

Introduction

The dataset contains 8368 rows and the following columns:

CPF, NASC, ESCOLARIDADE, ESTADO_CIVIL, RENDA_PRESUMIDA, QTD_Veiculos, SEXO
COD_PODER_AQUISITIVO, Regional, Filial, inicio_treinamento, termino_treinamento
Entrega_operacao, Instrutor, Fonte, Nascimento, Idade, DataDem, Status Discovery
Tipo Demissao, DataAdm, Dias Desligue, Causa Desligue, Escolaridade_Jobs
Distancia

Column Info

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 8368 entries, 0 to 8367
```

```
Data columns (total 25 columns):
```

#	Column	Non-Null Count	Dtype
0	CPF	8368 non-null	int64
1	NASC	8049 non-null	datetime64[ns]
2	ESCOLARIDADE	3033 non-null	object
3	ESTADO_CIVIL	456 non-null	object
4	RENDA_PRESUMIDA	6197 non-null	float64
5	QTD_Veiculos	118 non-null	object
6	SEXO	7894 non-null	object
7	COD_PODER_AQUISITIVO	5234 non-null	object
8	Regional	8368 non-null	object
9	Filial	8368 non-null	object
10	inicio_treinamento	8368 non-null	object
11	termino_treinamento	8368 non-null	object
12	Entrega_operacao	8368 non-null	object
13	Instrutor	8368 non-null	object
14	Fonte	8368 non-null	object
15	Nascimento	8368 non-null	object
16	Idade	8368 non-null	object
17	DataDem	8368 non-null	object
18	Status Discovery	8368 non-null	object
19	Tipo Demissao	8368 non-null	object
20	DataAdm	8368 non-null	object
21	Dias Desligue	8368 non-null	object
22	Causa Desligue	8368 non-null	object
23	Escolaridade_Jobs	8368 non-null	object
24	Distancia	8368 non-null	object

dtypes: datetime64[ns](1), float64(1), int64(1), object(22)
memory usage: 1.6+ MB

Column Info Transformed

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 8368 entries, 0 to 8367  
Data columns (total 25 columns):  
#   Column                Non-Null Count  Dtype  
---  ---  
0   CPF                    8368 non-null   int64  
1   NASC                   8049 non-null   datetime64[ns]  
2   ESCOLARIDADE          3033 non-null   object  
3   ESTADO_CIVIL          456 non-null    object  
4   RENDA_PRESUMIDA       6197 non-null   float64  
5   QTD_Veiculos          118 non-null    float64  
6   SEXO                   7894 non-null   object  
7   COD_PODER_AQUISITIVO  5234 non-null   float64  
8   Regional              8368 non-null   object  
9   Filial                8368 non-null   object  
10  inicio_treinamento    8368 non-null   int64  
11  termino_treinamento  8368 non-null   int64  
12  Entrega_operacao      8368 non-null   int64  
13  Instrutor             8368 non-null   object  
14  Fonte                 8368 non-null   object  
15  Nascimento            8368 non-null   object  
16  Idade                 8368 non-null   object  
17  DataDem               8368 non-null   object  
18  Status_Discovery      8368 non-null   object  
19  Tipo_Demissao         8368 non-null   object  
20  DataAdm               8368 non-null   object  
21  Dias_Desligue         8368 non-null   object  
22  Causa_Desligue        8368 non-null   object  
23  Escolaridade_Jobs     8368 non-null   object  
24  Distancia             8368 non-null   object  
dtypes: datetime64[ns](1), float64(3), int64(4), object(17)  
memory usage: 1.6+ MB
```

Transformed Columns

Converted to datetime: inicio_treinamento, termino_treinamento, Entrega_operacao
Nascimento, DataDem, DataAdm

Converted to int: QTD_Veiculos
COD_PODER_AQUISITIVO, inicio_treinamento, termino_treinamento
Entrega_operacao

Converted to bool:

Integer Columns

CPF, inicio_treinamento, termino_treinamento, Entrega_operacao

Float Columns

RENDA_PRESUMIDA, QTD_Veiculos, COD_PODER_AQUISITIVO

String Columns

ESCOLARIDADE, ESTADO_CIVIL, SEXO, Regional, Filial, Instrutor, Fonte, Nascimento
Idade, DataDem, Status Discovery, Tipo Demissao, DataAdm, Dias Desligue
Causa Desligue, Escolaridade_Jobs, Distancia

Boolean Columns

Table Unique Values Count

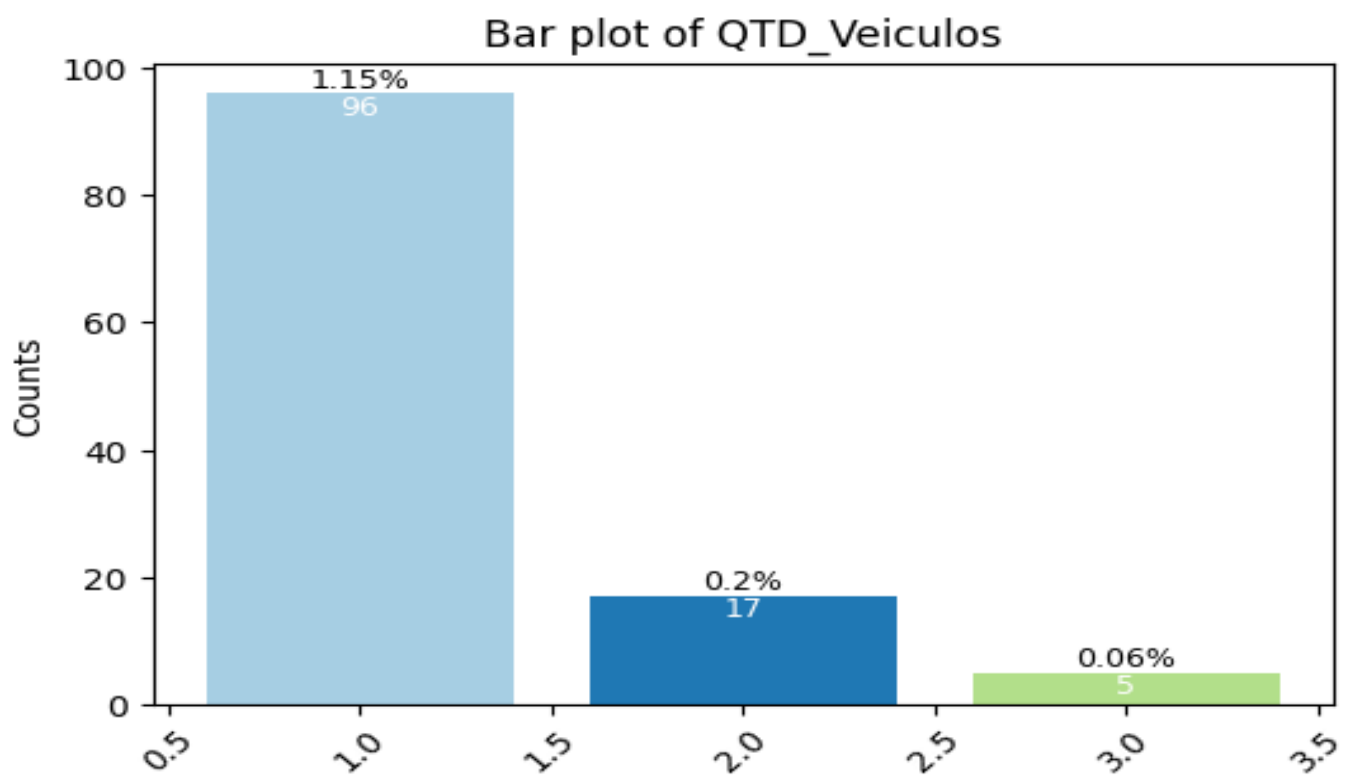
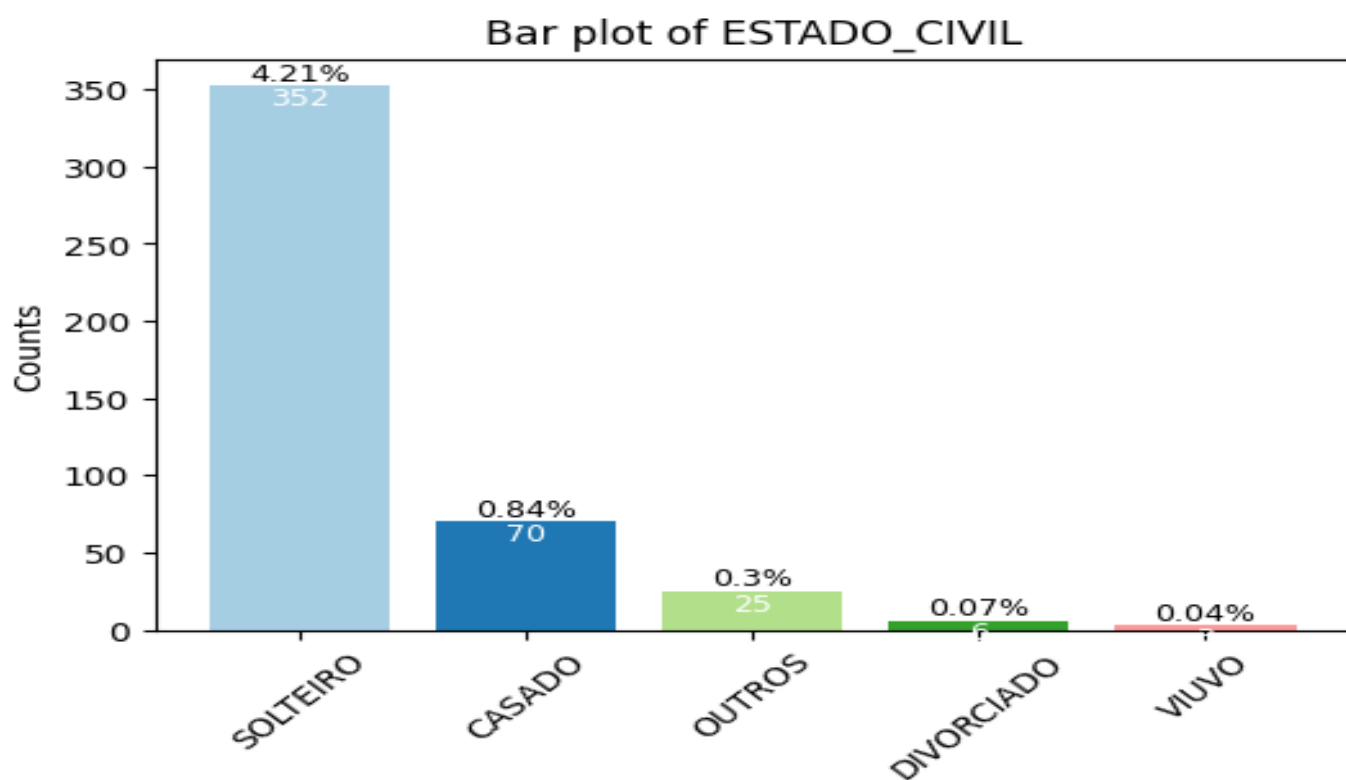
Column Name	Unique Values Count
CPF	8027
NASC	4490
ESCOLARIDADE	16
ESTADO_CIVIL	5
RENDAPRESUMIDA	506
QTD_Veiculos	3
SEXO	3
COD_PODER_AQUISITIVO	7
Regional	79
Filial	8
inicio_treinamento	67
termino_treinamento	75
Entrega_operacao	70
Instrutor	173
Fonte	20
Nascimento	3357
Idade	3375
DataDem	192
Status Discovery	10
Tipo Demissao	2
DataAdm	197
Dias Desligue	168
Causa Desligue	4
Escolaridade_Jobs	5
Distancia	1865

Possible IDs

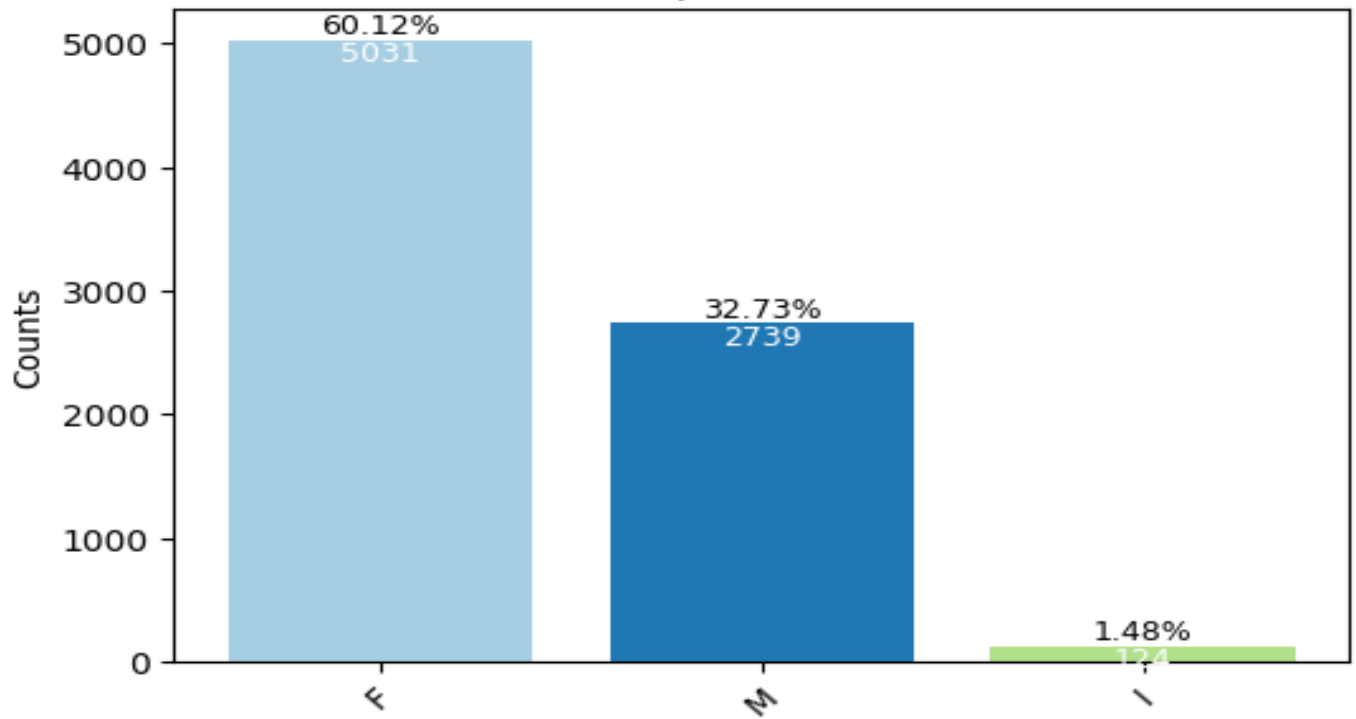
CPF

Table Education Levels

Nível de educação	Quantidade
ENSINO MEDIO COMPLETO	1774
SUPERIOR COMPLETO	473
ENSINO MEDIO INCOMPLETO	232
ENSINO FUNDAMENTAL COMPLETO	225
SUPERIOR INCOMPLETO	217
TECNOLOGO INCOMPLETO	20
TECNICO COMPLETO	19
POS-GRADUACAO COMPLETO	16
TECNOLOGO COMPLETO	15
TECNICO INCOMPLETO	13
SUPERIOR COMPLETO	8
SUPERIOR INCOMPLETO	8
PRIMARIO COMPLETO	6
ENSINO FUNDAMENTAL INCOMPLETO	4
PRIMARIO COMPLETO	2
PRIMARIO INCOMPLETO	1



Bar plot of SEXO



Bar plot of COD_PODER_AQUISITIVO

