Dataset Description for:

Brambilla, Marco; Javadian Sabet, Alireza, 2020, "Social Media Data about the Big Four Fashion Weeks - Fall-Winter 2018", https://doi.org/10.7910/DVN/8BNXES, Harvard Dataverse.

The dataset that we provide is composed of two Comma-Separated Values (CSV) files:

Posts and User Profiles authoring the posts on the social media platform named Instagram (www.instagram.com) covering the **Big Four Fashion Weeks, for the Fall-Winter seasons 2018**.

The provided dataset comprises **905,726** posts and **171,078** correspondent unique user profiles.

Details on their attributes are provided in the following lists:

- **Posts Dataset.** Columns and Descriptions:
  - **Post's PK**: ID of the post.
  - **User's PK**: Anonymized ID of the post author.
  - **Likes Count**: Total number of likes.
  - **Comments** Count: Total number of comments.
  - **Time_In**: 1 if the post published in the target event period.
  - **Time_Other**: 1 if the post published in other events periods.
  - **Time_None**: 1 if the post published in none of the events period.
  - **Caption Length**: Number of characters in the post caption.
  - **Hashtags Count**: Number of used hashtags in the post caption.
  - **Event_Milan**: 1 if the post is about Milan FW.
  - **Event_Paris**: 1 if the post is about Paris FW.
  - **Event_London**: 1 if the post is about London FW.
  - **Event_NewYork**: 1 if the post is about NY FW.
  - **Brand_X**: For each of the following 21 brands we provided a Boolean column which each of them can take 1 if the post caption, contains their relevant hashtag.
    - *Gucci, Chanel, Dior, Fendi, Burberry, D&G, Balenciaga, Versace, Prada, Louisvuitton, Tommy, Nike, Valentino, Adidas, Zara, CalvinKlein, VictoriaSecret, Miumiu, Bvlgari, H&M', Armani.*

While, due to the copyright and privacy regulations by Instagram and posts authors, we only publish the attributes that we prepared, still it is possible to access the posts (if it is publicly available at the time of request) through the post's identifier (PK).

- **Users Dataset.** Columns and Descriptions:
  - **User's PK**: Anonymized ID of the user.
  - **Event Posts Count**: Number of posts by the user in the dataset.
  - **Event Likes_X**: Highest, Sum, Average, Median of likes of user's posts regarding the event.
  - **Event Comments_X**: Highest, Sum, Average, Median of Comments of user's posts regarding the event.
  - **Event Geo-tagged Percent**: The percent of the user Geo posts.