Rider Happiness Benchmarking Data

PRE-ANALYSIS

The survey responses provided to you here have already gone through steps to prepare them for analysis.

Translation

The responses to multiple-choice questions have been translated to English. For example, "Non" has been translated to "No." While questions in the dataset shared with you are all in English, the actual question shown to the respondent may have been in English, Spanish, or French.

Deduplication

In some rare cases, a respondent can reuse the survey link to answer the survey twice. We are able to detect those instances and only retain the most recent survey response.

Assigned agency

Each respondent is assigned an agency based on their answer to the following question: "What is the public transit agency you ride the most? Throughout this survey, questions about 'my agency' will refer to the agency you write here."

In case the question was not answered, or the answer is ambiguous, Transit's team uses app data, such as the area the user most frequently opens the app, or transit lines they view most often in the app, to determine the respondent's transit agency.

DATASET CONTENT

RAW RESPONSES

The raw responses provided are those assigned to your agency. Data is provided in CSV format, with the following columns:

- *submitted_at*: Date and time at which the response to the survey was submitted, in ISO 8601 format

- *language*: Language of the survey to which the user responded. This affects the language of questions and choices presented to the respondent, which they have all been translated for simpler analysis.

- *Assigned Feed*: The GTFS feed to which the response relates. The column is only present if your agency is split into, for example, separate GTFS feeds for bus and rail.

- Other columns: All other columns are simply the questions from the survey. The content of those columns is simply the respondent's answer to that question. In the case of questions for which multiple choices are allowed, the responses are presented in the form of a JSON array. For example: ['Public transit', 'Walking']. These can be analysed in Excel with a formula such as =COUNTIF(range, "*Walking*")

UNIVARIATE STATISTICS FOR YOUR AGENCY

A CSV file with univariate statistics about your agency is provided to help you save time. The same statistics can be calculated from the raw responses. If your agency has multiple feeds (such as one for bus and one for rail), this file will contain combined statistics for responses across all of your agency's feeds.

UNIVARIATE STATISTICS FOR EACH OF YOUR FEEDS

In the "Assigned Feeds" directory, you will find CSV files containing the univariate statistics for each of your feeds separately. If your agency uses a single feed, this directory does not exist.

UNIVARIATE STATISTICS FOR EACH AGENCY

In the "Agencies" directory, you will find a CSV file for each agency for which we had a sufficient number of responses to perform univariate statistics. The structure of those files is the same as the one about your agency, and allows for comparison with other transit agencies.

transit

UNIVARIATE STATISTICS FOR EACH COUNTRY

In the "Countries" directory, you will find the CSV files containing the univariate statistics for all survey respondents within the United States and Canada.