

- **fl_matriz**: boolean value, `true` if the CNPJ corresponds to the "matriz".
- **de_natureza_juridica**: character, juridic nature of the company.
- **sg_uf**: character, initials of the name of the state of the company.
- **natureza_juridica_macro**: a macro legal category for "natureza jurídica".
- **de_ramo**: description of a macro activity category/branch based on the CNAE code.
- **setor**: character, economic sector of the company (Industry, Services, Agrobusiness, Etc.)
- **idade_empresa_anos**: numeric value, age of the company.
- **idade_emp_cat**: character, age of the company by ranges
- **fl_me**: boolean value, true if the company has the term 'ME' in the end of its legal name.
- **fl_sa**: boolean value, true if the company has the term 'SA' in the end of its legal name.
- **fl_epp**: boolean value, true if the company has the term 'EPP' in the end of its legal name.
- **fl_mei**: boolean value, true if the company has the term 'MEI' in the end of its legal name.
- **fl ltda**: boolean value, true if the company has the term 'LTDA' in the end of its legal name.
- **dt_situacao**: date when the "de_situacao" was registered by the IRS.
- **fl_st_especial**: boolean value, true if `dt_situacao_especial` is not null. If it is not null means that some extraordinary situation is identified by the IRS (*ESPOLIO DE EMPRESARIO EMPRESA INDIVIDUAL OU EIRELI, FALIDO, EM LIQUIDACAO, LIQUIDACAO JUDICIAL, LIQUIDACAO EXTRA JUDICIAL, REGISTRO NA JUNTA COMERCIAL EM ANDAMENTO, EM LIQUID EXTRA JUDICIAL, RECUPERACAO JUDICIAL, INTERVENCAO*)
- **fl_email**: boolean value, true if the cnpj has an email registered at the IRF database.
- **fl_telefone**: boolean value, true if the company has a phone number registered in IRS database.
- **fl_rm**: character, true if the company address is located in an metropolitan area.
- **nm_divisao**: character, name description of the primary economic activity of the company
- **nm_segmento**: character, name description of the primary economic activity cluster of the company
- **fl_spa**: boolean, true if the company has it's own fuel station
- **fl_antt**: boolean, true if the company is certified by ANTT for transportation purposes
- **fl_veiculo**: boolean, true if the company owns at least one vehicle
- **vl_total_tancagem**: numeric, total capacity of fuel storage of the company
- **vl_total_veiculos_antt**: integer, total number of vehicles of the company
- **vl_total_veiculos_leves**: integer, total number of light weight vehicles of the company
- **vl_total_veiculos_pesados**: integer, total number of heavy weight vehicles of the company
- **fl_optante_simples**: boolean, true if the company is taxed according to SIMPLES regime
- **qt_art**: integer, number of constructions(buildings) authorized by an architect (ART - Anotação de Responsabilidade Técnica)
- **vl_total_veiculos_pesados_grupo**: numeric, economic value of the heavy weight vehicles of the economic group of the company
- **vl_total_veiculos_leves_grupo**: numeric, economic value of the light weight vehicles of the economic group of the company
- **vl_total_tancagem_grupo**: numeric, total capacity of fuel storage of the economic group of the company
- **vl_total_veiculos_antt_grupo**: numeric, total number of vehicles of the economic group of the company that are registered with ANTT certification
- **vl_potenc_cons_oleo_gas**: numeric, potential consumption of oil and gas
- **fl_optante_simei**: boolean, true if the company is taxed as an MEI - Individual Micro Entrepreneur (Micro Empreendedor Individual)

- **sg_uf_matriz**:character, state of the main office location
- **de_saude_tributaria**:character, indicator of health tax status, Green if all tax are OK, Red if none are OK
- **de_saude_rescencia**:character, indicates time of update of the most lagged input of the indicator.
- **nu_meses_rescencia**:character, number of months since the last update of the most lagged input from saude_tributaria
- **de_nivel_atividade**:character, probability of being operating, ALTA high probability, BAIXA, low probability.
- **de_indicador_telefone**:character, propability to reach the company using the phone numbers provided by Neoway
- **fl_simples_irregular**: boolean, true if the company is taxed under the SIMPLES regime and has a impeditive CNAE (economic activities that are not allowed to be taxed according to the simples regime) revenue above the limit of the regime.
- **vl_frota**: numeric, aggregate value of all vehicles of the company
- **empsetorcensitariofaixarendapopulacao**: numeric, average income from a sample of residents in a census unity (smallest territory area according to IBGE)
- **nm_meso_regiao**:character, name of the meso region where the company is located.
- **nm_micro_regiao**:character, name of the micro region where the company is located.
- **fl_passivel_iss**:boolean, true if the company performs any activity that is taxable under ISS - Tax under services.
- **qt_socios**: integer value, quantity of partners/shareholders of the cnpj
- **qt_socios_pf**: integer value, quantity of partners/shareholders of the cnpj that are persons
- **qt_socios_pj**: integer value, quantity of partners/shareholders of the cnpj that are companies
- **idade_media_socios**: numeric value, average age of the partners of the cnpj(considers only the cpf partners)
- **idade_maxima_socios**: same as above but with max metric
- **idade_minima_socios**: same as above but wiht min metric
- **qt_socios_st_regular**: integer value, quantity of partners with regular situation under IRS - Receita Federal
- **qt_socios_st_suspensa**: integer value, quantity of partners with suspended situation under IRS - Receit Federal
- **qt_socios_masculino**: integer value, quantity of male partners
- **qt_socios_feminino**: integer value, quantity of female partners (womens)
- **qt_socios_pep**: integer value, quantity of partners of the cnpj that are politically exposed
- **qt_alteracao_socio_total**: integer value, quantity of changes in the QSA
- **qt_alteracao_socio_90d**: integer value, quantity of changes in the QSA in the last 90 days
- **qt_alteracao_socio_180d**: integer value, quantity of changes in the QSA in the last 180 days
- **qt_alteracao_socio_365d**: integer value, quantity of changes in the QSA in the last 365 days
- **qt_socios_pj_ativos**: integer value, quantity of PJ partners that are ATIVA in the RF.
- **qt_socios_pj_nulos**: integer value, quantity of PJ partners that are NULA in the RF.
- **qt_socios_pj_baixados**: integer value, quantity of PJ partners that are BAIXADA in the RF.
- **qt_socios_pj_suspensos**: integer value, quantity of PJ partners that are SUSPENSA in the RF.
- **qt_socios_pj_inaptos**: integer value, quantity of PJ partners that are INAPTA in the RF.
- **vl_idade_media_socios_pj**: float value, avarage age of the PJ partners.
- **vl_idade_maxima_socios_pj**: float value, maximum age of the PJ partners.
- **vl_idade_minima_socios_pj**: float value, minimum age of the PJ partners.

- **qt_coligados:** integer value, quantity of connected companies
- **qt_socios_coligados:** integer value, sum of partners from connected companies
- **qt_coligados_matriz:** integer value, quantity of connected companies that are a matrix company
- **qt_coligados_ativo:** integer value, quantity of connected companies that are considered as active according to the IRS - receita federal
- **qt_coligados_baixada:** integer value, quantity of connected companies that are considered as closed according to the IRS
- **qt_coligados_inapta:** integer value, quantity of connected companies that are considered as inapt/unfit according to the IRS
- **qt_coligados_suspensa:** integer value, quantity of connected companies that are considered suspended according to the IRS
- **qt_coligados_nula:** integer value, quantity of connected companies that are considered as nule companies according to the IRS
- **idade_media_coligadas:** float, average age (months) of connected companies
- **idade_maxima_coligadas:** float, maximum age (months) of connected companies
- **idade_minima_coligadas:** float, minimum age (months) of connected companies
- **coligada_mais_nova_ativa:** float, age (months) of youngest coligate that is active
- **coligada_mais_antiga_ativa:** float, age (months) of oldest coligate that is active
- **idade_media_coligadas_ativas:** float, average age (months) of active connected companies
- **coligada_mais_nova_baixada:** float, age (months) of yougest closed coligate
- **coligada_mais_antiga_baixada:** float, age (months) of oldest closed coligate
- **idade_media_coligadas_baixadas:** float, average (months) of closed coligate
- **qt_coligados_sa:** integer value, quantity of connected companies that has the term SA in the end of the business name
- **qt_coligados_me:** integer value, quantity of connected companies that has the term ME in the end of the business name
- **qt_coligados_mei:** integer value, quantity of connected companies that has the term MEI in the end of the business name
- **qt_coligados ltda:** integer value, quantity of connected companies that has the term LTDA in the end of the business name
- **qt_coligados_epp:** integer value, quantity of connected companies that has the term EPP in the end of the business name
- **qt_coligados_norte:** integer value, quantity of connected companies that are located in the north region of the country
- **qt_coligados_sul:** integer value, quantity of connected companies that are located in the south region of the country
- **qt_coligados_nordeste:** integer value, quantity of connected companies that are located in the northeast region of the country
- **qt_coligados_centro:** integer value, quantity of connected companies that are located in the center-west (centro-oeste) region of the country
- **qt_coligados_sudeste:** integer value, quantity of connected companies that are located in the southeast of the country
- **qt_coligados_exterior:** integer value, quantity of connected companies that are located abroad the country

- **qt_ufs_coligados:** integer value, quantity of distinct states where the connected companies are located
- **qt_regioes_coligados:** integer value, quantity of distinct ufs where the connected companies are located
- **qt_ramos_coligados:** integer value, quantity of distinct economic branches from connected companies
- **qt_coligados_industria:** integer value, quantity of connected companies from the industry sector
- **qt_coligados_agropecuaria:** integer value, quantity of connected companies from the agrobusiness sector
- **qt_coligados_comercio:** integer value, quantity of connected companies
- **qt_coligados_servico:** integer value, quantity of connected companies from the service sector
- **qt_coligados_ccivil:** integer value, quantity of connected companies from the construction sector
- **qt_funcionarios_coligados:** integer value, quantity of employees from the connected companies
- **qt_funcionarios_coligados_gp:** integer value, quantity of employees from connected companies and their subsidiaries
- **media_funcionarios_coligados_gp:** numeric value, average of employees from the connected companies and their subsidiaries
- **max_funcionarios_coligados_gp:** numeric value, maximum number of employees from connected companies and their subsidiaries
- **min_funcionarios_coligados_gp:** numeric value, minimum number of employees from connected companies and their subsidiaries
- **vl_folha_coligados:** numeric value, sum of payroll value from connected companies
- **media_vl_folha_coligados:** numeric value, average of payroll value from connected companies
- **max_vl_folha_coligados:** numeric value, maximum of payroll value from connected companies
- **min_vl_folha_coligados:** numeric value, minimum of payroll value from connected companies
- **vl_folha_coligados_gp:** numeric value, sum of payroll value from connected companies and their subsidiaries
- **media_vl_folha_coligados_gp:** numeric value, average of payroll value from connected companies and their subsidiaries
- **max_vl_folha_coligados_gp:** numeric value, maximum of payroll value from connected companies and their subsidiaries
- **min_vl_folha_coligados_gp:** numeric value, minimum of payroll value from connected companies and their subsidiaries
- **faturamento_est_coligados:** numeric value, sum of estimated revenue from connected companies
- **media_faturamento_est_coligados:** numeric value, average of estimated revenue from connected companies
- **max_faturamento_est_coligados:** numeric value, maximum value of estimated revenue from connected companies
- **min_faturamento_est_coligados:** numeric value, minimum value of estimated revenue from connected companies
- **faturamento_est_coligados_gp:** numeric value, sum of estimated revenue from connected companies and their subsidiaries
- **media_faturamento_est_coligados_gp:** numeric value, average of estimated revenue from connected companies and their subsidiaries
- **max_faturamento_est_coligados_gp:** numeric value, maximum value of estimated revenue from connected companies and their subsidiaries
- **min_faturamento_est_coligados_gp:** numeric value, minimum value of estimated revenue from

connected companies and their subsidiaries

- **total_filiais_coligados**: integer value, quantity of subsidiaries from connected companies
- **media_filiais_coligados**: numeric value, average of subsidiaries from connected companies
- **max_filiais_coligados**: numeric value, maximum number of subsidiaries from connected companies
- **min_filiais_coligados**: numeric value, minimum number of subsidiaries from connected companies
- **qt_coligados_atividade_alto**: integer value, quantity of connected companies with estimated activity level classifier equal to 'ALTO'
- **qt_coligados_atividade_medio**: integer value, quantity of connected companies with estimated activity level classifier equal to 'Medio'
- **qt_coligados_atividade_baixo**: integer value, quantity of connected companies with estimated activity level classifier equal to 'Baixo'
- **qt_coligados_atividade_mt_baixo**: integer value, quantity of connected companies with estimated activity level classifier equal to 'Muito Baixo'
- **qt_coligados_atividade_inativo**: integer value, quantity of connected companies with estimated activity level classifier equal to 'INATIVO'
- **qt_coligadas**: number of linked companies, i.e., companies where one of the main owners (referenced CJNPJ) has participation.
- **sum_faturamento_estimado_coligadas**: the sum of all *coligadas* estimated
- **de_faixa_faturamento_estimado**: character value, class of the estimated revenue of the company
- **de_faixa_faturamento_estimado_grupo**: character value, class of the sum of estimated revenue for the matrix company and branches
- **vl_faturamento_estimado_aux**: numeric value, value of the estimated revenue
- **vl_faturamento_estimado_grupo_aux**: numeric value, sum of the estimated revenue for the matrix company and the branches
- **qt_ex_funcionarios_cnseg**: integer value, number of employees that have left the company.
- **qt_funcionarios_grupo_cnseg**: integer value, quantity of active employees considering the branches and matrix company.
- **percent_func_genero_masc**: numeric value, share of male employees in the company
- **percent_func_genero_fem**: numeric value, share of female employees in the company
- **idade_ate_18**: integer value, quantity of employees under age of 18 years old.
- **idade_de_19_a_23**: integer value, quantity of employees with age between 19 and 23 years old.
- **idade_de_24_a_28**: integer value, quantity of employees with age between 24 and 28 years old.
- **idade_de_29_a_33**: integer value, quantity of employees with age between 29 and 33 years old.
- **idade_de_34_a_38**: integer value, quantity of employees with age between 34 and 38 years old.
- **idade_de_39_a_43**: integer value, quantity of employees with age between 39 and 43 years old.
- **idade_de_44_a_48**: integer value, quantity of employees with age between 44 and 48 years old.
- **idade_de_49_a_53**: integer value, quantity of employees with age between 49 and 53 years old.
- **idade_de_54_a_58**: integer value, quantity of employees with age between 54 and 58 years old.
- **idade_acima_de_58**: integer value, quantity of employees with age above
- **grau_instrucao_macro_analfabeto**: integer value, quantity of employees in positions with no education level required.
- **grau_instrucao_macro_escolaridade_fundamental**: integer value, quantity of employees in positions with required fundamental level (complete and incomplete)
- **grau_instrucao_macro_escolaridade_media**: integer value, sum of grau_instrucao_medio_incompl and

grau_instrucao_medio_compl

- **grau_instrucao_macro_escolaridade_superior**: integer value, sum of grau_instrucao_sup_incompl, grau_instrucao_sup_compl, grau_instrucao_mestrado and grau_instrucao_doutorado
- **grau_instrucao_macro_desconhecido**: integer value, same as grau_instrucao_desconhecido
- **total**: integer value, actual number of employees of the company
- **meses_ultima_contratacao**: numeric, months since the last hire.
- **qt_admitidos_12meses**: integer, quantity of employees hired in the last 12 months
- **qt_desligados_12meses**: integer, quantity of employees dismissed in the last 12 months
- **qt_desligados**: integer, quantity of employees dismissed by the company
- **qt_admitidos**: integer, quantity of employees hired by the company
- **media_meses_servicos_all**: numeric, average number of months worked by all the employees of the company (active and dismissed)
- **max_meses_servicos_all**: numeric, maximum number of months worked by all the employees of the company (active and dismissed)
- **min_meses_servicos_all**: numeric, minimum number of months worked by all the employees of the company (active and dismissed)
- **media_meses_servicos**: numeric, average number of months worked by the active employees of the company
- **max_meses_servicos**: numeric, maximum number of months worked by the active employees of the company
- **min_meses_servicos**: numeric, minimum number of months worked by the active employees of the company
- **qt_funcionarios**: integer, quantity of actual employees
- **qt_funcionarios_12meses**: integer, quantity of employees active 12 months before
- **qt_funcionarios_24meses**: integer, quantity of employees active 24 months before
- **tx_crescimento_12meses**: integer, growth of employees in relation to the number of employees 12 months before
- **tx_crescimento_24meses**: integer, growth of employees in relation to the number of employees 24 months before
- **tx_rotatividade**: integer, measures the overall stability of employees considering hiring and dismissals and the actual number of employees.
- **qt_filiais**: self-explanatory.