Relatorio

Augusto Noronha Leite;
David Leong Ho;
Pedro Maximo;

Relatório Técnico: Análise dos repositórios mais populares de acordo com a quantidade de "stars" recebidas

Objetivo: O objetivo deste experimento é encontrar e analisar padrões e características dos repositórios mais populares do GitHub, baseando-se na quantidade de estrelas (`stars`) que eles receberam. Os principais parâmetros analisados incluem **linguagem principal**, **data do último commit**, **tamanho**, **branch principal**, **tópicos**, **forks** e **watchers**.

Linguagem de programação: O Pyhton3 foi escolhido para desenvolver o experimento.

Dependências: Para executar este experimento, as seguintes dependências foram utilizadas:

- **requests** (para fazer requisições HTTP à API do GitHub)
- **csv** (para exportar os dados coletados para um arquivo CSV)
- **time** (para controlar a taxa de requisições e evitar o limite da API)
- **urllib3** (para gerenciar avisos de segurança de requisições)

API utilizada: Foi utilizada a GitHub API para coletar os dados necessários dos repositórios. Mais detalhes sobre a API podem ser encontrados na documentação do GitHub REST API.

- https://docs.github.com/pt/rest?apiVersion=2022-11-28

Metodologia:

- **Coleta de dados:** Foi utilizada a GitHub API (`get_top_starred_repos_paginated`) para obter informações detalhadas dos repositórios. A coleta é baseada na quantidade de estrelas (`q=stars:>0&sort=stars&order=desc`), focando nos repositórios mais populares.

- **Filtragem e Paginação:** A coleta de dados envolve a utilização da paginação da API do GitHub (`per_page=100&page={page}`) para obter um grande número de repositórios (até 1000). O processo lida com múltiplas requisições sequenciais.

- **Extração de Detalhes:** Para cada repositório encontrado, uma requisição de detalhes separada é feita (`get_repo_details`), coletando dados como linguagem principal, licença, tamanho, data do último commit e tópicos.

- **Salvamento dos Dados:** Os dados coletados são estruturados e salvos em um arquivo CSV (`repos_info.csv`) usando a biblioteca `csv` .

Resultados:

Considerando a quantidade de estrelas, os seguintes repositórios foram ranqueados como os 10 mais populares: 1) freeCodeCamp; 2) build-your-own-x; 3) sindresorhus/awesome; 4) free-programming-books; 5) public-apis; 6) developer-roadmap; 7) coding-interview-university; 8) system-design-primer; 9) 996.ICU; 10) awesome-python-vinta.

Conclusão:

Este experimento demonstrou a viabilidade de coletar e analisar dados de repositórios populares no GitHub. A metodologia utilizada permitiu identificar e ranquear os principais repositórios de código aberto com base em critérios objetivos, sendo a popularidade determinada pelo número de **estrelas** e complementada pela análise de características como **linguagem principal** e **licença**.

Repository	Owner	URL	Stars	Forks	Watchers	Last Commit Date	Main Language	License	Size (KB)	Main Branch	Topics
freeCodeCamp	freeCodeCamp	https://github.co	426319	41284	426319	2025-08-23T18:19:20Z	TypeScript	BSD 3-Clause "?	511220	main	careers, certification, community, curriculum, d3, education, freecodecamp, hacktoberfest,
build-your-own-x	codecrafters-io	https://github.co	413538	38785	413538	2025-07-29T21:02:41Z	Markdown	Sem licença	1079	master	awesome-list, free, programming, tutorial-code, tutorial-exercises, tutorials
awesome	sindresorhus	https://github.co	394472	31323	394472	2025-07-18T18:37:33Z		Creative Commo	1500	main	awesome, awesome-list, lists, resources, unicorns
free-programming-books	EbookFoundation	https://github.co	365952	63946	365952	2025-08-23T12:50:50Z	Python	Creative Commo	19603	main	books, education, hacktoberfest, list, resource
public-apis	public-apis	https://github.co	362688	38063	362688	2025-05-20T15:56:34Z	Python	MIT License	4771	master	api, apis, dataset, development, free, list, lists, open-source, public, public-api, public-apis,
developer-roadmap	kamranahmedse	https://github.co	334431	42626	334431	2025-08-23T00:12:35Z	TypeScript	Other	309093	master	angular-roadmap, backend-roadmap, blockchain-roadmap, computer-science, dba-roadma
coding-interview-university	jwasham	https://github.co	324752	79921	324752	2024-12-05T23:47:30Z		Creative Commo	22600	main	algorithm, algorithms, coding-interview, coding-interviews, computer-science, data-structur
system-design-primer	donnemartin	https://github.co	317342	51936	317342	2025-05-21T11:13:33Z	Python	Other	11239	master	design, design-patterns, design-system, development, interview, interview-practice, intervi
996.ICU	996icu	https://github.co	274140	21069	274140	2025-08-22T06:01:29Z		Other	189351	master	
awesome-python	vinta	https://github.co	256467	26185	256467	2025-07-17T16:35:51Z	Python	Other	6839	master	awesome, collections, python, python-framework, python-library, python-resources
awesome-selfhosted	awesome-selfhosted	https://github.co	243422	11272	243422	2025-08-22T08:51:14Z		Other	16591	master	awesome, awesome-list, cloud, free-software, hosting, privacy, self-hosted, selfhosted
project-based-learning	practical-tutorials	https://github.co	240911	31502	240911	2024-08-15T05:33:54Z		MIT License	438	master	beginner-project, cpp, golang, javascript, project, python, tutorial, webdevelopment
react	facebook	https://github.co	238363	49187	238363	2025-08-23T00:37:40Z	JavaScript	MIT License	1103856	main	declarative, frontend, javascript, library, react, ui
vue	vuejs	https://github.co	209294	33742	209294	2024-10-10T07:24:15Z	TypeScript	MIT License	32152	main	framework, frontend, javascript, vue
Python	TheAlgorithms	https://github.co	204318	47369	204318	2025-08-24T12:08:52Z	Python	MIT License	15444	master	algorithm, algorithm-competitions, algorithms-implemented, algos, community-driven, educ
linux	torvalds	https://github.co	200579	57422	200579	2025-08-23T15:30:56Z	C	Other	5657637	master	
javascript-algorithms	trekhleb	https://github.co	193069	30857	193069	2025-02-12T16:19:59Z	JavaScript	MIT License	13564	master	algorithm, algorithms, computer-science, data-structures, interview, interview-preparation,
tensorflow	tensorflow	https://github.co	191301	74790	191301	2025-08-24T09:40:57Z	C++	Apache License	1193563	master	deep-learning, deep-neural-networks, distributed, machine-learning, ml, neural-network, py
computer-science	ossu	https://github.co	190901	23837	190901	2025-08-23T18:48:52Z	HTML	MIT License	1767	master	awesome-list, computer-science, courses, curriculum
the-book-of-secret-knowledge	trimstray	https://github.co	183471	11410	183471	2024-11-19T14:00:38Z		MIT License	1748	master	awesome, awesome-list, bsd, cheatsheets, devops, guidelines, hacking, hacks, howtos, lin
You-Dont-Know-JS	getify	https://github.co	183017	33658	183017	2025-05-20T14:22:36Z		Other	15060	2nd-ed	async, book, book-series, closures, education, es2015, es6, javascript, learn-to-code, prog
CS-Notes	CyC2018	https://github.co	181546	51165	181546	2024-08-21T09:40:10Z		Sem licença	116179	master	algorithm, computer-science, cpp, interview, java, leetcode, python, system-design
ohmyzsh	ohmyzsh	https://github.co	180899	26179	180899	2025-08-19T10:46:13Z	Shell	MIT License	12923	master	cli, cli-app, hacktoberfest, oh-my-zsh, oh-my-zsh-plugin, oh-my-zsh-theme, ohmyzsh, plug
AutoGPT	Significant-Gravitas	https://github.co	178032	45945	178032	2025-08-24T07:03:50Z	Python	Other	275748	master	ai, artificial-intelligence, autonomous-agents, gpt-4, llama-api, openai, python
vscode	microsoft	https://github.co	176037	34560	176037	2025-08-24T05:25:56Z	TypeScript	MIT License	1054811	main	editor, electron, microsoft, typescript, visual-studio-code
bootstrap	twbs	https://github.co	172898	79125	172898	2025-08-22T22:16:00Z	MDX	MIT License	258879	main	bootstrap, css, css-framework, html, javascript, sass, scss
flutter	flutter	https://github.co	172065	29071	172065	2025-08-24T03:35:22Z	Dart	BSD 3-Clause "?	403759	master	android, app-framework, cross-platform, dart, dart-platform, desktop, flutter, flutter-package
Python-100-Days	jackfrued	https://github.co	171204	54403	171204	2025-03-28T10:29:23Z	Jupyter Noteboo	Sem licença	396875	master	
gitignore	github	https://github.co	168849	83010	168849	2025-08-22T23:36:39Z		Creative Commo	2941	main	git, gitignore
the-art-of-command-line	jlevy	https://github.co	157869	14746	157869	2024-06-25T18:13:44Z		Sem licença	2772	master	bash, documentation, linux, macos, unix, windows
stable-diffusion-webui	AUTOMATIC1111	https://github.co	155889	28911	155889	2025-05-03T06:17:03Z	Python	GNU Affero Gen	36427	master	ai, ai-art, deep-learning, diffusion, gradio, image-generation, image2image, img2img, pytor
JavaGuide	Snailclimb	https://github.co	151391	45979	151391	2025-08-20T02:17:55Z	Java	Apache License	177406	main	algorithms, interview, java, jvm, mysql, redis, spring, system-design, zookeeper
ollama	ollama	https://github.co	150818	12896	150818	2025-08-24T01:44:38Z	Go	MIT License	49022	main	deepseek, gemma, gemma3, gemma3n, go, golang, gpt-oss, llama, llama2, llama3, llava,