Data Visualization Project using Tableau Desktop

Executive Summary:

The scope of this project revolves around the analyzing the real time real data provided by a company called Canines Inc., a Delaware based corporation known for its product called **Dognition.com**. Dognition is basically a product designed for dog owners around the world to know their loving dogs even better. They provide a set of 20 games that can be played by the dog owners with their respective dogs and depending upon the pattern of completion of these games **Dognition** provides dog owners with some canine dimension like **Socialite, charmer, protodog, ace, expert, maverick** and so on.

Now the task bestowed to me as a candidate of the **Duke University** course for Tableau: "**Data Visualization and Communication with Tableau**" (offered in association in Coursera.org) was to figure the attributes of the dataset that may be effecting the attribute of importance in the dataset i.e. "**Total Test Completed**". This attribute **Total Test Completed** is very important metric for Dognition as because in practice it provides its Customers (i.e. Dog owners) with 20 free games for their dogs but all other games thereafter are paid and thus completion of 20 games is of significant importance for their revenue growth.

That is why the business question here is what are the various factors that leads to increase number of Total test completions for their customers? Whether it's the dog that effects the completion rate or it depends upon the attitude of the dog owners?

Data Visualizations associated:

All the data visualizations are being prepared to the answer the above questions using Big Data Visualization tool called **Tableau Desktop** and can be viewed using the below link of **Tableau Public server**:

https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT

or

Please see the README.md file for the Link to Tableau Desktop

Applicable tab names**: TTC vs_BG & BT,**
TTC vs. BT and DNA_Tested
TTC vs. BT, Dog_Fixed & DNA_Tested
Country vs. Dog_IDs
TTC vs. states of USA
TTC vs. Mean/Median ITI (minutes) (w/o outliers)

Acronym Notes: TTC – Total Test Completed
BG – Breed Group
BT – Breed Type

Few others questions answered as part of the project are as follows:

Question: What property is common to almost all the data points that had "Sign in Counts" of greater than 175 in the Dognition_aggregated_over_DogID data set?

Visualization Tab: **Q3**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT

or

Please see the README.md file for the Link to Tableau Desktop

Solution: The median weight of the dogs of the Breed Shih Tzu is 190 lbs. is highest amongst all available breeds of dog

Question Which Personality Dimension that has the highest average number ofcompleted tests in the Dognition_aggregated_over_DogID data set?

Visualization Tab: **Q4**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT
or
Please see the README.md file for the Link to Tableau Desktop

Solution: All of the personality dimensions have very similar completion rates.

Question: what personality type has the strongest representation (greatest number of records) in the sporting breed group?

Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT
or
Please see the README.md file for the Link to Tableau Desktop

Solution: Dogs with '**Socialite'** personality type has the strongest representation in the sporting breed group.

Question: What is the median number of tests dogs of different breed types complete?

Visualization Tab: **Q7**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT
or
Please see the README.md file for the Link to Tableau Desktop

Solution: All breed types complete a **median of 7** tests.

Question: In the Dognition_aggregated_over_DogID data set, how do the average number of tests dogs complete compare for fixed vs. not fixed dogs across different breed types?

Visualization Tab: **Q8**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT
or
Please see the README.md file for the Link to Tableau Desktop

Solution: Below are the inferences that can be drawn out of the above mentioned visualization

1. Fixed dogs complete more tests than non-fixed dogs in all breed types
2. The greatest difference between the average number of tests completed by fixed vs. non-fixed dogs occurs in the Mixed Breed/ I Don't Know breed category

Question What inference can be drawn from the average number of tests dogs complete when comparing DNA vs. not DNA-tested dogs who were fixed vs. not fixed across different breed types?

Visualization Tab: **Q9**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT
or
Please see the README.md file for the Link to Tableau Desktop

Solution Below are the inferences that can be drawn out of the above mentioned visualization

1. There was only one dog in the Popular Hybrid breed category who was DNA tested but not fixed
2. The Cross-Breed dogs that were DNA tested but NOT fixed were mostly Labrador Retriever-Golden Retriever Mixes

Question which state within the United States has the most Dognition customers?

Visualization Tab: **Q11**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT
or
Please see the README.md file for the Link to Tableau Desktop

Solution California is having the highest number of dognition customers followed by New York state which comes the second.

Question in which of the states in US did customers complete a median number of tests that was greater than 13?

Visualization Tab: **Q12**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT
or
Please see the README.md file for the Link to Tableau Desktop

Solution: All the below mentioned states are having Median number of Test completed greater than 13:

1. Maine (ME)
2. North Carolina (NC)
3. North Dakota (ND)
4. South Dakota (SD)
5. Wyoming (WY)
6. Arkansas (AK)
7. Iowa (IA)

Question what Inferences can be drawn from the relationship between Mean/Median Inter-Test-Intervals (ITIs) and Total Number of Test Completed?

Visualization Tab: **Q14**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT

or

Please see the README.md file for the Link to Tableau Desktop

Solution: Below are the inferences that can be drawn out of the above mentioned visualization

1. **There was a significant negative ($p < .05$) correlation between median ITIs and number of tests completed**
2. **There was a significant positive ($p < .05$) correlation between average ITIs and number of tests completed**

Question what is consistent about the relationship between breeding group and number of tests completed, regardless of whether you aggregate the variable representing the number of tests completed by the median or the average of the breeding group?

Visualization Tab: **Q5**
Link: https://public.tableau.com/profile/manish.bordoloi#!/vizhome/Book2_13281/TTCvs_BGBT

or

Please see the README.md file for the Link to Tableau Desktop

Solution: Toy dogs complete the least number of tests