

# Data Analysis Reporting Tool

## Contents

- [1. Resumo Geral do Dado](#)
- [2. Resumo de Qualidade do Dado](#)
- [3. Variável Alvo](#)
- [4. Variáveis Individuais](#)
- [5. Classificação de Variáveis](#)
- [6. Relação entre variável explicável e de resposta](#)

O DART (Data Analysis Reporting Tool) é uma ferramenta desenvolvida com o objetivo de ajudar a documentar a qualidade e realizar a exploração de um conjunto de maneira interativa.

## 1. Resumo Geral do Dado

### 1.1. Dimensão dos dados (linhas, colunas)

O conjunto de dados possui **891** linhas e **12** colunas

### 1.2. Inferir tipos das variáveis

Variáveis	
Identificadores	
Numéricas	SibSp, Parch, Age, Fare
Catóricas	PassengerId, Name, Ticket, Cabin, Embarked, Pclass, Sex
Datas	
Alvo	Survived

*Fig. 1.1* Tabela com o tipo das variáveis com base no arquivo de configuração.

### 1.3. Nome e tipo de dado das colunas

Coluna	
Tipo	
float64	Age, Fare
int64	Parch, PassengerId, Pclass, SibSp, Survived
object	Cabin, Embarked, Name, Sex, Ticket

*Fig. 1.2* Tabela com o tipo de dado das colunas no conjunto de dados.

## 2. Resumo de Qualidade do Dado

### 2.1. Dados faltantes

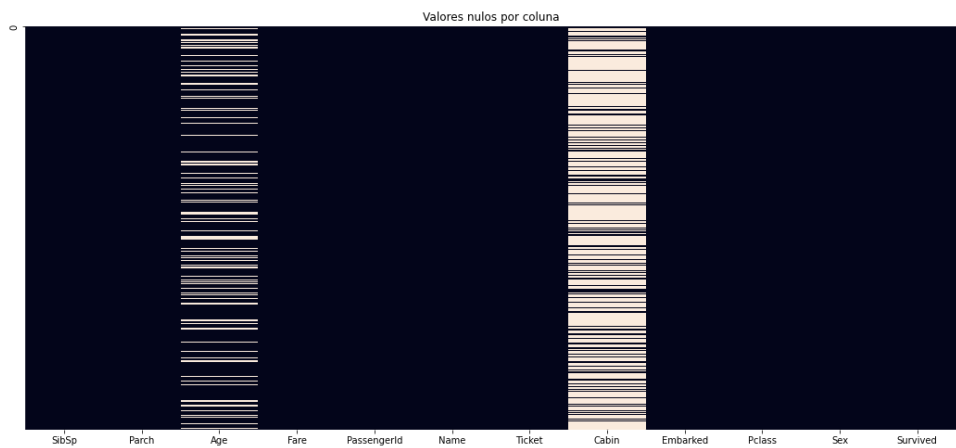


Fig. 2.1 Dados faltantes por coluna no conjunto de dados.

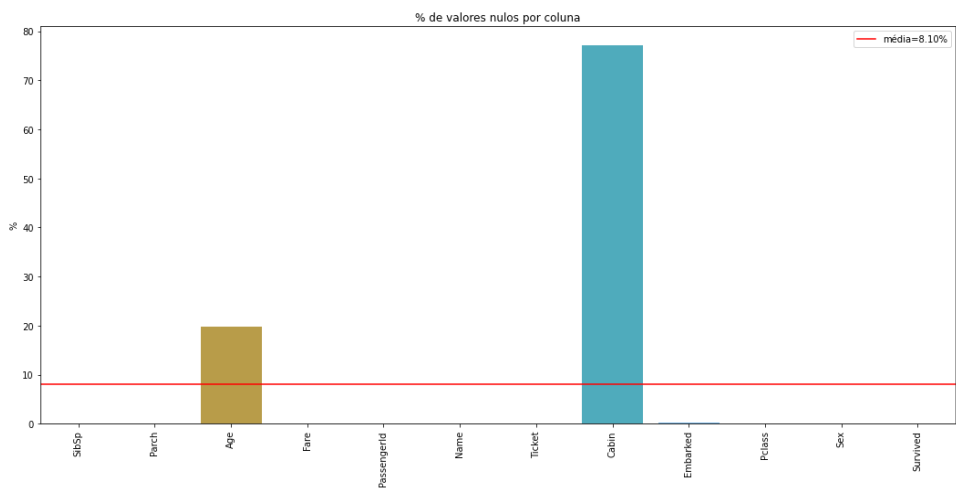


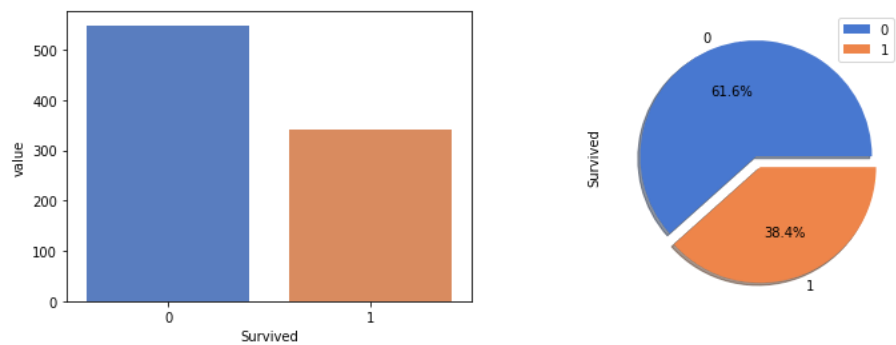
Fig. 2.2 Percentual de dados faltantes por coluna no conjunto de dados.

As variáveis **Age**, **Cabin**, **Embarked** possui(m) valores nulos

### 2.2. Linhas duplicadas

Existem **0** linhas duplicadas no geral

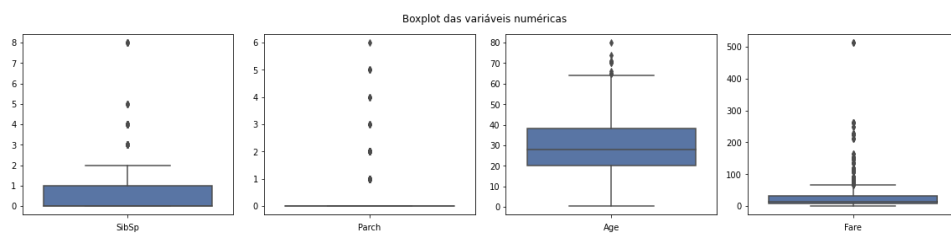
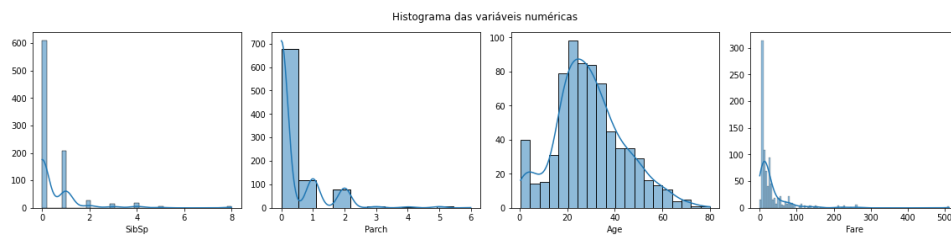
## 3. Variável Alvo



## 4. Variáveis Individuais

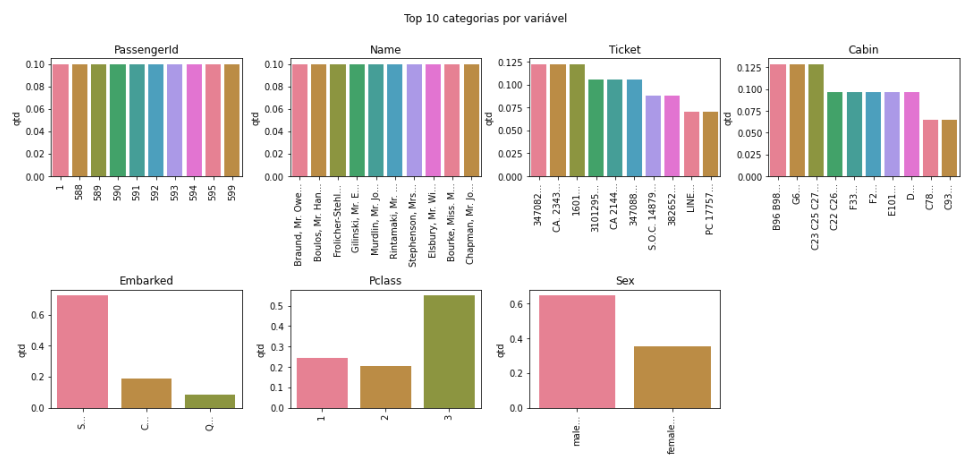
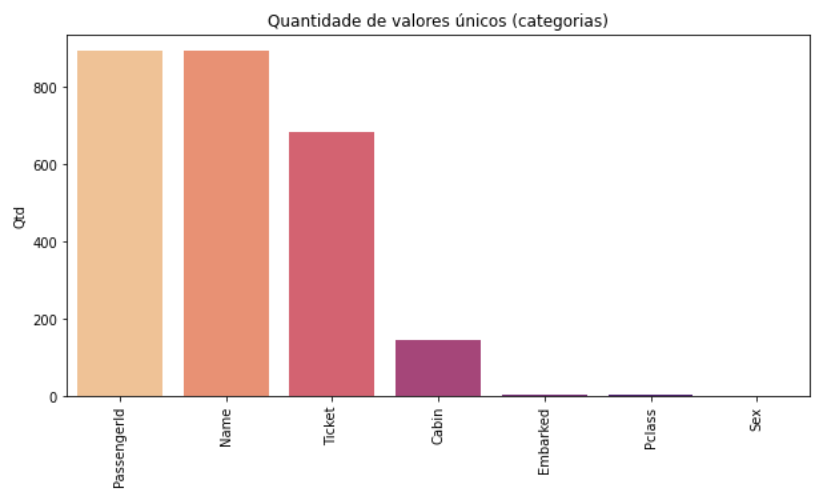
### 4.1. Numéricas

	Qtd	Média	Desvio Padrão	Min	25%	50%	75%	Max
<b>Age</b>	714.0	29.699118	14.526497	0.42	20.1250	28.0000	38.0	80.0000
<b>Fare</b>	891.0	32.204208	49.693429	0.00	7.9104	14.4542	31.0	512.3292
<b>Parch</b>	891.0	0.381594	0.806057	0.00	0.0000	0.0000	0.0	6.0000
<b>SibSp</b>	891.0	0.523008	1.102743	0.00	0.0000	0.0000	1.0	8.0000



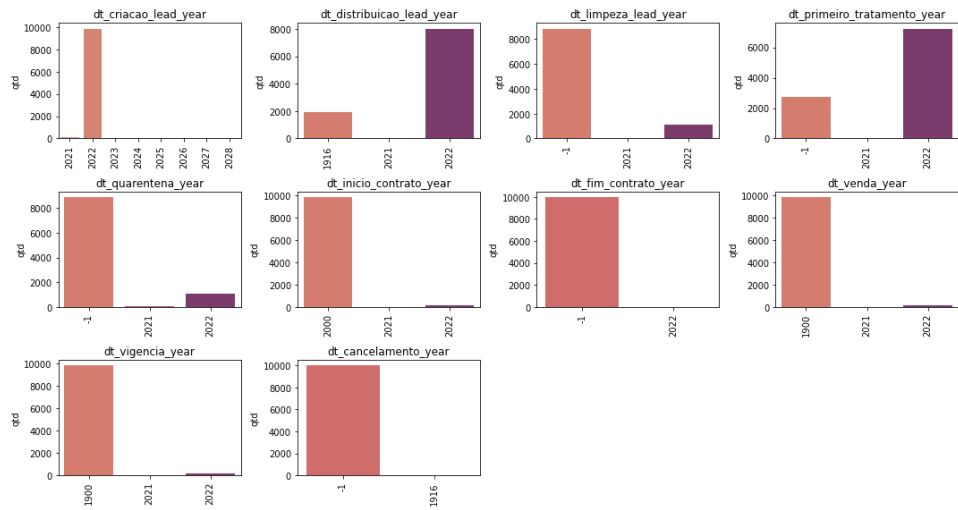
### 4.2. Categóricas

	Qtd	Único	Moda	Frequência
Cabin	204	147	C23 C25 C27	4
Embarked	889	3	S	644
Name	891	891	Abbing, Mr. Anthony	1
PassengerId	891	891	1	1
Pclass	891	3	3	491
Sex	891	2	male	577
Ticket	891	681	1601	7

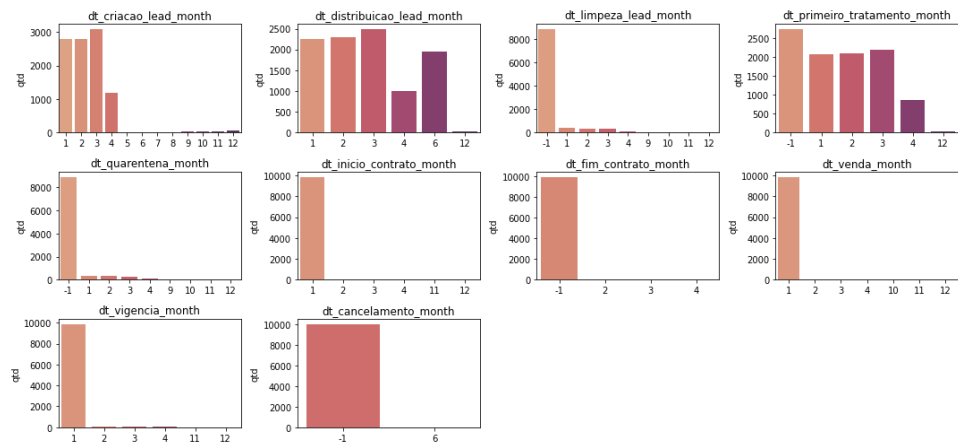


4.3. Data

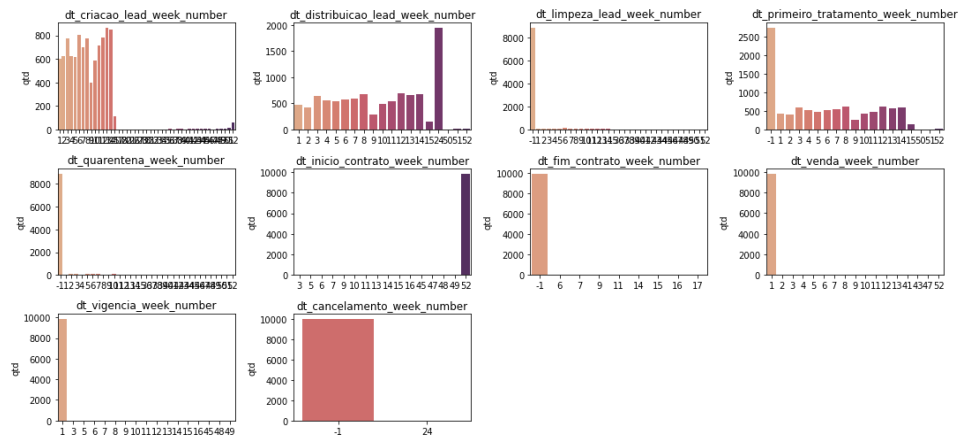
4.3.1. Ano



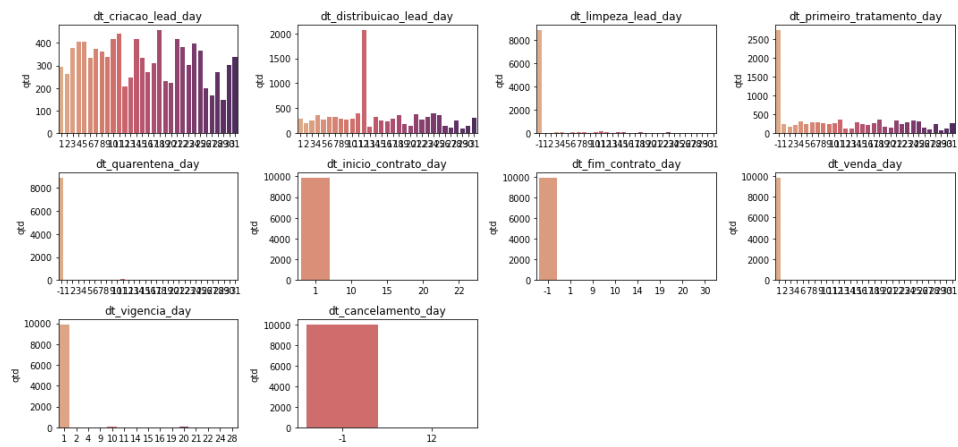
#### 4.3.2. Mês



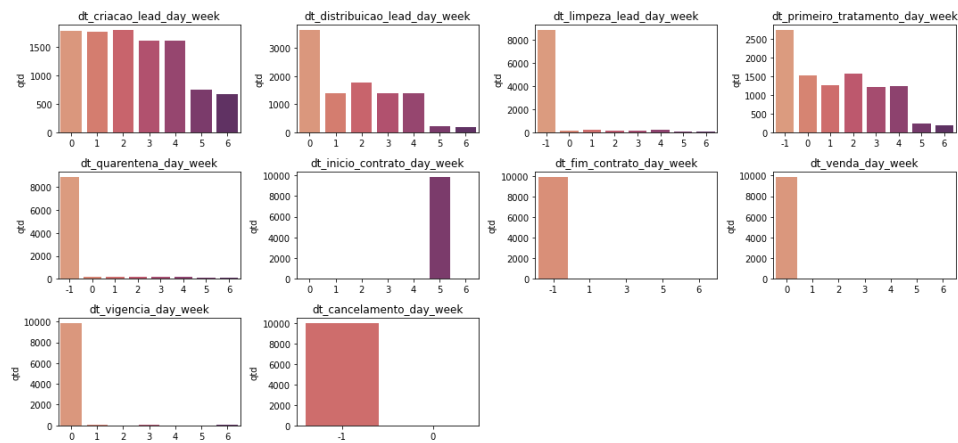
#### 4.3.3. Semana



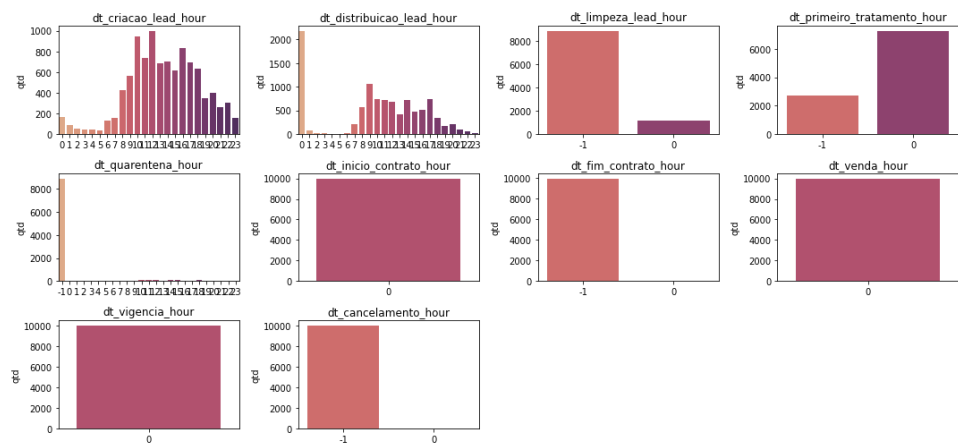
#### 4.3.4. Dia



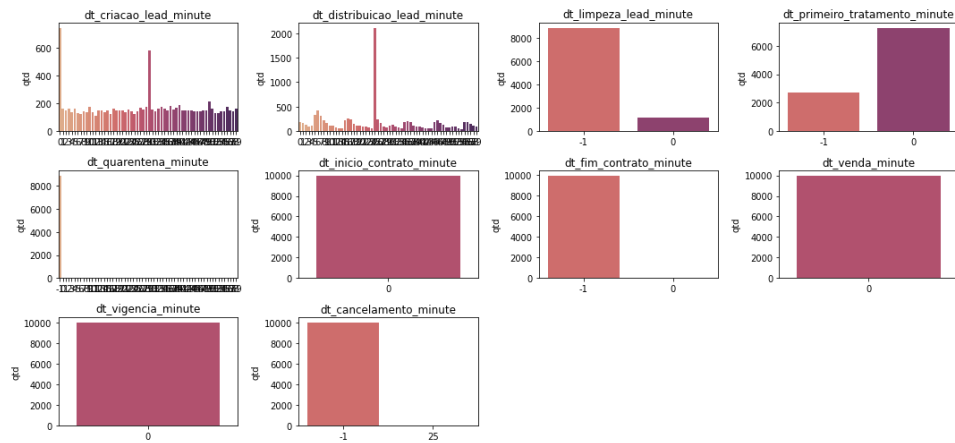
#### 4.3.5. Dia da Semana



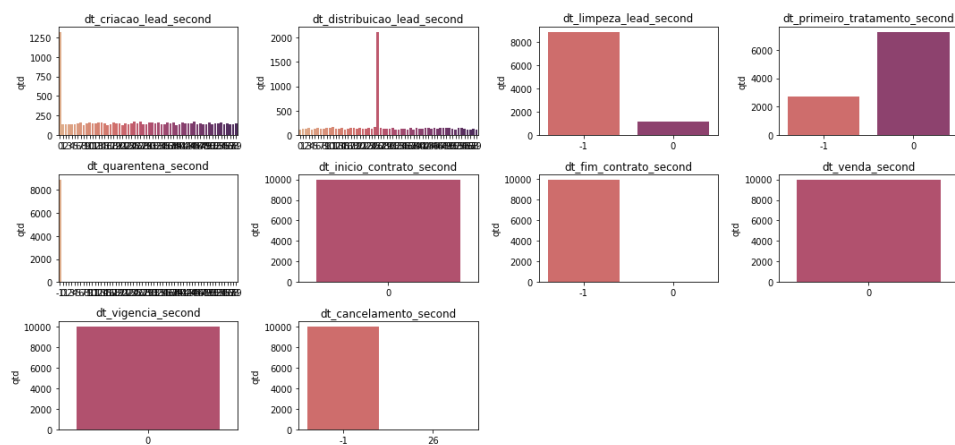
#### 4.3.6. Hora



#### 4.3.7. Minuto

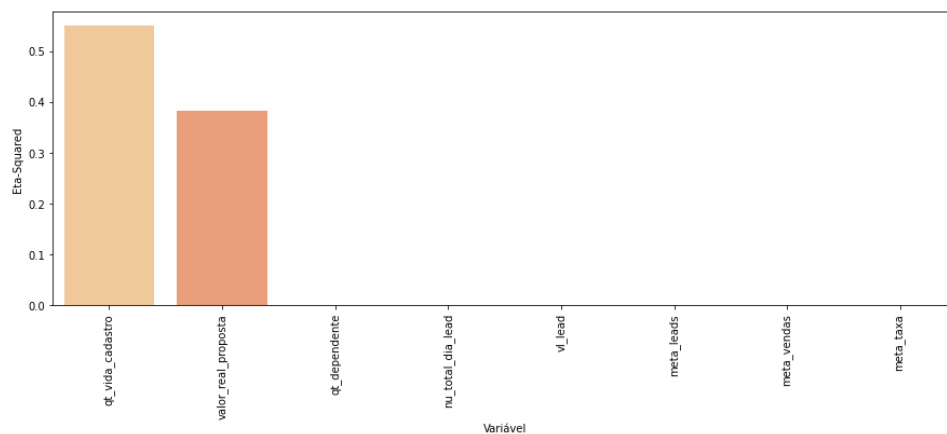


#### 4.3.8. Segundo

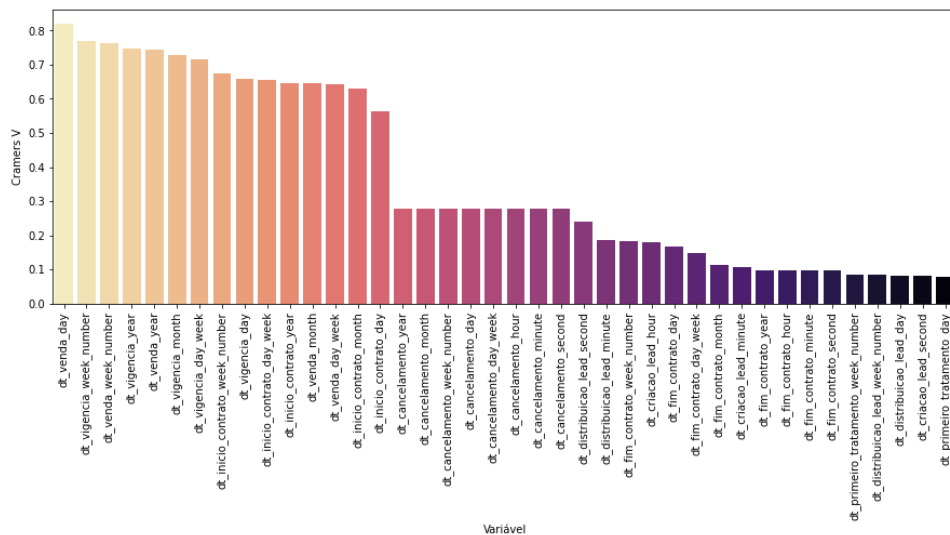
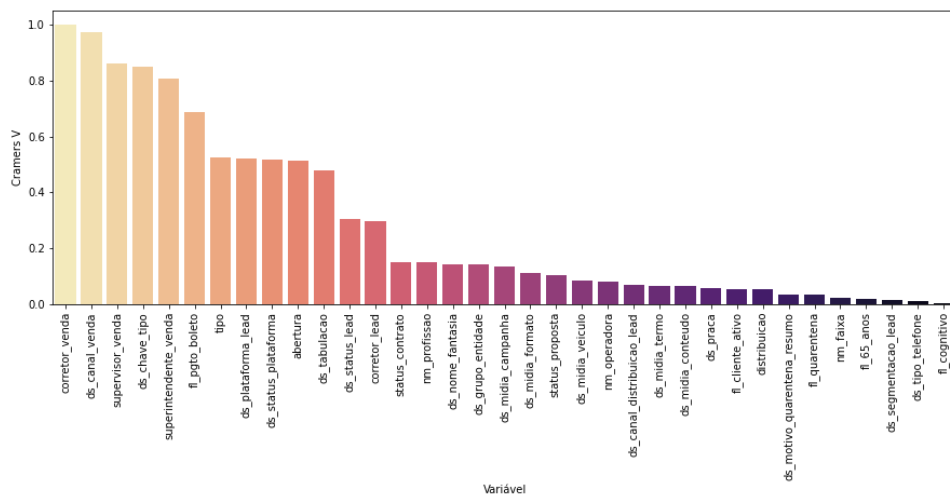


## 5. Classificação de Variáveis

### 5.1. Top 40 variáveis numéricas com maior associação com o alvo

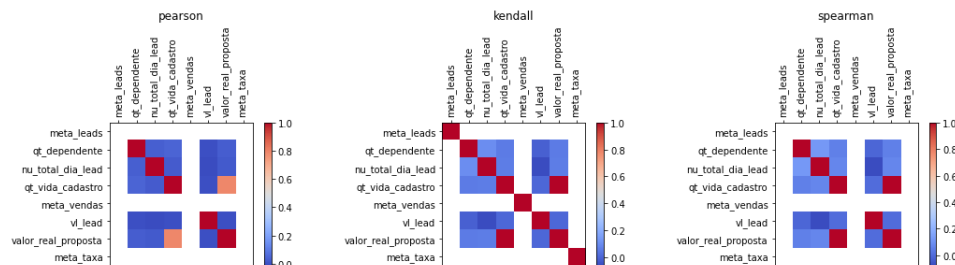


### 5.2. Top 40 variáveis categóricas com maior associação com o alvo



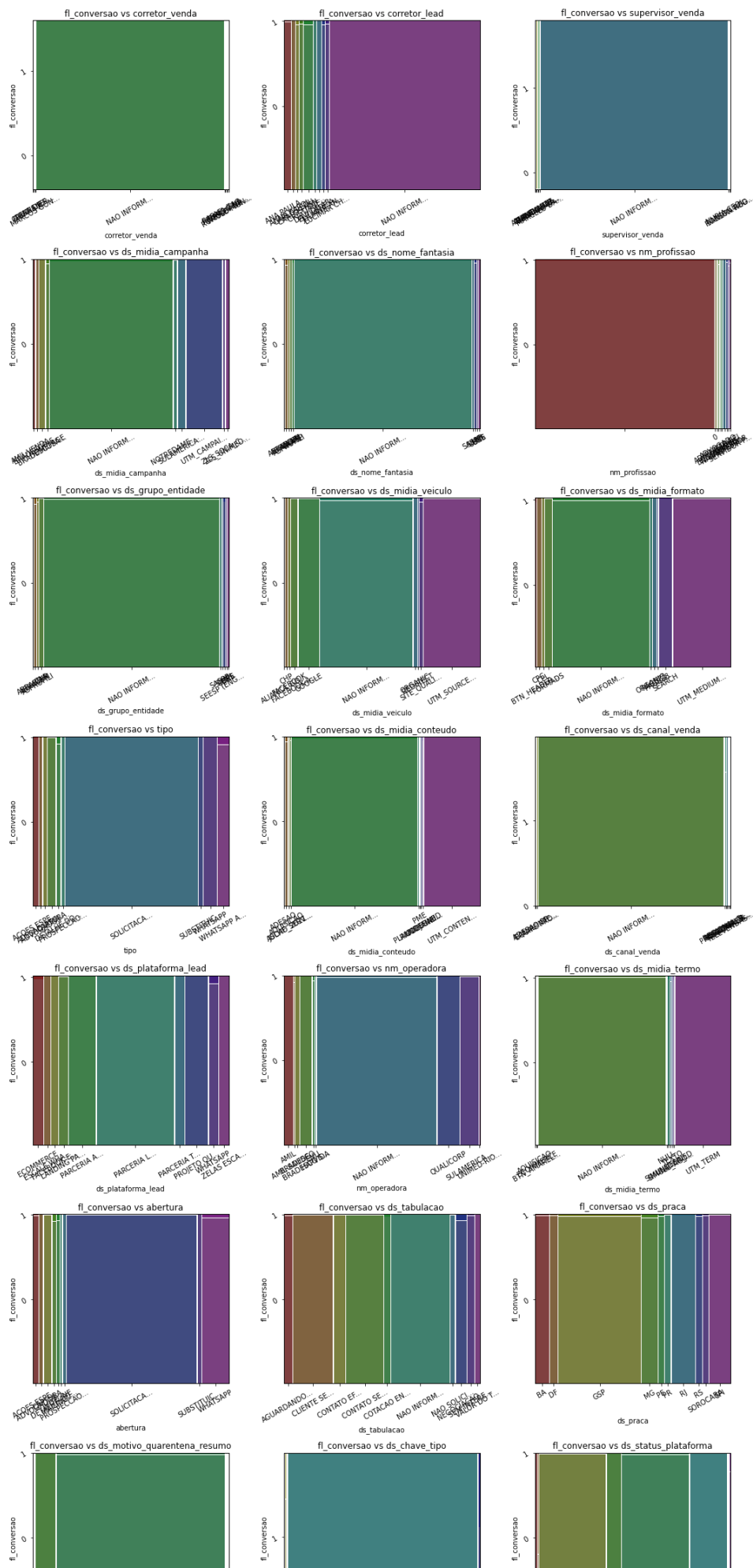
## 6. Relação entre variável explicável e de resposta

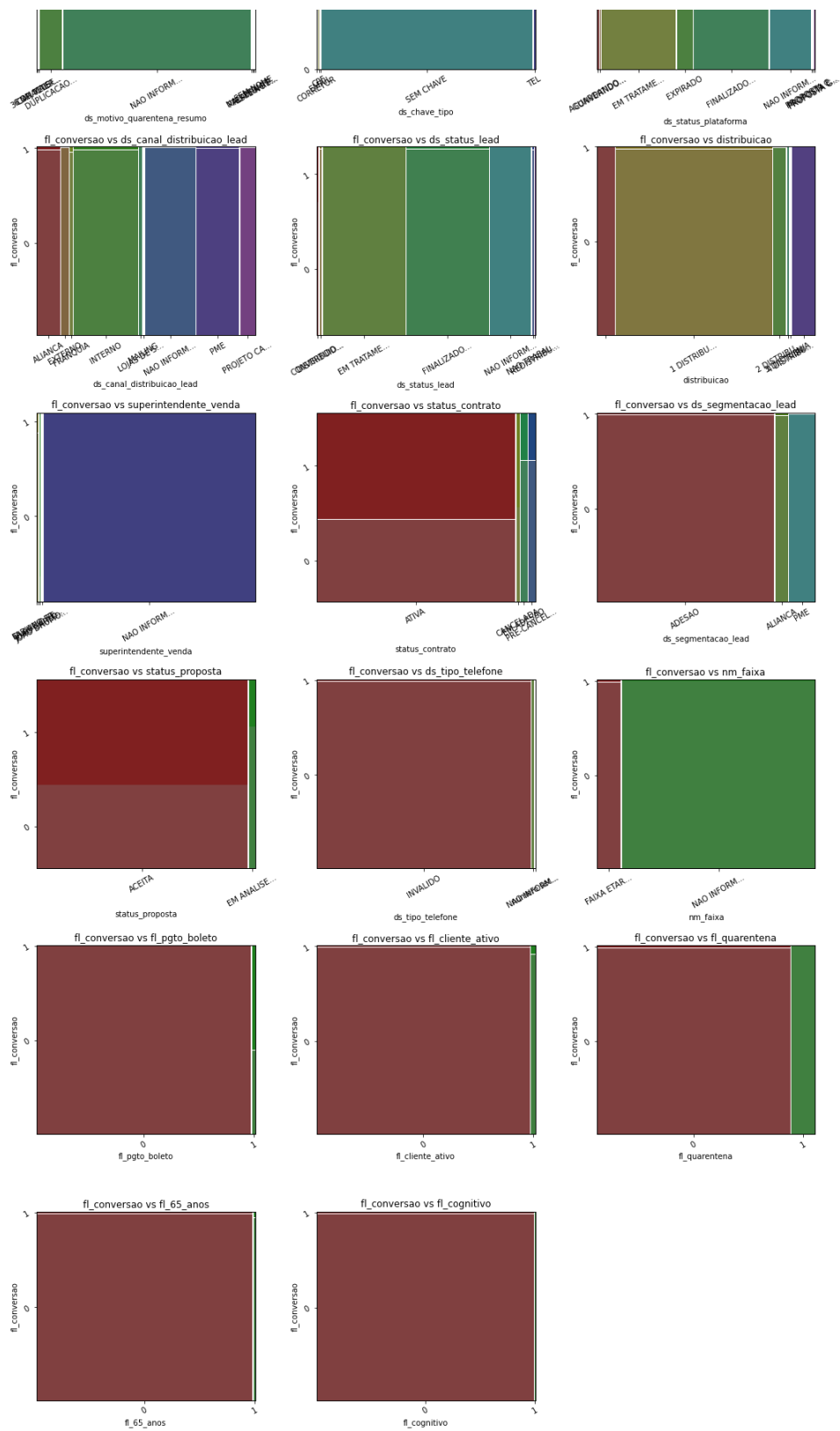
### 6.1. Numéricas



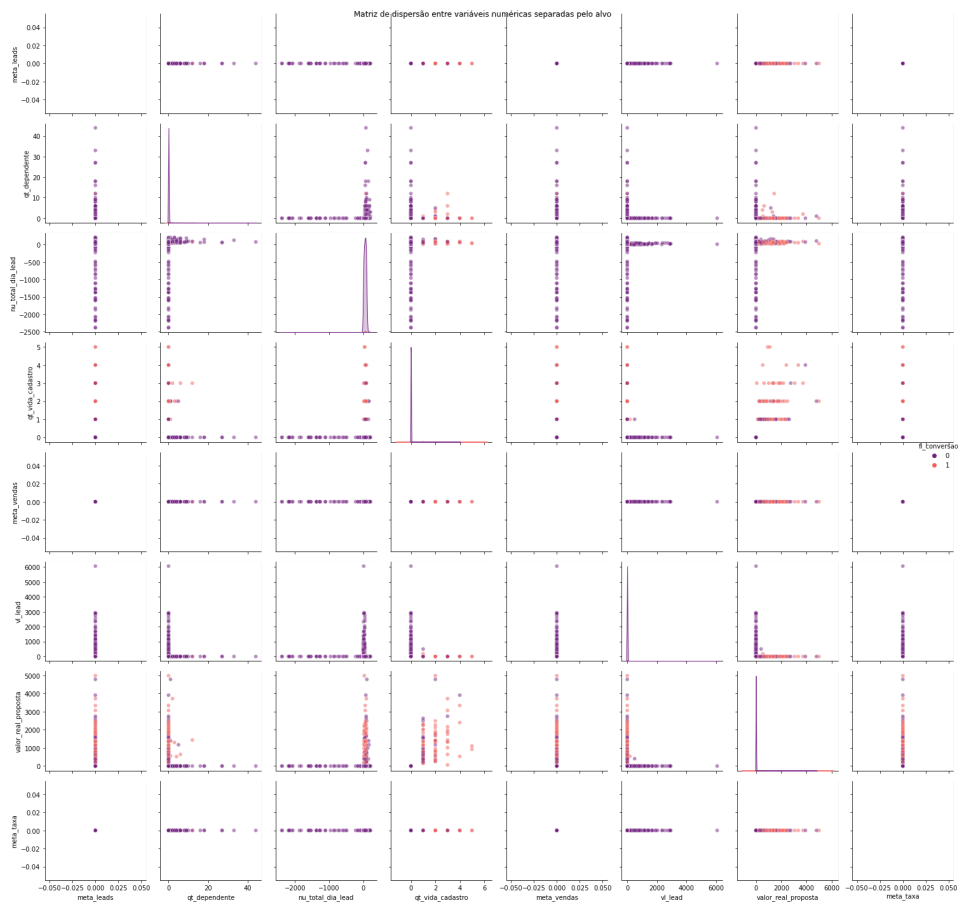
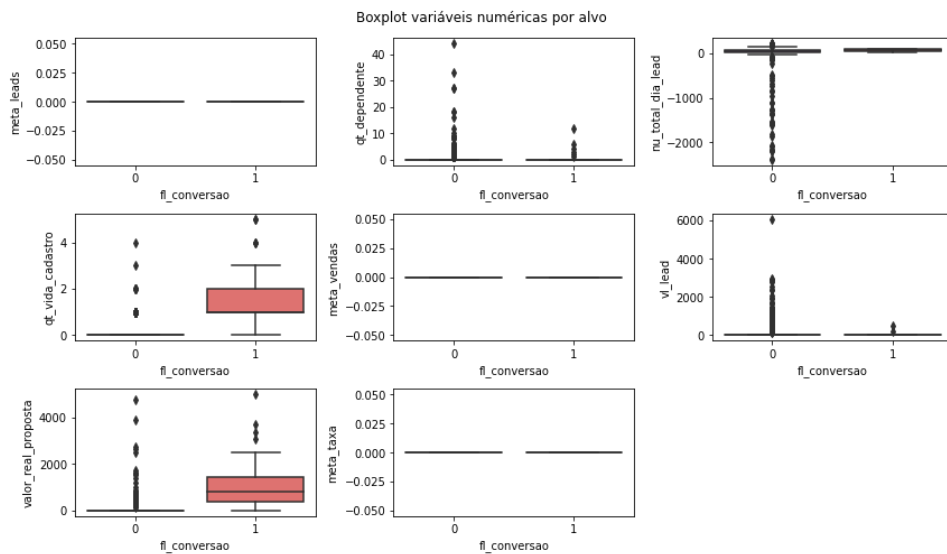
### 6.2. Categóricas







### 6.3. Numéricas e Categóricas



## 6.4. Análise de Componentes Principais

By Matheus Buarque

© Copyright 2021.