

## **DAP V1.1 Data Release Notes**

This document provides details about the updates in the V1.1 DAP 2020 Curated Data, which was released on 5/10/2021, from the V1.0 DAP Data that was released on 3/29/2021.

### **Updated values in the dd\_activities\_xx variables**

#### **Background:**

The dd\_activities\_xx variables (e.g., dd\_activities\_companion\_animal, dd\_activities\_obedience, dd\_activities\_show etc.) are based on two questions from the HLES Dog Demographics survey instrument. Dog owners are first asked to select, from a list of 12 possible choices, all of the dog's roles and/or formal activities they are involved in. If more than one activity is selected, respondents are shown a follow-up question that asks them to indicate which activity is the dog's primary and secondary role. If only one activity is selected, the respondent is not shown this follow-up question. With these questions, the dd\_activities\_xx variables have values of 1, 2, and 3 that correspond to the labels of Primary, Secondary, and Additional activity (the value of 3 was formerly labelled "Other" in the V1.0 data), respectively.

#### **Problem in the V1.0 data:**

It should be assumed that if only one activity is selected (which was the case for 20,502 dogs), then it is the primary activity and should have been assigned a value of 1, which corresponds to the label of Primary. However, in the V1.0 data, if only one activity was selected, that activity instead received a value of 3.

#### **Correction in the V1.1 data:**

All of the instances where the dog only has one activity now has the value of 1 (rather than a value of 3 as in the V1.0 data), corresponding to the label of Primary.

### **Added records to the HLES Health Conditions table**

#### **Background:**

The HLES Health Conditions table is based on questions from the HLES Health Status survey instrument. In the Health Status form, a user is first asked if their dog has ever been diagnosed with a condition of type X (e.g., eye condition) in a Yes or No type question. These "condition type" gating questions are all required to be answered. For each condition type that the participant indicates "Yes", they are then presented with a set of possible specific conditions (e.g., blindness, cataracts, etc. for the eye condition type). However, these questions asking about specific conditions are NOT required to be answered.

#### **Problem in the V1.0 data:**

In the V1.0 data, records were created in the HLES Health Conditions table based on each specific condition (e.g., blindness, cataracts, etc.) indicated by the study participants. This ignored the records where a respondent answered the gating Yes or No question about the condition type (e.g., eye), but then did not select any specific conditions (e.g., blindness, cataracts, etc.). For example, 7916 dogs are reported to have a skin condition based on the gating yes/no question, but 39 of them did not show up in the HLES Health Conditions table because they did not answer the non-required questions about what specific skin condition their dog had.

This also led to a mismatch in numbers between the general health condition type variables (e.g., hs\_health\_conditions\_eye) that are in the HLES dog-owner table and the number of records in the HLES Health Conditions table that matched that health condition type..

#### Correction in the V1.1 data:

We have updated the HLES Health Conditions table to make sure that we are pulling all the respondents who answered the required gating Yes or No condition type question, regardless of whether or not they indicated a specific condition. This, in turn, results in rows in the HLES Health Conditions table that have values for only the following three columns: dog\_id, hs\_condition\_type, and hs\_condition\_is\_congenital. As a result, these rows would have missing values for the other 9 columns in the HLES Health Conditions table (i.e., hs\_condition, hs\_condition\_cause\_other\_description, hs\_condition\_other\_description, hs\_diagnosis\_month, hs\_diagnosis\_year, hs\_eye\_condition\_cause, hs\_follow\_up\_ongoing, hs\_neurological\_condition\_vestibular\_disease\_type, hs\_required\_surgery\_or\_hospitalization). This change ultimately ensures the number of dogs that have a broad health condition (i.e., skin, eye, ear etc.) is the same across the HLES Health Conditions table and the HLES dog-owner table.

#### **Removed one duplicated record in the data**

##### Problem in the V1.0 data:

The V1.0 data had a total of 27,542 dogs. However, it was discovered that one record (dog\_id = 86510) was a duplicate DAP Pack record caused by an owner misunderstanding. A DAP team member deleted the record once they identified the issue.

##### Correction in the V1.1 data:

Given the removal of the duplicated record, we now have a total of 27,541 dogs that obtained DAP Pack status on or before December 31st 2020.

#### **Updated value labels for a few variables**

In the V1.0 data, the labels for three variables - df\_feedings\_per\_day, de\_neighborhood\_has\_sidewalks, and dd\_activities\_xx - were either incorrect or ambiguous. Below details the problems in the V1.0 data and the correction in the V1.1 data.

##### *df\_feedings\_per\_day*

- In the V1.0 data, the value labels were incorrectly given the following values: "Never", "Seldom", "Sometimes", "Usually", "Always".
- In the V1.1 data, we updated the value labels to "Once", "Twice", "Three or More", and "Free feed" that correspond to 1, 2, 3, and 4 values, respectively

##### *de\_neighborhood\_has\_sidewalks*

- In the V1.0 data, the value labels had the following labels which were unclear: "Yes", "Yes", and "No".
- In the V1.1 data, we updated the value labels to "Yes, on all the streets", "Yes, on some streets", and "No" that correspond to 1, 2, and 0 values, respectively.

##### *dd\_activities\_xx*

- In the V1.0 data, if multiple activities (e.g., companion animal, show etc.) were selected for a dog, but the owner did not identify the activity as Primary or Secondary, then the activity was given a value of 3, which was originally assigned the value of "Other." However, this may cause confusion with the "Other" activities option (dd\_activities\_other) that owners could select.
- In the V1.1 data, we updated the value labels to "Primary", "Secondary", and "Additional activity" that correspond to 1, 2, and 3 values, respectively.

