Nicholas Keeley (ngk3pf@virginia.edu)
DS 5001
04 May 2022

Manifest Document

Provenance

The source files used for this project consist of 4,215 earnings call transcripts publicly available on *The Motley Fool's* website, a financial publication that provides free and premium investing-related services. *The Motley Fool's* earnings call page can be accessed here: https://www.fool.com/earnings-call-transcripts/ .

Location

This project's source files are stored in a .zip file titled "final_project_SOURCE_FILES.zip", which can be found at the following shared Box link: https://virginia.box.com/s/kp22cn73zmhp90non1btxrcd2fjk4f2p . A separate file is also included, titled "final_project_URLS.csv", which contains the URL for each source used.

Description

Earnings calls are periodic reports by a company's chief officials (i.e. CEO, CFO, COO, etc.) to the company's shareholders. These reports are conducted on a quarterly basis, and typically consist of two sections: a prepared statement, as well as a question/answer session in which external financial analysts are allowed to probe the company's leadership for more information on the company's anticipated performance. These calls are conducted virtually (e.g. a public teleconference), but certain websites like *The Motley Fool* have transcribed these calls into text format. For this project, ~35,000 earnings calls between 2017 and 2022 were scraped from *The Motley Fool*, but only 4,215 between 2018 and 2022 were chosen for analysis. These calls were chosen because they belonged to companies that had at least one year's worth of earnings call transcripts available, and published earnings calls in 2022. This allowed for sentiment analysis over time. In addition to the text of each company's earnings calls, this project scraped each company's stock ticker, company name, and the year/quarter of the call. Python's "yfinance" package was used to supplement this metadata with each company's designated industry sector.

Format

The shared Box link listed above leads to a .zip file, which contains each investigated company's earnings call transcript text in CSV format. These texts were scraped from HTML-format "documents" embedded within *The Motley Fool's* website.