

```
In [1]:
import os
import requests
from pathlib import Path

# Recuperar o diretório onde salvaremos o arquivo csv
current_dir = Path.cwd()
parent_dir = current_dir.parent
data_folder = parent_dir/'data'

# URL do CSV
url = "https://dados.agricultura.gov.br/dataset/6c913699-e82e-4da3-a0a1-fb6c431e367f/resource/d30b30d7-e256-484e-9ab8-cd40974e1238/download/agrofitprodutosformulados.csv"

# Headers para simular um navegador
headers = {
    "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 Safari/537.36"
}

# Fazer download
response = requests.get(url, headers=headers)

# Salvar no caminho correto
csv_path = os.path.join(data_folder, "agrofitprodutosformulados.csv")
In [2]:
import pandas as pd
from datetime import datetime

# Importando arquivo csv e convertendo tipos
df = pd.read_csv(csv_path, sep=';', engine='pyarrow')
df = df.convert_dtypes()

# Limpando os nomes das empresas para facilitar a leitura
df['TITULAR'] = df['TITULAR_DE_REGISTRO'].str.split().str[0]
```

Dados utilizados na análise

Dados obtidos no dia 21/02/2026

Nota: Onde houver um asterisco (*), os dados foram obtidos na data acima e podem estar desatualizados de acordo com a evolução do banco de dados.

```
In [3]:
df
```

Out[3]:

	NR_REGISTRO	MARCA_COMERCIAL	FORMULACAO	INGREDIENTE_ATIVO	TITULAR_DE_REGISTRO	CLASSE	MODO_DE_ACAO	CULTURA	PRAGA_NOME_CIENTIFICO	PRAGA_NOME_COMUM	EMPRESA_PAIS_
0	35523	KBR-829M1-02	Nematóides vivos	Heterorhabditis bacteriophora (Nematóides ento...	Koppert do Brasil Holding S.A. - Piracicaba/SP	Agente Biológico de Controle	<NA>	Todas as culturas	Scaptocoris castanea	Percevejo-castanho	(Koppert do Brasil Holding Piracic
1	35523	KBR-829M1-02	Nematóides vivos	Heterorhabditis bacteriophora (Nematóides ento...	Koppert do Brasil Holding S.A. - Piracicaba/SP	Agente Biológico de Controle	<NA>	Todas as culturas	Sphenophorus levis	Bicudo da cana-de-açúcar; Gorgulho-da-cana	(Koppert do Brasil Holding Piracic
2	35523	KBR-829M1-02	Nematóides vivos	Heterorhabditis bacteriophora (Nematóides ento...	Koppert do Brasil Holding S.A. - Piracicaba/SP	Agente Biológico de Controle	<NA>	Todas as culturas	Spodoptera frugiperda	Lagarta-militar	(Koppert do Brasil Holding Piracic
3	33723	KBR-S39M1-02	Nematóides vivos	Steinemema carpocapsae (Nematóides entomopato...	Koppert do Brasil Holding S.A. - Piracicaba/SP	Agente Biológico de Controle	<NA>	Todas as culturas	Bradysia matogrossensis	fungus gnats	(Koppert do Brasil Holding Piracic
4	33723	KBR-S39M1-02	Nematóides vivos	Steinemema carpocapsae (Nematóides entomopato...	Koppert do Brasil Holding S.A. - Piracicaba/SP	Agente Biológico de Controle	<NA>	Todas as culturas	Sphenophorus levis	Bicudo da cana-de-açúcar; Gorgulho-da-cana	(Koppert do Brasil Holding Piracic
...	
272444	37019	ÍmparBR	FS - Suspensão Concentrada p/ Trat. Sementes	tiametoxam (neonicotinóide) (350 g/L)	Ouro Fino Química S.A. - Uberaba	Inseticida	Sistêmico	Sorgo	Dichelops melacanthus	percevejo-barriga-verde	(Ningbo Sunjoy Agrosience Co <CHIN
272445	37019	ÍmparBR	FS - Suspensão Concentrada p/ Trat. Sementes	tiametoxam (neonicotinóide) (350 g/L)	Ouro Fino Química S.A. - Uberaba	Inseticida	Sistêmico	Trigo	Dichelops furcatus	Percevejo-barriga-verde	(Ningbo Sunjoy Agrosience Co <CHIN
272446	37019	ÍmparBR	FS - Suspensão Concentrada p/ Trat. Sementes	tiametoxam (neonicotinóide) (350 g/L)	Ouro Fino Química S.A. - Uberaba	Inseticida	Sistêmico	Trigo	Rhopalosiphum graminum	Pulgão-da-espiga; Pulgão-verde-dos-cereais	(Ningbo Sunjoy Agrosience Co <CHIN
272447	7697	Óleo Vegetal Nortox	EC - Concentrado Emulsionável	óleo vegetal (ésteres de ácidos graxos) (930 g/L)	Nortox S.A. - Arapongas	Inseticida	De contato	Citros	Selenaspidus articulatus	Cochonilha-pardinha	(Nortox Arapongas<BRASIL>FORMULA
272448	7697	Óleo Vegetal Nortox	EC - Concentrado Emulsionável	óleo vegetal (ésteres de ácidos graxos) (930 g/L)	Nortox S.A. - Arapongas	Inseticida	De contato	<NA>	<NA>	<NA>	(Nortox Arapongas<BRASIL>FORMULA

272449 rows × 16 columns

Visão geral do mercado/concentração

Quantas empresas distintas existem no sistema?

```
In [4]:
# Criando um dataframe sem duplicatas de produtos com os nomes das empresas detentoras do registro
produtos_df = df[['NR_REGISTRO', 'TITULAR']].drop_duplicates(subset=['NR_REGISTRO'])
```

```
# Retornando o número de registros únicos dos titulares
produtos_df['TITULAR'].nunique()
```

Out[4]:
307

Como podemos ver, são **307** * empresas com CNPJ distintos inscritos no banco de dados do Agrofit.

Vale ressaltar que, algumas poucas empresas atuam sob diferentes inscrições.

Top 10 empresas com mais produtos registrados

```
In [5]:
# Ordenar as 10 empresas com mais produtos registrados
produtos_df["TITULAR"].value_counts().head(10)

Out[5]:
TITULAR
Syngenta    249
UPL         192
Adama       175
Nortox      167
Sumitomo    166
Rainbow     160
CTVA        144
BASF        118
FMC         108
Bayer        99
Name: count, dtype: int64
```

Acima, podemos ver as empresas e a quantidade de produtos registrados por cada uma em ordem decrescente.

Note que, dentre as empresas listada, **apenas 1** delas é genuinamente brasileira: a Nortox S.A., ocupando a quarta colocação com 167* produtos registrados.

Porcentagem de participação das top 10 empresas

```
In [6]:
# Porcentagem arredondada do número total de produtos das top 10 empresas comparadas ao total de produtos
round(produtos_df["TITULAR"].value_counts().head(10).sum() / produtos_df["NR_REGISTRO"].value_counts().sum() * 100, 2)

Out[6]:
np.float64(37.11)
In [7]:
# Contagem de quantos produtos as top 10 empresas têm no cadastro
produtos_df["TITULAR"].value_counts().head(10).sum()

Out[7]:
np.int64(1578)
In [8]:
# Contagem total de todos os produtos
produtos_df["NR_REGISTRO"].value_counts().sum()

Out[8]:
np.int64(4252)
```

Se formos contar a participação que as top 10 empresas têm sob o mercado nacional, elas juntas somam 37,11%* do mercado, ou ainda, 1578* produtos dos 4252*.

Quantos produtos cada empresa tem por classe das quatro grandes classes (herbicida, fungicida, inseticida e biológico)?

```
In [9]:
```

```
# Seleciona colunas relevantes e remove duplicatas de produtos
classes_df = df[['NR_REGISTRO', 'TITULAR', 'CLASSE']].drop_duplicates(subset=['NR_REGISTRO'])

# Cria colunas booleanas para as 4 principais classes
classes_df['HERBICIDA'] = classes_df['CLASSE'].str.contains('Herbicida', case=False, na=False)
classes_df['FUNGICIDA'] = classes_df['CLASSE'].str.contains('Fungicida', case=False, na=False)
classes_df['INSETICIDA'] = classes_df['CLASSE'].str.contains('Inseticida', case=False, na=False)
classes_df['(MICRO)BIOLÓGICO'] = classes_df['CLASSE'].str.contains('Biológico', case=False, na=False)

# Agrupa por empresa e soma a quantidade de produtos por classe
classes = classes_df.groupby('TITULAR', observed=True)[['HERBICIDA', 'FUNGICIDA', 'INSETICIDA', '(MICRO)BIOLÓGICO']].sum()
```

```
# Adiciona coluna de total para facilitar ordenação
classes['TOTAL'] = classes.sum(axis=1)
```

```
# Ordena pelas empresas com mais produtos e exibe as top 10
classes = classes.nlargest(n=10, columns='TOTAL')
```

classes

Out[9]:

	HERBICIDA	FUNGICIDA	INSETICIDA	(MICRO)BIOLÓGICO	TOTAL
TITULAR					
Syngenta	48	114	85	7	254
UPL	86	54	40	3	183
Adama	88	43	35	2	168
Nortox	84	38	42	0	164
Rainbow	114	27	17	0	158
Sumitomo	71	23	54	10	158
CTVA	77	28	30	0	135
Basf	30	53	29	5	117
FMC	31	17	50	12	110
Bayer	20	41	31	2	94

Como podemos ver acima, a participação dos grandes players no mercado de (micro)biológicos ainda é irrisória frente aos produtos químicos.

Apesar disso, podemos ver a Syngenta ocupando o posto de primeira colocada em número de fungicidas e inseticidas registrados, tendo próximo do dobro de produtos quando comparado às demais empresas.

Quando olhamos para os herbicidas, o mercado é mais competitivo, com 6* de 10 tendo mais 70* produtos registrados, com a chinesa Rainbow encabeçando a lista.

Ingredientes ativos mais usados

Ingredientes ativos são as substâncias principais presentes nos agroquímicos responsáveis pela ação sobre organismos-alvo, como pragas, doenças e plantas daninhas. Eles são o componente essencial da formulação e determinam o modo de ação, eficácia e toxicidade do produto. Exemplos comuns incluem **glifosato** (herbicida), **2,4-D** (herbicida), **clorotalonil** (fungicida), **fipronil** (inseticida) e **acefato** (inseticida).

Esses ingredientes são combinados com **ingredientes inertes** (como solventes ou dispersantes) para melhorar a aplicação, estabilidade e eficiência do produto. A legislação brasileira exige que o ingrediente ativo seja claramente identificado no rótulo, pois é ele que define o risco à saúde humana e ao meio ambiente.

Quais são os 10 ingredientes ativos mais frequentes nos produtos formulados?

In [10]:

```
# Seleciona colunas relevantes e remove duplicatas de produtos
ingredientes_df = df[['NR_REGISTRO', 'INGREDIENTE_ATIVO']].drop_duplicates(subset='NR_REGISTRO')
```

```
# Separa os ingredientes ativos de produtos que tenham mais de 1
ingredientes_df['INGREDIENTE_ATIVO'] = ingredientes_df['INGREDIENTE_ATIVO'].str.split(' + ')
```

```
# Separa em linhas diferentes a lista gerada pelo split
ingredientes_df = ingredientes_df.explode('INGREDIENTE_ATIVO')
```

```
# Remove os espaços, conta quantas ocorrências tiveram de cada ativo e retorna somente os 10 primeiros
ingredientes_df['INGREDIENTE_ATIVO'].str.strip().value_counts().head(10)
```

```
Out[10]:
INGREDIENTE_ATIVO
Glufosinato - sal de amônio (homoalanina substituída) (200 g/L)    51
glifosato-sal de isopropilamina (glicina substituída) (480 g/L)   49
cletodim (oxima ciclohexanodiona) (240 g/L)                     37
fipronil (pirazol) (800 g/kg)                                     34
sulfentrazona (triazolona) (500 g/L)                             33
Cotesia flavipes                                                  31
clorotalonil (isoftalonitrila) (720 g/L)                         31
glifosato-sal de amônio (glicina substituída) (792.5 g/kg)       31
clorpirifós (organofosforado) (480 g/L)                         29
acefato (organofosforado) (750 g/kg)                             27
Name: count, dtype: int64
```

Analisando apenas os 10 ingredientes ativos mais utilizados, vemos um domínio do **glifosato**, com 80* produtos contendo alguma formulação do ativo, seguido pelo **Glufosinato**, com 51* produtos utilizando o ativo.

Fipronil* segue sendo o primeiro ativo inseticida mais utilizado, mas é seguido por um agente inesperado: *Cotesia flavipes*. Este inseto é uma vespa parasitóide, amplamente utilizada no controle biológico da *broca-da-cana-de-açúcar*.

Dentre os ingredientes ativos, apenas 1* é um fungicida: o **clorotanonil**, ativo multissítio de contato.

Quantos % dos produtos são alguma formulação de Glifosato?

```
In[11]:
# Gera uma variável "glifosato", onde é um agregado de todos os produtos que tenham alguma fórmula desse ativo
glifosato = ingredientes_df['INGREDIENTE_ATIVO'].str.contains('glifosato', case=False, na=False)
```

```
# Retorna a porcentagem de produtos que são glifosato dentro do banco de dados
round(ingredientes_df.value_counts(glifosato, normalize=True) * 100, 2)
```

```
Out[11]:
INGREDIENTE_ATIVO
False    95.65
True      4.35
Name: proportion, dtype: float64
```

Diante da enorme expressividade do **Glifosato**, uma análise curiosa revela que 4,35%* dos produtos cadastrados são alguma espécie de Glifosato.

Quais ingredientes ativos podem ser utilizados no maior número de culturas diferentes?

```
In[12]:
# Gera uma variável "culturas_df", que é o agrupamento dos valores únicos dos ingredientes ativos por cultura
culturas_df = df.groupby('INGREDIENTE_ATIVO', observed=True)['CULTURA'].nunique()
```

```
# Retorna os 10 ingredientes ativos com mais ocorrência dentro das culturas
culturas_df.nlargest(10)
```

Out[12]:

INGREDIENTE_ATIVO			
Azoxistrobina (estrobilurina) (200 g/L) + difenoconazol (triazol) (125 g/L)	124		
difenoconazol (triazol) (250 g/L)	120		
fluxapiroxade (carboxamida) (167 g/L) + piraclostrobina (estrobilurina) (333 g/L)	100		
lambda-cialotrina (piretróide) (250 g/L)	91		
glifosato-sal de isopropilamina (glicina substituída) (480 g/L)	88		
ciprodinil (anilinopirimidina) (750 g/kg)	87		
diafentiurom (feniltioureia) (500 g/L)	86		
Azoxistrobina (estrobilurina) (500 g/kg)	85		
Swinglea glutinosa (Terpenos), Extrato de (Derivado vegetal) (868.4 g/L)	83		
azadiractina (Tetranortriterpenóide) (12 g/L)	83		
Name: CULTURA, dtype: int64			

Quando olhamos para os ingredientes ativos mais utilizados no espectro das culturas, vemos uma predominância dos fungicidas*, ocupando os 3 primeiros lugares no ranking. Isso se deve ao amplo espectro de ação dos mesmos, sendo capaz de atuar nas mais diversas culturas cultivadas.

A lambda-cialotrina* (inseticida) e o glifosato - sal de isopropilamina* (herbicida) compõem o restante dos top 5 ingredientes ativos mais versáteis.

Há destaque para o top 9 e 10*, onde ambos são óleos de origem vegetal (um fungicida/bactericida e outro inseticida).

Classe de uso

Os **agroquímicos** são classificados de acordo com o tipo de praga ou organismo que atuam, sendo os principais tipos de ação:

- **Inseticidas**: Combatem insetos e pragas que atacam as plantas, como pulgões, lagartas e percevejos.
- **Fungicidas**: Controlam fungos causadores de doenças em culturas, como mofo, ferrugem e manchas foliares.
- **Herbicidas**: Destinam-se ao controle de plantas daninhas, eliminando ou inibindo seu crescimento.
- **Acaricidas**: Atuam contra ácaros, pequenos artrópodes que danificam folhas e frutos.
- **Nematicidas**: Combatem nematoides (vermes parasitas do solo) que prejudicam as raízes das plantas.
- **Rodenticidas**: Eliminam roedores como ratos e toupeiras que danificam lavouras.
- **Bactericidas**: Inibem ou matam bactérias que causam infecções em plantas.
- **Reguladores de crescimento**: Ajustam o desenvolvimento vegetal, como acelerar o amadurecimento ou controlar o crescimento de plantas.

Essa classificação é baseada na **natureza do organismo-alvo**, sendo essencial para o uso correto e eficaz dos produtos na agricultura.

Os **produtos biológicos** são produtos derivados de organismos vivos, utilizados no controle de pragas e doenças de forma mais sustentável. Diferentemente dos agroquímicos, atuam com maior seletividade e menor impacto ambiental.

Principais tipos de ação dos produtos biológicos:

1. **Inimigos naturais**
Insetos predadores ou parasitoides (como *Trichogramma* spp.) que controlam populações de pragas, como lagartas e pulgões.
2. **Micro-organismos**
Bactérias (ex: *Bacillus thuringiensis*), fungos (ex: *Beauveria bassiana*) e vírus que infectam e eliminam insetos-praga.
3. **Antagonistas microbianos**
Micro-organismos que competem ou inibem patógenos do solo, como fungos causadores de murchas.

Esses produtos são essenciais na **Agricultura Orgânica** e no **Manejo Integrado de Pragas (MIP)**, promovendo equilíbrio ecológico e reduzindo a dependência de produtos sintéticos.

Distribuição de produtos por classe de uso

```
In [13]:
# Seleciona colunas relevantes e remove duplicatas de produtos
distribuicao_df = df[['NR_REGISTRO', 'CLASSE']].drop_duplicates('NR_REGISTRO')

# Retorna a proporção em que as classes de ativos estão separadas, incluindo combinações de classes
round(distribuicao_df.value_counts(subset='CLASSE', normalize=True).head(10) * 100, 2)
```

```
Out[13]:
CLASSE
Herbicida      33.80
Fungicida      17.50
Inseticida     15.26
Inseticida Microbiológico    7.76
Acaricida/Inseticida    3.74
Fungicida microbiológico    3.43
Regulador de Crescimento    2.56
Nematicida Microbiológico    2.35
Agente Biológico de Controle    2.19
Acaricida Microbiológico/Inseticida Microbiológico    1.76
Name: proportion, dtype: float64
```

```
In [14]:
# Cria uma coluna CLASSE onde as combinações de ativos são separadas
distribuicao_df['CLASSE'] = distribuicao_df['CLASSE'].str.split('/')
```

```
# Separa a lista criada pelo split em diferentes linhas para cada classe obtida
distribuicao_explodido_df = distribuicao_df.explode(column='CLASSE')
```

```
# Retorna a proporção de cada classe de ativo, separando as combinações
round(distribuicao_explodido_df.value_counts(subset='CLASSE', normalize=True).head(10) * 100, 2)
```

```
Out[14]:
CLASSE
Herbicida      30.13
Inseticida     19.21
Fungicida      18.39
Inseticida Microbiológico    8.45
Acaricida       6.03
Fungicida microbiológico    3.77
Nematicida Microbiológico    2.73
Regulador de Crescimento    2.46
Agente Biológico de Controle    1.94
Acaricida Microbiológico    1.60
Name: proportion, dtype: float64
```

Ao analisar as diferentes classes dos ingredientes ativos, percebemos a predominância dos herbicidas, sendo aproximadamente **1 a cada 3** * produtos um herbicida.

Quando separamos os produtos cujas classes abrangiam mais de um tipo de ação, vemos um **aumento** no número de inseticidas, significando que diversos produtos têm ação inseticida, mas também são classificados com outros tipos de ação.

Outro ponto a se destacar é a participação de produtos biológicos/microbiológicos, com aproximadamente **17,49%** * de participação na classificação das top 10 classes. Quando separados, a proporção sobe para **18,49%** *.

```
In [15]:
# Resetando a variável
distribuicao_df = df[['NR_REGISTRO', 'CLASSE']].drop_duplicates('NR_REGISTRO')

e_biologico = distribuicao_df['CLASSE'].str.contains('biológico', case=False, na=False)
distribuicao_df.loc[e_biologico].value_counts().sum()

Out[15]:
np.int64(792)
In [16]:
round(distribuicao_df.loc[e_biologico].value_counts().sum() / produtos_df['NR_REGISTRO'].value_counts().sum() * 100, 2)

Out[16]:
np.float64(18.63)
```

Ao todo, existem **792*** produtos cadastrados como (micro)biológicos, ou seja, **18,63%** * dos produtos da base de dados são produtos (micro)biológicos.

```
In [17]:
```

```
# Selecciona columnas relevantes e remove duplicatas de produtos
biologico_df= df[['NR_REGISTRO', 'TITULAR', 'CLASSE']].drop_duplicates(subset=['NR_REGISTRO'])
```

```
# Cria uma coluna condicional onde se conter a palavra "biológico" na classe, ele retorne True
biologico_df['(MICRO)BIOLÓGICO'] = biologico_df['CLASSE'].str.contains('biológico', case=False, na=False)
```

```
# Cria uma variável com o agrupamento por titular de registro da coluna de (micro)biológicos
biologico = biologico_df.groupby("TITULAR", observed=True)['(MICRO)BIOLÓGICO'].sum()
```

```
# Mostra as 10 empresas com mais produtos (micro)biológicos cadastrados na base
biologico.nlargest(10)
```

```
Out[17]:
TITULAR
Koppert      48
Total        42
Simbiose     34
Vittia       25
Bioma        22
Ballagro     19
Biota        18
Solubio      18
Agrobiológica  17
Genica       17
Name: (MICRO)BIOLÓGICO, dtype: Int64
In [18]:
biologico.nlargest(10).sum()

Out[18]:
np.int64(260)
In [19]:
round(biologico.nlargest(10).sum() / distribuicao_df.loc[e_biologico].value_counts().sum() * 100, 2)

Out[19]:
np.float64(32.83)
```

Aqui vemos as 10 empresas com mais produtos (micro)biológicos cadastrados na base de dados.

Tanto a Koppert quanto a Biotrop (Total) possuem mais de **40** produtos com essa característica, totalizando **90** * produtos ao todo.

Quando somados os 10 maiores players no mercado de produtos (micro)biológicos, temos **260** produtos ao todo, ou ainda, **32,83%** *, quase 1 a cada 3.

Toxicologia e classe ambiental

Os agroquímicos no Brasil são classificados principalmente com base em sua **toxicidade para a saúde humana** e em seu **potencial de periculosidade ambiental**, conforme estabelecido pela Anvisa e regulamentado pelo Sistema Globalmente Harmonizado (GHS).

Classificação Toxicológica (baseada em risco à saúde humana)

A classificação é indicada no rótulo por cores e categorias:

- **Categoria 1 – Produto Extremamente Tóxico** (faixa **vermelha**): produtos com DL50 oral em ratos ≤ 5 mg/kg (formulações sólidas) ou ≤ 20 mg/kg (líquidas).
- **Categoria 2 – Produto Altamente Tóxico** (faixa **vermelha**): DL50 oral entre 5–50 mg/kg (sólidas) ou 20–200 mg/kg (líquidas).
- **Categoria 3 – Produto Moderadamente Tóxico** (faixa **amarela**): DL50 oral entre 50–500 mg/kg (sólidas) ou 200–2.000 mg/kg (líquidas).
- **Categoria 4 – Produto Pouco Tóxico** (faixa **azul**): DL50 oral > 500 mg/kg (sólidas) ou > 2.000 mg/kg (líquidas).
- **Categoria 5 – Produto Improvável de Causar Dano Agudo** (faixa **azul**): produtos que não causam efeitos tóxicos agudos em doses normais.
- **Não classificado** (faixa **verde**): produtos cuja toxicidade não pôde ser definida ou que não foram avaliados.

Nota: A reclassificação realizada pela Anvisa em 2019, com base no GHS, alterou a classificação de 1.924 produtos, com 43 enquadrados na Categoria 1 e 899 na Categoria 5.

Classificação Ambiental (Potencial de Periculosidade Ambiental – PPA)

- **Classe I – Altamente Perigoso**
- **Classe II – Muito Perigoso**
- **Classe III – Perigoso**
- **Classe IV – Pouco Perigoso**

Essa classificação considera fatores como persistência no ambiente, bioacumulação e ecotoxicidade a organismos não-alvo.

Quantos % dos produtos são categoria 1 (extremamente tóxico), 2, 3, 4 ou 5?

```
In [20]:
# Dicionário com o fim de padronizar as classificações toxicológicas dos produtos
mapa_toxico = {
    'Extremamente Tóxico': 'Categoria 1 / Extremamente Tóxico',
    'Categoria 1 Produto Extremamente Tóxico': 'Categoria 1 / Extremamente Tóxico',

    'Altamente Tóxico': 'Categoria 2 / Altamente Tóxico',
    'Categoria 2 Produto Altamente Tóxico': 'Categoria 2 / Altamente Tóxico',

    'Medianamente Tóxico': 'Categoria 3 / Moderadamente Tóxico',
    'Categoria 3 Produto Moderadamente Tóxico': 'Categoria 3 / Moderadamente Tóxico',

    'Pouco Tóxico': 'Categoria 4 / Pouco Tóxico',
    'Categoria 4 Produto Pouco Tóxico': 'Categoria 4 / Pouco Tóxico',

    'Categoria 5 Produto Improvável de Causar Dano Agudo': 'Categoria 5 / Improvável de Causar Dano',

    'Não Classificado - Produto Não Classificado': 'Não Classificado',
    'NÃO DETERMINADO DEVIDO À NATUREZA DO PRODUTO': 'Não Determinado',
    'BAIXA EXPOSIÇÃO PARA USO RESTRITO EM ARMADILHAS': 'Uso Restrito (Armadilhas)',
}

# Criação de coluna com a classificação padronizada
df['CLASSE_TOXICOLOGICA_PADRAO'] = df['CLASSE_TOXICOLOGICA'].map(mapa_toxico)

# Seleciona colunas relevantes e remove duplicatas de produtos
toxicidade_df = df[['NR_REGISTRO', 'CLASSE_TOXICOLOGICA_PADRAO']].drop_duplicates('NR_REGISTRO')

# Retorna a proporção de cada classificação no contexto geral da base de dados
round(toxicidade_df['CLASSE_TOXICOLOGICA_PADRAO'].value_counts(normalize=True) * 100, 2)
```

```
Out[20]:
CLASSE_TOXICOLOGICA_PADRAO
Categoria 5 / Improvável de Causar Dano    52.59
Categoria 4 / Pouco Tóxico                 25.42
Não Classificado                         11.27
Categoria 3 / Moderadamente Tóxico         5.36
Categoria 2 / Altamente Tóxico              3.25
Categoria 1 / Extremamente Tóxico          1.15
Não Determinado                           0.89
Uso Restrito (Armadilhas)                  0.07
Name: proportion, dtype: float64
```

Podemos notar algo bastante positivo com essa tabela: cerca de **78%** * estão classificados ou como *improváveis de causar dano* ou como *pouco tóxicos*.

Apenas **4,4%***** são classificados ou como *extremamente tóxicos* ou como *altamente tóxico*, significando uma baixa proporção frente ao universo do banco de dados.

Qual a distribuição toxicológica das classes de produtos?

```
In [21]:
```

```
# Seleciona colunas relevantes e remove duplicatas de produtos
classes_toxicologia_df = df[['NR_REGISTRO', 'CLASSE', 'CLASSE_TOXICOLOGICA_PADRAO']].drop_duplicates(subset=['NR_REGISTRO'])

# Cria colunas booleanas para as 5 classes toxicológicas
for x in range(1, 6):
    classes_toxicologia_df[f'CATEGORIA_{x}'] = classes_toxicologia_df['CLASSE_TOXICOLOGICA_PADRAO'].str.contains(f'Categoria {x}', case=False, na=False)

# Cria coluna booleana para a classe de "Não Classificado"
classes_toxicologia_df['NAO_CLASSIFICADO'] = classes_toxicologia_df['CLASSE_TOXICOLOGICA_PADRAO'].str.contains('Não Classificado', case=False, na=False)

# Separa as classes em cada linha para melhor compreensão
classes_toxicologia_df['CLASSE'] = classes_toxicologia_df['CLASSE'].str.split('/')
classes_toxicologia = classes_toxicologia_df.explode(column='CLASSE')

# Agrupa por classe e soma a quantidade de produtos por classe
classes_toxicologia = classes_toxicologia.groupby('CLASSE', observed=True)[['CATEGORIA_1', 'CATEGORIA_2', 'CATEGORIA_3', 'CATEGORIA_4', 'CATEGORIA_5', 'NAO_CLASSIFICADO']].sum()

# Adiciona coluna de total para facilitar ordenação
classes_toxicologia['TOTAL'] = classes_toxicologia.sum(axis=1)

# Ordena pelas classes com mais produtos e exibe as top 10
classes_toxicologia = classes_toxicologia.nlargest(n=10, columns='TOTAL')
```

classes_toxicologia								
Out[21]:								
	CATEGORIA_1	CATEGORIA_2	CATEGORIA_3	CATEGORIA_4	CATEGORIA_5	NAO_CLASSIFICADO	TOTAL	
CLASSE								
Herbicida	6	20	27	422	908	65	1448	
Inseticida	25	83	152	339	265	59	923	
Fungicida	5	37	44	268	485	45	884	
Inseticida Microbiológico	0	0	1	26	297	82	406	
Acaricida	1	22	31	103	123	10	290	
Fungicida microbiológico	0	0	1	1	138	41	181	
Nematicida Microbiológico	0	0	0	0	84	47	131	
Regulador de Crescimento	1	0	2	22	62	31	118	
Acaricida Microbiológico	0	0	0	2	69	6	77	
Agente Biológico de Controle	0	0	0	1	0	54	55	

Herbicidas são os mais numerosos, mas também os menos tóxicos, onde a maioria deles registrados hoje são de **baixa toxicidade** (Categorias 4 e 5). Ou seja, apesar de representarem 1/3 dos produtos, herbicidas têm perfil toxicológico predominantemente baixo.

Inseticidas têm a **maior proporção de produtos nas categorias mais tóxicas** (1, 2 e 3) entre as classes principais, onde quase 30% dos inseticidas registrados são classificados como moderadamente a extremamente tóxicos. Isso se deve ao fato dos mecanismos de ação dos produtos atuarem por vias que são análogas nos seres humanos.

Produtos microbiológicos são **extremamente seguros** com praticamente **zero** nas categorias tóxicas. Por exemplo, 100% dos nematicidas microbiológicos são classificados como baixa toxicidade ou não classificados.

Produtos biológicos frequentemente **não se enquadram** nas categorias toxicológicas tradicionais, o que faz sentido, já que são organismos vivos, não moléculas químicas.

Biológicos são **proporcionalmente muito mais seguros** que químicos. Menos de 1%* dos produtos microbiológicos estão nas categorias de maior toxicidade, contra ~10-15%* dos químicos.

Fungicidas são majoritariamente de baixa toxicidade, mas não tanto quanto herbicidas.

Reguladores de crescimento são predominantemente seguros ou não classificados.