

Variable name	Details
LANCS_ID	Firm-specific identifier (name changes accounted for)
REP_ID	Report-specific identifier
FirmName	Firm name at report publication date
Pages	Number of pages in annual report_raw extraction
ReportNumberPages_2	Number of pages in annual report_refined
Wordcount_all	Wordcount for entire annual report (includes text from tables, graphics, etc. in addition to running text)
Pagecountsum_Front_3	Number of pages in narrative component of annual report (see notes for further details)
wordcountsum_Front_3	Wordcount for narrative component of annual report (see notes for further details)
nsections_Front_3	Number of sections in narrative component of annual report (see notes for further details)
fleschmu_Front_3	Readability (Flesch) for narrative component of annual report (see notes for further details)
fogmu_Front_3	Readability (Fog) for narrative component of annual report (see notes for further details)
causal_Front_3	Causal reasoning keywords / wordcount for narrative component of annual report (see notes for further details)
forwardlooking_Front_3	Forward-looking keywords / wordcount for narrative component of annual report (see notes for further details)
net_tone_Front_3	Net positive keywords / (positive + negative) keywords for narrative component of annual report (see notes for further details)
uncertainty_Front_3	Uncertainty keywords / wordcount for narrative component of annual report (see notes for further details)
Pagecountsum_Rear_3	Number of pages in financial statement component of annual report (see notes for further details)
Wordcountsum_Rear_3	Wordcount for financial statement component of annual report (see notes for further details)
nsections_Rear_3	Number of sections in financial statement component of annual report (see notes for further details)
ChairFlag	Indicator variable for presence of chairman's statement: 1 = chairman's statement extracted; 0 = no chairman's statement extracted (see notes for further details)
Pagecountsum_Chair	Number of pages in chairman's statement (see notes for further details)
wordcountsum_Chair	Wordcount for chairman's statement (see notes for further details)
nsections_Chair	Number of sections in chairman's statement (see notes for further details)
fleschmu_Chair	Readability (Flesch) for chairman's statement (see notes for further details)
fogmu_Chair	Readability (Fog) for chairman's statement (see notes for further details)
causal_Chair	Causal reasoning keywords / wordcount for chairman's statement (see notes for further details)
forwardlooking_Chair	Forward-looking keywords / wordcount for chairman's statement (see notes for further details)
net_tone_Chair	Net positive keywords / (positive + negative) keywords for chairman's statement (see notes for further details)
uncertainty_Chair	Uncertainty keywords / wordcount for chairman's statement (see notes for further details)
CEOFlag	Indicator variable for presence of CEO review: 1 = CEO review extracted; 0 = no CEO review extracted (see notes for further details)
Pagecountsum_CEO	Number of pages in CEO review (see notes for further details)
wordcountsum_CEO	Wordcount for CEO review (see notes for further details)
nsections_CEO	Number of sections in CEO review (see notes for further details)
fleschmu_CEO	Readability (Flesch) for CEO review (see notes for further details)
fogmu_CEO	Readability (Fog) for CEO review (see notes for further details)
causal_CEO	Causal reasoning keywords / wordcount for CEO review (see notes for further details)
forwardlooking_CEO	Forward-looking keywords / wordcount for CEO review (see notes for further details)
net_tone_CEO	Net positive keywords / (positive + negative) keywords for CEO review (see notes for further details)
uncertainty_CEO	Uncertainty keywords / wordcount for CEO review (see notes for further details)
CFOFlag	Indicator variable for presence of CFO review: 1 = CFO review extracted; 0 = no CFO review extracted (see notes for further details)
Pagecountsum_CFO	Number of pages in CFO review (see notes for further details)
wordcountsum_CFO	Wordcount for CFO review (see notes for further details)
nsections_CFO	Number of sections in CFO review (see notes for further details)
fleschmu_CFO	Readability (Flesch) for CFO review (see notes for further details)
fogmu_CFO	Readability (Fog) for CFO review (see notes for further details)
causal_CFO	Causal reasoning keywords / wordcount for CFO review (see notes for further details)
forwardlooking_CFO	Forward-looking keywords / wordcount for CFO review (see notes for further details)
net_tone_CFO	Net positive keywords / (positive + negative) keywords for CFO review (see notes for further details)
uncertainty_CFO	Uncertainty keywords / wordcount for CFO review (see notes for further details)
RevFlag	Indicator variable for aggregate business/financial/operating reviews: 1 = review(s) extracted; 0 = no review(s) extracted (see notes for further details)
Pagecountsum_REV	Number of pages in aggregate business/financial/operating reviews (see notes for further details)
wordcountsum_REV	Wordcount for aggregate business/financial/operating reviews (see notes for further details)
nsections_REV	Number of sections in aggregate business/financial/operating reviews (see notes for further details)
fleschmu_REV	Readability (Flesch) for aggregate business/financial/operating reviews (see notes for further details)
fogmu_REV	Readability (Fog) for aggregate business/financial/operating reviews (see notes for further details)
causal_REV	Causal reasoning keywords / wordcount for aggregate business/financial/operating reviews (see notes for further details)
forwardlooking_REV	Forward-looking keywords / wordcount for aggregate business/financial/operating reviews (see notes for further details)
net_tone_REV	Net positive keywords / (positive + negative) keywords for aggregate business/financial/operating reviews (see notes for further details)
uncertainty_REV	Uncertainty keywords / wordcount for aggregate business/financial/operating reviews (see notes for further details)
performanceFlag_2	Indicator variable for aggregate performance commentary: 1 = performance commentary extracted; 0 = no performance commentary extracted (see notes for further details)
Pagecountsum_Perform	Number of pages in aggregate performance commentary (see notes for further details)
wordcountsum_Perform	Wordcount for aggregate performance commentary (see notes for further details)
nsections_Perform	Number of sections in aggregate performance commentary (see notes for further details)
fleschmu_Perform	Readability (Flesch) for aggregate performance commentary (see notes for further details)
fogmu_Perform	Readability (Fog) for aggregate performance commentary (see notes for further details)
causal_Perform	Causal reasoning keywords / wordcount for aggregate performance commentary (see notes for further details)
forwardlooking_Perform	Forward-looking keywords / wordcount for aggregate performance commentary (see notes for further details)
net_tone_Perform	Net positive keywords / (positive + negative) keywords for performance commentary (see notes for further details)
uncertainty_Perform	Uncertainty keywords / wordcount for aggregate performance commentary (see notes for further details)
strategyFlag	Indicator variable for strategy section: 1 = strategy section extracted; 0 = no strategy section extracted (see notes for further details)
Pagecountsum_Strategy	Number of pages in strategy section (see notes for further details)
wordcountsum_Strategy	Wordcount for strategy section (see notes for further details)
nsections_Strategy	Number of sections in strategy section (see notes for further details)
fleschmu_Strategy	Readability (Flesch) for strategy section (see notes for further details)
fogmu_Strategy	Readability (Fog) for strategy section (see notes for further details)
causal_Strategy	Causal reasoning keywords / wordcount for strategy section (see notes for further details)
forwardlooking_Strategy	Forward-looking keywords / wordcount for strategy section (see notes for further details)
net_tone_Strategy	Net positive keywords / (positive + negative) keywords for strategy section (see notes for further details)
uncertainty_Strategy	Uncertainty keywords / wordcount for strategy section (see notes for further details)
govFlag	Indicator variable for aggregate governance sections: 1 = governance sections extracted; 0 = no governance sections extracted (see notes for further details)
Pagecountsum_Gov	Number of pages in aggregate governance sections (see notes for further details)
wordcountsum_Gov	Wordcount for aggregate governance sections (see notes for further details)
nsections_Gov	Number of sections in aggregate governance sections (see notes for further details)
fleschmu_Gov	Readability (Flesch) for aggregate governance sections (see notes for further details)
fogmu_Gov	Readability (Fog) for aggregate governance sections (see notes for further details)
causal_Gov	Causal reasoning keywords / wordcount for aggregate governance sections (see notes for further details)
forwardlooking_Gov	Forward-looking keywords / wordcount for aggregate governance sections (see notes for further details)
net_tone_Gov	Net positive keywords / (positive + negative) keywords for aggregate governance sections (see notes for further details)
uncertainty_Gov	Uncertainty keywords / wordcount for aggregate governance sections (see notes for further details)
remFlag	Indicator variable for aggregate remuneration report sections: 1 = remuneration report sections extracted; 0 = no remuneration report sections extracted (see notes for further details)
Pagecountsum_Rem	Number of pages in aggregate remuneration report sections (see notes for further details)
wordcountsum_Rem	Wordcount for aggregate remuneration report sections (see notes for further details)
nsections_Rem	Number of sections in aggregate remuneration report sections (see notes for further details)
fleschmu_Rem	Readability (Flesch) for aggregate remuneration report sections (see notes for further details)
fogmu_Rem	Readability (Fog) for aggregate remuneration report sections (see notes for further details)
causal_Rem	Causal reasoning keywords / wordcount for aggregate remuneration report sections (see notes for further details)
forwardlooking_Rem	Forward-looking keywords / wordcount for aggregate remuneration report sections (see notes for further details)
net_tone_Rem	Net positive keywords / (positive + negative) keywords for aggregate remuneration report sections (see notes for further details)
uncertainty_Rem	Uncertainty keywords / wordcount for aggregate remuneration report sections (see notes for further details)

NOTES:

Variable names for annual report structure and textual features use the following two-part naming convention: *textualfeature_annualreportsection*

1. Definitions for the *textualfeature* component of the variable name are as follows:

Pagecountsum	Number of pages for annual report section s or defined aggregations of multiple annual sections (e.g., performance commentaries)
wordcountsum	Number of words for annual report section s or defined aggregations of multiple annual sections (e.g., performance commentaries)
nsections	Number of component sections from the annual report table of contents associated with section s or defined aggregations of multiple annual sections (e.g., performance commentaries)
fleschmu	Readability for section s or defined aggregations of multiple annual sections, computed using the Flesch-Kincaid measure. (Higher values of fleschmu are associated with less readable content.)
fogmu	Readability for section s or defined aggregations of multiple annual sections, computed using the Fog index measure. (Higher values of fogmu are associated with less readable content.)
causal	Number of causal reasoning keywords in section s (or defined aggregation) scaled by the total number of words in the corresponding section s (or defined aggregation).
forwardlooking	Number of forward-looking keywords in section s (or defined aggregation) scaled by the total number of words in the corresponding section s (or defined aggregation).
net_tone	Number of positive keywords minus number of negative keywords, scaled by aggregate number of positive and negative keywords, for section s (or defined aggregation).
uncertainty	Number of uncertainty keywords in section s (or defined aggregation) scaled by the total number of words in the corresponding section s (or defined aggregation).

2. Definitions for the *annualreportsection* component of the variable name are as follows:

Front_3	The narrative component of the annual report, defined as all annual sections other than: auditors' report, statement of directors' responsibilities, financial statements and notes, statutory shareholder information, statutory five-year summaries, subsidiaries and their locations, and information regarding the annual general meeting.
Rear_3	The financial statements component of the annual report, equal to as the residual sections after defining Front_3.
Chair	The Chairman's Statement or synonyms thereof (e.g., Chairman's Letter, President's Letter, Letter to Shareholders, etc.). For firms with a dual chairman and CEO, the combined Chairman and CEO review is classified as the chairman's statement.
CEO	The Chief Executive Officer's (CEO) Review or synonyms thereof (e.g., Managing Director's Review). Business reviews and operating reviews are classified separately unless the section title in the annual report table of contents refers to the CEO or synonyms thereof.
CFO	The Chief Financial Officer's (CFO) Review or synonyms thereof (e.g., Financial Review, Finance Director's Review, etc.). Business reviews and operating reviews are classified separately unless the section title in the annual report table of contents refers to the CFO or synonyms thereof.
REV	Aggregate business review sections, including CEO reviews, CFO reviews, business reviews, and operating reviews
Perform	Aggregate performance commentaries, comprising all sections defined as part of Front_3 from the annual report table of contents that discuss any aspect of operating activities or financial performance. This category explicitly excludes corporate social responsibility reports, directors' reports, risk reports, governance statements, and remuneration statements.
Strategy	Aggregate strategy-related commentary, comprising all sections defined as part of Front_3 from the annual report table of contents that include the keywords strategy, strategic, business model, key performance indicator or KPI.
Gov	Corporate governance statements, comprising all sections defined as part of Front_3 from the annual report table of contents containing governance-related content (e.g., chairman's introduction to governance, governance statements, audit committee reports, nomination committee reports, internal control and risk management reports). Corporate social responsibility and directors' reports are explicitly excluded from this category.
Rem	Remuneration reports, comprising all sections defined as part of Front_3 from the annual report table of contents that form part of the report on executive compensation practices.