

# Acidentes em Rodovias - Ano 2019 e 2020

Bonifácio, Leonardo de Holanda

200095056@aluno.unb.br

9 de março de 2021

## 1 Requisitos

Deseja-se criar um banco de dados contendo todas as ocorrências de acidentes em rodovias federais. Para isso, informações como local de ocorrência (Unidade Federativa, Município, BR), causa do acidente, tipo de acidente, entre outros devem ser levados em conta.

Todo acidente deve conter:

- BR, Km, Município, UF onde ocorreu;
- Número de pessoas envolvidas, mortas, feridas levemente, feridas gravemente, ilesos, ignorados e feridos;
- Número de veículos;
- Data da ocorrência;
- Traçado da via;
- Causa do acidente;
- Fase do dia;
- Condições meteorológicas;
- Tipo de pista;
- Classificação do acidente;
- Sentido da via;
- Tipo de acidente.

## 2 Modelo lógico

O modelo lógico, feito no MySQL Workbench, pode ser visto no Anexo [A](#). A lógica utilizada foi de um centralizador (Tabela ACIDENTE), que tem todas as informações necessárias para realizar

as consultas desejadas. As tabelas secundárias são compostas de variáveis categóricas com suas respectivas *Primary Keys*. Essas foram transformadas em *Foreign Keys* na tabela principal no processo de ETL.

### 3 PrintScreen das Tabelas

O *dataset* foi baixado no [Kaggle](#) e, antes dos tratamentos, era composto de 83154 registros com 30 colunas no total. Após tratamento e limpeza, a tabela final ACIDENTE ficou com 83033 registros.

Todas as tabelas podem ser vistas no Anexo B. Algumas imagens foram cortadas para melhorar a visualização. Os dados foram tratados utilizando Python e a biblioteca Pandas e Numpy.

Dentre os processos de ETL pode-se citar:

- Transformação de *string*: *lower case*, retirada de acentos caracteres especiais, remoção de espaços excedentes;
- Remoção ou substituição de dados faltantes;
- Criação dos *dataframes* para inserção;
- Inserção no banco de dados com a biblioteca Python *mysql*.

### 4 Consultas Álgebra Relacional

As consultas foram descritas de forma manual por falta de uma ferramenta mais adequada.

- Consulta C.1: Listagem de todos acidentes que aconteceram em Brasília (*ID\_MUNICIPIO* = 27) com seus dados de localização, número de envolvidos e data.
- Consulta C.2: Listagem de todos acidentes por tipo de acidente, com seus dados de localização, número de envolvidos e data.
- Consulta C.3: Listagem de todos acidentes com fatalidade com seus dados de localização, número de envolvidos e data.

### 5 Consultas SQL

- Consulta D.1: Listagem de todos acidentes que aconteceram em Brasília (*ID\_MUNICIPIO* = 27) com seus dados de localização, número de envolvidos e data.
- Consulta D.2: Listagem de todos acidentes por tipo de acidente, ordenados pelo número de pessoas envolvidas, com seus dados de localização, número de envolvidos e data.
- Consulta D.3: Número total de acidentes no período, separado por Unidade Federativa.
- Consulta D.4: Número total pessoas envolvidas em acidentes no período, separado por Município.

- Consulta [D.5](#): Número de acidentes para cada mês para cada Unidade Federativa.

## A Modelo Lógico

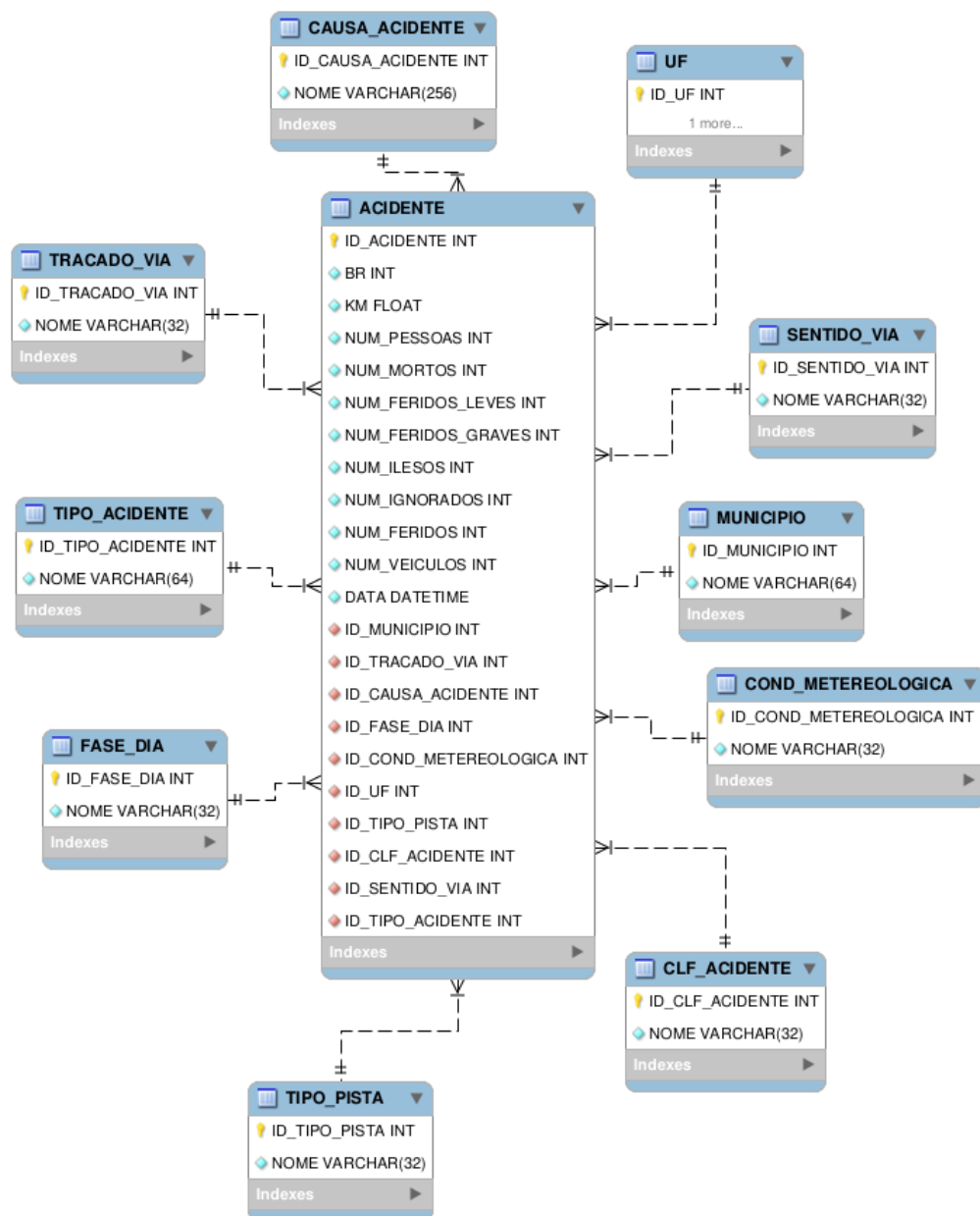


Figura 1: Modelo lógico no MySQL Workench. Fonte: Autor.

## B PrintScreen das Tabelas

1	ID_ACIDENTE	BR	KM	NUM_PESSOAS	NUM_MORTOS	NUM_FERIDOS	LEVES	NUM_FERIDOS_GRAVES	NUM_ILESOS	NUM_IGNORADOS	NUM_FERIDOS	NUM_VEICULOS	DATA	ID_MUNICIPIO	ID
2	260031	153	678.1	5	0	0	0	5	0	0	5	1	2020-01-01 01:00:00	1	
3	260036	116	178	2	1	0	0	0	1	0	0	1	2020-01-01 01:00:00	2	
4	260037	101	206.9	2	0	1	0	0	1	0	1	1	2020-01-01 01:52:00	3	
5	260038	364	236	1	0	0	0	1	0	0	1	1	2020-01-01 01:15:00	4	
6	260039	153	360	6	0	2	4	0	0	0	6	1	2020-01-01 01:30:00	5	
7	260040	319	64	2	0	0	0	2	0	0	0	2	2020-01-01 00:40:00	6	
8	260041	101	99.9	1	0	0	0	0	1	0	0	1	2020-01-01 02:20:00	7	
9	260042	369	155	3	0	1	0	2	0	1	1	2	2020-01-01 02:15:00	8	
10	260043	104	37	1	1	0	0	0	0	0	0	1	2020-01-01 02:30:00	9	
11	260044	290	97	2	0	2	0	0	0	0	2	2	2020-01-01 00:20:00	10	
12	260045	470	116.3	4	0	2	2	0	0	0	4	4	2020-01-01 02:50:00	11	
13	260046	282	509	3	0	0	1	1	1	1	1	1	2020-01-01 02:15:00	12	
14	260050	101	49.5	4	0	0	2	1	1	1	2	2	2020-01-01 03:40:00	13	
15	260051	364	265.3	2	0	2	0	0	0	0	2	2	2020-01-01 02:30:00	14	
16	260052	101	174	2	0	1	1	0	0	0	2	2	2020-01-01 03:10:00	15	
17	260053	101	132	2	1	0	0	1	0	0	0	1	2020-01-01 04:00:00	16	
18	260054	156	406	1	0	0	0	1	0	0	0	1	2020-01-01 00:05:00	17	
19	260055	60	164.8	4	0	4	0	0	0	0	4	4	2020-01-01 03:20:00	18	
20	260057	20	21.8	3	0	2	0	1	0	2	2	2	2020-01-01 04:55:00	19	

Figura 2: Tabela ACIDENTE. Fonte: Autor.

1	ID_CAUSA_ACIDENTE	NOME
2	20	agressao externa
3	1	animais na pista
4	10	avarias e/ou desgaste excessivo no pneu
5	24	carga excessiva e/ou mal acondicionada
6	7	condutor dormindo
7	14	defeito mecanico no veiculo
8	16	defeito na via
9	12	deficiencia ou nao acionamento do sistema de iluminacao/sinalizacao do veiculo
10	6	desobediencia as normas de transito pelo condutor
11	23	desobediencia as normas de transito pelo pedestre
12	5	falta de atencao a conducao
13	2	falta de atencao do pedestre
14	21	fenomenos da natureza
15	3	ingestao de alcool
16	13	ingestao de alcool e/ou substancias psicoativas pelo pedestre
17	17	ingestao de substancias psicoativas
18	9	mal subito
19	11	nao guardar distancia de seguranca
20	19	objeto estatico sobre o leito carrocavel
21	8	pista escorregadia
22	22	restricao de visibilidade
23	18	sinalizacao da via insuficiente ou inadequada
24	15	ultrapassagem indevida
25	4	velocidade incompativel

Figura 3: Tabela CAUSA ACIDENTE. Fonte: Autor.

1	ID_CLF_ACIDENTE	NOME
2	2	com vitimas fatais
3	1	com vitimas feridas
4	3	sem vitimas

Figura 4: Tabela CLF ACIDENTE. Fonte: Autor.

1	ID_COND_METEREOLOGICA	NOME
2	2	ceu claro
3	4	chuva
4	3	garoa/chuvisco
5	7	ignorado
6	8	nevoeiro/neblina
7	1	nublado
8	5	sol
9	6	vento

Figura 5: Tabela COND METEREOLOGICA. Fonte: Autor.

1	ID_FASE_DIA	NOME
2		2 amanhecer
3		4 anoitecer
4		1 plena noite
5		3 pleno dia

Figura 6: Tabela FASE DIA. Fonte: Autor.

1	ID_MUNICIPIO	NOME
2		1 gurupi
3		2 nova iguacu
4		3 sao jose
5		4 cacoal
6		5 reboucas
7		6 porto velho
8		7 barra velha
9		8 londrina
10		9 uniao dos palmares
11		10 porto alegre
12		11 ibirama
13		12 xanxere
14		13 abreu e lima
15		14 sao pedro da cipa
16		15 feira de santana
17		16 balneario camboriu

Figura 7: Tabela MUNICIPIO. Fonte: Autor.

1	ID_SENTIDO_VIA	NOME
2		1 crescente
3		2 decrescente

Figura 8: Tabela SENTIDO VIA. Fonte: Autor.

ID_TIPO_ACIDENTE	NOME
1	atropelamento de animal
2	atropelamento de pedestre
5	capotamento
14	colisao com objeto em movimento
7	colisao com objeto estatico
9	colisao frontal
3	colisao lateral
6	colisao transversal
11	colisao traseira
15	danos eventuais
16	derramamento de carga
13	engavetamento
12	incendio
10	queda de ocupante de veiculo
4	saida de leito carrocavel
8	tombamento

Figura 9: Tabela TIPO ACIDENTE. Fonte: Autor.

ID_TIPO_PISTA	NOME
2	dupla
3	multipla
1	simples

Figura 10: Tabela TIPO PISTA. Fonte: Autor.

ID_TRACADO_VIA	NOME
2	curva
7	desvio temporario
5	intersecao de vias
3	nao informado
9	ponte
1	reta
8	retorno regulamentado
6	rotatoria
10	tunel
4	viaduto

Figura 11: Tabela TRACADO VIA. Fonte: Autor.

ID_UF	NOME
24	ac
6	al
27	am
11	ap
10	ba
21	ce
13	df
23	es
12	go
22	ma
17	mg
20	ms
9	mt
18	pa
15	pb

Figura 12: Tabela UF. Fonte: Autor.

## C Consultas Álgebra Relacional

### C.1 Listagem de acidentes em Brasília

#### C.1.1 Consulta

$$\pi_{(ID\_ACIDENTE, BR, KM, NUM\_PESSOAS, DATA)}(\sigma_{ID\_MUNICIPIO=27}(ACIDENTE))$$

Figura 13: Fórmula em álgebra relacional. Fonte: Autor.

#### C.1.2 Resultado

Presente em [D.1](#)

### C.2 Listagem de acidentes por tipo de acidente

#### C.2.1 Consulta

$$\pi_{(ID\_ACIDENTE, BR, KM, NUM\_PESSOAS, DATA, NOME)}(ACIDENTE \text{ leftjoin}_{(ACIDENTE.ID\_TIPO\_ACIDENTE=TIPO\_ACIDENTE.ID\_TIPO\_ACIDENTE)} TIPO\_ACIDENTE)$$

Figura 14: Fórmula em álgebra relacional. Fonte: Autor.

#### C.2.2 Resultado

Presente em [D.2](#)

### C.3 Listagem de acidentes com fatalidade

#### C.3.1 Consulta

$$\pi_{(ID\_ACIDENTE, UF, BR, KM, NUM\_PESSOAS, NUM\_MORTOS, DATA)}(\sigma_{NUM\_MORTOS>0}(ACIDENTE)) \\ (ACIDENTE \text{ leftjoin}_{(ACIDENTE.ID\_TIPO\_ACIDENTE=TIPO\_ACIDENTE.ID\_TIPO\_ACIDENTE)} TIPO\_ACIDENTE)$$

Figura 15: Listagem de acidentes com fatalidade com UF. Fonte: Autor.

#### C.3.2 Resultado



ID_ACIDENTE	NOME	BR	KM	NUM_PESSOAS	NUM_MORTOS	DATA
212266	to	153	306	2	1	2019-05-28 18:30:00
216200	to	153	191.9	3	1	2019-06-17 00:00:00
196863	to	153	289	2	1	2019-03-14 06:50:00
229589	to	242	376.4	1	1	2019-08-18 18:00:00
266699	to	153	257	4	1	2020-02-02 18:30:00
186256	to	230	49	3	1	2019-01-19 20:30:00
188555	to	153	442.2	4	1	2019-01-31 21:40:00
248681	to	153	163.3	1	1	2019-11-11 04:45:00
191820	to	153	526.3	4	1	2019-02-16 22:15:00
276968	to	153	707.7	4	1	2020-03-23 14:00:00
247610	to	153	635.3	2	1	2019-07-01 19:00:00
211554	to	230	20	1	1	2019-05-24 21:00:00
213479	to	153	673.8	2	1	2019-06-03 20:50:00
225993	to	242	480	1	1	2019-08-01 15:00:00
199925	to	153	620.7	2	1	2019-03-29 16:00:00
200358	to	153	487.2	2	1	2019-03-31 13:48:00
259892	to	235	167	2	1	2019-12-31 18:30:00
207357	to	153	148.1	2	2	2019-05-05 00:05:00

Figura 16: Listagem de acidentes com fatalidade. Fonte: Autor.

## D Consultas SQL

### D.1 Listagem de acidentes em Brasília

#### D.1.1 Consulta

```
select ID_ACIDENTE, BR, KM, NUM_PESSOAS, DATA
from acidentes_db.ACIDENTE as t1
where t1.ID_MUNICIPIO = 27;
```

#### D.1.2 Resultado

A	B	C	D	E
ID_ACIDENTE	BR	KM	NUM_PESSOAS	DATA
182243	40	6	2	2019-01-01 07:00:00
182254	70	16	2	2019-01-01 08:20:00
182357	20	12	2	2019-01-01 18:30:00
182391	251	40	5	2019-01-01 21:10:00
182587	20	7.3	3	2019-01-02 19:40:00
183211	70	13	3	2019-01-03 15:10:00
183222	70	7.8	2	2019-01-03 16:30:00
183249	20	4.1	2	2019-01-03 19:00:00
183287	60	4.5	1	2019-01-04 00:45:00
183428	80	31.8	2	2019-01-04 18:50:00
183519	70	9.2	2	2019-01-05 08:20:00
183562	20	21	1	2019-01-05 11:30:00
183651	60	2	1	2019-01-05 07:00:00
183733	20	8	1	2019-01-06 13:20:00
183740	70	8.7	2	2019-01-06 14:25:00
183975	70	11	2	2019-01-07 23:45:00
183976	70	13	1	2019-01-08 04:00:00
184101	20	19.4	2	2019-01-08 22:40:00
184104	20	33	2	2019-01-08 23:00:00
184207	60	29.6	2	2019-01-09 15:50:00
184214	20	53	2	2019-01-09 16:00:00
184222	60	12	2	2019-01-09 17:50:00

Figura 17: Listagem de acidentes em Brasília com seus dados de localização, número de envolvidos e data. Fonte: Autor.

### D.2 Listagem de acidentes por tipo de acidente

#### D.2.1 Consulta

```
select ID_ACIDENTE, BR, KM, NUM_PESSOAS, DATA, t2.NOME
from acidentes_db.ACIDENTE as t1
inner join acidentes_db.TIPO_ACIDENTE as t2
on t1.ID_TIPO_ACIDENTE = t2.ID_TIPO_ACIDENTE
order by NUM_PESSOAS desc;
```

## D.2.2 Resultado

ID_ACIDENTE	BR	KM	NUM_PESSOAS	DATA	NOME
200728	365	295.1		74 2019-03-02 02:30:00	colisao traseira
274483	364	366		73 2020-02-29 21:25:00	incendio
230311	116	409.6		62 2019-08-22 03:00:00	incendio
274666	153	18		61 2020-02-01 17:15:00	colisao frontal
228659	365	387		56 2019-08-14 05:40:00	colisao traseira
215894	365	320.2		55 2019-02-11 21:17:00	colisao lateral
260817	251	465.2		54 2020-01-04 15:00:00	colisao traseira
192368	153	500.1		52 2019-01-27 02:30:00	saida de leito carrocavel
221710	393	102.5		51 2019-07-13 03:35:00	saida de leito carrocavel
264482	40	96.7		51 2019-10-27 06:00:00	colisao com objeto estatico
260743	381	749.6		50 2020-01-04 03:30:00	colisao traseira
202460	381	525.5		50 2019-03-24 06:00:00	saida de leito carrocavel
255614	116	719.2		49 2019-12-12 19:55:00	colisao frontal
255874	153	107.1		49 2019-12-14 01:20:00	colisao frontal
251894	262	172		48 2019-11-25 18:30:00	colisao frontal
246908	386	209		48 2019-11-03 05:55:00	colisao frontal
243713	381	845.4		47 2019-10-19 18:40:00	colisao lateral
190002	262	361.5		47 2019-02-08 07:15:00	colisao traseira

Figura 18: Listagem de acidentes por tipo de acidente ordenado por número de pessoas envolvidas.

Fonte: Autor.

## D.3 Contagem de acidentes por UF

### D.3.1 Consulta

```
select NOME, count(*)
from acidentes_db.ACIDENTE as t1
inner join acidentes_db.UF as t2
on t1.ID_UF = t2.ID_UF
group by NOME
order by 2 desc
limit 10;
```

### D.3.2 Resultado

NOME	count(*)
mg	10782
sc	10319
pr	9460
rj	5615
rs	5575
sp	5426
ba	4299
go	4202
pe	3401
es	3394

Figura 19: Contagem de acidentes por UF. Fonte: Autor.

## D.4 Contagem de pessoas acidentadas por Município

### D.4.1 Consulta

```
select NOME, sum(NUM_PESSOAS)
from acidentes_db.ACIDENTE as t1
inner join acidentes_db.MUNICIPIO as t2
on t1.ID_MUNICIPIO = t2.ID_MUNICIPIO
group by NOME
order by 2 desc
limit 10;
```

### D.4.2 Resultado

NOME	sum(NUM_PESSOAS)
brasil	3344
curitiba	3082
sao jose	2170
guarulhos	2059
serra	1777
palhoca	1704
duque de caxias	1658
betim	1573
porto velho	1361
recife	1309

Figura 20: Contagem de pessoas acidentadas por Município. Fonte: Autor.

## D.5 Contagem de acidentes mensais pra cada UF

### D.5.1 Consulta

```
select NOME, count(*), DATE_FORMAT(DATA, "%Y-%m-01")
from acidentes_db.ACIDENTE as t1
left join acidentes_db.UF as t2
on t1.ID_UF = t2.ID_UF
group by NOME, DATE_FORMAT(DATA, "%Y-%m-01");
```

### D.5.2 Resultado

NOME	count(*)	DATE_FORMAT(DATA, "%Y-%m-01")
ac	17	2019-01-01
ac	19	2019-02-01
ac	16	2019-03-01
ac	13	2019-04-01
ac	21	2019-05-01
ac	24	2019-06-01
ac	21	2019-07-01
ac	32	2019-08-01
ac	20	2019-09-01
ac	28	2019-10-01
ac	20	2019-11-01
ac	21	2019-12-01
ac	19	2020-01-01
ac	26	2020-02-01
ac	25	2020-03-01
al	53	2019-01-01
al	60	2019-02-01
al	48	2019-03-01
al	52	2019-04-01
al	50	2019-05-01
al	43	2019-06-01
al	61	2019-07-01
al	51	2019-08-01
al	52	2019-09-01
al	52	2019-10-01
al	51	2019-11-01
al	56	2019-12-01
al	36	2020-01-01
al	59	2020-02-01
al	54	2020-03-01
am	8	2019-01-01
am	7	2019-02-01
am	4	2019-03-01
am	7	2019-04-01
am	6	2019-05-01
am	5	2019-06-01
am	11	2019-07-01

Figura 21: Contagem de acidentes mensais pra cada UF. Fonte: Autor.