Initial Results for Meeting.

— using RandomForest.

Before Balancing:

- *Baseline Function Accuracy*:
  - 0.4603775 — sum(dog_total$Outcome_Type == Adoption)/nrows(dog_total)
- *Random Forest Accuracy*:
  - 0.9413287

```
Call:
 randomForest(formula = Outcome.Type ~ ., data = train)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 3

        OOB estimate of  error rate: 5.36%
Confusion matrix:
                Adoption Disposal Euthanasia Missing Return to Owner Rto-Adopt Transfer
Adoption           15775        0          0       0              384         0        0
Disposal               5        1          0       0                4         0        0
Euthanasia             2        0       1071       0                2         0        0
Missing                1        0          0       0                3         0        4
Return to Owner     1330        0          1       0             8897         0        0
Rto-Adopt            132        0          0       0               15         0        0
Transfer               0        0          0       0                0         0     7530
                class.error
Adoption         0.02376385
Disposal         0.90000000
Euthanasia       0.00372093
Missing          1.00000000
Return to Owner  0.13013297
Rto-Adopt        1.00000000
Transfer         0.00000000

Purebred.or.Mixed            Gender  Intake.Condition    Found.Location         Gen.Color
         36.57362          61.81315         216.83973         220.43282         342.08906
        AKC.Group     Outcome.Month  Age.upon.Outcome  Sex.upon.Outcome    Sex.Difference
        411.89071         524.43941         999.82782        1002.50222        1450.98899
      Intake.Type   Days.to.Outcome   Outcome.Subtype
       2704.68630        4393.47602        9656.09670
```

^^ MeanDecreaseGini

After Balancing:

- *Random Forest Accuracy*:
  - 0.9386536

Distribution of Classification Classes:

Class	Total Entries
Missing	16
Disposal	15
RTO	14571
RTO-Adopt	216
Transfer	10747
Euthanasia	1537
Adopt	23122

```
Call:
 randomForest(formula = Outcome.Type ~ ., data = train_smote)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 3

        OOB estimate of  error rate: 3.11%
Confusion matrix:
                Adoption Disposal Euthanasia Missing Return to Owner Rto-Adopt Transfer
Adoption            7517       23          0       0              134         0        0
Disposal               7     8289          0       0               20         0        0
Euthanasia             3        1        497       0                1         0        0
Missing                0        0          0       2                2         0        1
Return to Owner      486       26          1       0             4253         0        0
Rto-Adopt             63        0          0       0                7        11        0
Transfer               0        0          0       0                0         0     3586
                class.error
Adoption        0.020458692
Disposal        0.003246753
Euthanasia      0.009960159
Missing         0.600000000
Return to Owner 0.107637432
Rto-Adopt       0.864197531
Transfer        0.000000000
```

```
> tree_test[[1]]$importance[order(tree_test[[1]]$importance),]
Purebred.or.Mixed            Gender    Found.Location         Gen.Color         AKC.Group
         47.03807         138.40770         414.12539         506.20640         621.11206
 Intake.Condition       Intake.Type Age.upon.Outcome    Sex.Difference     Outcome.Month
       1051.52949        1245.95766        1296.16407        1421.49660        1445.26952
  Days.to.Outcome Sex.upon.Outcome   Outcome.Subtype
       2461.48005        2487.01568        4716.51258
```