

Estudio de las tendencias en el desarrollo de software en GitHub.
José Casadiego.

- **Objetivo:**
 - Identificar las tendencias de los lenguajes usados en el desarrollo de software en GitHub.
- **Alcance:**
 - El proyecto analizará un conjunto de datos de GitHub para identificar las tendencias en el uso de lenguajes de programación.
- **Usuario final:**
 - El usuario final del proyecto es cualquier persona que esté interesada en aprender sobre las tendencias en el desarrollo de software en GitHub. Esto podría incluir desarrolladores, gerentes de proyectos, investigadores y otros profesionales que trabajan en la industria de la tecnología.
- **Nivel de aplicación:**
 - El nivel de aplicación del análisis es estratégico. El objetivo del proyecto es identificar las tendencias en el desarrollo de software en GitHub que puedan ayudar a los usuarios finales a tomar mejores decisiones sobre sus propios proyectos de desarrollo de software.
- **Tabla developers:**
 - Descripción: Esta tabla almacena información sobre los desarrolladores.
 - Columnas:
 - dev_id: El ID del desarrollador.
 - username: El nombre de usuario del desarrollador.
 - name: El nombre del desarrollador.
 - bio: La biografía del desarrollador.
 - company: La empresa del desarrollador.
 - location: La ubicación del desarrollador.
 - public_repositories: El número de repositorios públicos del desarrollador.
 - followers: El número de seguidores del desarrollador.
 - following: El número de desarrolladores a los que sigue el desarrollador.
 - list_of_pinned_repos: La lista de repositorios fijados del desarrollador.
 - starred_repo_count: El número de repositorios estrellados del desarrollador.
 - yearly_contributions: El número de contribuciones del desarrollador en el último año.
 - repos_name_list: La lista de nombres de los repositorios públicos del desarrollador.
 - starred_repos_name_list: La lista de nombres de los repositorios estrellados del desarrollador.
 - followers_name: La lista de nombres de usuario de los seguidores del desarrollador.
 - following_names: La lista de nombres de usuario de los desarrolladores a los que sigue el desarrollador.

- achievements: La lista de todos los nombres de insignias de logros del desarrollador.
 - status: El contenido de la sección de estado del desarrollador.
- Llaves primarias: dev_id
- Llaves foráneas: N/A
- Relaciones:
 - La tabla developers está relacionada con la tabla repositories a través de la columna dev_id. Es una relación de muchos a muchos
 - La tabla developers está relacionada con la tabla edgelist a través de la columna dev_id. Es una relación de uno a muchos.
- **Tabla repositories:**
 - Descripción: Esta tabla almacena información sobre los repositorios.
 - Columnas:
 - repo_id: El ID del repositorio.
 - owner_username: El nombre de usuario del propietario del repositorio.
 - repo_name: El nombre del repositorio.
 - description: La descripción del repositorio.
 - created_at: La fecha y hora de creación del repositorio.
 - pushed_at: La fecha y hora del último commit en el repositorio.
 - size: El tamaño del repositorio en KB.
 - stargazers_count: El número de estrellas del repositorio.
 - has_projects: True si el repositorio tiene proyectos, False en caso contrario.
 - has_wiki: True si el repositorio tiene un wiki, False en caso contrario.
 - has_pages: True si el repositorio tiene páginas, False en caso contrario.
 - forks_count: El número de forks del repositorio.
 - open_issues_count: El número de problemas abiertos del repositorio.
 - license: El nombre de la licencia del repositorio.
 - is_template: True si el repositorio es una plantilla, False en caso contrario.
 - topics: La lista de temas del repositorio.
 - watching: El número de personas que están observando el repositorio.
 - contributors_count: El número de colaboradores del repositorio.
 - commits_count: El número de commits del repositorio.
 - languages: La lista de lenguajes utilizados en el repositorio.
 - readme: El contenido del archivo readme.
 - Llaves primarias: repo_id
 - Llaves foráneas: N/A
 - Relaciones:
 - La tabla repositories está relacionada con la tabla developers a través de la columna owner_username. Es una relación de muchos a muchos
 - La tabla repositories está relacionada con la tabla edgelist a través de la columna repo_id. Es una relación de uno a muchos.

- Tabla edgelist:
 - Descripción: Esta tabla almacena información sobre las relaciones entre desarrolladores y repositorios.
 - Columnas:
 - isForked: Un valor booleano que indica si el desarrollador ha bifurcado el repositorio.
 - isTopContributor: Un valor booleano que indica si el desarrollador es un colaborador principal del repositorio.
 - dev_id: El ID del desarrollador.
 - repo_id: El ID del repositorio.
 - Llaves foráneas: dev_id (tabla developers), repo_id (tabla repositories)
 - Relaciones:
 - La tabla edgelist está relacionada con la tabla developers a través de la columna dev_id. Es una relación de uno a muchos.
 - La tabla edgelist está relacionada con la tabla repositories a través de la columna repo_id. Es una relación de uno a muchos.

Data	SQL Definition	Nullable	Unique
dev_id	int	No	Sí
username	varchar(255)	No	Sí
Name	varchar(255)	Sí	No
Bio	varchar(255)	Sí	No
Company	varchar(255)	Sí	No
Location	varchar(255)	Sí	No
Public Repositories	int	Sí	No
Followers	int	Sí	No
Following	int	Sí	No
listOfPinnedRepos	varchar(255)	Sí	No
starredRepoCount	int	Sí	No
yearly_contributions	int	Sí	No
repos_name_list	varchar(255)	Sí	No
starred_repos_nameList	varchar(255)	Sí	No
followers_name	varchar(255)	Sí	No
following_names	varchar(255)	Sí	No
achievements	varchar(255)	Sí	No
status	varchar(255)	Sí	No

Data	SQL Definition	Nullable	Unique
repo_id	int	No	Sí
owner_username	varchar(255)	No	No
repo_name	varchar(255)	No	Sí
description	varchar(255)	Sí	No
created_at	datetime	No	No
pushed_at	datetime	No	No
size	int	Sí	No
stargazers_count	int	Sí	No
has_projects	boolean	Sí	No
has_wiki	boolean	Sí	No
has_pages	boolean	Sí	No
forks_count	int	Sí	No
open_issues_count	int	Sí	No
license	varchar(255)	Sí	No
is_template	boolean	Sí	No
topics	varchar(255)	Sí	No
watching	int	Sí	No
contributors_count	int	Sí	No
commits_count	int	Sí	No
languages	varchar(255)	Sí	No
readme	varchar(255)	Sí	No

Data	SQL Definition	Nullable	Unique
isForked	boolean	No	No
isTopContributor	boolean	No	No
dev_id	int	No	No
repo_id	int	No	No