

Introduction

The dataset contains 5926 rows, 29 columns, and a total of 171854 data points in Dados_Oper_FTZ.csv. The columns are:

ADMISSÃO, DEMISSÃO, NASCIMENTO, SEXO, RECEBE VTR, CLIENTE, TÍTULO
CLASSIFICAÇÃO CCU, DESCRIÇÃO, RUA, BAIRRO, CIDADE, CEP, ESTADO, ESCOLARIDADE
ESTADO_CIVIL, FILHOS, ESCALA, CLASSE, CARGA, 1º Emprego, SITE, ID, NIVEL CARGO
Dias_trabalhados, ano_ADMISSÃO, mes_ADMISSÃO, Idade, Mais_de_1_ano

Column Info

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5926 entries, 0 to 5925
Data columns (total 29 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   ADMISSÃO              5926 non-null   object
1   DEMISSÃO              5926 non-null   object
2   NASCIMENTO            5926 non-null   object
3   SEXO                  5926 non-null   object
4   RECEBE VTR           5926 non-null   object
5   CLIENTE               5926 non-null   object
6   TÍTULO                5926 non-null   object
7   CLASSIFICAÇÃO CCU    5926 non-null   object
8   DESCRIÇÃO             5926 non-null   object
9   RUA                   5926 non-null   object
10  BAIRRO                5926 non-null   object
11  CIDADE                5926 non-null   object
12  CEP                   5926 non-null   int64
13  ESTADO               5926 non-null   object
14  ESCOLARIDADE         5926 non-null   object
15  ESTADO_CIVIL         5926 non-null   object
16  FILHOS               5926 non-null   int64
17  ESCALA               5926 non-null   object
18  CLASSE               5926 non-null   object
19  CARGA                5926 non-null   object
20  1º Emprego           5926 non-null   object
21  SITE                 5926 non-null   object
22  ID                   5926 non-null   int64
23  NIVEL CARGO          5926 non-null   object
24  Dias_trabalhados     5926 non-null   int64
25  ano_ADMISSÃO         5926 non-null   int64
26  mes_ADMISSÃO         5926 non-null   int64
27  Idade                5926 non-null   int64
28  Mais_de_1_ano        5926 non-null   int64
dtypes: int64(8), object(21)
memory usage: 1.3+ MB
```

Column Info Transformed

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5926 entries, 0 to 5925
Data columns (total 29 columns):
#   Column                Non-Null Count  Dtype
---  -
0   ADMISSÃO              5926 non-null  object
1   DEMISSÃO              5926 non-null  object
2   NASCIMENTO            5926 non-null  object
3   SEXO                  5926 non-null  object
4   RECEBE VTR            5926 non-null  object
5   CLIENTE               5926 non-null  object
6   TÍTULO                5926 non-null  object
7   CLASSIFICAÇÃO CCU     5926 non-null  object
8   DESCRIÇÃO             5926 non-null  object
9   RUA                   5926 non-null  object
10  BAIRRO                5926 non-null  object
11  CIDADE                5926 non-null  object
12  CEP                   5926 non-null  int64
13  ESTADO                5926 non-null  object
14  ESCOLARIDADE          5926 non-null  object
15  ESTADO_CIVIL          5926 non-null  object
16  FILHOS                5926 non-null  int64
17  ESCALA                5926 non-null  object
18  CLASSE                5926 non-null  int64
19  CARGA                 5926 non-null  int64
20  1º Emprego            5926 non-null  object
21  SITE                  5926 non-null  object
22  ID                    5926 non-null  int64
23  NIVEL CARGO           5926 non-null  object
24  Dias_trabalhados      5926 non-null  int64
25  ano_ADMISSÃO          5926 non-null  int64
26  mes_ADMISSÃO          5926 non-null  int64
27  Idade                 5926 non-null  int64
28  Mais_de_1_ano         5926 non-null  bool
dtypes: bool(1), int64(9), object(19)
memory usage: 1.3+ MB
```

Transformed Columns

Converted to datetime: ADMISSÃO, DEMISSÃO, NASCIMENTO

Converted to int: CLASSE
CARGA

Converted to bool: Mais_de_1_ano

Integer Columns

CEP, FILHOS, CLASSE, CARGA, ID, Dias_trabalhados, ano_ADMISSÃO, mes_ADMISSÃO
Idade

Float Columns

String Columns

ADMISSÃO, DEMISSÃO, NASCIMENTO, SEXO, RECEBE VTR, CLIENTE, TÍTULO
CLASSIFICAÇÃO CCU, DESCRIÇÃO, RUA, BAIRRO, CIDADE, ESTADO, ESCOLARIDADE
ESTADO_CIVIL, ESCALA, 1º Emprego, SITE, NIVEL CARGO

Boolean Columns

Mais_de_1_ano

Table Unique Values Count (Part 1)

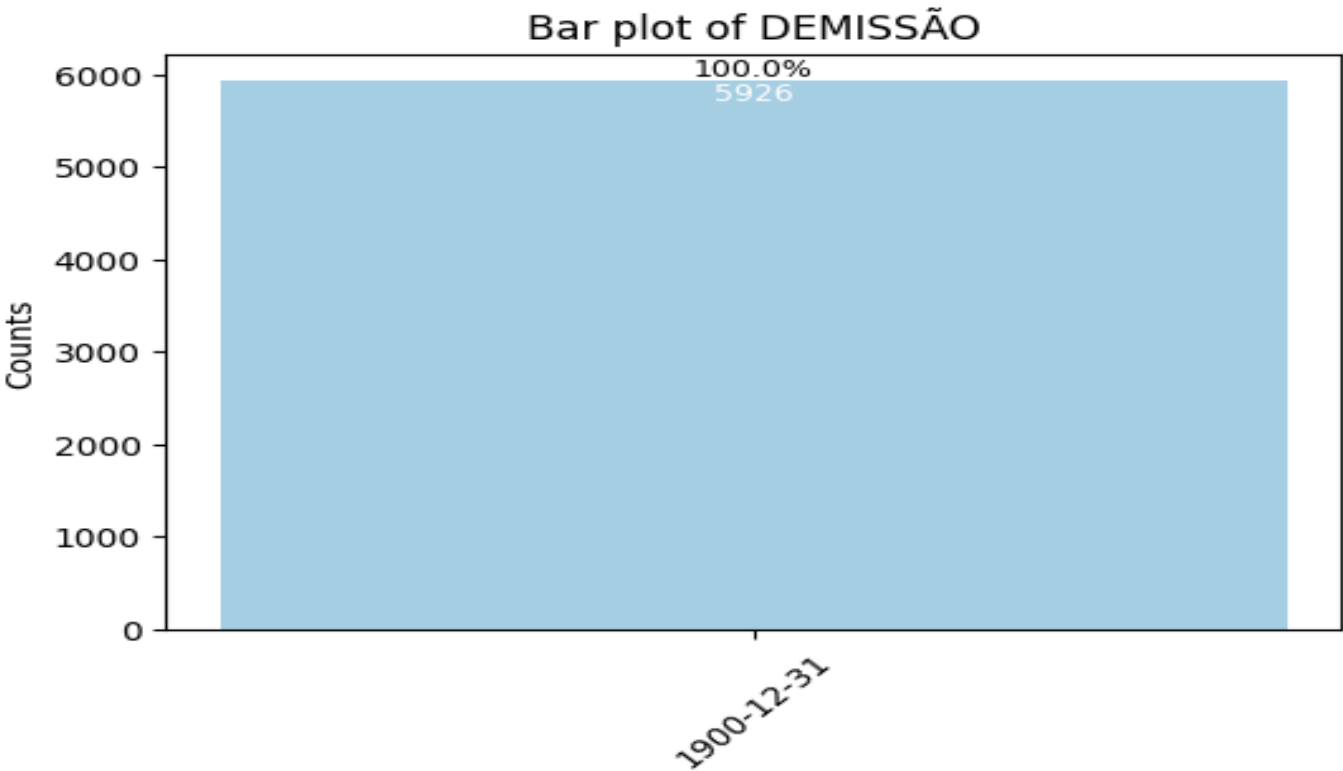
Column Name	Unique Values Count
ADMISSÃO	368
DEMISSÃO	1
NASCIMENTO	3841
SEXO	2
RECEBE VTR	2
CLIENTE	11
TÍTULO	4
CLASSIFICAÇÃO CCU	2
DESCRIÇÃO	13
RUA	3015
BAIRRO	536
CIDADE	19
CEP	3247
ESTADO	5
ESCOLARIDADE	4
ESTADO_CIVIL	5
FILHOS	6
ESCALA	127
CLASSE	6
CARGA	4
1º Emprego	2
SITE	1
ID	5926
NIVEL CARGO	1
Dias_trabalhados	368
ano_ADMISSÃO	13
mes_ADMISSÃO	12
Idade	44
Mais_de_1_ano	2

Possible IDs

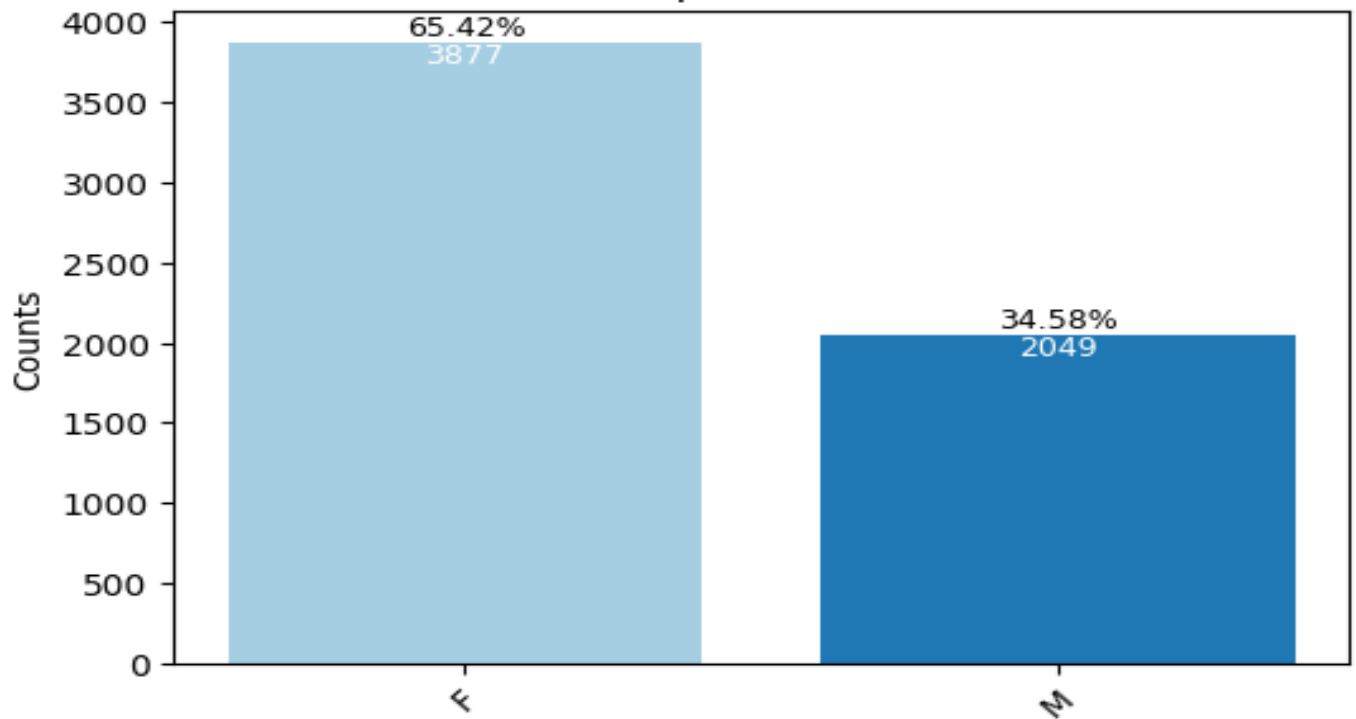
ID

Table Education Levels

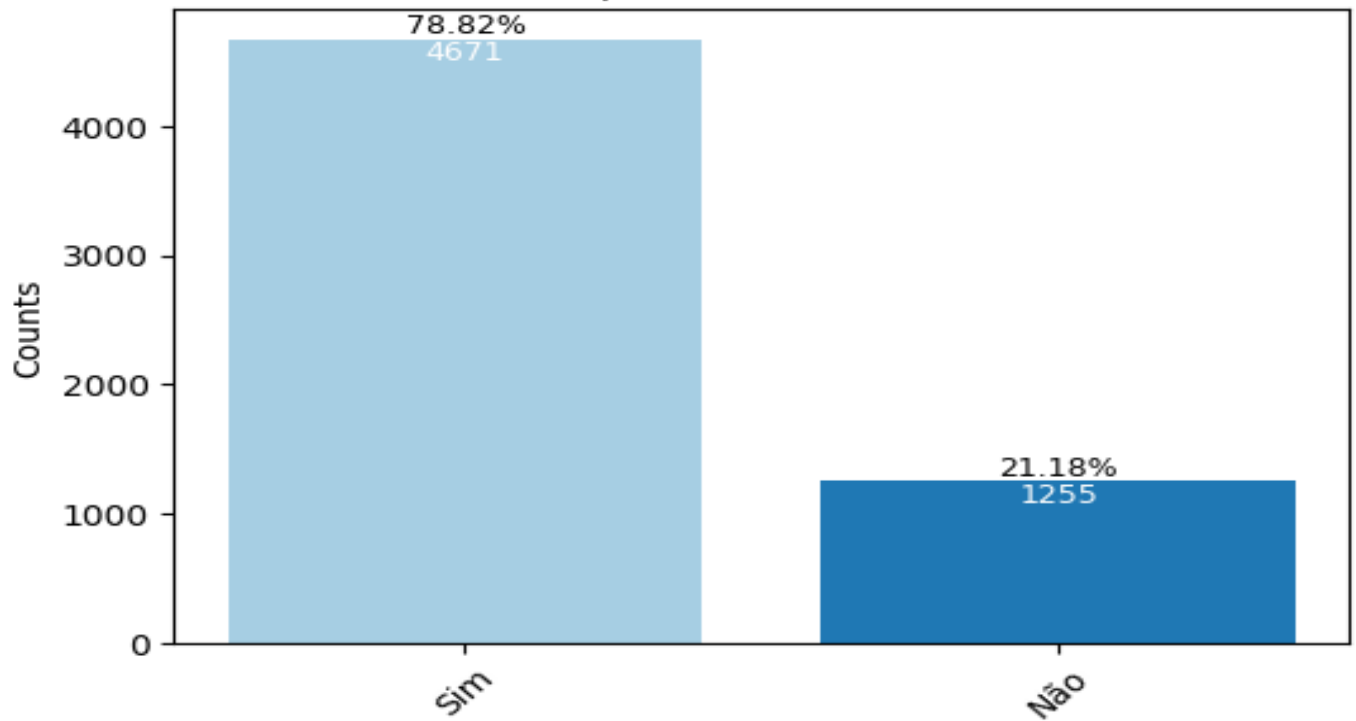
Nível de educação	Quantidade
ENSINO MEDIO COMPLETO	5309
SUPERIOR INCOMPLETO	338
SUPERIOR COMPLETO	260
ENSINO MEDIO INCOMPLETO	19

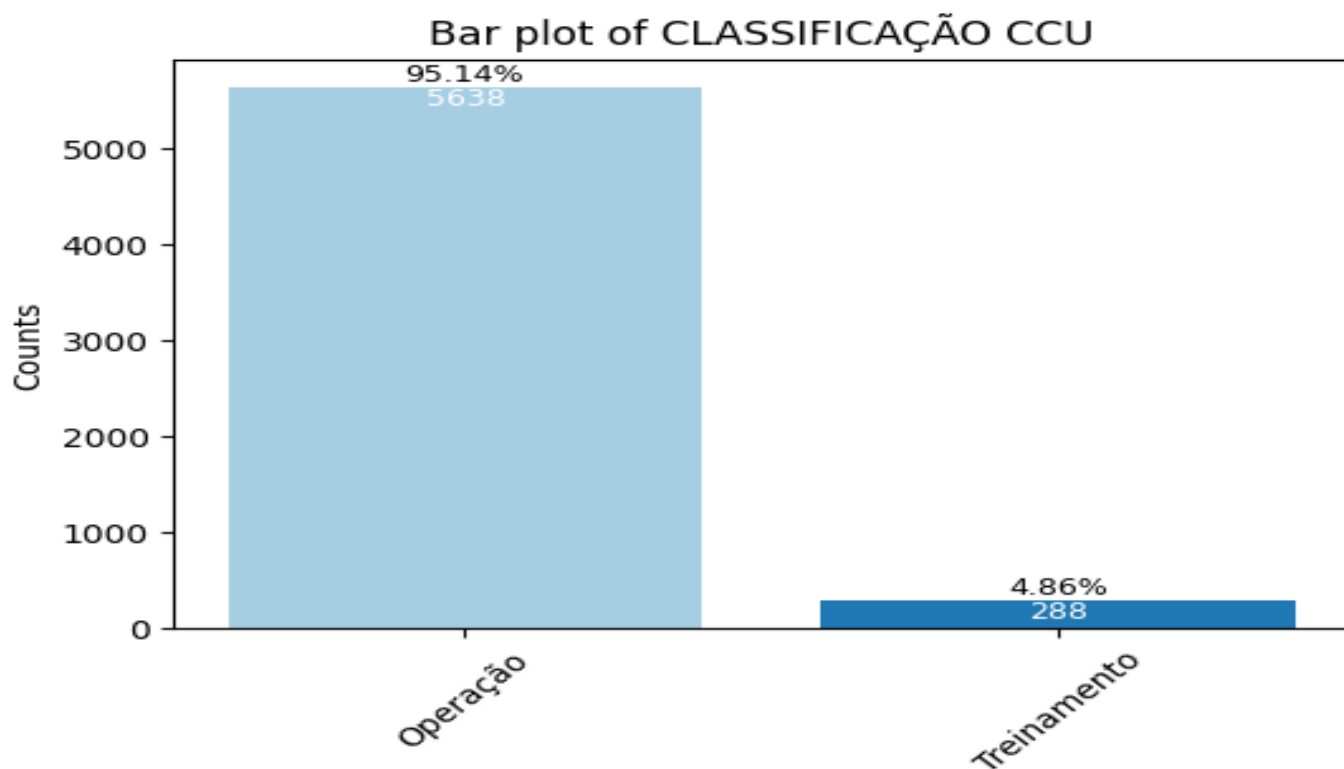
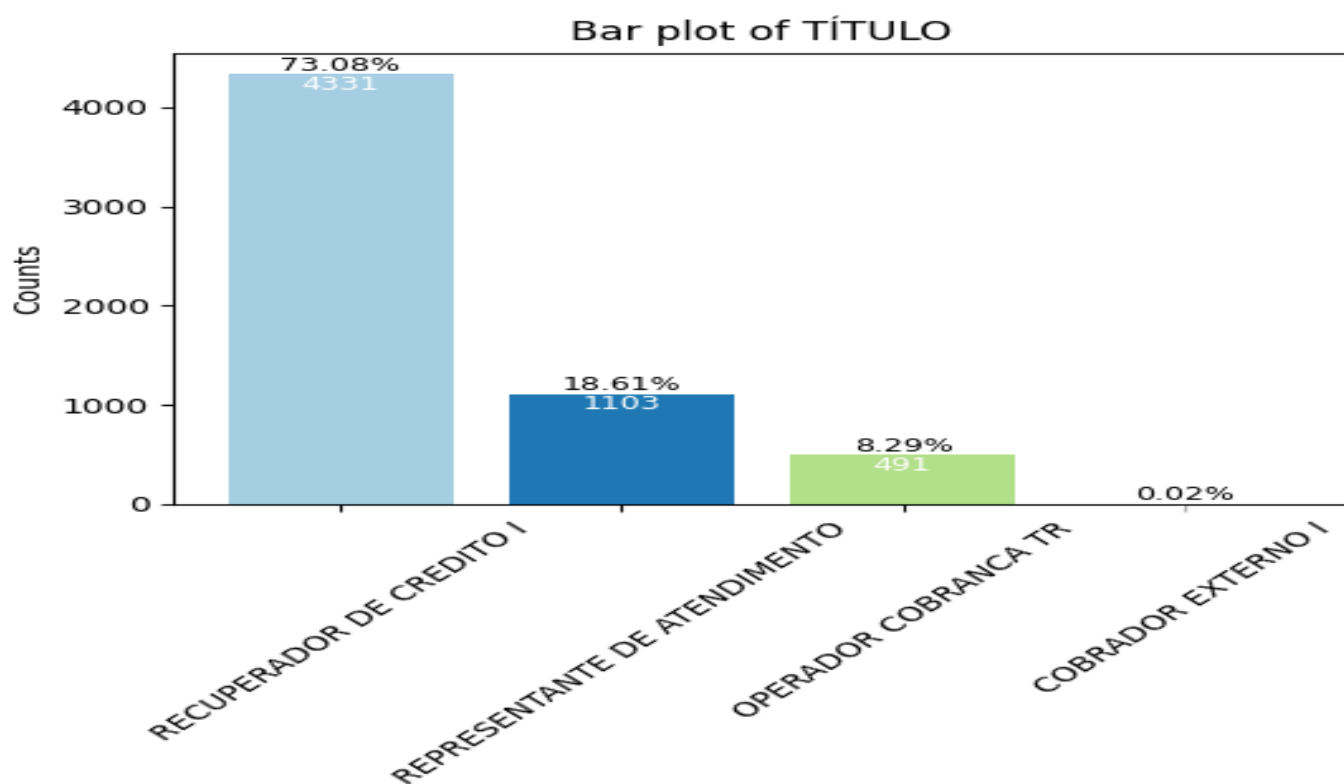


Bar plot of SEXO

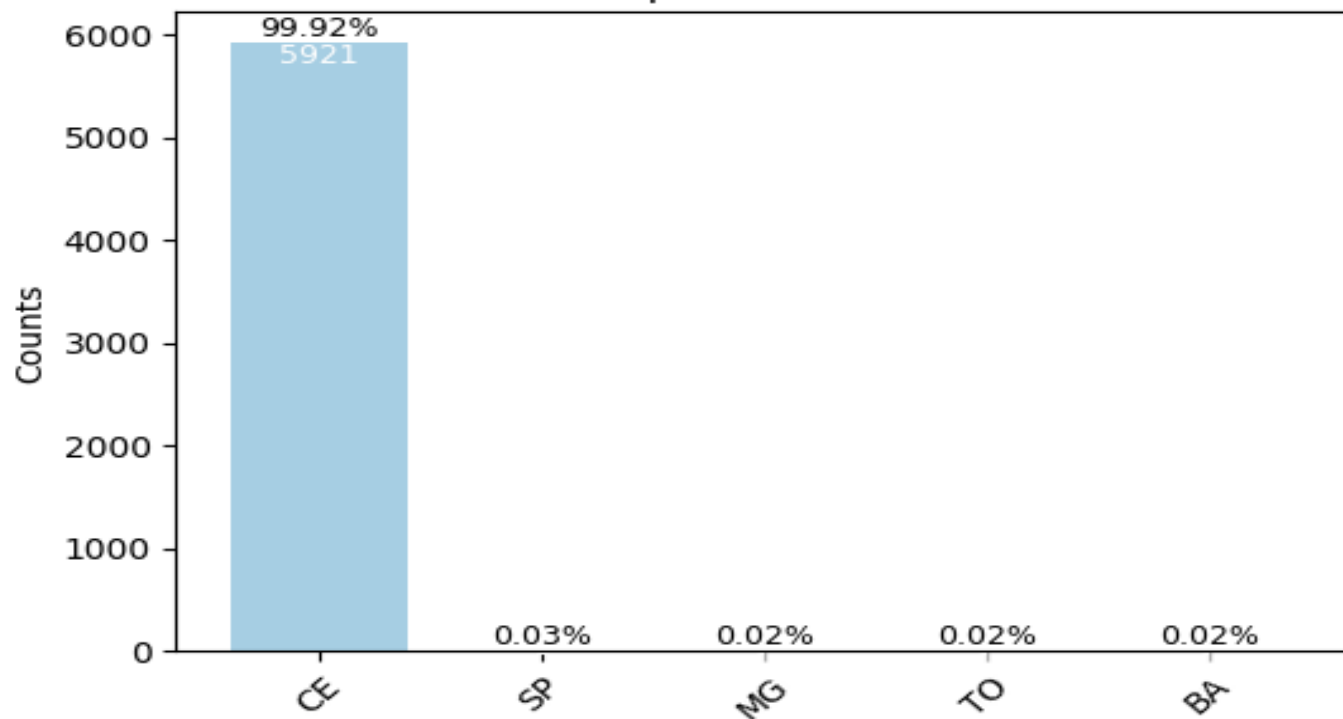


Bar plot of RECEBE VTR

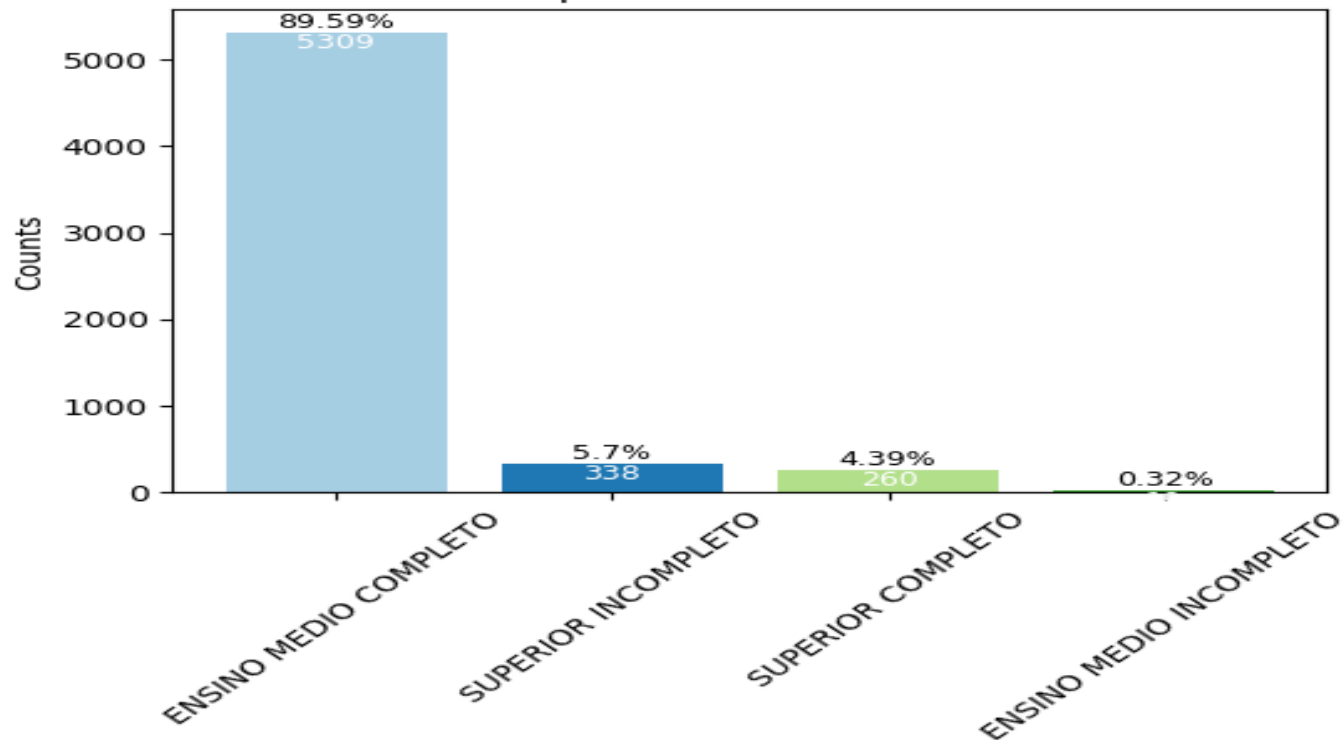




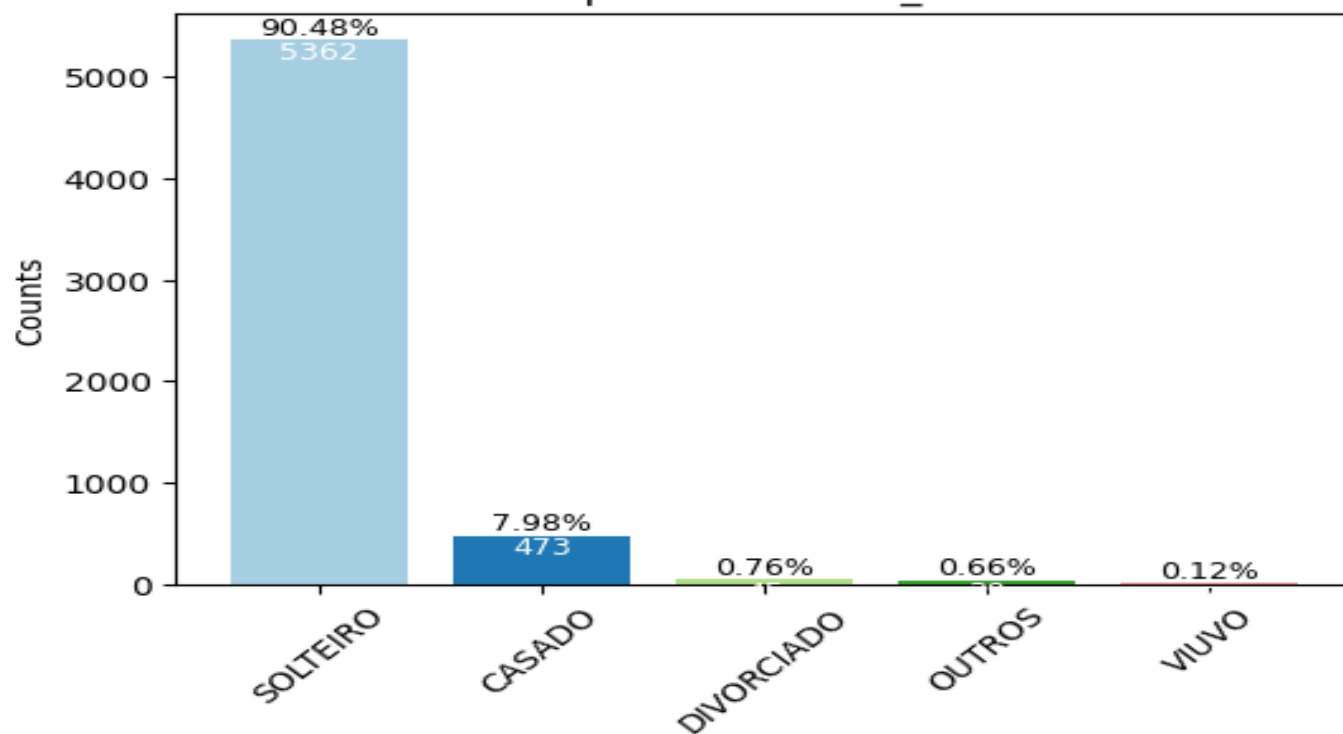
Bar plot of ESTADO



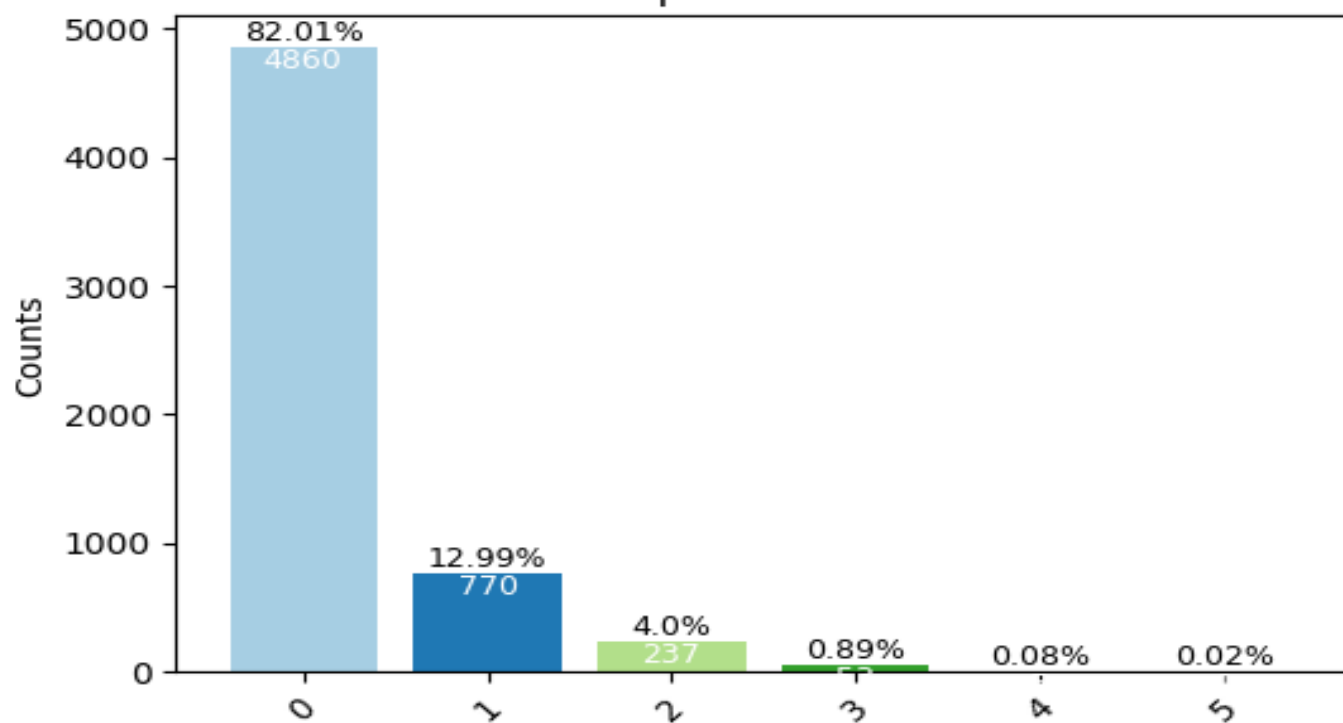
Bar plot of ESCOLARIDADE

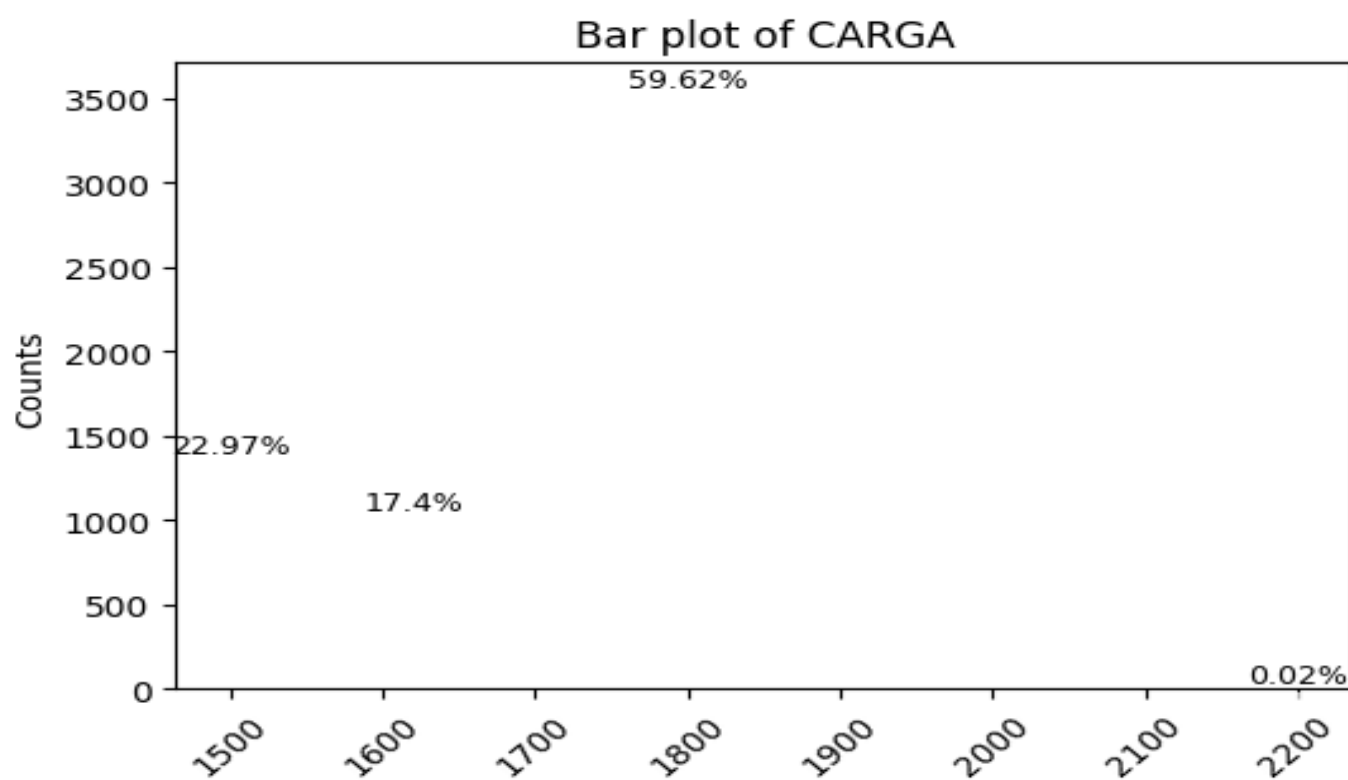
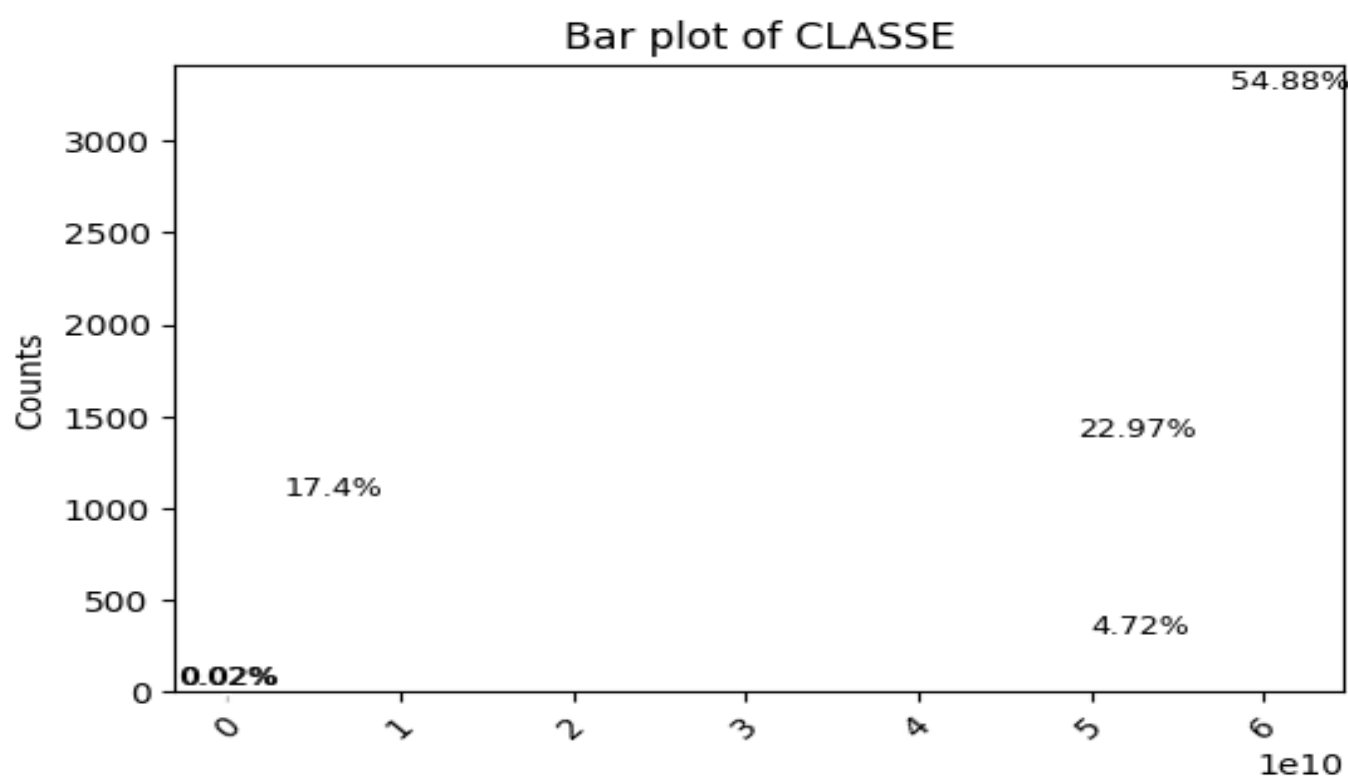


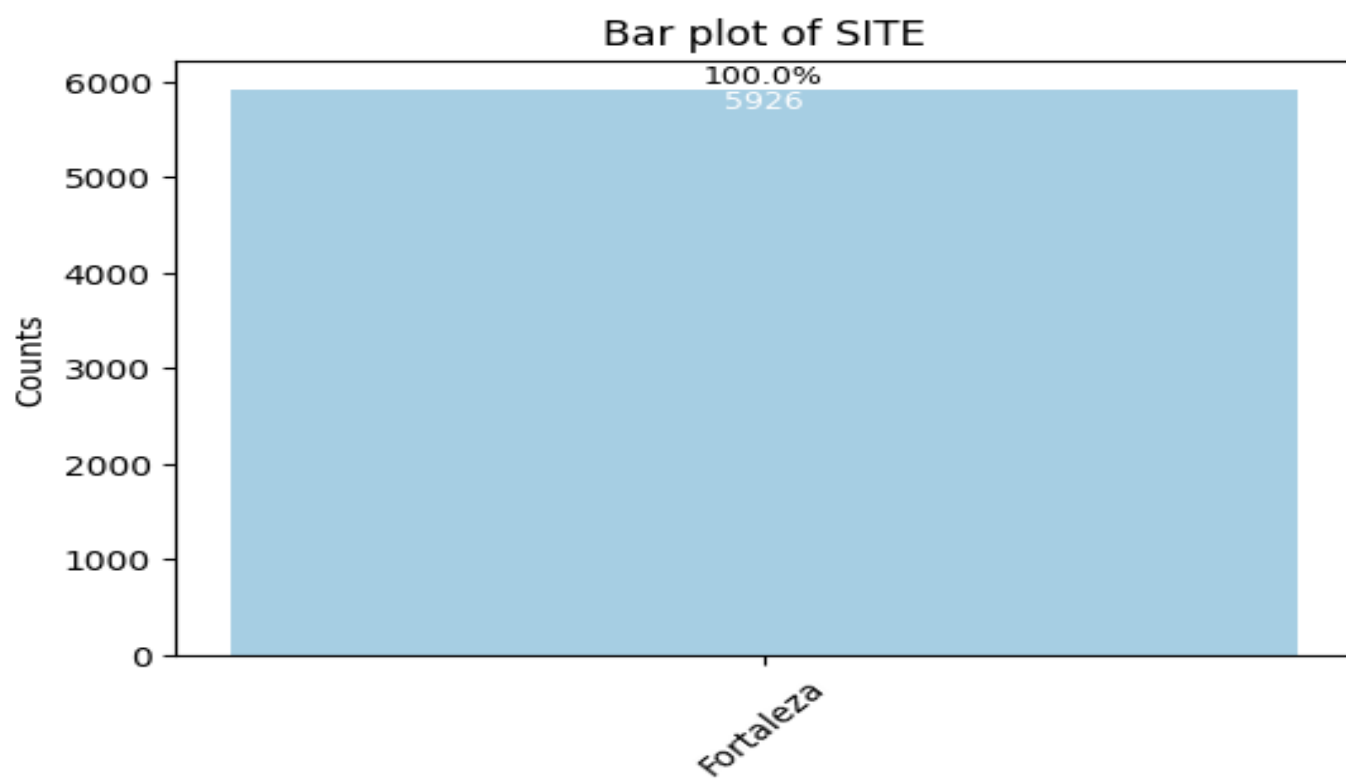
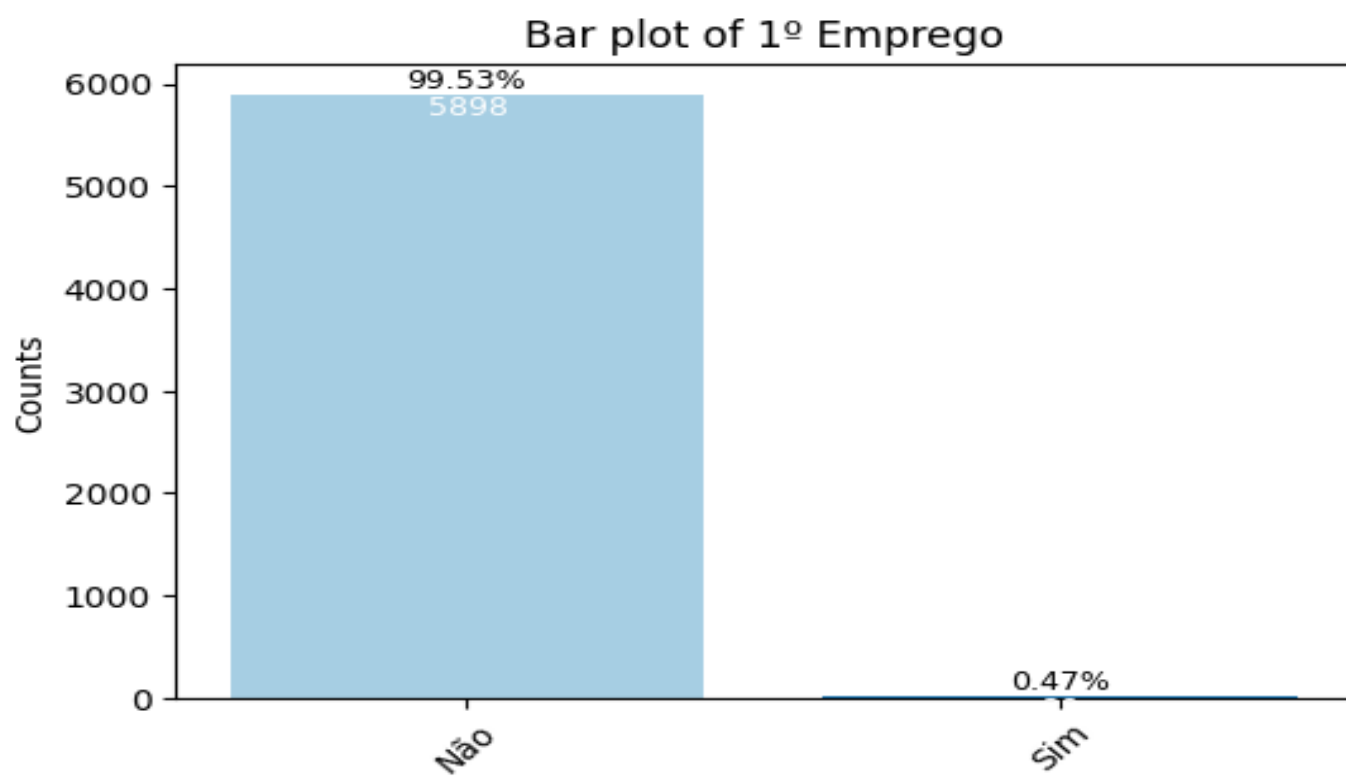
Bar plot of ESTADO_CIVIL

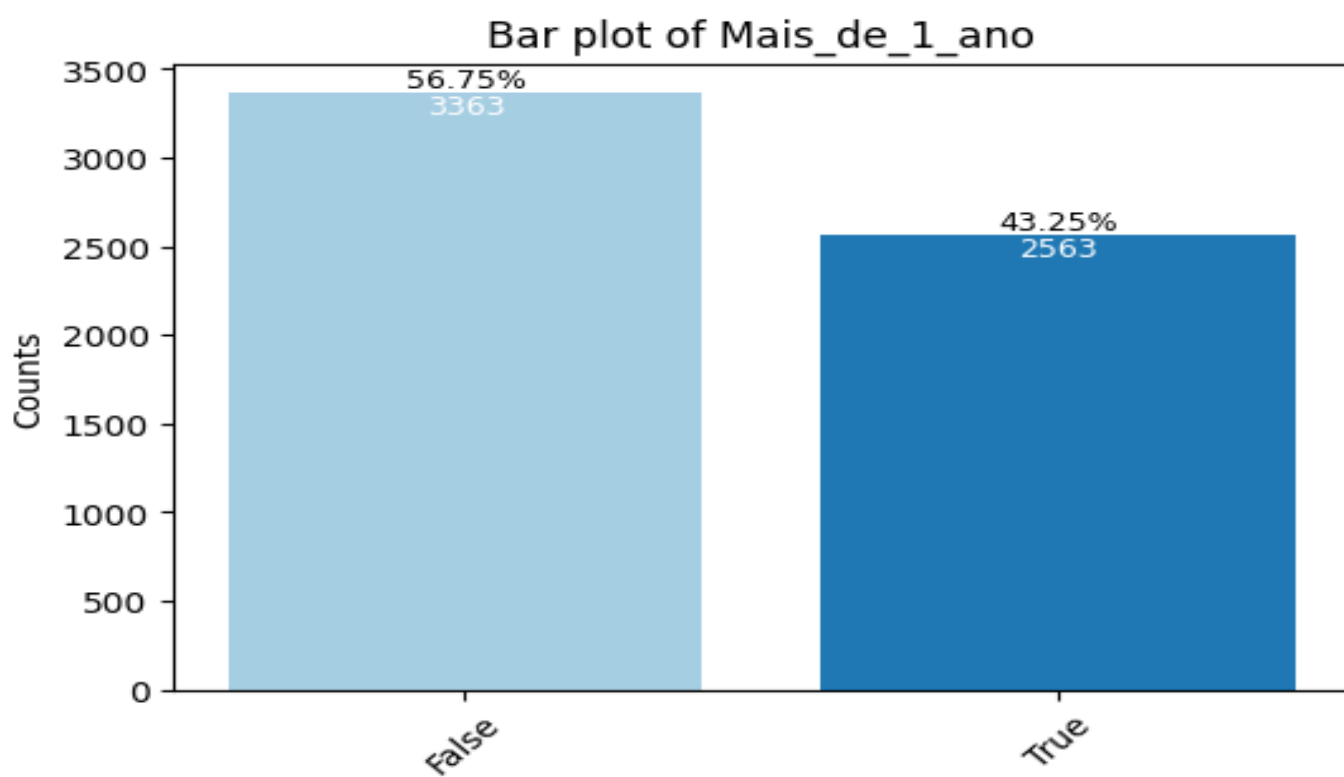
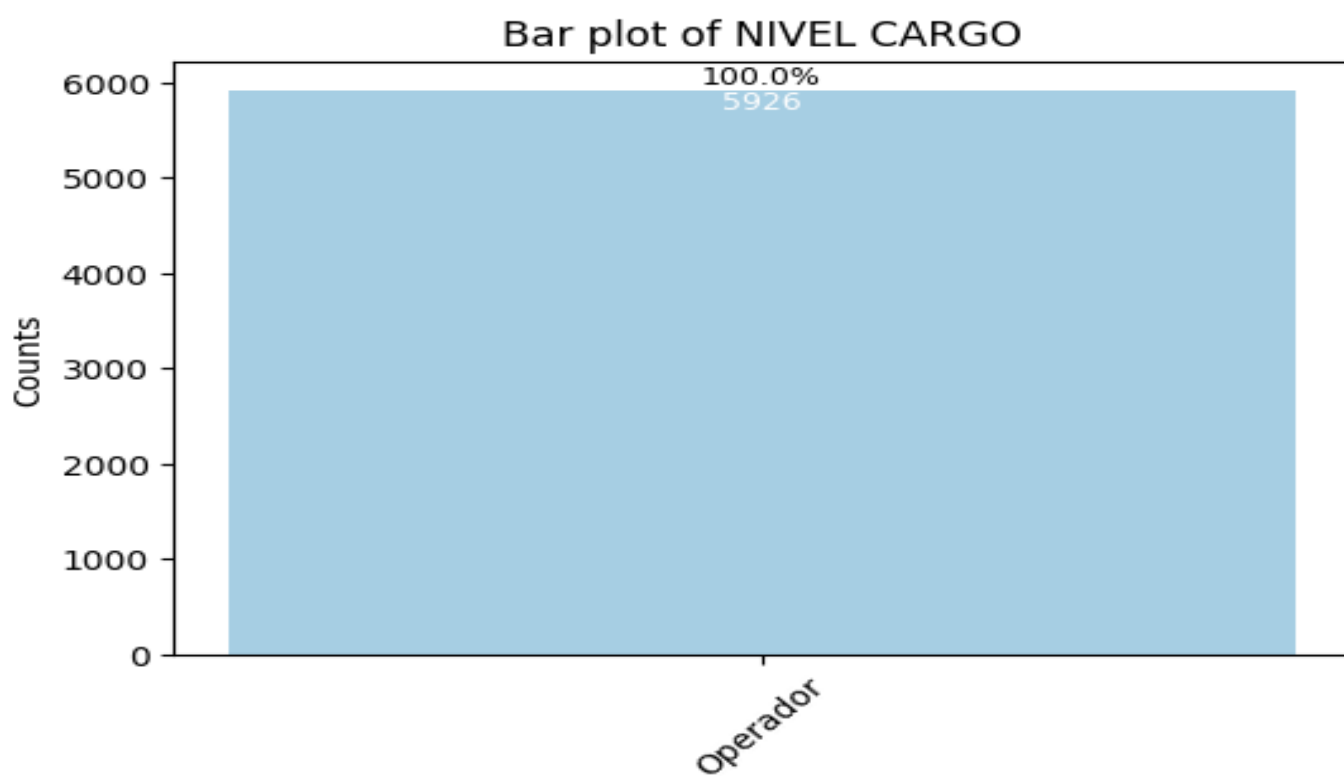


Bar plot of FILHOS

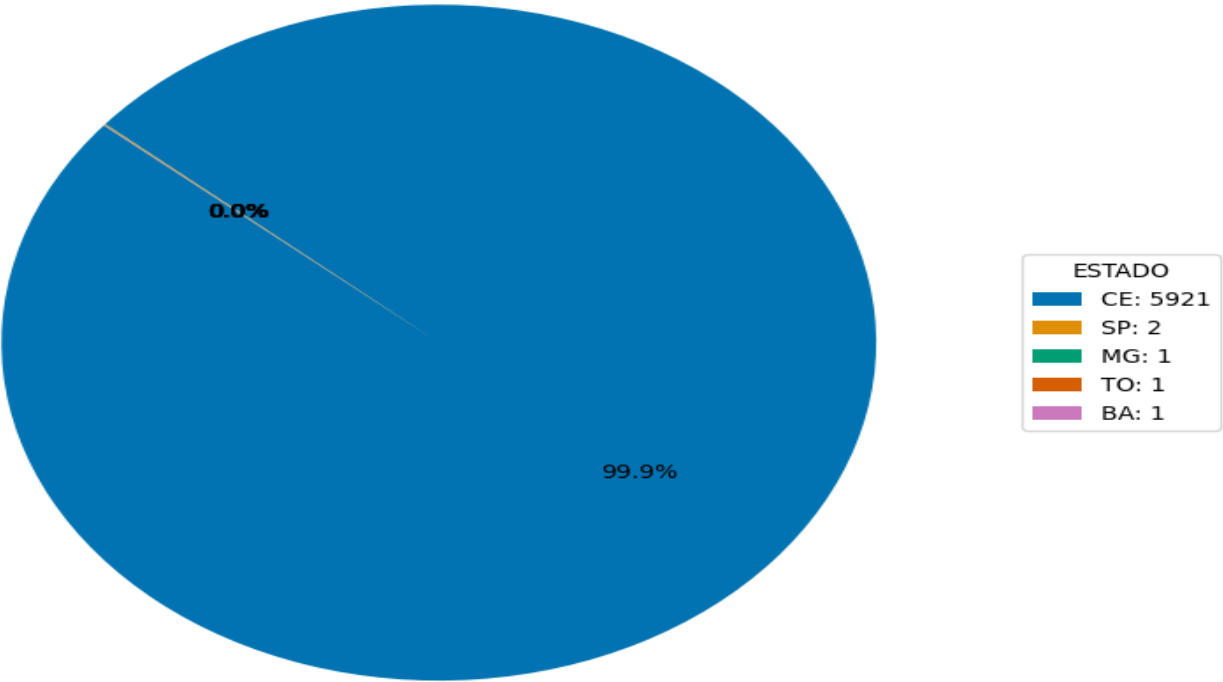








ESTADO Distribution



ESTADO Distribution ($\geq 5\%$)

