**Final Report**

***Problem Statement and Questions***

**Motivation:**

In the last few years, Airbnb has gained popularity in the United States has increased significantly ([pet adoption stats](https://news.orvis.com/dogs/pet-adoption-statistics-the-numbers-behind-the-need)). Nowadays families are preferring to have pets instead of kids (anyone watched Boss Baby – an animated movie??).

Considering this, a pet venture can be a good idea to start a business.

**Study Limitation:**

For the analysis purpose we are limiting our study related to Cats or Dogs as pets. We are not taking into account for other types of pets.

**Problem Context:**

“Petwale” is a new startup aiming to provide pets all over the US. They are planning to open multiple outlets across the states but in order to start the business, they would like to invest in one state in the beginning and then based on the response they would like to take next steps.

Analytics team of “Petwale” has following key questions to answer before they start their venture:

* Location: Which state they should invest in first to begin their venture?
* Pet: Whether they should start with Dogs or Cats?
* Pet Popularity:  Understand which pet (Cats or Dogs) is popular across the state?
* Consumers: Like Number of Households (HH) in each state, Number HH owning pets, cats and dogs?
* Opportunity Available: Which state has the potential of adopting pets?

**Data Granularity:**

The data contains state level information about the total households, households with Pets, household with Cats/Dogs and their corresponding proportion. In the data, we have total 48 rows where each row represents 1 state\*

Source: [Cat vs Dog Popularity in the US](https://public.tableau.com/en-us/s/resources?qt-overview_resources=1)

Considering this a pet venture can be a good idea to start a business.

**Variables in Data:**

|  |  |
| --- | --- |
| **Variables** | **Description** |
| Location | States in United States |
| State | 2 letter state code |
| Number of Households (in 1000) | Number of households in each state |
| Number of Pet Households (in 1000) | Count of households owning pets |
| Percentage of Households with Pets | Number of households in each state/Count of households owning pets |
| Dog Owning Households (1000s) | Count of households owning dog pets |
| Percentage of Dog Owners | Dog Owning Households (1000s)/Number of Households (in 1000) |
| Mean Number of Dogs per household | Dog Population (in 1000)/Dog Owning Households (1000s) |
| Dog Population (in 1000) | Total Dogs population in each state |
| Percentage of Cat Owners | Cat Owning Households/Number of Households (in 1000) |
| Cat Owning Households | Count of households owning cat pets |
| Mean Number of Cats | Cats Population (in 1000)/Cats Owning Households (1000s) |
| Cats Population | Total Dogs population in each state |

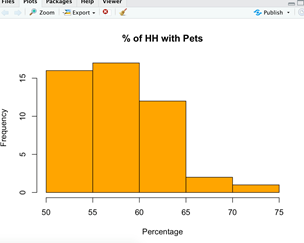
\*Removed state “District of Columbia” as it has only 287 HH’s, which significantly low as compared to other states HH’s count

***Descriptive Statistics***

Cats vs. Dogs Comparison

From the preliminary comparison analysis between Cats and Dogs across all the states, we see that:

1. Number of HH with Dogs are higher than Cats
2. But, the Total Population of Cats is more than Dogs







Comparison of % of HH’s with Pets, Cats and Dogs

From the histograms we can observe that:

1. 33 States out of 48 have 50-60% of HH’s with Pets
2. 25 Sates out of 48 states have 30-40% of HH’s with Dogs only
3. 37 Sates have 30-45% of HH’s with Dogs
4. 40 States have 25-35% of HH with Cats

**From this we can confirm that initial analysis shows that the Dogs are more popular among the HH’s across all the States.**

Comparison of avg. number of Dogs or Cats per HH across the States

From the histogram charts we notice that:

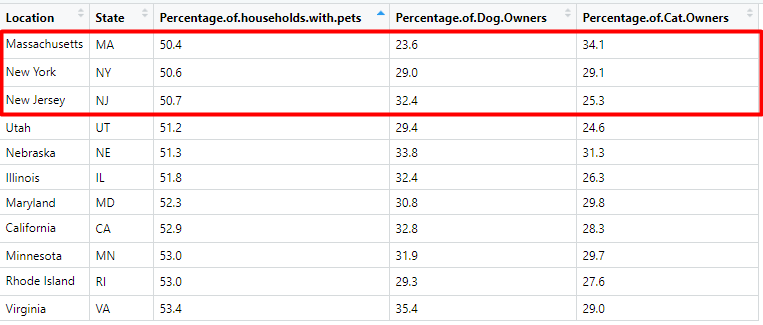
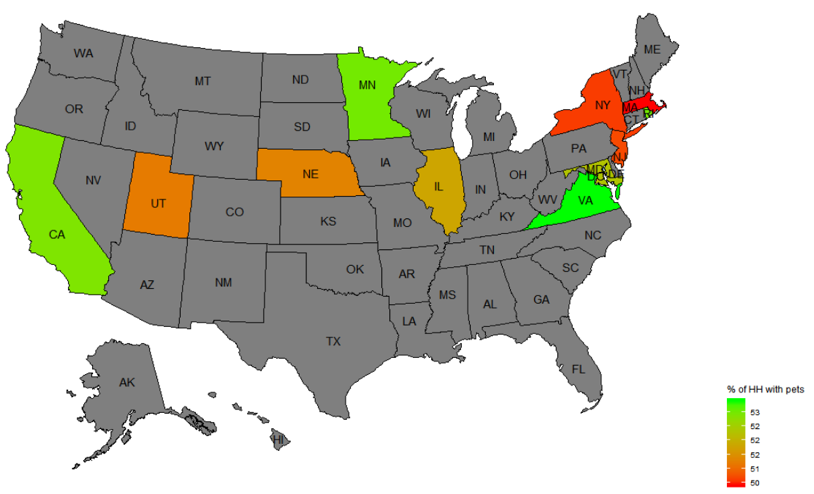
1. Avg. Cats varied from 1.7 to 2, whereas Dogs varied from 1.1 to 2.1
2. Major proportion of States (42 out of 48) have avg. no. of Cats between 1.8 to 2.2
3. Whereas, 32 States have avg. no. of Dogs between 1.2 to 1.6

We can conclude that per HH there are more number of Cats than Dogs across the United States.

So far, we analyzed the data at overall level including all the states. Now, we would like to do some analysis and comparison at State level.

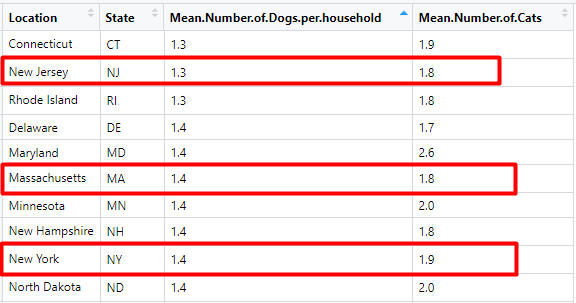
1. Bottom 10 States with the least % of HH’s having Pets:

* As objective is to start the pet venture, we would first like to understand what the states with the least are % of HH’s having Pets.
* This would help us to scope out the potential states to start the venture
* From the below charts, we can observe that states like Massachusetts, New York and New Jersey have least % of HH with Pets. Also, Massachusetts and New York have least % of HH with Dogs as well.



Bottom 10 States with average number of Dogs and Cats per HH:

* In the below data table, we are comparing the states with the least avg. number of Cats vs. Dogs per HH



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Top 10 pet-owning states**  **(percentage of households that owned a pet)** | **Top 10 states with the most dog owners**  **(percentage of households that owned a dog)** | **Top 10 states with the most cat owners**  **(percentage of households that owned a cat)** | **Bottom 10 pet-owning states**  **(percentage of households that owned a pet)** | **Bottom 10 pet-owning states**  **(percentage of households that owned a dog)** | **Bottom 10 pet-owning states**  **(percentage of households that owned a cat)** |
| Vermont: 71% | Arkansas: 48% | Vermont: 50% | Massachusetts: 50.4 % | Massachusetts: 23.6 % | Utah: 24.6 % |
| New Mexico: 68% | New Mexico: 46% | Maine: 46% | New York: 50.6 % | Connecticut: 28.3% | New Jersey: 25.3% |
| South Dakota: 66% | Kentucky: 46% | Oregon: 40% | New Jersey: 50.7% | New York: 29 % | Louisiana: 25.9% |
| Oregon: 64% | Missouri: 46% | South Dakota: 39% | Nebraska: 51.2% | Rhode Island: 29.3% | Illinois: 26.3% |
| Maine: 63% | West Virginia: 46% | Washington: 39% | Utah: 51.3 % | Utah: 29.4 % | Florida: 27.3 |
| Washington: 63% | Mississippi: 45% | West Virginia: 38% | Illinois: 51.8% | New Hampshire: 30.3% | Georgia: 27.3 % |
| Arkansas: 62% | Alabama: 44% | Kentucky: 37% | Maryland: 52.3% | Maryland: 30.8% | Alabama: 27.4% |
| West Virginia: 62% | Tennessee: 44% | Idaho: 35% | California: 52.9% | Minnesota: 31.9% | Rhode Island: 27.6% |
| Idaho: 62 % | Alabama: 44% | Indiana: 34.4% | Minnesota: 53% | Illinois: 32.4% | South Carolina: 27.8% |
| Wyoming: 61.8% | Texas: 44 % | New Hampshire: 34.2 % | Rhode Island: 53% | New Jersey: 32.4% | California: 28.3% |

**Overall, we observe that West Virginia is common in top states owning pets- both dogs as well as cats. Kentucky is common between dogs and cats states whereas it is not appearing in top pet owning states as dogs and cats may be popular in Kentucky, on the other hand Vermont seems to have cats popularity than dogs, also it is on the top of pet owning states. New Mexico is popular in dogs than cats and it counts to 68% of HH that own a pet.**

**Also the states like - Massachusetts, New York and New Jersey have both low Percent of households owning Pets (Cats and Dog) and Average number of Cats & Dogs owned per household.**

***Sampling Distribution Analysis***

We observed that in order to start the Pet Venture we needed to do deep-dive analysis on two main attributes to make decision about which type of pet to go with in the market. The two main attributes are:

- Proportion of HH with Cats/Dogs/Pets

- Mean number of Cats/Dogs per HH

Before starting with the deep-dive analysis, we looked for some stats available online to understand about the population values

https://lh4.googleusercontent.com/iYLiSX8Z156418UhY_WrnTXMKQ_cZxc2wRWrfLIEWRo7xg_YnfFHGa6cwIS1hYE1i9RKG4Li7rLLMvtzMnU-2f8Cnk2Zmbq6TuZTHZXgeUf8U1W-ZsdXjtipVHiX6p6-Pm2xXtgY

**Source:** [American Veterinary Medical Association](https://www.avma.org/KB/Resources/Statistics/Pages/Market-research-statistics-US-pet-ownership.aspx)

Now, we would like to verify how close our sample data is to the information mentioned above and look at some other attributes:

**Terminology:**

**Interval Range**: Let say, if we obtain 100 different samples with same attributes, then we would like to know what the range of values is shown by 95 samples and how close are that to above values

|  |  |  |
| --- | --- | --- |
| **Attributes** | **Result from Sample Data** | **Interval Range** |
| **Total % of HH with Dogs** | 36% | 22.9% to 50.1% |
| **Total % of HH with Cats** | 30% | 17.3% to 43.3% |

**Conclusion:**From above intervals, we can observe that although the range of interval is same across Cats and Dogs but interval of % of HH with Dogs is higher than Cats.

|  |  |
| --- | --- |
| **Attributes** | **Interval Range** |
| **Diff in # of HH with Dogs & Cats** | 78 to 226 |

**Conclusion:** From the interval we can conclude that if we look at the across all the states, the number of HH with Dogs are more than the Cats.

**Mean number of Cats and Dogs per HH:**

**Mean number of Dogs per HH across all the states:**

Sample mean number of Dogs per HH = **1.59**

95% confidence interval of number of Dogs per HH is **1.54 to 1.65,** where the population mean is 1.6 (as shown above from source)

**Mean number of Cats per HH across all the states:**

Sample mean number of Cats per HH = **2.04**

95% confidence interval of number of Cats per HH is **1.99 to 2.11,** where the population mean is 2.1 (as shown above from source)

**Conclusion:** From above CI’s of Dogs and Cats, we can observe that the point estimates i.e. the average number of Cats per HH is higher than Dogs and its corresponding interval is also higher.

**“Mean of the difference” between Avg. number of Dogs & Cats per HH calculated at State level (matched sample case):**

Here we are taking the Matched Sample case to understand the difference between Avg. number of Cats & Dogs per HH. Although we take into account the factor of number of HH in states while taking an average of Dogs/Cats per HH, but there are other factors like:

- Different states have different pet related rules and regulations

- The household income and job type are different

- Weather conditions

These factors can also impact the number of HH with Dogs or Cats. Therefore, instead of taking the difference of the mean (Independent sample), we took the mean of the difference (matched sample).

Sample Mean of difference between the mean number of Dogs and Cats per HH= **0.45**

95% confidence interval is: **-0.50 to -0.038**

**Conclusion:** As the above CI is completely negative, the mean number of Dogs per HH is less than the Cats.

We got 2 important insights from above analysis of Sampling Distribution between Dogs and Cats:

* % of HH with Dogs is higher than Cats, there are some states where HH proportion of Cats is also higher.
* Average number of Dogs per HH is less than the Cats.

**Hypothesis testing that number of HH with Dogs is higher than Cats across the states:**

**H0:**  HH with Dogs is lower or equal to Cats

**H1:** HH with Dogs is higher than Cats

Upper-tailed test

Taking Significance level - Alpha = 0.05

**p-value = 0.00008 <0.05, reject Ho, means HH with Dogs is higher than Cats**

**Hypothesis testing of average number of Dogs is less than Cats per HH:**

**H0:**  Mean number of Dogs per HH is equal or greater than Cats

**H1:**  Mean number of Dogs per HH is lower than Cats

Lower-tailed, Matched Sample t-test

Taking Significance level - Alpha = 0.05

**p-value = almost 0 < 0.05, can reject Ho, means Ha is correct i.e. Mean number of Dogs per HH is lower than Cats**

**Interpretation and Decision:**

From the hypothesis testing we can confirm that:

1. More HH’s are there with Dogs as compared to Cats
2. But, the average number of Cats per HH is higher than Dogs.

**Interpretation:**

From above analysis, we notice that although the Dogs are more popular in the HH’s but the avg. number of Cats per HH is higher. There could be following reasons for having avg. of Cats being higher than Dogs per HH:

1. In actual, people like to have more than one cat at home whereas in case of Dogs they prefer only 1
2. Life span of pet cats and dogs could be another factor. Study shows that avg. lifespan of Indoor Cats is 13 to 17 years ([Source](https://www.petcarerx.com/article/what-can-affect-a-cats-lifespan/1248)) whereas for Dogs its 8 to 11 years ([Source](https://www.petmd.com/dog/wellness/evr_dg_how_long_do_dogs_live))

**Recommendation:**

**Therefore, considering all the factors and results obtained above we recommended to go ahead with selling the Dogs in the state.**

***Final Analysis***

Since from previous analysis we came to know that dog is the pet to start business (compared to cats), we will now target which state to start “Petwale” venture with. To do so we are considering these factors to answer our question:

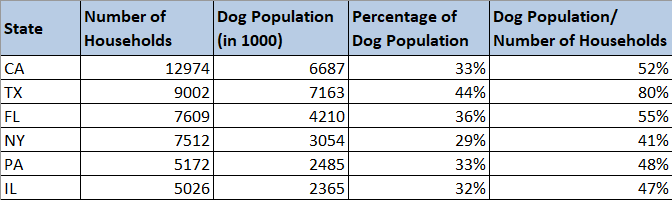
* Total Number of Households
* Dog Population
* Dog Households
* Mean % of Dog Population

These 4 factors will help us know:

* What is the population of Dogs with respect to total HH in that state?
* Which state is still left with potential to adopt Dogs as pet?

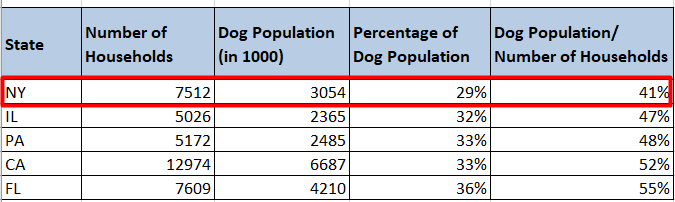
**Steps:**

1. Below table shows the top 6 states with the highest Number of Households. To start our analysis to select the state, we have taken the mean of Percentage of Dog Population.



2. Next, we as calculated earlier, the % of HH with Dogs at overall across the States is 36.51%

3. Now, we will be looking at the states with Percentage of Dog Population below the overall value of 36.51%. We get following states:



4. Now as seen from the above table state New York is the least with Percentage of Dog Population/Total HH i.e. 29, it is very less. This also depicts that 71% of dog population is yet to be adopted in New York.

5. If we look at the average % of dog population across the New York households, it again depicts the least (from the list of top states) with just 41%. This depicts the open market potential in New York to start “Petwale” venture.

**Finally, after the whole analysis and result, we can conclude that New York state is the state with great market potential and ideal to start because of the demand and trend towards adoption of dogs as pets all over the US.**