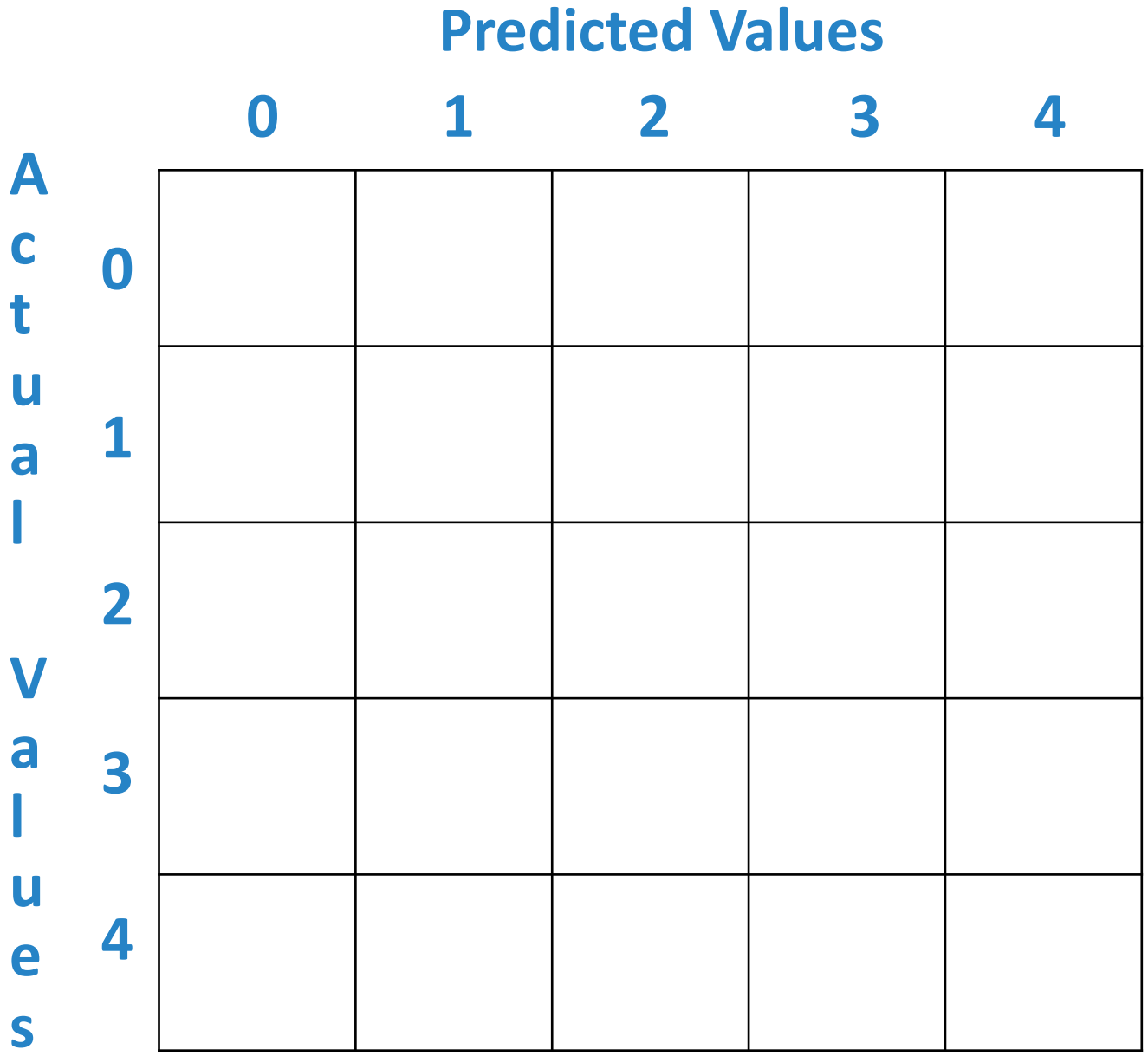


Modeling the Data

Creating our Predictions

Model the Data

Confusion Matrix



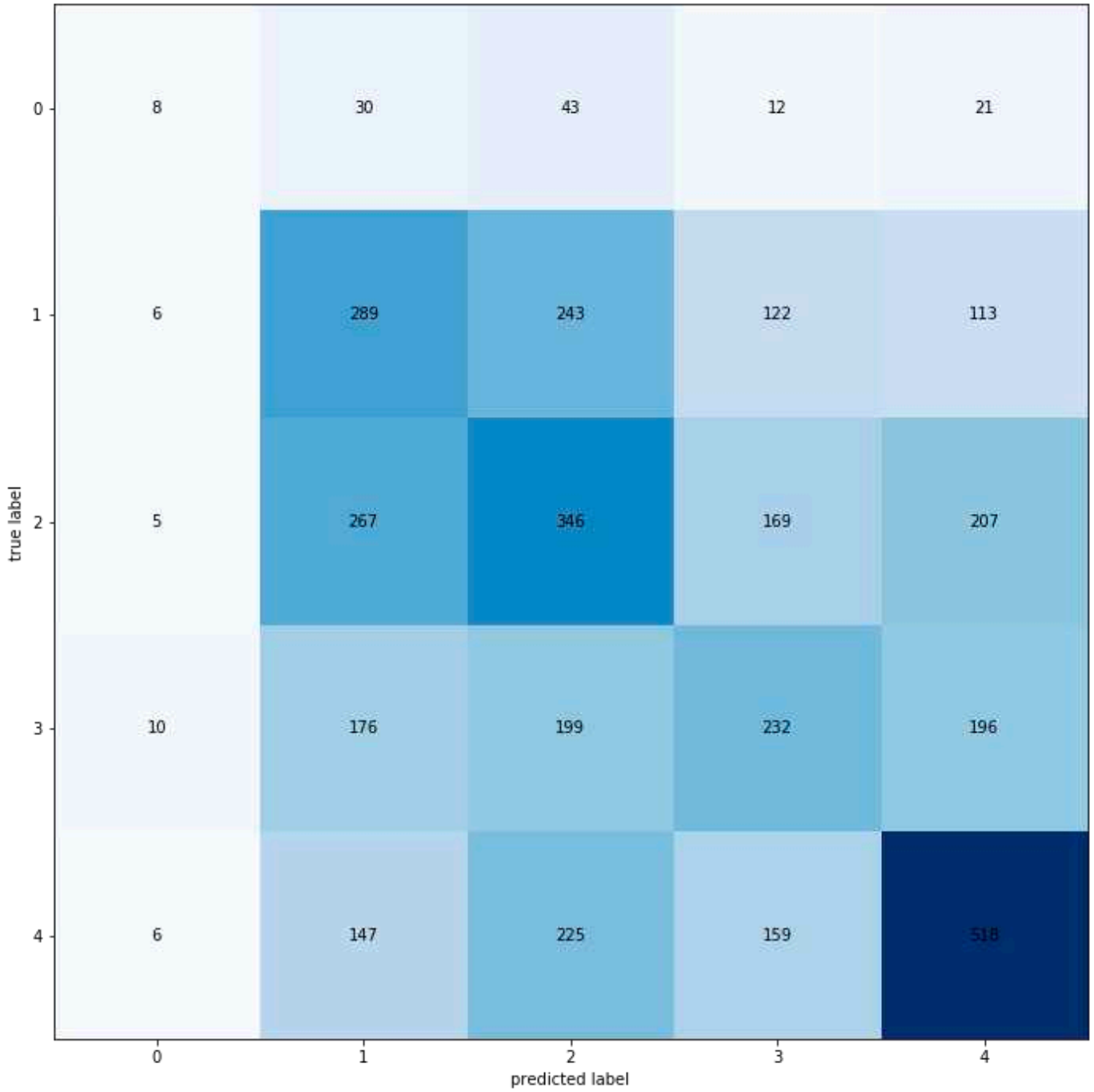
Model the Data

Confusion Matrix

Actual Values \ Predicted Values	0	1	2	3	4
0	100	0	0	0	0
1	0	500	0	0	0
2	0	0	1000	0	0
3	0	0	0	500	0
4	0	0	0	0	1000

Model the Data

Confusion Matrix



Evaluate the Model

Attempted Five Varieties of Classifiers

Random Forest | Decision Tree | Logistic Regression | K Nearest Neighbor (KNN) | Bagging

GENERAL ASSEMBLY

Evaluate the Model

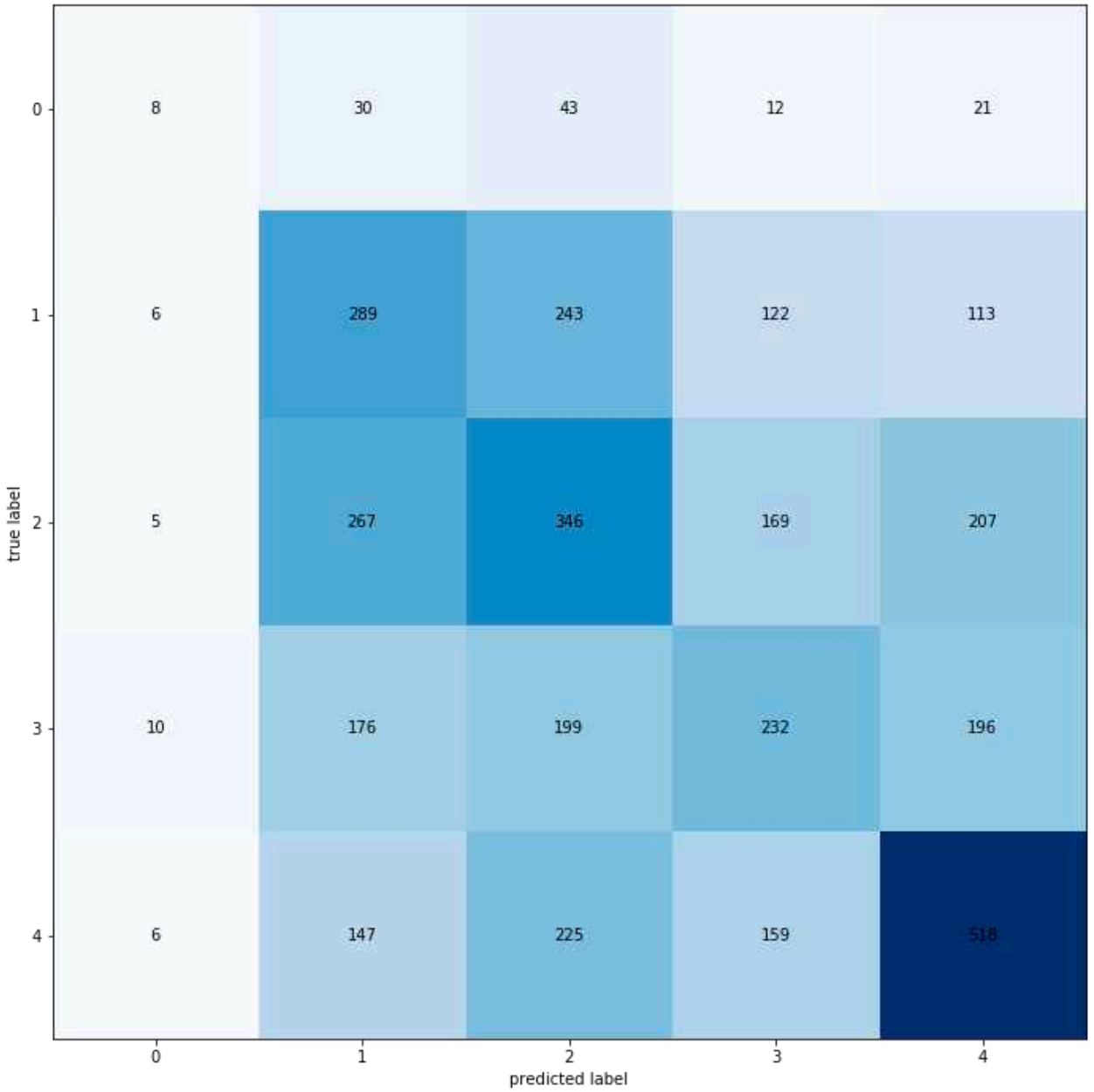
Confusion Matrix

Actual Values \ Predicted Values	0	1	2	3	4
0	1000	0	0	0	0
1	0	3000	0	0	0
2	0	0	3000	0	0
3	0	0	0	3000	0
4	0	0	0	0	4000

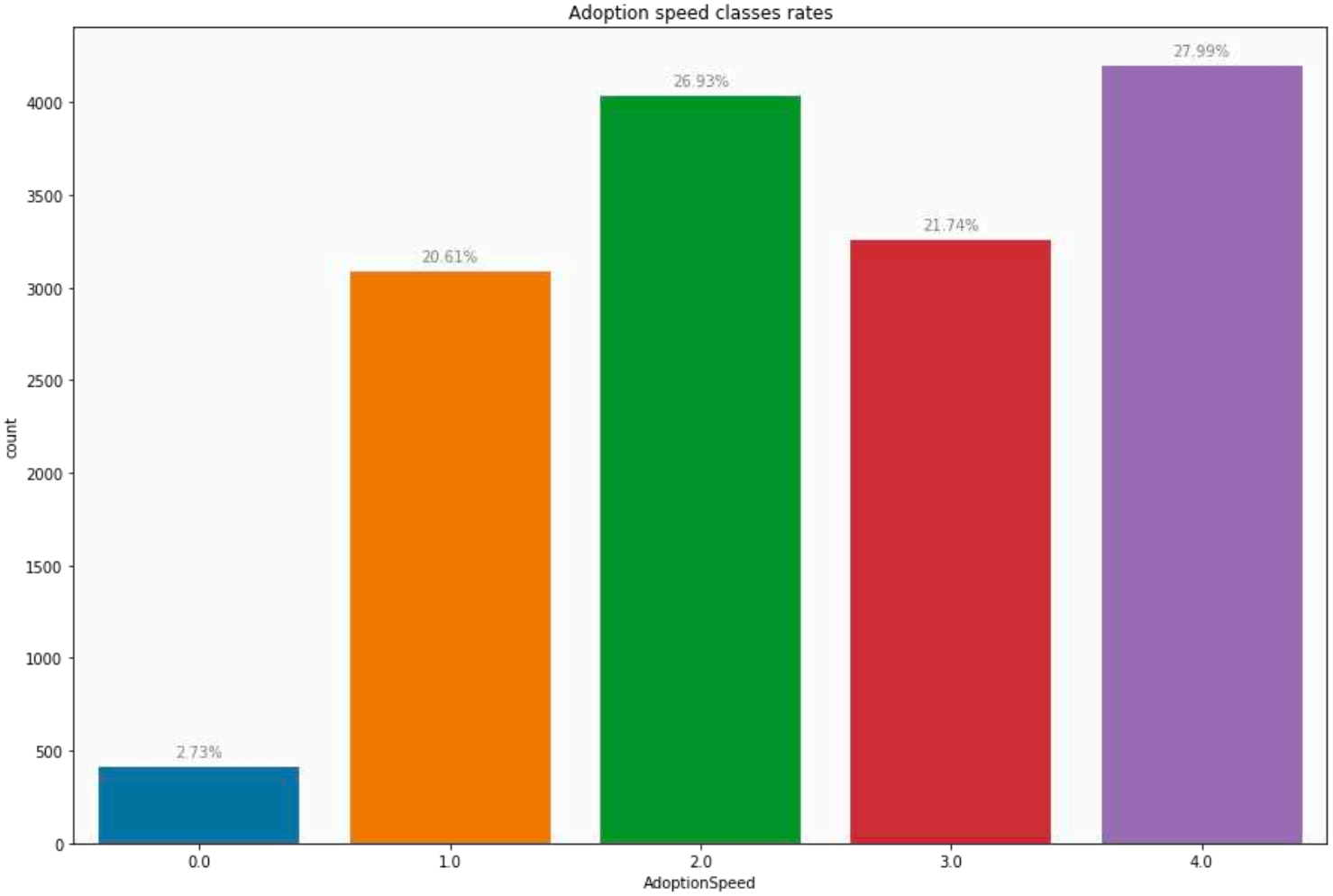
Evaluate the Model

Confusion Matrix

Bagging Classifier



Evaluate the Model



- **0** Adopted on the same day as it was listed
- **1** Adopted within the first week it was listed
- **2** Adopted within the first month it was listed
- **3** Adopted within 90 days of being listed
- **4** No adoption after 100 days of being listed

GENERAL ASSEMBLY

Evaluate the Model

Bagging Metrics

	precision	recall	f1-score	support
Class 0	0.19	0.07	0.10	114
Class 1	0.32	0.36	0.34	773
Class 2	0.33	0.37	0.35	994
Class 3	0.34	0.27	0.30	813
Class 4	0.49	0.51	0.50	1055
micro avg	0.38	0.38	0.38	3749
macro avg	0.33	0.32	0.32	3749
weighted avg	0.37	0.38	0.37	3749

Recommendations

Where do we go from here?


COMING SOON...!

Thank you