DASC 1104 Project Proposal

Braden Fowler

My blog is available at https://dasc1104-blog-project.netlify.app

```
##    season episode             title imdb_rating total_votes   air_date
## 1       1       1             Pilot         7.6        3706 2005-03-24
## 2       1       2     Diversity Day         8.3        3566 2005-03-29
## 3       1       3       Health Care         7.9        2983 2005-04-05
## 4       1       4      The Alliance         8.1        2886 2005-04-12
## 5       1       5        Basketball         8.4        3179 2005-04-19
## 6       1       6          Hot Girl         7.8        2852 2005-04-26
## 7       2       1       The Dundies         8.7        3213 2005-09-20
## 8       2       2 Sexual Harassment         8.2        2736 2005-09-27
## 9       2       3   Office Olympics         8.4        2742 2005-10-04
## 10      2       4          The Fire         8.4        2713 2005-10-11
```

For this project, I am going to be looking at Office Ratings data set contained in the office_ratings.csv file on Tidy Tuesdays website. This data set contains ratings of each episode from the famous US adapted sitcom television series. The data has 188 observation and 6 variables. The variable season tells us which season of the show is being selected from. The variable episode tells us which episode of the season is being selected. The variable title tells us the title of the episode being selected. The variable imdb_rating tell us the rating given for the episode by imdb. The variable total_votes tells us how many people contributed to the imdb_rating; and finally, the variable air_data tells us when that specific episode was aired.

Question 1: What is the highest rated episode? Question 2: Have the ratings of the episodes changed over time? Question 3: Which season contains the highest average rating?

```
## Observations: 58,180
## Variables: 36
## $ id              <int> 46042150, 46042002, 46040898, 46039877, 46039306, 46039304...
## $ org_id          <fct> NV163, NV163, NV99, NV202, NV184, NV184, NV184, NV184, NV1...
## $ url             <fct> https://www.petfinder.com/dog/harley-46042150/nv/las-vegas...
## $ species         <fct> Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog...
## $ breed_primary   <fct> American Staffordshire Terrier, Pit Bull Terrier, Shepherd...
## $ breed_secondary <fct> Mixed Breed, Mixed Breed, NA, NA, NA, Beagle, Chihuahua, N...
## $ breed_mixed     <lgl> TRUE, TRUE, FALSE, FALSE, FALSE, TRUE, TRUE, TRUE, TRUE, T...
## $ breed_unknown   <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FA...
## $ color_primary   <fct> "White / Cream", "Brown / Chocolate", "Brindle", NA, NA, N...
## $ color_secondary <fct> Yellow / Tan / Blond / Fawn, White / Cream, NA, NA, NA, NA...
## $ color_tertiary  <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ age             <fct> Senior, Adult, Adult, Baby, Young, Baby, Baby, Baby, Baby,...
## $ sex             <fct> Male, Male, Male, Female, Male, Male, Female, Male, Female...
## $ size            <fct> Medium, Large, Large, Large, Small, Medium, Small, Medium,...
## $ coat            <fct> Short, Short, Short, NA, Long, Short, Short, Medium, Mediu...
## $ fixed           <lgl> TRUE, TRUE, TRUE, FALSE, TRUE, TRUE, TRUE, TRUE, TRUE, TRU...
## $ house_trained   <lgl> TRUE, TRUE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALS...
```

```
## $ declawed        <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ special_needs   <lgl> FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FALSE, FA...
## $ shots_current   <lgl> TRUE, TRUE, TRUE, FALSE, TRUE, TRUE, TRUE, TRUE, TRUE, TRU...
## $ env_children    <lgl> NA, NA, NA, NA, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, ...
## $ env_dogs        <lgl> NA, NA, NA, NA, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, ...
## $ env_cats        <lgl> NA, NA, NA, NA, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, TRUE, ...
## $ name            <fct> HARLEY, BIGGIE, Ziggy, Gypsy, Theo, Oliver, Macadamia, Dod...
## $ tags            <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ photo           <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ status          <fct> adoptable, adoptable, adoptable, adoptable, adoptable, ado...
## $ posted          <fct> 2019-09-20T16:37:59+0000, 2019-09-20T16:24:57+0000, 2019-0...
## $ contact_city    <fct> Las Vegas, Las Vegas, Mesquite, Pahrump, Henderson, Hender...
## $ contact_state   <fct> NV, NV, NV, NV, NV, NV, NV, NV, NV, NV, NV, NV, NV, NV, NV...
## $ contact_zip     <fct> 89147, 89147, 89027, 89048, 89052, 89052, 89052, 89052, 89...
## $ contact_country <fct> US, US, US, US, US, US, US, US, US, US, US, US, US, US, US...
## $ stateQ          <fct> 89009, 89009, 89009, 89009, 89009, 89009, 89009, 89009, 89...
## $ accessed        <fct> 2019-09-20, 2019-09-20, 2019-09-20, 2019-09-20, 2019-09-20...
## $ type            <fct> Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog, Dog...
## $ description     <fct> "Harley is not sure how he wound up at shelter in his seni...
```

```
## Observations: 90
## Variables: 5
## $ location <fct> Texas, Alabama, North Carolina, South Carolina, Georgia, Puerto R...
## $ exported <int> 635, 268, 158, 139, 137, 131, 130, 76, 66, 57, 55, 54, 53, 50, 49...
## $ imported <int> NA, 2, 14, 12, 19, NA, 3, NA, 20, 4, NA, NA, NA, NA, 32, 1, NA, N...
## $ total    <int> 566, 1428, 2627, 1618, 3479, NA, 1664, NA, 1769, 1123, 510, NA, 9...
## $ inUS     <lgl> TRUE, TRUE, TRUE, TRUE, TRUE, FALSE, TRUE, FALSE, TRUE, TRUE, TRU...
```

```
## Observations: 6,194
## Variables: 8
## $ id            <int> 44520267, 44698509, 45983838, 44475904, 43877389, 43082511, ...
## $ contact_city  <fct> Anoka, Groveland, Adamstown, Saint Cloud, Pueblo, Manchester...
## $ contact_state <fct> MN, FL, MD, MN, CO, CT, OH, OH, AL, AL, AL, AL, AL, AL, AL, ...
## $ description   <fct> "Boris is a handsome mini schnauzer who made his long trek u...
## $ found         <fct> Arkansas, Abacos, Adam, Adaptil, Afghanistan, Afghanistan, A...
## $ manual        <fct> NA, Bahamas, Maryland, NA, NA, NA, Ohio, Ohio, NA, NA, NA, N...
## $ remove        <lgl> NA, NA, NA, TRUE, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA...
## $ still_there   <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, ...
```

Using the data from TidyTuesdays website, I am going to look at dog descriptions, exporting and importing of dgos, and dogs that have traveled. The first data set contains 58,180 observations and 36 variables. The second data set contains 90 observations and 5 variables. The third data set contains 6,194 observations and 8 variables. From these data sets I can see which dogs are better able to perform tricks, which dogs travel more than other dogs, whether these certain dogs are mixed, or purebred, their colors, sex, size, coat length, and many others. Using this data you can find out quite a bit about certain dog species.

Question 1: Which state exports the most dogs? Which state imports the most dogs? Question 2: Where do dogs travel the most? What state has the most traveling dogs? Question 3: What proportion of dogs are pure-bred? What proportion of dogs have long coats? Question 4: What proportion of dogs are seniors? What breed is put up for adoption the most?