

Recommended content

1. Introductory information

Title: Relevant Data on the Impact of Visual Elements on Click Behavior When Bilibili (a Chinese social media platform) Users Proactively Search for Popular Science Videos.

The file "data" is an Excel spreadsheet containing statistical variables from 41,224 original video records used in the research. It includes variables such as the number of followers, image links, image quality, brightness, color, as well as repost count, danmaku count, and view count.

The "py" file corresponds to its name and can be opened with Jupyter or Spyder for code execution and operation.

2. Methodological information

Methodology: The data was collected using Python to scrape Bilibili's popular science channel data from 2010 to 2022. Python version 3.9 was used for code execution and processing.

Specific Data Information:

UP 主 (Blogger): Name of the video uploader.

UP 主主页 (Blogger Homepage): Link to view the blogger's fan information.

标题 (Title): Title of the video.

视频时长 (Video Duration): Duration of the video, later converted to seconds using Python software.

图片链接 (Image Links): Links to images, to be used for brightness and color parameter recognition using Python.

视频链接 (Video Links): Links to the videos.

标题类型 (Title Type): Statistics on the types of titles, including questions, statements, exclamations, etc.

时间秒数 (Time in Seconds): Duration of the video in seconds.

时间分类 (Time Category): Categorization of video duration into short, medium, and long videos.

标题长度 (Title Length): The number of Chinese characters in the title, used for classification.

标题长度分类 (Title Length Category): Classification of title length into short, medium, and long titles.

播放量 (Views): Statistics of view count obtained from scraping.

实际播放量 (Actual Views): View count minus reposts, excluding non-user-initiated video views.

点赞量 (Likes): Number of likes as a variable.

粉丝量 (Followers): Number of followers as a variable.

转发量 (Reposts): Repost count.

弹幕 (Bullet screen comments): Bullet screen comments

缩略图是否有符合主题的视觉元素或文字 (Thumbnail Relevance): Categorization of thumbnails as having images only, text only, both images and text, etc.

色彩 (Color Richness Index): Parameter related to color richness.

亮度 (Brightness Index): Parameter related to brightness.

图像质量 (Image Quality Parameters): Parameter related to image quality.

3. Author Information

Zhang Weifeng, School of Journalism and Communication, Chongqing Normal University.

Research Focus: Science Communication.

Corresponding Author: He Yi

Contact Information: 15918728137@163.com

4. General Introduction

This data includes an EXCEL table with variables related to user-initiated searches for popular science videos, as well as Python code for data scraping and cleaning. It can provide reference for future researchers and serves as supplementary data for publications and papers. It also allows other researchers to use the data in their own work and validate the methods used in their papers. The dataset comprises all video-related information from Bilibili's popular science channel between 2010 and 2022.

Purpose of the Test Activity: The purpose is to measure the factors influencing users' click behavior when actively searching for popular science videos, in order to improve the design and innovation of visual appeal for future popular science video creators.

Test Equipment: Macbook 2020 Pro

Description of Data in this Dataset: The data in this dataset has been organized by samples and follows a specific naming convention.