

Data Report

Title	Stock Values and Earnings Call Transcripts: a Sentiment Analysis Dataset
-------	--

Organization

Author	Dexter Roozen, Francesco Lelli
Affiliation	F. Lelli : Prof. at Tilburg University, Department of Management D. Roozen, Master Student at Tilburg University, Tilburg School of Economics and Management
Email	f.elli@tilburguniversity.edu
Distributor	DataverseNL
Ethical Clearance	Research has the approval of IRB of TiSEM " (IRB EXE 2021-002) as indicated in the submission form version 1, which was submitted on February 23rd, 2021. The execution of this research is in accordance with the Netherlands Code of Conduct for Research Integrity (VSNU, 2018)
Preregistration	N.A.

Research context

Description	Stock values and earnings call transcripts from 10 NASDAQ companies and the NASDAQ sector index to perform sentiment analysis with.
Kind of data	Text data, stock prices
Publication	Roozen, D.; Lelli, F. Stock Values and Earnings Call Transcripts: a Dataset Suitable for Sentiment Analysis. Preprints 2021, 2021020424 (doi: 10.20944/preprints202102.0424.v1).
Keywords	Tilburg University, Economics, Stock Values, Earnings Call Transcripts, NLP, Natural language processing

Data production

Producer	Tilburg University
Production date	02-10-2020 – 14-1-2020
Method	Stock values generated through the use of Yahoo Finance Earnings Call transcripts generated through the use of Thomson Reuters Eikon
Universe	NASDAQ companies
Data sources	N.A.
Country / Nation	USA

Rights and restrictions

Restrictions	Public
License	CC0
Rights	No restrictions
Data retention period	10 years

Read-me

Data files	Raw, unaltered documents: <ul style="list-style-type: none">- 10x Microsoft Excel Comma Separated Values File from NASDAQ companies- 1x Microsoft Excel Comma Separated Values File from NASDAQ sector index- 188x Text Document from NASDAQ companies
Supplemental material	<ul style="list-style-type: none">- Data_Publication_Package.pdf (this file)- Detailed description of the data has been release in the following preprint: doi: 10.20944/preprints202102.0424.v1
Structure data package	<p>The data publication package is in the root folder that contains 2 subfolder as well. The folder “Stock Values and Sector Index” contains a set of .csv files related to prices during the period of consideration. The name of the file is the tick symbol at the stock exchange.</p> <p>The folder “Transcripts” contains a set of folders, one for each company taken into account. Each subfolder contains the transcripts related to their earning call. The name of the file contains the date of origin as well.</p>