

JOSÉ CASADIEGO

ESTUDIO DE LAS TENDENCIAS EN EL DESARROLLO DE SOFTWARE EN GITHUB

CODERHOUSE

TEMÁTICA

El conjunto de datos de GitHub contiene información sobre los desarrolladores y los repositorios. El conjunto de datos se puede utilizar para medir el número de desarrolladores, el número de repositorios, las relaciones entre desarrolladores y repositorios, y las tendencias en el desarrollo de software. El conjunto de datos es relevante para la investigación y la práctica del desarrollo de software, y se puede utilizar para identificar desarrolladores y repositorios prometedores.



HIPÓTESIS

Con el conjunto de datos se plantea identificar a los desarrolladores mas activos, los repositorios mas populares y las tendencias en el lenguaje de programación utilizado en los repositorios, con lo que se puede identificar a los desarrolladores y repositorios prometedores que puedan potencialmente contribuir significativamente al desarrollo de software.

ANALIZAR LOS DATOS

Los resultados del análisis se utilizarán para probar la hipótesis y para identificar desarrolladores y repositorios prometedores.

Los resultados del análisis tendrán implicaciones importantes para la investigación y la práctica del desarrollo de software. Los resultados pueden utilizarse para ayudar a los investigadores a comprender mejor las tendencias en el desarrollo de software y para ayudar a los desarrolladores a identificar oportunidades para contribuir al desarrollo de software.



DATASET

1

INFORMACIÓN

El conjunto de datos contiene información sobre 990 desarrolladores y 2954 repositorios.

La DB cuenta con tres tablas, developers, repositories y edgelist.

2

DESCRIPCIÓN

developers.csv contiene información sobre desarrolladores, incluyendo su ID de desarrollador, entre otros.

repositories.csv contiene información sobre repositorios, incluyendo su ID de repositorio, nombre de usuario del propietario, nombre de repositorio, entre otros.

3

RELACIÓN

edgelist.csv contiene información sobre las relaciones entre developers y repositories, incluyendo si el repositorio está bifurcado, si el desarrollador es un colaborador principal, el ID del repositorio y el ID del desarrollador.