

# Anàlisi descriptiva

Anna Salazar

2022-07-01



UNIVERSITAT POLITÈCNICA DE CATALUNYA  
BARCELONATECH

Facultat de Matemàtiques i Estadística



UNIVERSITAT DE  
BARCELONA

# Contents

<b>Descripció de la base de dades</b>	<b>1</b>
<b>Anàlisi univariant</b>	<b>2</b>
Variables categòriques . . . . .	2
Variables numèriques . . . . .	12

## Descripció de la base de dades

La base de dades està formada per les següents tres taules:

**accident** és un llistat d'accidents de tràfic ocorreguts al desembre de 2015 als Estats Units:

- ST\_CASE: codi de l'accident (PK)
- DAY: dia de l'accident (de l'1 al 31)
- HOUR: hora de l'accident (99 = desconeguda)
- MINUTE: minut de l'accident (99 = desconegut)
- RUR\_URB: informació sobre la localització (1 = Rural, 2 = Urbà, 6 = Via no classificada, 8 = No registrat, 9 = Desconegut)
- DAY\_WEEK: dia de la setmana (1 = Diumenge, 2 = Dilluns, ..., 7 = Dissabte)
- FATALS: nombre de ferits a l'accident
- DRUNK\_DR: nombre de conductors beguts involucrats a l'accident

**person** és un llistat de totes les persones (conductors, passatgers o vianants) involucrades als accidents

- ST\_CASE: codi de l'accident al qual està involucrada la persona (PK)
- PER\_NO: número de persona dins de cada accident (PK)
- AGE: edat de la persona (998 = No registrada, 999 = Desconeguda)
- SEX: sexe de la persona (1 = home, 2 = dona, 8 = No registrat, 9 = Desconegut)
- PER\_TYP: tipus de persona (1 = conductor, 2 = ocupant, resta de codis = altres)
- DOA: tipus de víctima (0 = sobrevis, 7 = mort a l'accident, 8 = mort al trasllat, 9 = Desconegut)

**vehicle** és un llistat de tots els vehicles involucrats als accidents

- ST\_CASE: codi de l'accident al qual està involucrat el vehicle (PK)
- VEH\_NO: número de vehicle dins de cada accident (PK)
- HIT\_RUN: identificador de vehicle fugit (0 = No, 1 = Sí, 9 = Desconegut)
- TRAV\_SP: velocitat estimada (mph)\* del vehicle quan va tenir l'accident (997,998 i 999 = Desconegut)
- PREV\_SP: indicador d'existència de límit de velocitat permesa just abans de l'accident (997,998 i 999 = Desconegut)

\*mph vol dir milles per hora. 100 mph=161 Km/h

# Anàlisi univariant

## Variables categòriques

### Dia de l'accident

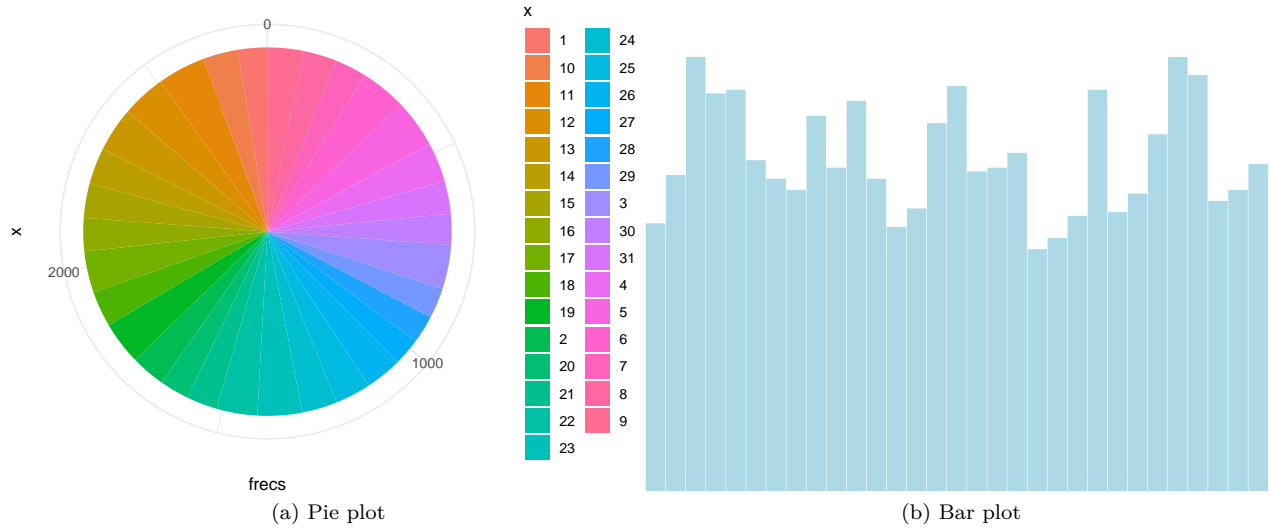


Figure 1: Anàlisi descriptiu de la variable Dia

### Hora de l'accident

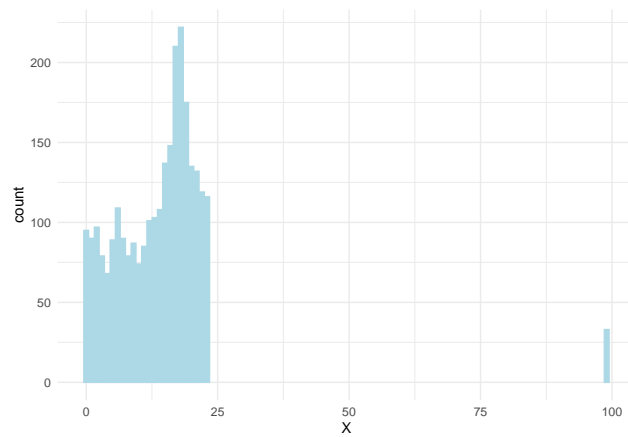


Figure 2: Histograma de la variable Hora de l'accident

Table 1: Resum numèric de la variable Hora de l'accident

N.Valid	Min	Q1	Median	Mean	Std.Dev	Q3	Max	IQR
2781	0	7	15	14.02877	11.49813	19	99	12

## Minut de l'accident

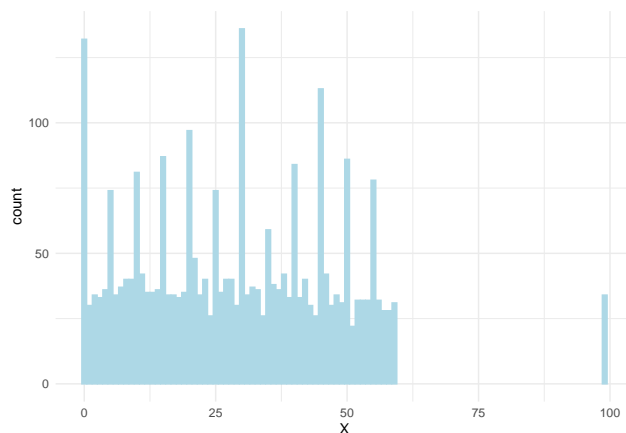


Figure 3: Histograma de la variable Minut de l'accident

Table 2: Resum numèric de la variable Minut de l'accident

N.Valid	Min	Q1	Median	Mean	Std.Dev	Q3	Max	IQR
2781	0	14	29	29.02229	18.79037	44	99	30

## Localització

Variable	Stats / Values	Freqs (% of Valid)	Missing
RUR_URB [factor]	1. 1	1174 (42.2%)	0
	2. 2	1289 (46.4%)	(0.0%)
	3. 6	10 ( 0.4%)	
	4. 8	301 (10.8%)	
	5. 9	7 ( 0.3%)	

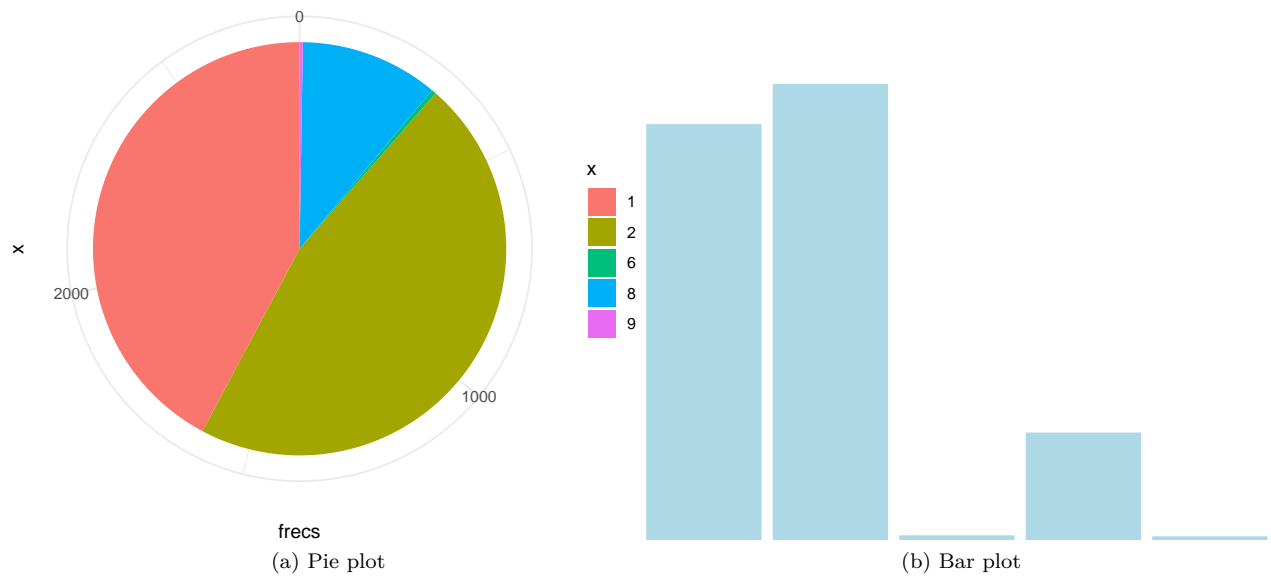


Figure 4: Anàlisi descriptiu de la variable Localització

#### Dia de la setmana

Variable	Stats / Values	Freqs (% of Valid)	Missing
DAY_WEEK [factor]	1. 1	356 (12.8%)	0 (0.0%)
	2. 2	311 (11.2%)	
	3. 3	410 (14.7%)	
	4. 4	437 (15.7%)	
	5. 5	460 (16.5%)	
	6. 6	387 (13.9%)	
	7. 7	420 (15.1%)	

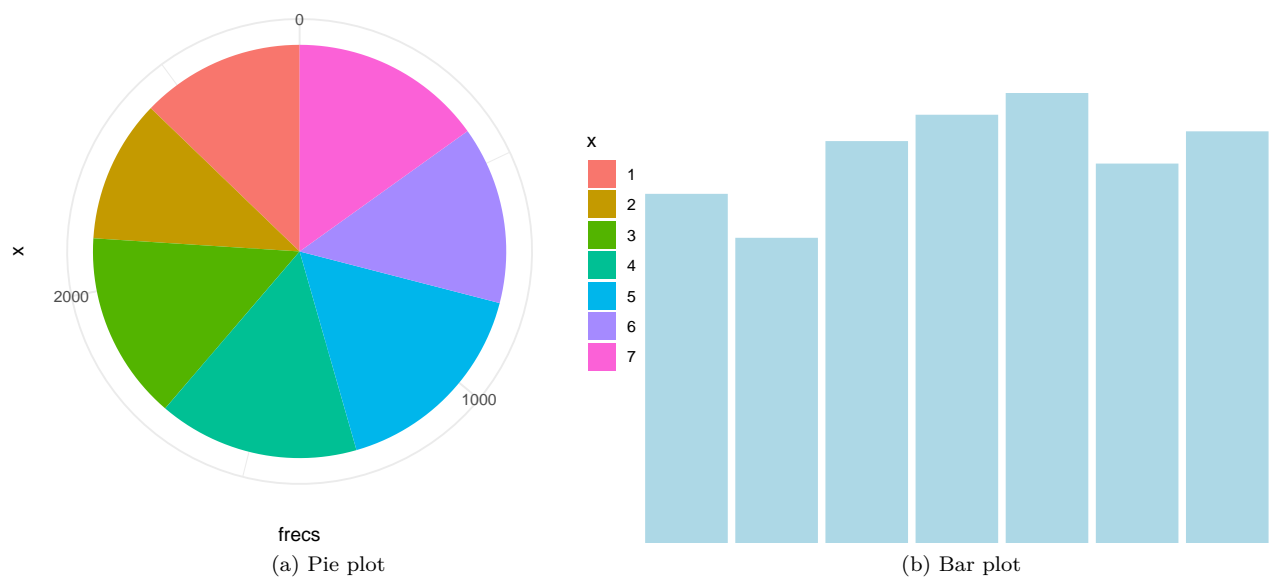


Figure 5: Anàlisi descriptiu de la variable Dia de la setmana

#### Nombre de ferits a l'accident

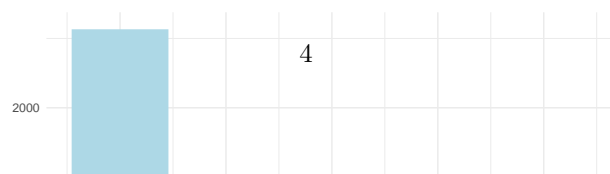


Table 5: Resum numèric de la variable Nombre de ferits a l'accident

N.Valid	Min	Q1	Median	Mean	Std.Dev	Q3	Max	IQR
2781	1	1	1	1.099964	0.3831946	1	5	0

### Nombre de conductors beguts

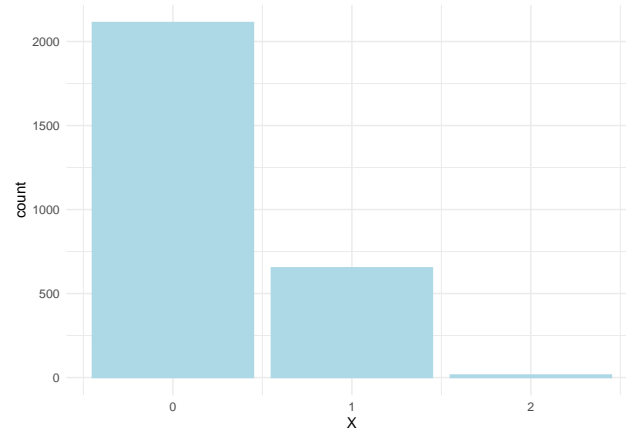


Figure 7: Histograma de la variable Nombre de conductors beguts

Table 6: Resum numèric de la variable Nombre de conductors beguts

N.Valid	Min	Q1	Median	Mean	Std.Dev	Q3	Max	IQR
2781	0	0	0	0.2455951	0.4428726	0	2	0

## Nombre de la persona dins de l'accident

Variable	Stats / Values	Freqs (% of Valid)	Missing
PER_NO [factor]	1. 1	4840 (68.3%)	0
	2. 2	1328 (18.7%)	(0.0%)
	3. 3	464 ( 6.5%)	
	4. 4	208 ( 2.9%)	
	5. 5	92 ( 1.3%)	
	6. 6	39 ( 0.6%)	
	7. 7	25 ( 0.4%)	
	8. 8	12 ( 0.2%)	
	9. 9	11 ( 0.2%)	
	10. 10	8 ( 0.1%)	
	[ 41 others ]	60 ( 0.8%)	

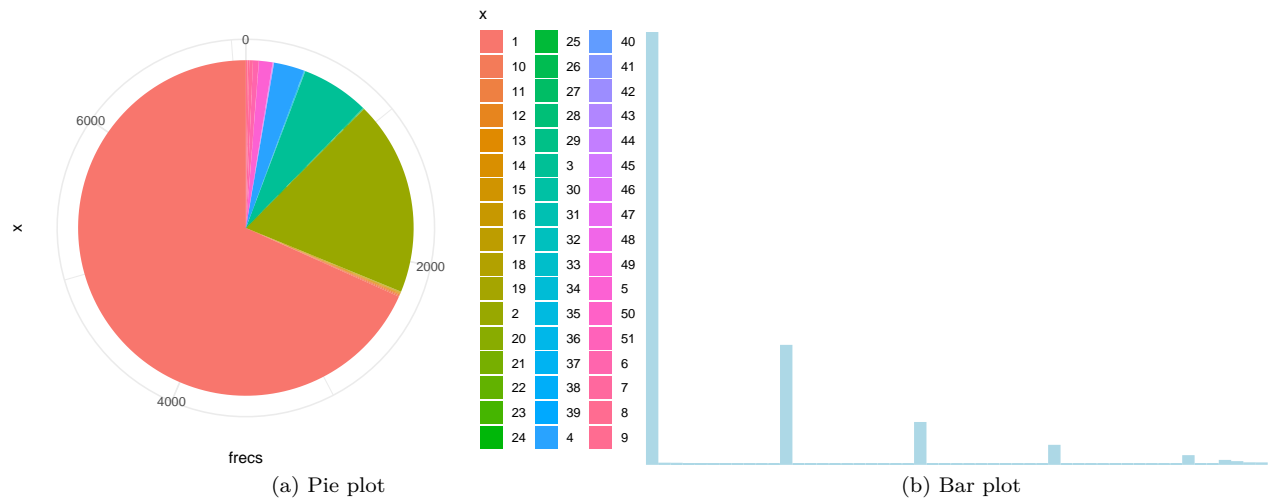


Figure 8: Anàlisi descriptiu de la variable Nombre de la persona



## Edat

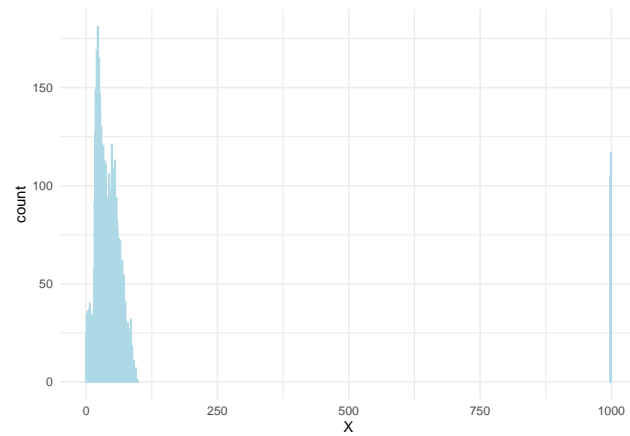


Figure 9: Histograma de la variable Edat

Table 8: Resum numèric de la variable Edat

N.Valid	Min	Q1	Median	Mean	Std.Dev	Q3	Max	IQR
7087	0	24	38	69.96487	168.1965	57	999	33

## Sexe

Variable	Stats / Values	Freqs (% of Valid)	Missing
SEX [factor]	1. 1	4575 (64.6%)	0 (0.0%)
	2. 2	2335 (32.9%)	
	3. 8	76 ( 1.1%)	
	4. 9	101 ( 1.4%)	

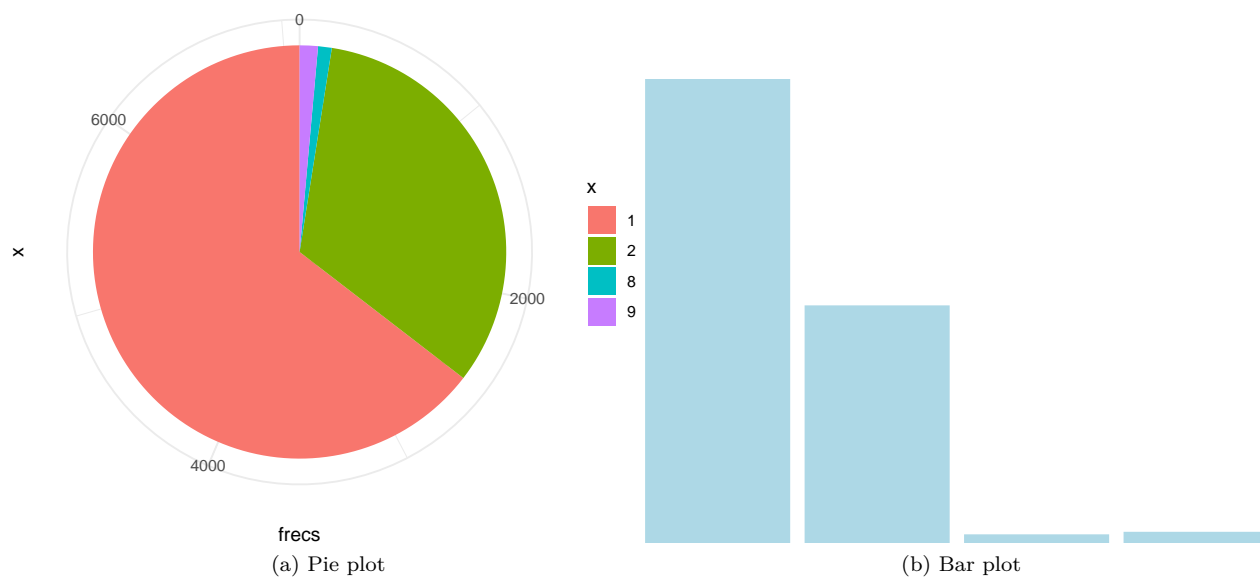


Figure 10: Anàlisi descriptiu de la variable Sexe

### Tipus de persona

Variable	Stats / Values	Freqs (% of Valid)	Missing
PER_TYP [factor]	1. 1	4161 (58.7%)	0
	2. 2	2168 (30.6%)	(0.0%)
	3. 3	12 ( 0.2%)	
	4. 4	23 ( 0.3%)	
	5. 5	643 ( 9.1%)	
	6. 6	52 ( 0.7%)	
	7. 8	14 ( 0.2%)	
	8. 9	14 ( 0.2%)	

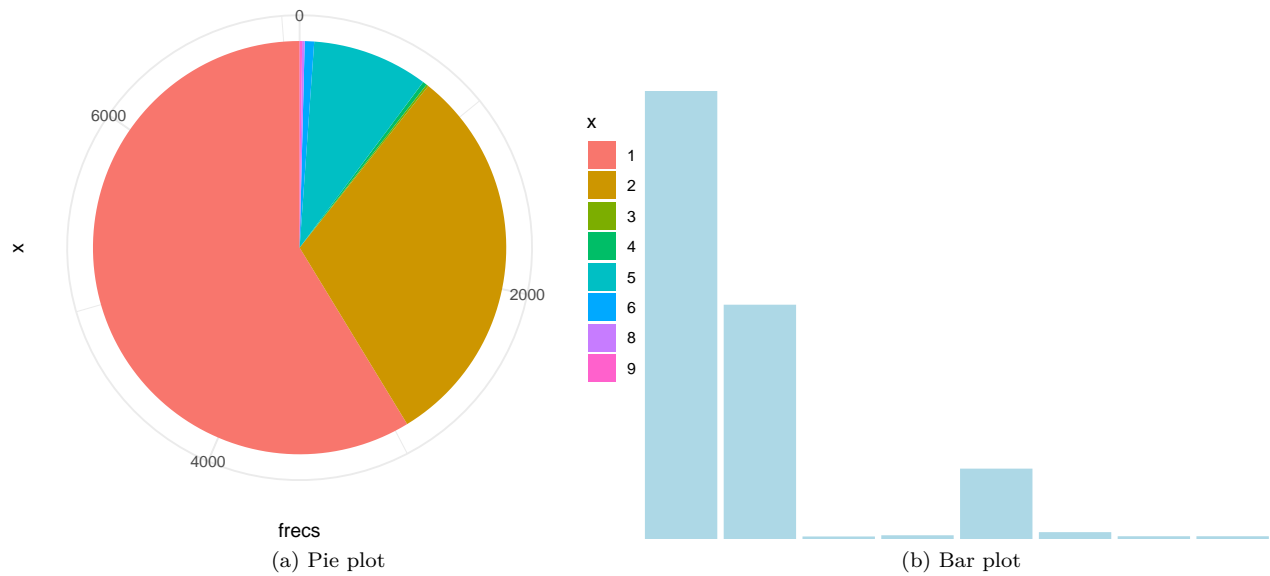


Figure 11: Anàlisi descriptiu de la variable Tipus de persona

### Tipus de víctima

Variable	Stats / Values	Freqs (% of Valid)	Missing
DOA [factor]	1. 0	5293 (74.7%)	0
	2. 7	1764 (24.9%)	(0.0%)
	3. 8	27 ( 0.4%)	
	4. 9	3 ( 0.0%)	

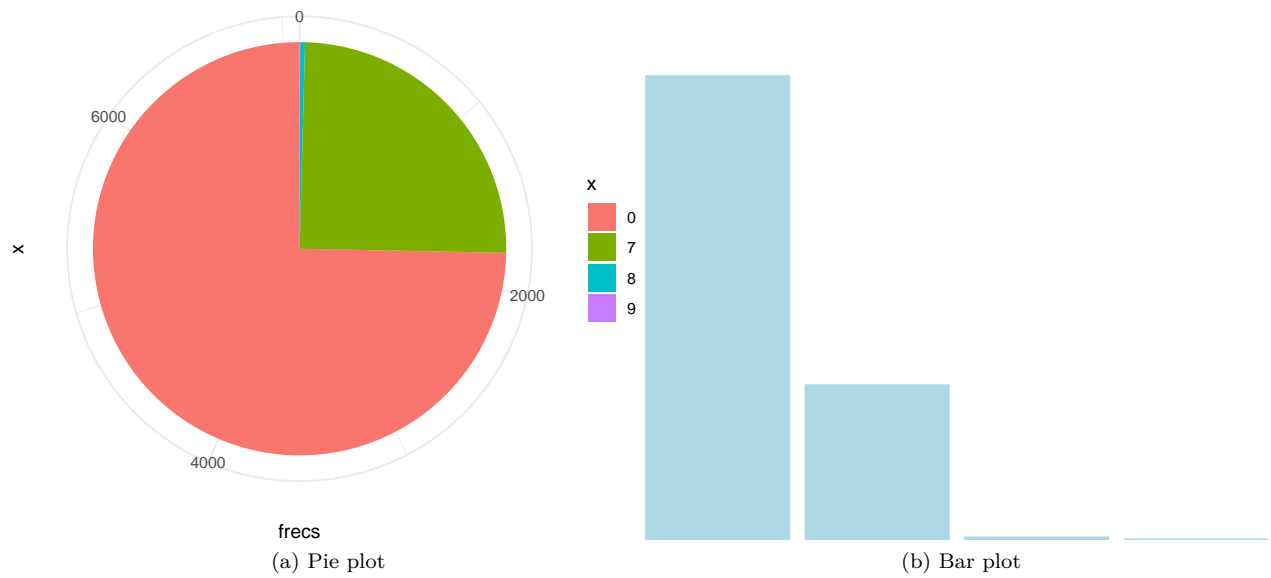


Figure 12: Anàlisi descriptiu de la variable Tipus de víctima

#### Nombre del vehicle dins de l'accident

Variable	Stats / Values	Freqs (% of Valid)	Missing
VEH_NO [factor]	1. 1	2777 (66.2%)	0 (0.0%)
	2. 2	1163 (27.7%)	
	3. 3	190 ( 4.5%)	
	4. 4	46 ( 1.1%)	
	5. 5	14 ( 0.3%)	
	6. 6	4 ( 0.1%)	

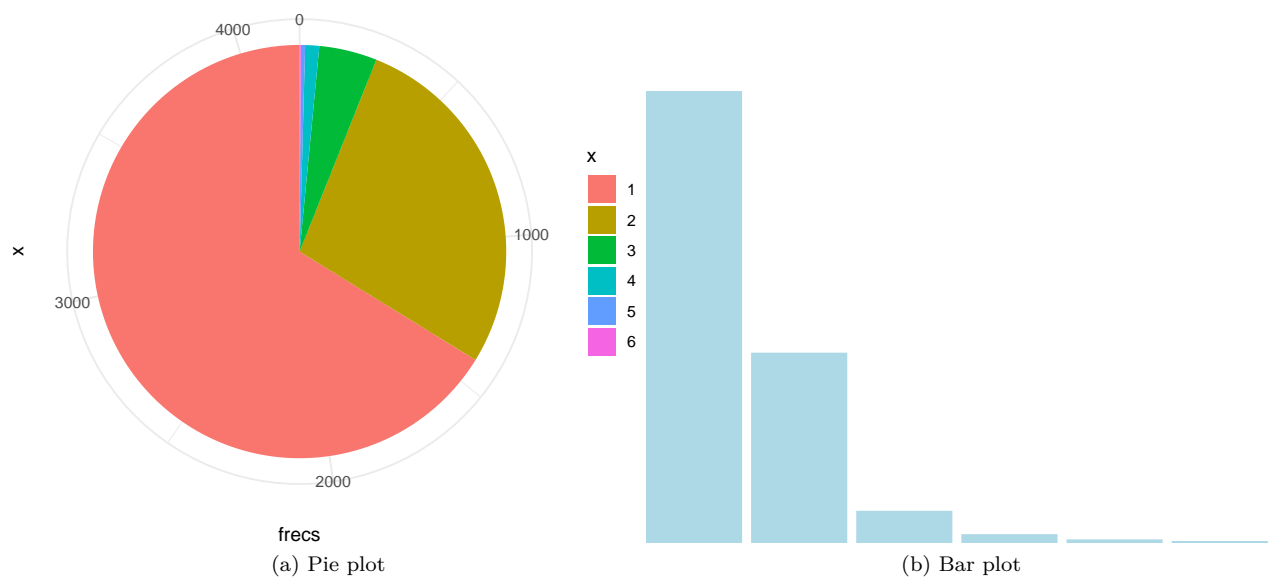


Figure 13: Anàlisi descriptiu de la variable Nombre del vehicle

#### Identificació del vehicle

Variable	Stats / Values	Freqs (% of Valid)	Missing
HIT_RUN [factor]	1. 0	4021 (95.9%)	0 (0.0%)
	2. 1	170 ( 4.1%)	
	3. 9	3 ( 0.1%)	

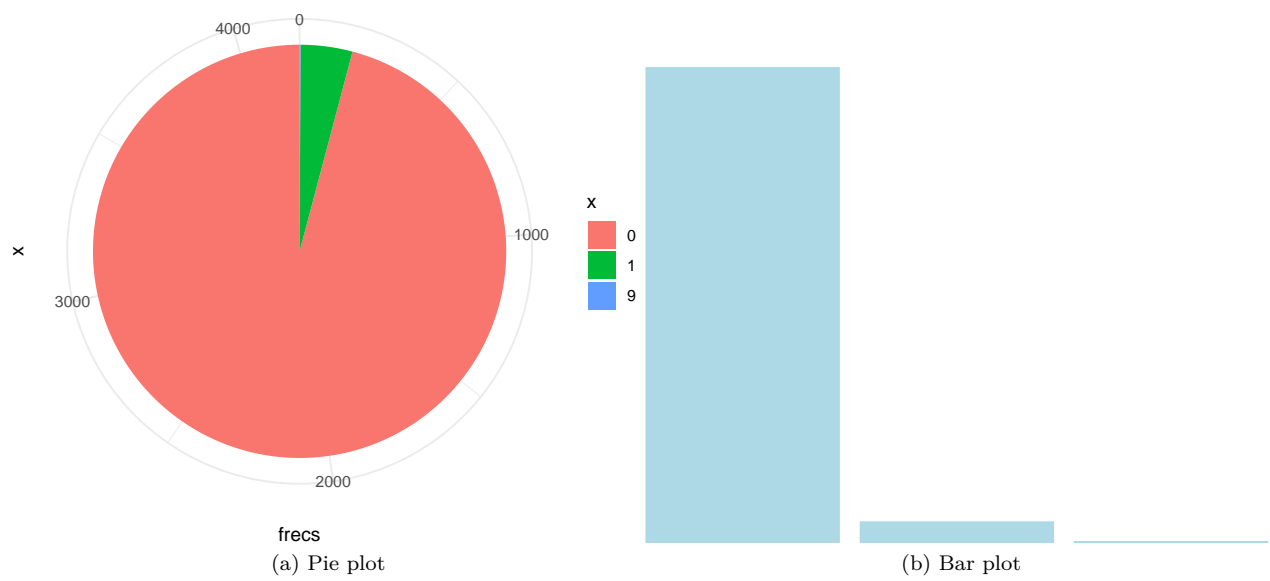


Figure 14: Anàlisi descriptiu de la variable Identificació del vehicle

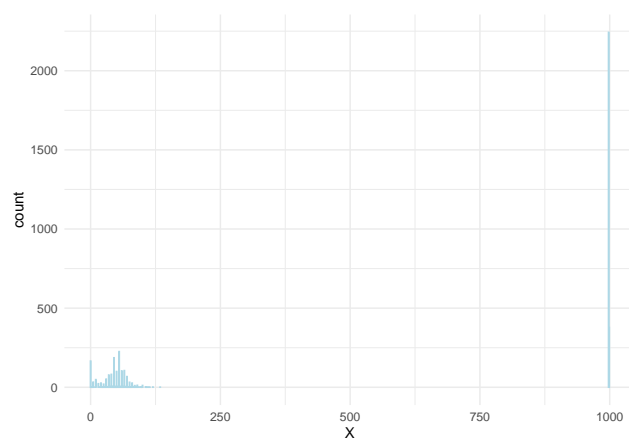


Figure 15: Histograma de la variable Velocitat estimada

Variable	Stats / Values	Freqs (% of Valid)	Missing
PREV_SP [factor]	1. 0	3167 (75.5%)	0 (0.0%)
	2. 1	549 (13.1%)	
	3. 2	164 ( 3.9%)	
	4. 3	61 ( 1.5%)	
	5. 4	21 ( 0.5%)	
	6. 5	4 ( 0.1%)	
	7. 7	1 ( 0.0%)	
	8. 9	1 ( 0.0%)	
	9. 99	198 ( 4.7%)	
	10. 998	28 ( 0.7%)	

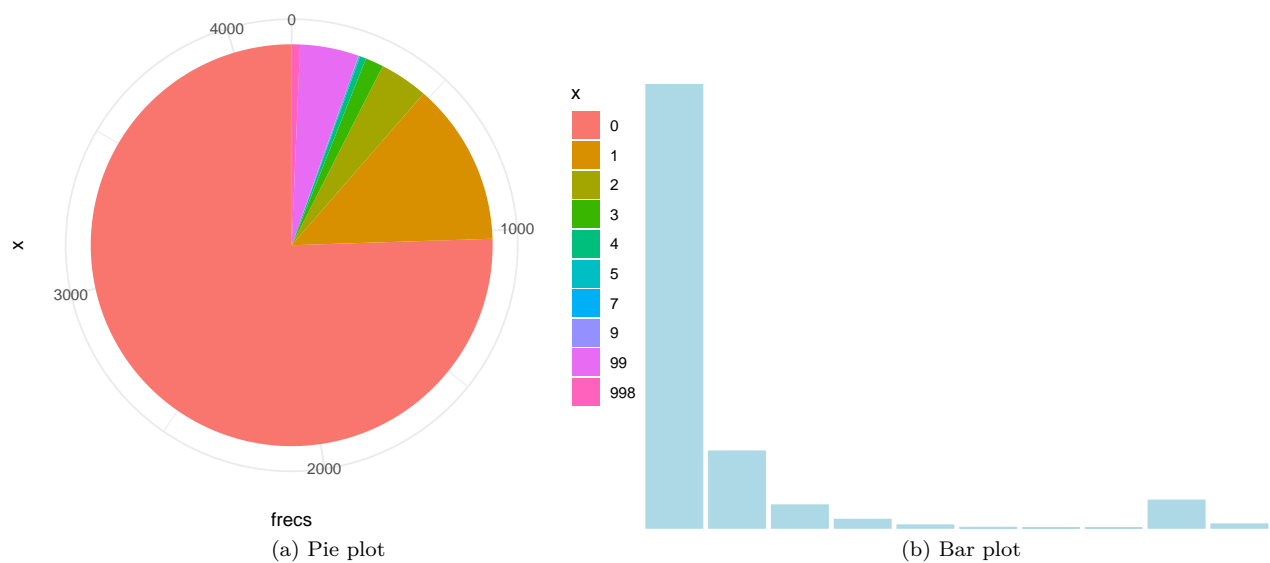


Figure 16: Anàlisi descriptiu de la variable Indicador d'existència de límit de velocitat

## Variables numèriques