

Projeto sobre os Acidentes de Trabalho no Brasil

Beatriz Ornelas Miranda

Jean Carlos Meurer

Resumo

Este projeto tem a finalidade de mostrar a quantidade de acidentes de trabalho em determinado período e quais profissões possuem maior índice de acidentes, seja pela falta de um supervisor, pela negligência, ou até mesmo pela profissão ser de alto risco. Acidentes estes que podem influenciar a saúde do trabalhador ou o mais importante, a vida.

1 Introdução

Muitas das vezes, em nosso ambiente de trabalho, não levamos em consideração os riscos ao qual estamos expostos. Seja por falta de conhecimento, falta de atenção ou até mesmo descaso. Com a quantidade de acidentes de trabalho ou doença laboral de profissionais brasileiros na atualidade, ao evidenciar a internet, especificamente em sites de notícias, temos ali ao nosso dispor, vários artigos sobre a crescente taxa de óbitos ou acidentes de trabalhos de diversos profissionais.

Para tanto, o nosso objetivo é mostrar através dos dados, os impactos negativos dos diferentes riscos pelos quais os trabalhadores possam estar expostos dentro de uma empresa. Temos também como intuito, analisar os dados de quais são as ocupações com mais registros de acidente de trabalho, onde elas se encontram, o gênero predominante e também a região no Brasil em que mais ocorrem acidentes de trabalho.

Para realizar este trabalho, foi preciso selecionar bancos de dados públicos brasileiros, de forma que os resultados obtidos sirvam para identificar e alertar as empresas e os profissionais que desempenham funções de risco ou simples acidentes de trabalhos ocorridos no território brasileiro.

2 Descrição das bases de dados utilizadas

Após termos definido o objetivo do projeto, buscamos por dados públicos que pudessem ser analisados para responder nossos questionamentos. Os dados foram encontrados na plataforma do governo por meio do site www.gov.br. Estes dados estão apresentados na forma de uma planilha, onde podemos associar as colunas trazendo informações de quantidade de acidentes no período de 2018 a 2020, distribuídas em: quantidades por gênero, idade, motivo, atividade econômica e região.

Os tópicos abaixo resumem o conteúdo das planilhas que se encontram disponíveis nos links citados na bibliografia:

- Quantidade de acidentes de trabalho com totalizador nacional;
- Quantidade de acidentes de trabalho na região Norte;
- Quantidade de acidentes de trabalho na região Nordeste;
- Quantidade de acidentes de trabalho na região Sudeste;
- Quantidade de acidentes de trabalho na região Sul;
- Quantidade de acidentes de trabalho na região Centro-Oeste;
- Quantidade de acidentes de trabalho por idade e gênero;
- Quantidade de acidentes de trabalho por ocorrência mensal.

Todos os tópicos citados acima foram acessados na data de 18 de Março de 2022, onde os dados foram transferidos para uma planilha com o intuito de deixá-los salvos externamente ao site do governo para que possamos manipulá-los no decorrer do projeto.

Além dos dados referentes à quantidade de acidentes divididos por atividade econômica sendo apresentados apenas com o código desta atividade, buscamos também uma tabela que pudesse traduzir o código da Classificação Nacional de Atividades Econômicas (CNAE). Abaixo está a figura 1 mostrando os dados obtidos em uma planilha baixada no site do IBGE:

Seção	Divisão	Grupo	Classe	Subclasse	
A					AGRICULTURA, PECUÁRIA, PRODUÇÃO FLORESTAL, PESCA E AQUICULTURA
	01				AGRICULTURA, PECUÁRIA E SERVIÇOS RELACIONADOS
		01.1			Produção de lavouras temporárias
			01.11-3		Cultivo de cereais
				0111-3/01	Cultivo de arroz
				0111-3/02	Cultivo de milho
				0111-3/03	Cultivo de trigo
				0111-3/99	Cultivo de outros cereais não especificados anteriormente
			01.12-1		Cultivo de algodão herbáceo e de outras fibras de lavoura temporária
				0112-1/01	Cultivo de algodão herbáceo
				0112-1/02	Cultivo de juta
				0112-1/99	Cultivo de outras fibras de lavoura temporária não especificadas anteriormente
			01.13-0		Cultivo de cana-de-açúcar
				0113-0/00	Cultivo de cana-de-açúcar
			01.14-8		Cultivo de fumo
				0114-8/00	Cultivo de fumo
			01.15-6		Cultivo de soja
				0115-6/00	Cultivo de soja
			01.16-4		Cultivo de oleaginosas de lavoura temporária, exceto soja
				0116-4/01	Cultivo de amendoim
				0116-4/02	Cultivo de girassol
				0116-4/03	Cultivo de mamona
				0116-4/99	Cultivo de outras oleaginosas de lavoura temporária não especificadas anteriormente
			01.19-9		Cultivo de plantas de lavoura temporária não especificadas anteriormente
				0119-9/01	Cultivo de abacaxi
				0119-9/02	Cultivo de alho
				0119-9/03	Cultivo de batata-inglesa
				0119-9/04	Cultivo de cebola
				0119-9/05	Cultivo de feijão

Figura 1: Planilha dos códigos de Classificação Nacional de Atividades Econômicas (CNAE).

Para um melhor entendimento dos dados que conseguimos, a seguir apresentamos na figura 2 os dados de quantidades de acidentes do trabalho com o totalizador nacional. Um totalizador de acidentes com Comunicação de Acidentes do Trabalho (CAT) registrado que, conforme informado pelo site Gov.br, trata-se de um serviço para o trabalhador ou empresa realizar a comunicação de um acidente de trabalho, e também dividindo as quantidades de acidentes pelo motivo.

CNAE	QUANTIDADE DE ACIDENTES DO TRABALHO																				
	Total			Com CAT Registrada															Sem CAT Registrada		
				Total com CAT			Motivo														
							Típico			Trajeto			Doença do Trabalho								
													2018	2019	2020	2018	2019	2020			
2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020				
Total	586.017	586.857	445.814	481.993	487.739	403.694	363.314	375.300	313.575	106.082	102.405	59.520	10.597	10.034	30.599	104.024	99.118	42.120			
111	448	555	926	433	537	917	395	479	846	35	51	63	3	7	8	15	18	9			
112	148	172	139	141	167	130	129	157	123	12	10	7	-	-	-	7	5	9			
113	1.939	1.817	1.338	1.727	1.641	1.285	1.629	1.539	1.226	89	93	55	9	9	4	212	176	53			
114	28	25	22	25	17	21	18	17	15	7	-	6	-	-	-	3	8	1			
115	1.450	1.645	1.335	1.375	1.573	1.309	1.234	1.414	1.196	139	147	105	2	12	6	75	72	26			
116	48	27	20	43	20	17	36	19	18	6	1	1	1	-	-	5	7	3			
119	406	389	312	382	359	296	330	309	255	47	49	39	5	1	2	24	30	16			
121	187	243	220	173	234	216	128	181	187	44	48	27	1	5	2	14	9	4			
122	86	109	65	85	109	64	65	95	55	19	13	9	1	1	-	1	-	1			
131	797	939	778	748	875	748	699	823	711	44	46	35	5	6	2	49	64	30			
132	649	677	679	628	644	654	541	580	622	85	62	28	2	2	4	21	33	25			
133	938	874	637	826	788	589	729	717	544	87	70	38	10	1	7	112	86	48			
134	494	452	245	478	429	237	439	402	222	38	27	15	1	-	-	16	23	8			
135	35	40	19	30	37	19	27	30	17	2	7	2	1	-	-	6	3	-			

Figura 2: Planilha de quantidade de acidentes do trabalho com dados nacionais.

Seguindo uma estrutura semelhante da figura 2, a figura 3 mostra os dados de acidentes de trabalho divididos por grupos de idade e sexo.

Idade e sexo		QUANTIDADE DE ACIDENTES DO TRABALHO																	
		Com CAT Registrada															Sem CAT Registrada		
		Total			Total			Típico			Motivo			Doença do Trabalho					
2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020		
Total	Total	586.017	586.857	445.814	481.993	487.739	403.694	363.314	375.300	313.575	106.082	102.405	59.520	10.597	10.034	30.599	104.024	99.118	42.120
	Masculino	387.346	386.601	294.912	321.017	324.542	268.177	250.672	257.992	219.359	64.308	60.870	38.580	6.037	5.680	10.238	66.329	62.059	26.735
	Feminino	198.650	200.111	149.595	160.955	163.053	134.210	112.621	117.181	93.235	43.774	41.522	20.734	4.560	4.350	20.241	37.695	37.058	15.385
	Ignorado	21	145	1.307	21	144	1.307	21	127	981	-	13	206	-	4	120	-	1	-
Até 19 anos	Total	12.585	11.675	8.711	11.424	10.791	8.403	8.322	8.070	6.601	3.056	2.675	1.664	46	46	138	1.161	884	308
	Masculino	9.449	8.697	6.770	8.539	8.012	6.526	6.501	6.286	5.264	2.007	1.698	1.203	31	28	59	910	685	244
	Feminino	3.136	2.938	1.941	2.885	2.739	1.877	1.821	1.748	1.337	1.049	974	461	15	17	79	251	199	64
	Ignorado	-	40	-	-	40	-	-	38	-	-	3	-	-	1	-	-	-	-
20 a 24 anos	Total	76.586	77.881	61.542	68.539	70.363	58.716	50.588	52.926	45.795	17.491	16.985	10.784	460	472	2.137	8.047	7.518	2.826
	Masculino	56.003	56.722	44.692	50.120	51.409	42.650	38.524	40.837	34.447	11.341	11.114	7.468	255	258	735	5.683	5.313	2.042
	Feminino	20.581	21.159	16.584	18.417	18.954	15.800	12.062	12.889	11.134	6.150	5.851	3.277	205	214	1.389	2.164	2.205	784
	Ignorado	2	-	266	2	-	266	2	-	214	-	-	39	-	-	13	-	-	-
25 a 29 anos	Total	88.255	87.982	68.213	76.999	77.191	64.027	56.997	58.597	50.064	19.138	17.666	10.414	864	928	3.549	11.296	10.791	4.186
	Masculino	60.405	60.020	46.830	52.762	52.938	44.015	40.743	41.739	36.176	11.540	10.672	6.706	479	527	1.133	7.643	7.082	2.815
	Feminino	27.883	27.912	20.786	24.230	24.203	19.415	16.247	16.815	13.470	7.598	6.988	3.589	385	400	2.356	3.653	3.709	1.371
	Ignorado	7	50	597	7	50	597	7	43	418	-	6	119	-	1	60	-	-	-

Figura 3: Planilha de quantidade de acidentes do trabalho por grupo de idade e sexo.

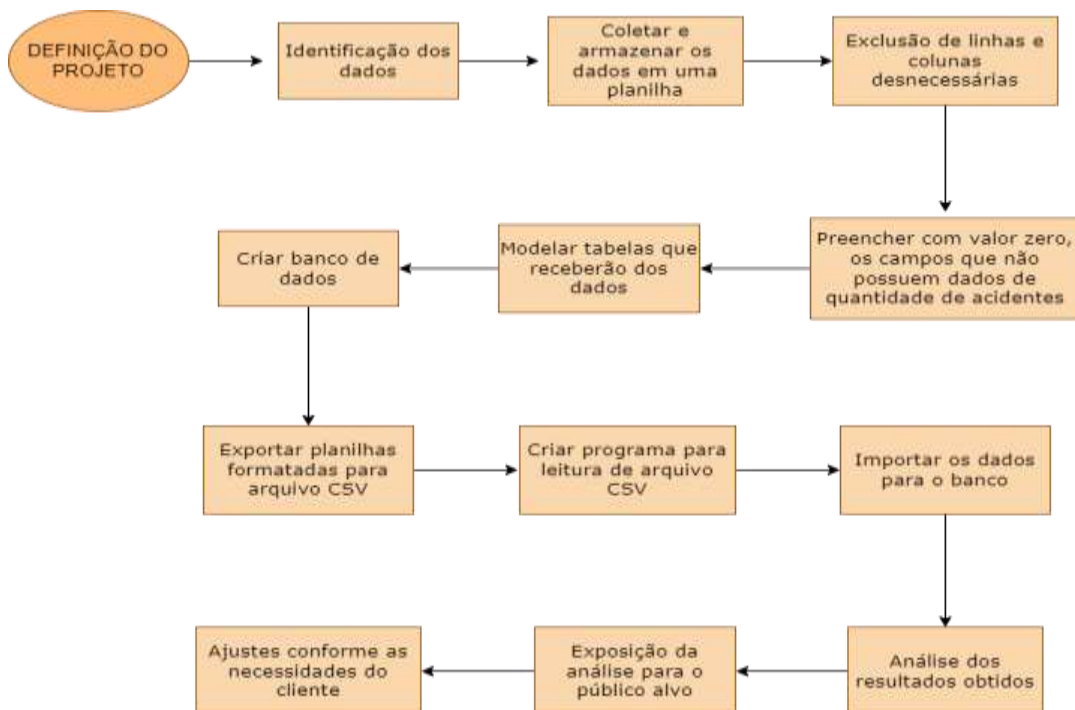
A apresentação dos dados da planilha com o Total Nacional e as planilhas divididas por regiões estão no mesmo formato, tendo ainda outra planilha com dados de idade e sexo, sendo mais fácil o entendimento no quadro abaixo:

Nome da coluna	Descrição	Tipo
CNAE	Nesta coluna, é apresentado os códigos da Classificação Nacional de Atividades Econômicas	Númérico
Quantidade	A maioria das colunas trazem dados que são referentes à	Númérico

	quantidade de acidentes, onde em algumas é apresentado um totalizado	
Motivo	Motivo do acidente de trabalho, sendo dividido em 3 específicos: Típico, Trajeto e Doença do Trabalho.	Caracteres
Sem CAT Registrada	Quantidade de acidentes que não ocorreram Comunicação de Acidentes do Trabalho.	Númérico
Idade	Quantidade de acidentes de trabalho separados por grupos de idade.	Númérico
sexo	Quantidade de acidentes de trabalho separados por sexo.	Númérico

3 ARQUITETURA

É possível observar como serão os passos no decorrer do projeto conforme a arquitetura mostrada no diagrama a seguir:



4 METODOLOGIA

Os dados que foram coletados no site do governo atualmente estão na mesma estrutura de planilha que foram encontrados e exemplificados na descrição das bases de dados, tendo totalizadores de registros junto com os dados únicos que serão analisados. Estes dados serão importados para um banco de dados por meio de um programa que leia planilhas em formato CSV.

Para facilitar a importação dos dados de acidentes de trabalho em um banco de dados, pensando em uma melhor leitura, visualização e organização das informações nas tabelas com os dados regionais, iremos remover as colunas totalizadoras, criar novas colunas que representam as informações do ano em que o dado foi coletado e qual o motivo deste acidente, onde atualmente estas informações estão dispersas em diversas colunas, o que dificulta a execução de consultas SQL. Sendo assim, as novas colunas terão os nomes conforme a tabela a seguir:

Nome da coluna	Descrição	Tipo
ANO	Na coluna, será informado o ano em que o dado correspondente foi coletado	Numérico
MOTIVO	Será informado o motivo do acidente, sendo as informações de “Trajeto”, “Típico”, “Doença do Trabalho” e “Sem registro” conforme dados coletados	Caracteres

A imagem a seguir (figura 4), representa com destaque em vermelho as colunas e linhas que serão excluídas, por se tratarem de dados totalizadores:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	QUANTIDADE DE ACIDENTES DO TRABALHO																		
2	Com CAT Registrada																		
3	Motivo																		
4																			
5	CNAE	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020
6	Total	25 536	24 816	19 286	19 581	18 151	17 264	14 488	14 127	13 062	4 046	3 532	2 472	547	452	1 670	6 855	5 865	2 082
7	111	26	29	43	24	26	43	20	21	32	3	7	10	1	-	1	2	1	-
8	113	15	22	8	12	19	8	10	19	8	2	-	-	-	-	3	3	-	-
9	115	61	73	74	67	78	71	44	62	67	13	6	14	-	2	-	4	3	3
10	119	3	2	4	3	2	4	3	1	4	-	1	-	-	-	-	-	-	-
11	121	1	4	1	1	4	1	1	1	-	-	2	1	-	1	-	-	-	-

Figura 4: Exemplo dos dados que serão excluídos das planilhas de acidentes de trabalho.

Tendo como resultado uma planilha apenas com os dados únicos de cada ano, qual CNAE está vinculado a cada registro e seu devido motivo. Também serão substituídos todos os registros que contenham o sinal de hífen pelo número zero para realizar a leitura correta dos registros sem ocorrência de acidentes. Outro ponto importante a se observar é que no final de cada planilha há uma linha de registros ignorados que iremos excluir também por não ser relevante ao nosso foco

de análise e por não ter uma referência de qual área econômica representa. A figura abaixo representa o resultado das alterações citadas previamente:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	CNAE	Típico			Trajeto			Doença do Trabalho			Sem CAT		
2		2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020
3	111	20	21	32	3	7	10	1	0	1	2	1	0
4	113	10	19	8	2	0	0	0	0	0	3	3	0
5	115	44	62	57	13	6	14	0	2	0	4	3	3
6	119	3	1	4	0	1	0	0	0	0	0	0	0
7	121	1	1	0	0	2	1	0	1	0	0	0	0

Figura 5: Planilha de quantidade de acidentes do trabalho por grupo de idade e sexo.

As alterações deverão ser feitas nas cinco tabelas de dados regionais, e após isto, serão exportados para arquivo de extensão CSV para que o programa que será criado possa fazer a leitura correta dos dados.

Para criar o programa, teremos como base o conhecimento adquirido em Algoritmos e Lógica de Programação, onde também será necessário pesquisar métodos e adquirir mais conhecimento externo ao curso.

Por fim, iremos criar um banco de dados PostgreSQL, o qual escolhemos por ser um gerenciador de banco de dados relacional de código aberto, o que representa uma ferramenta sem custo, onde será criada uma tabela para armazenar os dados de quantidade de acidentes individualmente em cada linha, diferentemente da planilha atual onde há vários registros de quantidade em uma única linha. A figura 6 a seguir exemplifica como ficarão os dados importados nesta tabela:

ID	CNAE	REGIAO	MOTIVO	ANO	QUANTIDADE
1	111	Norte	Típico	2018	20
2	111	Norte	Típico	2019	21
3	111	Norte	Típico	2020	32
4	111	Norte	Trajeto	2018	6
5	111	Norte	Trajeto	2019	8
6	111	Norte	Trajeto	2020	13
7	111	Norte	Doença do trabalho	2018	1
8	111	Norte	Doença do trabalho	2019	1

Figura 6: exemplo da tabela que será criada ao importar os dados ao banco.

Conforme é possível observar na figura anterior, há uma coluna chamada CNAE, onde iremos criar uma outra tabela no banco de dados para armazenar as informações referentes ao

CNAE que serão úteis para o projeto, assim podendo fazer uma relação entre o código CNAE da tabela de acidentes por região, e qual área econômica determinado registro representa.

Observamos no site do governo (conforme figura 7) que a coluna que representa o CNAE é apresentada por um número de 4 dígitos, então buscando entender a estrutura do código CNAE que encontramos no site do IBGE, encontramos no site “contabilix.com” uma informação mostrando que o código CNAE é dividido em cinco níveis hierárquicos, sendo eles: seção, divisão, grupo, classe e subclasse.

Comparando a estrutura dos códigos CNAEs das planilhas de acidentes do site do governo com os códigos da planilha de CNAEs, identificamos que eles representam o código do nível de classe sem o dígito verificador. Sendo assim, para simplificar a planilha e identificar corretamente as áreas de atividade econômica que cada código representa, filtramos apenas os dados de classe como é possível observar na figura 8.

CNAE	QUANTIDADE DE ACIDENTES DO TRABALHO					
	Total			Com CAT Registrada		
				Total		
	2018	2019	2020	2018	2019	2020
Total	47.141	47.608	38.236	39.039	39.598	35.037
O111	70	119	260	69	118	259
O112	105	130	110	102	127	106
O113	330	292	239	277	247	228

Figura 7: Exemplo da coluna de códigos CNAE no site do governo.

Seção	Divisão	Grupo	Classe	Subclasse
			01.11-3	Cultivo de cereais
			01.12-1	Cultivo de algodão herbáceo e de outras fibras de lavoura temporária
			01.13-0	Cultivo de cana-de-açúcar
			01.14-8	Cultivo de fumo
			01.15-6	Cultivo de soja
			01.16-4	Cultivo de oleaginosas de lavoura temporária, exceto soja
			01.19-9	Cultivo de plantas de lavoura temporária não especificadas anteriormente
			01.21-1	Horticultura
			01.22-9	Cultivo de flores e plantas ornamentais
			01.31-8	Cultivo de laranja
			01.32-6	Cultivo de uva
			01.33-4	Cultivo de frutas de lavoura permanente, exceto laranja e uva
			01.34-2	Cultivo de café
			01.35-1	Cultivo de cacau
			01.39-3	Cultivo de plantas de lavoura permanente não especificadas anteriormente
			01.41-5	Produção de sementes certificadas

Figura 8: Tabela CNAE com filtro de classe

Comparando as figuras 7 e 8, é possível identificar um padrão na numeração, porém a formatação está diferente, tendo um dígito extra no final do código da figura 8. Para resolver isto, precisamos formatar os números dos códigos de CNAE, utilizando a ferramenta de localizar e substituir disponível no editor de planilhas, e também remover todos os pontos e hífens dos registros. Também utilizamos uma fórmula para obter apenas os dígitos correspondentes à classe, sem o último dígito, dessa forma obtendo os códigos no mesmo padrão conforme a figura 9:

	A	B
1	CODIGO_CNAE	ATIVIDADE
2	111	Cultivo de cereais
3	112	Cultivo de algodão herbáceo e de outras fibras de lavoura temporária
4	113	Cultivo de cana-de-açúcar
5	114	Cultivo de fumo
6	115	Cultivo de soja
7	116	Cultivo de oleaginosas de lavoura temporária, exceto soja
8	119	Cultivo de plantas de lavoura temporária não especificadas anteriormente
9	121	Horticultura
10	122	Cultivo de flores e plantas ornamentais
11	131	Cultivo de laranja
12	132	Cultivo de uva
13	133	Cultivo de frutas de lavoura permanente, exceto laranja e uva
14	134	Cultivo de café
15	135	Cultivo de cacau
16	139	Cultivo de plantas de lavoura permanente não especificadas anteriormente
17	141	Produção de sementes certificadas
18	142	Produção de mudas e outras formas de propagação vegetal, certificadas

Figura 9: Tabela CNAE com formatação completa.

Os dados da planilha de idade e sexo também terão um tratamento de exclusão nas linhas e colunas de totais e dados que serão ignorados, com o objetivo de preparar os dados em uma planilha de formato CSV para poder importar em outra tabela no nosso banco de dados. Assim como no exemplo das tabelas de dados regionais, será criada uma coluna para ter a informação do ano de ocorrência também.

Na figura 10 a seguir, está destacado em vermelho as linhas de totais e linhas ignoradas que serão excluídas, e na figura 11 a representação da planilha formatada que será utilizada para importar os dados ao banco de dados.

		QUANTIDADE DE ACIDENTES DO TRABALHO Com CAT Registrada																				
		Total						Motivo						Doença do Trabalho						Sem CAT Registrada		
		2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020
Idade e sexo	Total	189.017	168.327	144.814	401.993	451.739	483.694	363.318	375.359	312.975	198.583	112.405	58.520	10.997	19.104	36.599	164.024	89.118	43.129			
	Masculinos	187.346	166.921	139.843	321.817	328.542	368.177	258.472	287.982	219.359	64.305	60.878	18.500	8.057	9.889	10.230	98.329	62.859	26.798			
	Femininas	199.858	200.111	149.945	168.955	163.892	134.216	112.621	117.181	89.235	43.774	41.622	28.754	4.560	8.350	26.341	37.698	37.659	15.388			
	Ignorado	21	145	1.367	31	844	1.507	31	127	801	13	206	13	206	13	119	1	1	1			
Até 19 anos	Total	12.569	11.879	8.711	11.424	19.791	8.403	8.322	8.070	6.991	3.058	2.979	1.864	46	46	119	1.191	554	100			
	Masculinos	9.448	8.991	6.779	8.539	8.812	6.626	6.591	6.256	5.264	2.587	1.990	1.293	31	28	91	916	485	244			
	Femininas	3.119	2.928	1.943	2.885	7.739	1.877	1.831	1.748	1.727	5.649	674	487	15	17	19	251	199	54			
	Ignorado	49	1	1	40	1	1	30	1	1	1	1	1	1	1	1	1	1	1			
20 a 24 anos	Total	16.388	17.261	61.542	68.539	78.363	58.718	58.508	62.926	45.785	17.491	18.969	18.754	488	472	2.137	6.047	7.618	2.820			
	Masculinos	96.903	99.723	44.992	90.129	81.409	42.990	38.524	40.931	34.447	11.341	11.114	7.488	255	250	739	5.953	6.313	2.042			
	Femininas	20.581	21.159	16.554	18.417	18.954	15.909	12.662	12.899	11.334	6.150	5.851	3.277	229	214	1.399	2.094	2.295	784			
	Ignorado	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2			
25 a 29 anos	Total	88.299	87.582	68.713	79.889	77.191	64.877	54.967	55.591	55.064	18.039	17.889	18.414	584	625	3.549	11.796	18.791	4.190			
	Masculinos	80.495	80.629	46.320	52.782	52.936	44.819	48.745	47.739	38.179	11.548	10.672	8.798	475	527	1.133	7.843	7.682	2.815			
	Femininas	27.853	27.612	28.756	24.233	24.203	19.419	16.247	16.815	13.479	7.595	6.988	3.988	385	400	2.258	3.953	3.799	1.371			
	Ignorado	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2			

Figura 10: Dados que serão excluídos da planilha de acidentes por grupo de idade e sexo.

		Típico			Trajeto			Doença do Trabalho			Sem CAT		
		2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020
Até 19 anos	Masculino	6.501	6.286	5.264	2.007	1.698	1.203	31	28	59	910	685	244
	Feminino	1.821	1.748	1.337	1.049	974	461	15	17	79	251	199	64
20 a 24 anos	Masculino	38.524	40.037	34.447	11.341	11.114	7.468	255	258	735	5.883	5.313	2.042
	Feminino	12.062	12.889	11.134	6.150	5.851	3.277	205	214	1.389	2.164	2.205	784
25 a 29 anos	Masculino	40.743	41.739	36.176	11.540	10.672	6.706	479	527	1.133	7.643	7.082	2.815
	Feminino	16.247	16.815	13.470	7.598	6.988	3.589	385	400	2.356	3.653	3.709	1.371
30 a 34 anos	Masculino	41.120	41.401	34.200	11.263	10.391	6.188	851	806	1.532	9.005	8.511	3.429
	Feminino	18.971	19.113	15.305	7.477	7.035	3.436	693	678	3.367	5.013	4.881	2.002
35 a 39 anos	Masculino	37.882	38.319	32.498	9.632	9.161	5.770	1.098	951	1.851	9.808	9.301	4.211
	Feminino	19.177	19.815	15.686	6.770	6.410	3.125	941	867	4.084	6.131	5.985	2.516
40 a 44 anos	Masculino	28.574	30.436	26.775	6.860	6.594	4.279	1.024	912	1.657	9.208	8.759	4.012
	Feminino	15.430	16.411	13.431	4.988	4.814	2.553	837	789	3.512	5.954	5.901	2.683
45 a 49 anos	Masculino	22.108	23.000	19.824	4.832	4.644	3.081	915	858	1.232	8.403	7.845	3.530
	Feminino	11.847	12.306	9.667	3.667	3.533	1.822	738	644	2.523	5.620	5.569	2.353
50 a 54 anos	Masculino	17.163	17.738	14.938	3.448	3.304	2.058	744	704	1.049	7.599	7.169	3.191
	Feminino	8.875	9.288	7.140	3.076	2.921	1.301	508	468	1.614	5.239	5.037	2.117
55 a 59 anos	Masculino	10.736	11.391	9.962	2.036	1.982	1.236	421	434	669	5.285	5.019	2.244
	Feminino	5.430	5.790	4.301	1.922	1.908	871	192	213	966	2.833	2.798	1.206
60 a 64 anos	Masculino	5.348	5.448	3.835	964	939	451	135	155	213	2.260	2.090	921
	Feminino	2.015	2.173	1.325	774	763	235	38	49	262	663	655	241
65 a 69 anos	Masculino	1.485	1.643	1.124	290	277	108	53	31	64	285	269	92
	Feminino	576	622	348	224	251	52	5	9	71	152	103	38
70 anos e mais	Masculino	487	552	314	95	93	32	31	16	44	39	13	4
	Feminino	169	211	90	79	73	12	3	2	18	21	15	10

Figura 11: dados que serão utilizados para importação no banco de dados.

A tabela de registros de idade e sexo será utilizada para comparativos gráficos com os totalizadores obtidos por meio de consultas realizadas na tabela de quantidades de acidentes. Não será possível relacionar estas tabelas para observar a idade e sexo das pessoas em comparativos regionais, onde será útil nos comparativos nacionais.

5 CAPTURA DOS DADOS

Para capturar os dados, primeiramente preparamos o banco de dados que irá recebê-los, o qual optamos por criar em Postgres por ser um SGBD bastante utilizado no mercado e também por ser uma tecnologia de código aberto, o que representa um software gratuito.

Com o Postgres instalado, criamos o banco de dados, tendo seu nome definido como 'PROJETO', e criamos as tabelas 'QTDACIDENTESREG', 'QTDIDADESEXO' e 'CNAE', que irão armazenar os registros individuais das quantidades de acidentes por região, os registros totais de acidentes divididos por idade e sexo e a descrição das atividades econômicas respectivamente.

A estrutura das três tabelas foi definida com base nos dados coletados para armazenar individualmente cada registro de acidente, tendo a necessidade de vincular a cada registro o número

do CNAE, o motivo, a região, o ano e a quantidade registrada, sendo cada um destes representado por uma coluna na tabela 'QTDACIDENTESREG'.

Para a tabela 'QTDIDADESEXO', no lugar das colunas CNAE e região, teremos as colunas 'Grupo_idade' e 'sexo', armazenando a quantidade de acidentes agrupadas por um grupo de idade e separando entre os sexos masculino e feminino. Por final a tabela CNAE possui apenas duas colunas sendo uma delas o próprio código CNAE e a outra contendo a descrição da atividade econômica. A figura 12 a seguir mostra os comandos completos para criação das tabelas com os campos previamente citados:

```
CREATE DATABASE PROJETO;

CREATE TABLE QTDACIDENTESREG (
    ID_ACIDENTE SERIAL,
    CODIGOCNAE SMALLINT NOT NULL,
    REGIAO VARCHAR(15) NOT NULL,
    MOTIVO VARCHAR(30) NOT NULL,
    ANO SMALLINT NOT NULL,
    QUANTIDADE INT NOT NULL,

    CONSTRAINT PK_ID_ACIDENTE PRIMARY KEY (ID_ACIDENTE)
);

CREATE TABLE QTDIDADESEXO (
    ID_IDADESEXO SERIAL,
    GRUPO_IDADE VARCHAR(15) NOT NULL,
    SEXO VARCHAR(10) NOT NULL,
    MOTIVO VARCHAR(30) NOT NULL,
    ANO SMALLINT NOT NULL,
    QUANTIDADE INT NOT NULL,

    CONSTRAINT PK_ID_IDADESEXO PRIMARY KEY (ID_IDADESEXO)
);

CREATE TABLE CNAE (
    CODIGOCNAE SMALLINT PRIMARY KEY NOT NULL,
    ATIVIDADE VARCHAR(100) NOT NULL
);
```

Figura 12: Script para criação do banco de dados e das tabelas.

Após definirmos a estrutura das tabelas, utilizamos um programa tendo seu código escrito em C#, no qual adaptamos a nossa necessidade para realizar a leitura dos arquivos CSV para poder gerar os comandos de inserção.

No exemplo a seguir (figura 13) está um trecho do código para importar a planilha de acidentes da região Norte, onde é possível visualizar que colocamos doze linhas de impressão para gerar os comandos, pois em cada linha do arquivo CSV contém um total de 12 registros de quantidade de acidentes que estão divididos pelo ano de ocorrência e motivo.

```

static void Main(string[] args)
{
    String linha;
    String[] valores;

    AbreArquivos();

    Console.WriteLine("Lendo valores do arquivo CSV e gerando script SQL...");

    while (!arquivo_in.EndOfStream)
    {
        linha = arquivo_in.ReadLine();

        valores = linha.Split(',');

        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Tipico', 2015, {valores[1]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Tipico', 2015, {valores[2]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Tipico', 2020, {valores[3]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Trajeto', 2018, {valores[4]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Trajeto', 2019, {valores[5]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Trajeto', 2020, {valores[6]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Doença do trabalho', 2018, {valores[7]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Doença do trabalho', 2019, {valores[8]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Doença do trabalho', 2020, {valores[9]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Sem Registro', 2018, {valores[10]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Sem Registro', 2019, {valores[11]})");
        arquivo_out.WriteLine($"INSERT INTO QTDACIDENTES VALUES(DEFAULT, {valores[0]}, 'norte', 'Sem Registro', 2020, {valores[12]})");

        Console.WriteLine("Script SQL gerado com sucesso!");

        FechaArquivos();
    }
}

```

Figura 13: Trecho de código para importar os dados da planilha da região norte.

Os dados que o programa fará a leitura podem ser observados na figura 14, o qual é o mesmo arquivo mostrado na figura 5, porém exportado para CSV, utilizando o separador de vírgula, a primeira coluna representa o código CNAE, e todas as outras doze colunas são quantidades de acidentes que ocorreram em um ano específico e com um determinado motivo, o qual foi configurado no programa para fazer a leitura correta de cada coluna e assim transformá-los em registros únicos:

1	111,20,21,32,3,7,10,1,0,1,2,1,0
2	113,10,19,8,2,0,0,0,0,0,3,3,0
3	115,44,62,57,13,6,14,0,2,0,4,3,3
4	119,3,1,4,0,1,0,0,0,0,0,0,0
5	121,1,1,0,0,2,1,0,1,0,0,0,0
6	122,1,1,0,0,0,1,0,1,0,0,0,0
7	131,1,4,2,0,1,1,0,0,0,0,2,0

Figura 14: Dados de quantidade de acidentes da planilha da região norte exportados em arquivo CSV.

Para executar o programa, utilizamos o site 'replit.com' que nos permite executar programas em C#, no qual enviamos os arquivos CSVs, executamos o programa, que gera o script SQL pronto para inserir os dados conforme exemplo das figuras 15 e 16. O trecho de código mostrado na figura 15 representa o mesmo programa exemplificado na figura 13.

```

1  using System;
2  using System.IO;
3
4  class Program
5  {
6      static StreamWriter arquivo_out;
7      static StreamReader arquivo_in;
8
9      static void AbreArquivos()
10     {
11         arquivo_in = new StreamReader("Idadesexo.csv");
12         arquivo_out = new StreamWriter("idadesexo.sql");
13     }
14
15     static void FechaArquivos()
16     {
17         arquivo_in.Close();
18         arquivo_out.Close();
19     }
20
21     static void Main(string[] args)
22     {
23         String linha;
24         String[] valores;
25
26         AbreArquivos();
27
28         Console.WriteLine("Lendo valores do arquivo CSV e gravando script SQL...");

```

Figura 15: Plataforma Replit para executar o programa e gerar os scripts de inserção de dados.

```

1  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Típico',2018,20);
2  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Típico',2019,21);
3  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Típico',2020,32);
4  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Trajeto',2018,3);
5  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Trajeto',2019,7);
6  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Trajeto',2020,10);
7  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Doença do trabalho',2018,1);
8  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Doença do trabalho',2019,0);
9  INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Doença do trabalho',2020,1);
10 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Sem Registro',2018,2);
11 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Sem Registro',2019,1);
12 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,111,'norte','Sem Registro',2020,0);
13 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Típico',2018,10);
14 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Típico',2019,19);
15 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Típico',2020,8);
16 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Trajeto',2018,2);
17 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Trajeto',2019,0);
18 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Trajeto',2020,0);
19 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Doença do trabalho',2018,0);
20 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Doença do trabalho',2019,0);
21 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Doença do trabalho',2020,0);
22 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Sem Registro',2018,3);
23 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Sem Registro',2019,3);
24 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,113,'norte','Sem Registro',2020,0);
25 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Típico',2018,44);
26 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Típico',2019,62);
27 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Típico',2020,57);
28 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Trajeto',2018,13);
29 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Trajeto',2019,6);
30 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Trajeto',2020,14);
31 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Doença do trabalho',2018,0);
32 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Doença do trabalho',2019,2);
33 INSERT INTO QTDACIDENTESREG VALUES (DEFAULT,115,'norte','Doença do trabalho',2020,0);

```

Figura 16: Trecho de arquivo com comandos de inserção de dados para a tabela 'QTDACIDENTESREG'.

Após criação do banco de dados e das tabelas (figura 17), executamos todos os scripts para inserir os dados, que é possível observar conforme a figura 18 que foram inseridos 36408 (trinta e seis mil quatrocentos e oito) registros de acidentes.

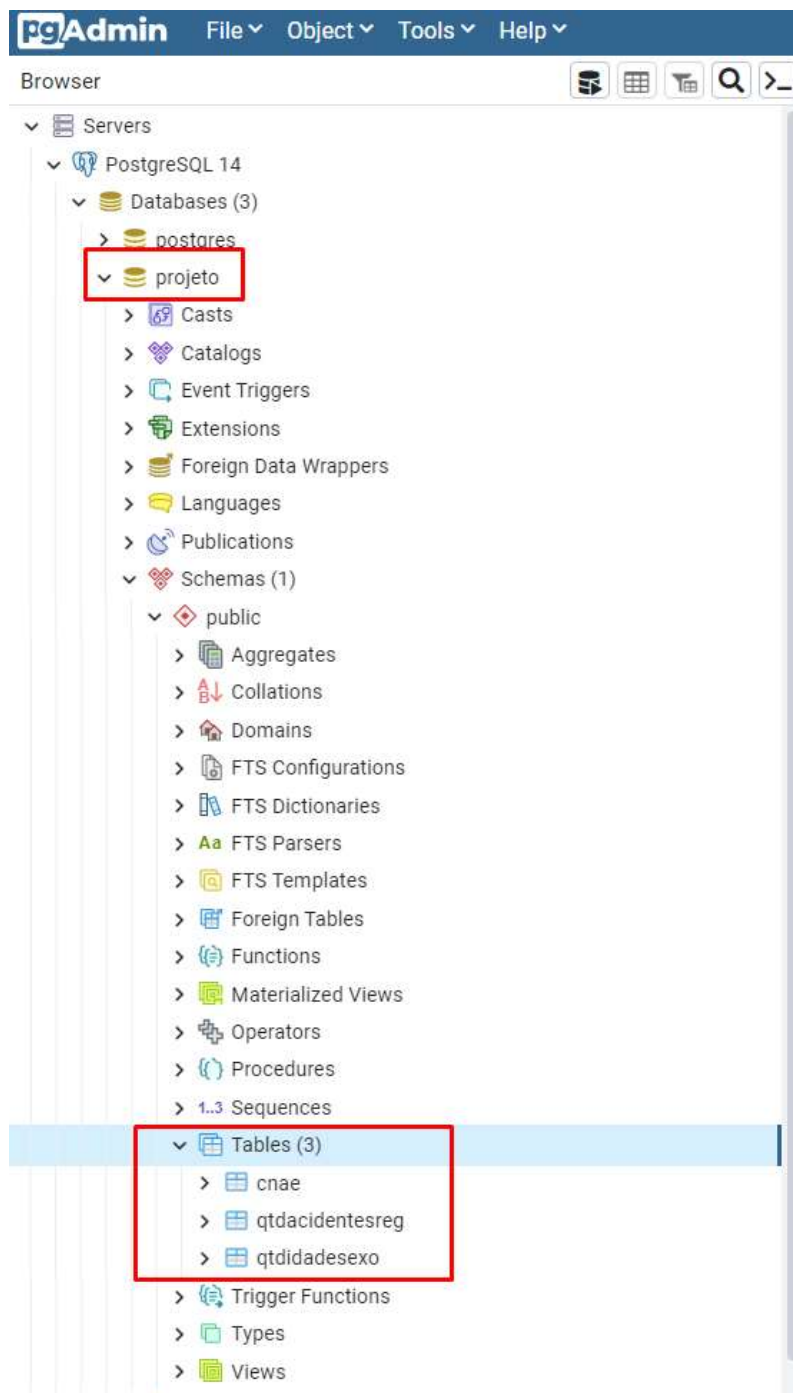


Figura 17: Interface gráfica do PgAdmin4 para o SGBD PostgreSQL.

projeto/postgres@PostgreSQL 14 ▾	
Query Editor Query History	
1	<code>select count(*) from qtdacidentesreg</code>
Data Output Explain Messages Notifications	
	count bigint
1	36408

Figura 18: comando de seleção para contar todos os registros da tabela 'QTDACIDENTESREG'

Na tabela 'QTDIDADESEXO', diferente da tabela anterior, obtemos apenas 288 registros pois os dados de quantidade estavam agrupados.

Query Editor Query History	
1	<code>select count(*) from qtdidadesexo</code>
Data Output Explain Messages Notifications	
	count bigint
1	288

Figura 19: comando de seleção para contar todos os registros da tabela 'QTDIDADESEXO'

Para a tabela CNAE, foram importados 673 registros, que compreendem a classificação correspondente ao padrão de código que foi utilizado nos dados de acidentes obtidos, dessa forma a tabela CNAE terá um relacionamento direto com a tabela 'QTDACIDENTESREG'.

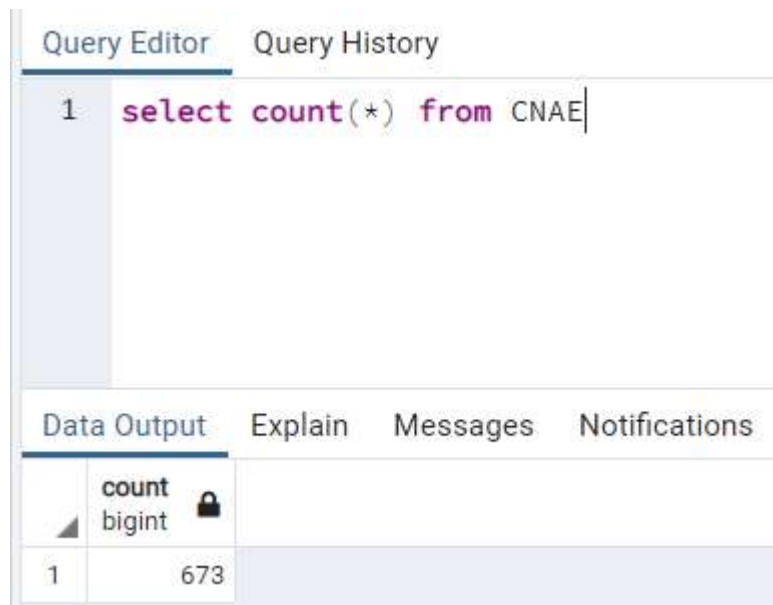


Figura 20: comando de seleção para contar todos os registros da tabela 'CNAE'

6 LIMPEZA E TRANSFORMAÇÃO

Nesta etapa do projeto, não haverá a necessidade de realizar a limpeza ou transformação dos dados, pois os mesmos já foram organizados de forma que pudessem ser manipulados no SGBD no momento em que foram importados, dessa forma podemos dizer que o processo de limpeza já foi efetuado no momento em que foram formatadas as planilhas. Conforme foi citado nos tópicos anteriores, os procedimentos foram:

- Exclusão de linhas e colunas desnecessárias:

Tanto nas planilhas de acidentes por região quanto nas planilhas de acidente por grupo e idade, haviam colunas e linhas totalizadoras, o qual não havia a necessidade de manter para que pudéssemos importar os dados, considerando que posteriormente as consultas realizadas irão trazer somatórios com os totais dos dados. As figuras 21 e 22 demonstram quais colunas e linhas foram deletadas:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1																			
2																			
3																			
4																			
5																			
6																			
7																			
8																			
9																			
10																			
11																			
12																			
13																			
14																			
15																			
16																			
17																			
18																			
19																			
20																			
21																			
22																			
23																			
24																			
25																			
26																			
27																			
28																			
29																			
30																			
31																			
32																			
33																			
34																			
35																			
36																			
37																			
38																			
39																			
40																			
41																			
42																			
43																			
44																			
45																			
46																			
47																			
48																			
49																			
50																			
51																			
52																			
53																			
54																			
55																			
56																			
57																			
58																			
59																			
60																			
61																			
62																			
63																			
64																			
65																			
66																			
67																			
68																			
69																			
70																			
71																			
72																			
73																			
74																			
75																			
76																			
77																			
78																			
79																			
80																			
81																			
82																			
83																			
84																			
85																			
86																			
87																			
88																			
89																			
90																			
91																			
92																			
93																			
94																			
95																			
96																			
97																			
98																			
99																			
100																			

Figura 21: Exemplo das colunas que foram excluídas das planilhas de acidentes de trabalho.

		QUANTIDADE DE ACIDENTES DO TRABALHO																			
		Com CAT Registrada																			
		Total					Total					Tipo de					Motivo				
		2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020	2018	2019	2020		
Idade e sexo	Total	584.071	586.557	445.874	481.993	481.739	483.684	363.318	375.388	313.575	338.683	182.406	88.550	8.597	19.034	36.599	164.004	80.118	42.128		
	Masculinos	387.346	386.901	294.813	321.871	324.542	268.177	268.672	217.882	219.359	64.305	40.873	18.588	8.007	5.683	10.238	68.329	62.959	26.718		
	Femininas	196.685	199.111	149.585	166.955	163.693	134.216	115.621	157.185	89.235	45.774	41.522	25.754	4.590	25.241	37.688	37.668	37.668	15.388		
	Ignorado	21	145	3.567	31	544	1.507	21	127	801		13	206		4	128					
Até 15 anos	Total	12.485	11.879	8.711	11.424	10.791	6.403	8.322	8.875	6.981	3.088	2.975	1.884	49	46	138	1.781	834	368		
	Masculinos	9.448	8.997	6.773	9.539	8.812	6.026	8.391	9.258	5.264	2.587	1.998	1.293	31	29	916	916	495	244		
	Femininas	3.136	2.938	1.943	1.885	2.739	1.877	1.824	1.746	1.337	5.649	914	487	95	17	19	254	199	84		
	Ignorado	49						30													
16 a 24 anos	Total	76.959	77.381	61.542	66.629	76.365	58.718	66.568	92.920	49.785	17.491	36.885	18.754	488	472	2.137	6.947	7.618	2.826		
	Masculinos	58.093	58.722	44.862	50.123	61.469	42.855	58.824	48.937	34.487	11.341	11.114	7.493	255	256	731	6.893	8.713	2.042		
	Femininas	20.581	21.159	18.564	18.417	18.954	19.393	12.062	12.859	11.124	8.150	5.851	3.217	239	214	1.369	2.194	2.239	718		
	Ignorado	2	289	2				268								12					
25 a 24 anos	Total	386.298	375.567	306.713	319.888	317.911	344.077	348.987	360.991	350.664	18.120	18.828	18.814	584	928	1.548	11.790	11.790	4.786		
	Masculinos	360.495	349.628	286.350	302.762	302.956	344.839	349.745	341.739	336.178	13.548	16.672	16.786	478	827	1.132	7.845	7.782	2.875		
	Femininas	27.893	27.912	28.758	24.239	24.200	39.419	16.243	18.815	17.479	7.586	5.000	3.589	365	400	2.568	3.653	3.653	1.371		

CNAE	QUANTIDADE DE ACIDENTES DO TRABALHO					
	Total			Com CAT Registrada		
				Total		
	2018	2019	2020	2018	2019	2020
Total	47.141	47.608	38.236	39.039	39.598	35.037
0111	70	119	260	69	118	259
0112	105	130	110	102	127	106
0113	330	292	239	277	247	228

Figura 23: Exemplo da coluna de códigos CNAE no site do governo.

Utilizando o filtro de classe na planilha com os códigos CNAEs, conseguimos identificar o padrão de código que foi utilizado no site do governo, tendo apenas a diferença do dígito final e a formatação com o ponto e hífen (figura 24), dessa forma removemos estes caracteres, e conseguimos igualar ao padrão, assim conseguindo formar uma coluna de relacionamento entre as tabela CNAE e QTDACIDENTESREG pelo campo CODIGOCNAE.

Seção	Divisão	Grupo	Classe	Subclasse
			01.11-3	Cultivo de cereais
			01.12-1	Cultivo de algodão herbáceo e de outras fibras de lavoura temporária
			01.13-0	Cultivo de cana-de-açúcar
			01.14-8	Cultivo de fumo
			01.15-6	Cultivo de soja
			01.16-4	Cultivo de oleaginosas de lavoura temporária, exceto soja
			01.19-9	Cultivo de plantas de lavoura temporária não especificadas anteriormente
			01.21-1	Horticultura
			01.22-9	Cultivo de flores e plantas ornamentais
			01.31-8	Cultivo de laranja
			01.32-6	Cultivo de uva
			01.33-4	Cultivo de frutas de lavoura permanente, exceto laranja e uva
			01.34-2	Cultivo de café
			01.35-1	Cultivo de cacau
			01.39-3	Cultivo de plantas de lavoura permanente não especificadas anteriormente
			01.41-5	Produção de sementes certificadas

Figura 24: Tabela CNAE com filtro de classe.

A figura 25 mostra como ficou o resultado após a remoção dos caracteres na tabela CNAE:

	A	B
1	CODIGO_CNAE	ATIVIDADE
2	111	Cultivo de cereais
3	112	Cultivo de algodão herbáceo e de outras fibras de lavoura temporária
4	113	Cultivo de cana-de-açúcar
5	114	Cultivo de fumo
6	115	Cultivo de soja
7	116	Cultivo de oleaginosas de lavoura temporária, exceto soja
8	119	Cultivo de plantas de lavoura temporária não especificadas anteriormente
9	121	Horticultura
10	122	Cultivo de flores e plantas ornamentais
11	131	Cultivo de laranja
12	132	Cultivo de uva
13	133	Cultivo de frutas de lavoura permanente, exceto laranja e uva
14	134	Cultivo de café
15	135	Cultivo de cacau
16	139	Cultivo de plantas de lavoura permanente não especificadas anteriormente
17	141	Produção de sementes certificadas
18	142	Produção de mudas e outras formas de propagação vegetal, certificadas

Figura 9: Tabela CNAE com formatação completa.

7 VISUALIZAÇÃO DOS DADOS

Após a importação dos dados de acidentes de trabalho ser efetuada nas etapas anteriores, agora podemos realizar consultas para responder aos questionamentos realizados na etapa inicial do projeto. O primeiro questionamento realizado era a respeito da quantidade de acidentes de trabalho ocorridos no período, o qual o gráfico 1 mostra um percentual da quantidade de acidentes ocorridos no período de 2018 a 2020.

quantidade por ano

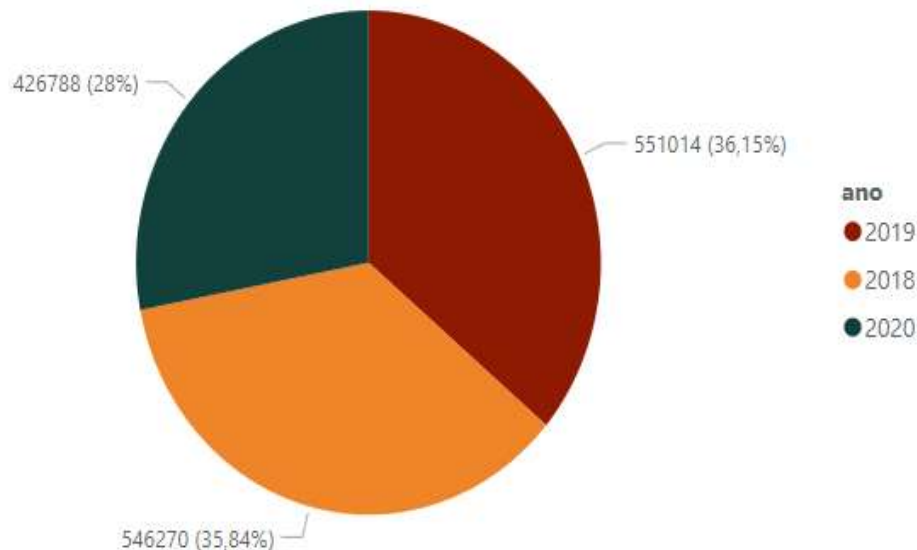


Gráfico 1: quantidade de acidentes por ano no período de 2018 a 2020.

Para gerar este gráfico, utilizamos a seguinte consulta SQL:

```
SELECT
    A.ANO,
    SUM(A.QUANTIDADE) AS QUANTIDADE
FROM QTDACIDENTESREG A
GROUP BY A.ANO
```

Nesta consulta, buscamos pela coluna ano e realizamos uma soma da quantidade, esta que foi agrupada por ano. Como é possível observar no gráfico, no ano de 2019 ocorreu um maior número de acidentes de trabalho, chegando ao total de 551.014 (Quinhentos e cinquenta e um mil e catorze) registros de acidente no Brasil.

Somando o resultado dos três anos, chegamos a um total de 1.524.072 (Um milhão, quinhentos e vinte e quatro mil e setenta e dois) registros de acidentes no período de 2018 a 2020. Esse resultado é obtido também ao realizar a consulta a seguir, que faz um somatório de todos os registros da coluna QUANTIDADE da tabela QTDACIDENTESREG:

```
SELECT
    SUM(A.QUANTIDADE) AS QUANTIDADE
FROM QTDACIDENTESREG A
```

Como estes registros são de amplitude nacional, podemos ainda dividi-los em regiões, utilizando a seguinte consulta:

```
SELECT
    A.REGIAO,
    SUM(A.QUANTIDADE) AS QUANTIDADE
FROM QTDACIDENTESREG A
GROUP BY A.REGIAO
ORDER BY QUANTIDADE desc
```

Nesta consulta, selecionamos a coluna da região e realizamos o somatório da coluna de quantidade, o qual é agrupada por região, por fim ordenando o resultado de forma decrescente, tendo o maior valor somatório primeiro, dessa forma obtemos o resultado que é mostrado no gráfico 2, e podemos observar uma diferença grande em acidentes de trabalho entre a região Sudeste e as demais regiões, tendo a região Sudeste um valor agregado de 828.920 (oitocentos e vinte e oito mil, novecentos e vinte) registros de acidentes no período de 2018 a 2020.

No mesmo período, a região Sul teve 346.975 (Trezentos e quarenta e seis mil, novecentos e setenta e cinco) registros, a região Nordeste 161.311 (Cento e sessenta e um mil, trezentos e onze)

registros, a região Centro-Oeste 123.295 (Cento e vinte e três mil, duzentos e noventa e cinco) registros, e por final, a região Norte teve 63.571 (Sessenta e três mil, quinhentos e setenta e um) registros de acidentes.

quantidade por regioao

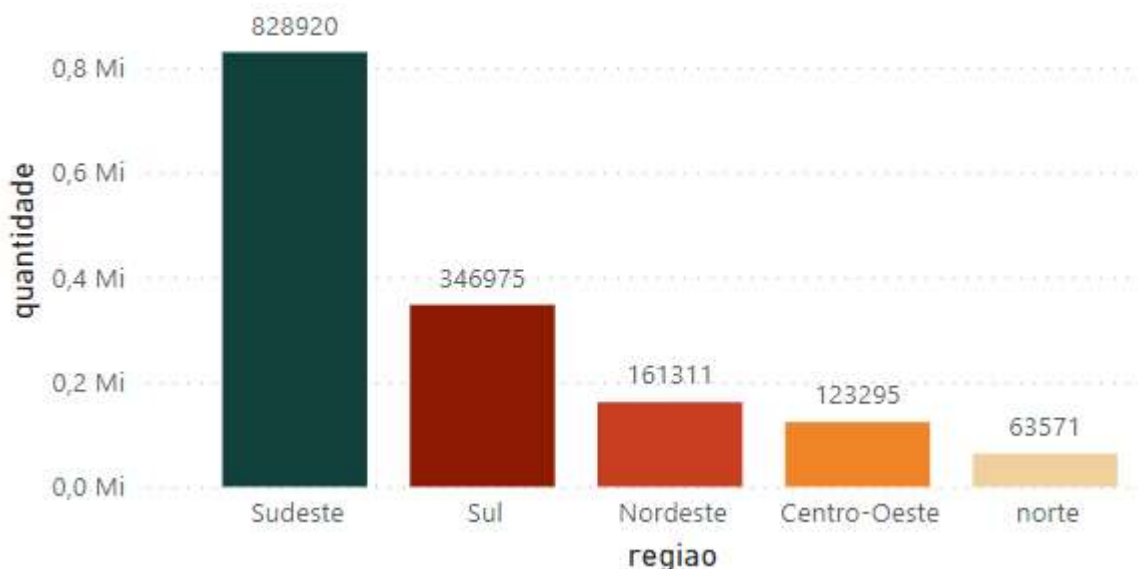


Gráfico 2: quantidade de acidentes divididos por região no período de 2018 a 2020.

Para responder ao nosso questionamento a respeito de quais atividades econômicas mais possuem registros de acidentes de trabalho, buscamos pelo nome da atividade que consta na tabela CNAE, fazendo uma junção com a tabela QTDACIDENTESREG por meio do campo CODIGOCNAE, somamos a quantidade de acidentes e agrupamos pela atividade. A consulta realizada é expressada abaixo:

```
SELECT
    A.CODIGOCNAE,
    B.ATIVIDADE,
    SUM(A.QUANTIDADE) AS QUANTIDADE
FROM QTDACIDENTESREG A
JOIN CNAE B ON A.CODIGOCNAE = B.CODIGOCNAE
GROUP BY A.CODIGOCNAE, B.ATIVIDADE
ORDER BY QUANTIDADE DESC
LIMIT 5
```

Utilizamos o comando LIMIT para trazer apenas as 5 atividades com maior ocorrência, o qual podem ser observadas na tabela 1 a seguir:

CODIGOCNAE	ATIVIDADE	QUANTIDADE
8610	Atividades de atendimento hospitalar	173.977
4711	Comércio varejista de mercadorias em geral	75.810
8411	Administração pública em geral	47.273
4930	Transporte rodoviário de carga	39.574
1012	Abate de suínos	37.621

Tabela 1: Atividades com maior ocorrência de acidentes de trabalho no período de 2018 a 2020.

Como podemos observar na tabela, a atividade de atendimento hospitalar lidera os registros de acidentes de trabalho no Brasil, totalizando 173.977 (Cento e setenta e três mil, novecentos e setenta e sete) registros no período de 2018 a 2020.

A atividade econômica de Comércio varejista de mercadorias em geral vem em segundo lugar, tendo um total de 75.810 (setenta e cinco mil, oitocentos e dez) registros de acidentes de trabalho, seguido da administração pública em geral com 47.273 (Quarenta e sete mil, duzentos e setenta e três) registros, transporte rodoviário de carga com 39.574 (Trinta e nove mil, quinhentos e setenta e quatro) registros e por fim, abate de suínos com 37.621 (Trinta e sete mil, seiscentos e vinte e um) registros.

Observando as atividades com maior ocorrência de acidentes de trabalho, utilizamos outra seleção para observar o motivo predominante dos acidentes, fazendo uma soma da quantidade de acidentes e agrupando pela atividade e pelo motivo, assim separando eles entre típico, doença do trabalho, trajeto e sem registro. A seleção buscou apenas os códigos CNAE das cinco atividades já apresentadas na tabela anterior. Abaixo é mostrada a seleção utilizada para gerar o gráfico 3:

```
SELECT
    B.ATIVIDADE,
    A.MOTIVO,
    SUM(A.QUANTIDADE) AS QUANTIDADE
FROM QTDACIDENTESREG A
JOIN CNAE B ON A.CODIGOCNAE = B.CODIGOCNAE
WHERE A.CODIGOCNAE IN (8610,4711,8411,4930,1012)
GROUP BY B.ATIVIDADE, A.MOTIVO
ORDER BY QUANTIDADE DESC
```

quantidade por atividade e motivo

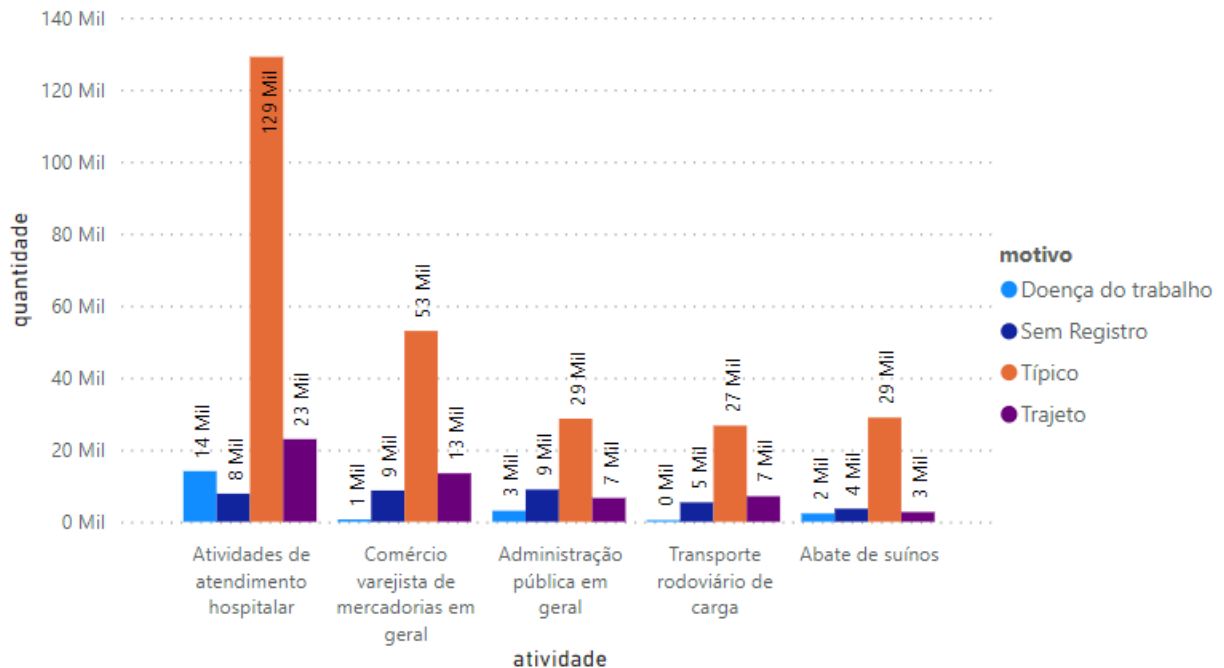


Gráfico 3: quantidade de acidentes divididos por atividade e motivo no período de 2018 a 2020.

Podemos observar no gráfico 3 que em todas as atividades, o motivo predominante é o típico, o qual representa que o acidente de trabalho ocorreu durante a execução da função o qual a pessoa exerce em sua atividade.

Voltando ao cenário nacional, temos outra tabela com registros de acidentes que são divididos por idade e sexo, onde selecionamos primeiramente qual o gênero da pessoa acidentada, para definirmos o percentual de homens e mulheres. Na seleção a seguir realizamos um somatório da quantidade de acidentes e agrupamos por sexo:

```
SELECT
    A.SEXO,
    SUM(A.QUANTIDADE) AS QUANTIDADE
FROM QTDIDADESEXO A
GROUP BY A.SEXO
```

A partir desta seleção, obtemos o gráfico 4, que mostra que 66,09% dos acidentes ocorrem com homens, contra 33,91% que ocorrem com mulheres.

quantidade por sexo

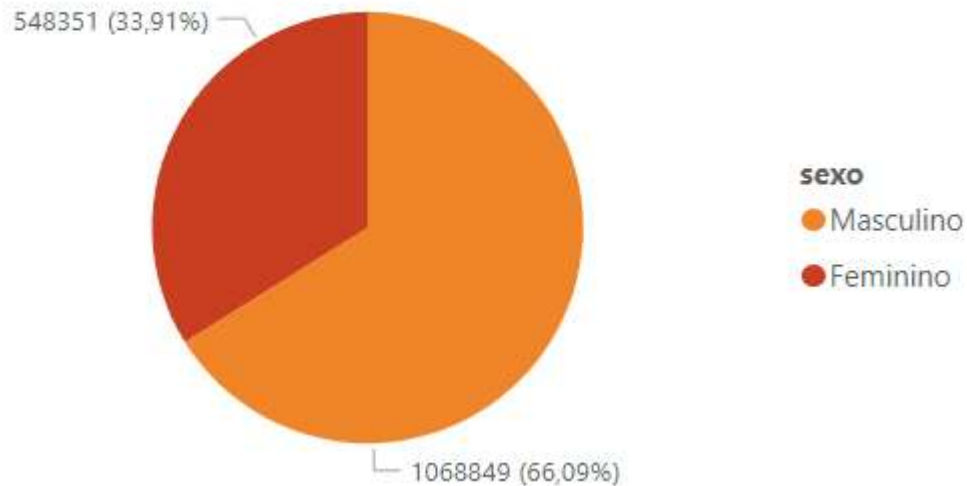


Gráfico 4: representação percentual de quantidade de acidentes entre homens e mulheres no período de 2018 a 2020..

Também realizamos uma seleção para identificar o grupo de idade que mais são registrados acidentes, para isto, utilizamos o comando abaixo, o qual faz um somatório da quantidade de acidentes e agrupa eles pela coluna GRUPO_IDADE da tabela QTDIDADESEXO.

```
SELECT
    A.GRUPO_IDADE,
    SUM(A.QUANTIDADE) AS QUANTIDADE
FROM QTDIDADESEXO A
GROUP BY A.GRUPO_IDADE
ORDER BY QUANTIDADE DESC
```

Com esta seleção, definimos a tabela 2, que apresenta o percentual de acidentes correspondente a cada grupo de idade, onde podemos observar que 15,87% dos acidentes ocorreram com pessoal de 30 a 34 anos de idade, próximo a isto, pessoas de 35 a 39 anos de idade representam 15,58% dos acidentes e pessoas de 25 a 29 anos de idade representam 15,08% dos acidentes registrados no período de 2018 a 2020.

Grupo de Idade	Percentual
30 a 34 anos	15,87%
35 a 39 anos	15,58%
25 a 29 anos	15,08%
20 a 24 anos	13,34%
40 a 44 anos	12,76%
45 a 49 anos	9,93%
50 a 54 anos	7,83%
55 a 59 anos	4,94%
Até 19 anos	2,04%
60 a 64 anos	1,98%
65 a 69 anos	0,51%
70 anos e mais	0,15%
Total	100,00%

Tabela 2: Atividades com maior ocorrência de acidentes de trabalho no período de 2018 a 2020.

8 CONCLUSÃO

Considerando os objetivos citados na introdução do projeto e análise realizada com a visualização dos dados de forma gráfica, conseguimos identificar as áreas econômicas e regiões com maior índice de acidentes. Porém com a falta de disponibilização de dados mais específicos não conseguimos nomear as profissões, consequentemente também não conseguimos definir precisamente se os acidentes ocorrem por algum tipo de negligência tão menos por profissões de alto risco.

Dessa forma, ao analisar os dados do período de 2018 a 2020, observamos que houveram mais acidentes de trabalho na região Sudeste do país, fato que corrobora com dados de uma publicação do site Brasil Escola, o qual cita que o estado de São Paulo concentra cerca de 22% da população brasileira conforme dados do IBGE de 2021.

Além disso, a maior ocorrência de acidentes de trabalho está entre o público masculino, concentrando 66% dos registros de acidentes. Considerando a idade das pessoas com maior índice, 15,87% das ocorrências são de pessoas com idades entre 30 e 34 anos. Ademais, como dito anteriormente nos dados obtidos, as cinco profissões de maior ocorrência, respectivamente são: Atividades de atendimento hospitalar, Comércio varejista de mercadorias em geral, Administração pública em geral, Transporte rodoviário de carga e Abate de suínos.

Apesar de não conseguirmos identificar precisamente as profissões, o projeto ajudou a identificar que quando o assunto é “acidente de trabalho”, nem sempre são as profissões de alto risco como a área de construção civil, elétrica, metalúrgica, por exemplo, lideram nos índices. Na verdade, a análise mostrou que áreas mais comuns são as dos profissionais da saúde e do comércio varejista.

9 BIBLIOGRAFIA

- G1.GLOBO.COM. Brasil é o 2º país do G20 em mortalidade por acidentes no trabalho, c2021. Disponível em: <https://g1.globo.com/economia/noticia/2021/05/01/brasil-e-2o-pais-do-g20-em-mortalidade-por-acidentes-no-trabalho.ghtml>. Acesso em: 18 de Março de 2022.
- GOV.BR. Ministério do Trabalho e Previdência: Acidentes do Trabalho - Brasil e Grandes Regiões, c2021. Disponível em: https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes. Acesso em: 18 de Março de 2022.
- Gov.br. Ministério do Trabalho e Previdência: Quantidade de acidentes do trabalho, por situação do registro e motivo, segundo a Classificação Nacional de Atividades Econômicas (CNAE), no Brasil - 2018/2020. Disponível em: https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-1-quantidade-de-acidentes-do-trabalho-por-situacao-do-registro-e-motivo-segundo-a-classificacao-nacional-de-atividades-economicas-cnae-no-brasil-2018-2019. Acesso em 18 de Março de 2022.
- Gov.br. Ministério do Trabalho e Previdência: Quantidade de acidentes do trabalho, por situação do registro e motivo, segundo a Classificação Nacional de Atividades Econômicas (CNAE), na Região Norte - 2018/2020. Disponível em: https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-2-quantidade-de-acidentes-do-trabalho-por-situacao-do-registro-e-motivo-segundo-a-classificacao-nacional-de-atividades-economicas-cnae-na-regiao-norte-2018-2019. Acesso em 18 de Março de 2022.
- Gov.br. Ministério do Trabalho e Previdência: Quantidade de acidentes do trabalho, por situação do registro e motivo, segundo a Classificação Nacional de Atividades Econômicas (CNAE), na Região Nordeste - 2018/2020. Disponível em: https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-3-quantidade-de-acidentes-do-trabalho-por-situacao-do-registro-e

[motivo-segundo-a-classificacao-nacional-de-atividades-economicas-cnae-na-regiao-nordeste-2018-2019](#). Acesso em 18 de Março de 2022.

- Gov.br. Ministério do Trabalho e Previdência: Quantidade de acidentes do trabalho, por situação do registro e motivo, segundo a Classificação Nacional de Atividades Econômicas (CNAE), na Região Sudeste - 2018/2020. Disponível em:
[https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-4-quantidade-de-acidentes-do-trabalho-por-situacao-do-registro-e-motivo-segundo-a-classificacao-nacional-de-atividades-economicas-cnae-na-regiao-sudeste-2018-2019](#). Acesso em 18 de Março de 2022.
- Gov.br. Ministério do Trabalho e Previdência: Quantidade de acidentes do trabalho, por situação do registro e motivo, segundo a Classificação Nacional de Atividades Econômicas (CNAE), na Região Sul - 2018/2020. Disponível em:
[https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-5-quantidade-de-acidentes-do-trabalho-por-situacao-do-registro-e-motivo-segundo-a-classificacao-nacional-de-atividades-economicas-cnae-na-regiao-sul-2018-2019](#). Acesso em 18 de Março de 2022.
- Gov.br. Ministério do Trabalho e Previdência: Quantidade de acidentes do trabalho, por situação do registro e motivo, segundo a Classificação Nacional de Atividades Econômicas (CNAE), na Região Centro-Oeste - 2018/2020. Disponível em:
[https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-6-quantidade-de-acidentes-do-trabalho-por-situacao-do-registro-e-motivo-segundo-a-classificacao-nacional-de-atividades-economicas-cnae-na-regiao-centro-oeste-2018-2019](#). Acesso em 18 de Março de 2022.
- Gov.br. Ministério do Trabalho e Previdência: Quantidade de acidentes do trabalho, por situação do registro e motivo, segundo os grupos de idade e sexo, no Brasil - 2018/2020. Disponível em:
[https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-7-quantidade-de-acidentes-do-trabalho-por-situacao-do-registro-e-](#)

[motivo-segundo-os-grupos-de-idade-e-sexo-no-brasil-2018-2019](#). Acesso em 18 de Março de 2022.

- Gov.br. Ministério do Trabalho e Previdência: Quantidade mensal de acidentes do trabalho, por situação do registro e motivo, no Brasil - 2018/2020. Disponível em:
https://www.gov.br/trabalho-e-previdencia/pt-br/assuntos/previdencia-social/saude-e-seguranca-do-trabalhador/dados-de-acidentes-do-trabalho/arquivos/AEAT_2020/secao-i-estatisticas-de-acidentes-do-trabalho/subsecao-a-acidentes-do-trabalho/capitulo-1-brasil-e-grandes-regioes/1-8-quantidade-mensal-de-acidentes-do-trabalho-por-situacao-do-registro-e-motivo-no-brasil-2018-2019. Acesso em 18 de Março de 2022.
- Cnae.ibge.gov.br. Página de download da Comissão Nacional de Classificações. Disponível em:
<https://cnae.ibge.gov.br/classificacoes/download-concla/8265-download>. Acesso em 17 de Abril de 2022.
- Contabilix. Tabela CNAE: o que é e como consultar. Disponível em:
<https://www.contabilix.com.br/contabilidade-online/tabela-cnae/>. Acesso em 17 de Abril de 2022.
- GUITARRARA, Paloma. "População do Brasil"; Brasil Escola. Disponível em:
<https://brasilecola.uol.com.br/brasil/a-populacao-brasileira.htm>. Acesso em 05 de julho de 2022.