Data Source Information

- **Collection Source**: Google Play Store
- **Collection Date Range**: September 19, 2023, to September 20, 2023
- **Platform URL**: https://play.google.com/store/apps

Data Collection Methodology

- **Method**: Automated scraping followed by manual verification
- **Selection Criteria**: Reviews with any length of words, covering various product categories

Anonymization Techniques

- **Personal Data Removal**: Username, Userimage (URL), Review ID, and any identifiable personal information well be replaced with generic placeholders or dropped.
- **Geographical Data**: Specific locations are generalized (e.g., city level)

Data Quality Information

- **Accuracy**: Data verified for completeness with a 2% margin of error
- **Limitations**: Reviews are not just in English and may contain subjective opinions

Usage Guidelines

- **Intended Use**: Academic research on consumer behavior
- **Restrictions**: Not to be used for commercial profiling or targeted marketing

Version Control Information

- **Initial Dataset Release**: Version 1.0, November 15, 2023
- **Update Frequency**: Bi-annually if needed (expected updates in January and July each year)

Metadata Schema

Review ID:

- **Description**: Unique identifier assigned to each individual review in the dataset.
- **Format**: Alphanumeric string that may contain hyphens, (**Note.** example is not provided to ensure user privacy.)

App Domain:

- **Description**: The general category or domain of the application being reviewed.
- **Format**: Text, e.g., Food Delivery.

App Name:

- **Description**: The specific name of the application that is being reviewed.
- **Format**: Text, e.g., talabat.

User Name:

- **Description**: The name of the user who wrote the review, as displayed on the platform.
- **Format**: Text, possibly including both first and last names, (**Note.** example is not provided to ensure user privacy.)

User Image:

- **Description**: URL link to the image of the user who wrote the review.
- **Format**: URL, (**Note.** example is not provided to ensure user privacy.)

Content:

- **Description**: The actual text content of the review submitted by the user.
- **Format**: Free text, variable length.

Score:

- **Description**: The numerical rating given by the user as part of their review.
- **Format**: Integer, typically on a scale (e.g., 1-5).

Thumbs Up Count:

- **Description**: The number of positive ratings or 'thumbs up' the review has received from other users.
- **Format**: Integer, e.g., 19.

Review Created Version:

- **Description**: The version of the app that was current when the review was written.
- **Format**: Alphanumeric, indicating major and minor version numbers, e.g., 10.13.0.

Review Date (at):

- **Description**: The date and time when the review was posted.
- **Format**: Date and time, e.g., 23-10-2023 12:41.

Reply Content:

- **Description**: The text content of any official reply to the user review.
- **Format**: Free text, variable length.

Replied At:

- **Description**: The date and time when the official reply was posted.
- **Format**: Date and time, e.g., 24-10-2023 10:13.

App Version:

- **Description**: The current version of the app at the time the dataset was compiled.
- **Format**: Alphanumeric, indicating major and minor version numbers, e.g., 10.13.0.