

# Desenvolvimento de módulo de Análise de Dados para identificação de situações de fraude e corrupção em processos de contratação pública

**Francisco Valente Pereira**

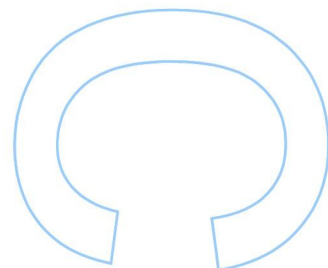
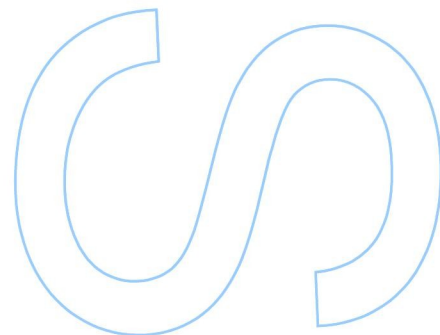
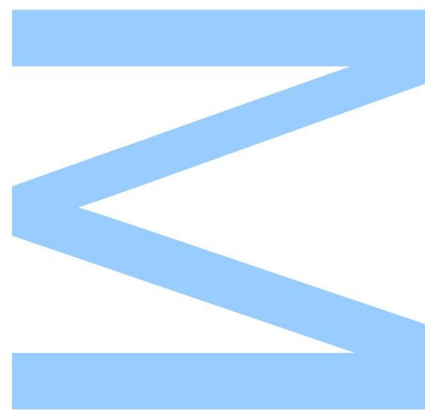
Mestrado em Estatística Computacional e Análise de Dados  
Departamento de Matemática  
Faculdade de Ciências da Universidade do Porto  
2024

## **Orientador**

Prof. Dr. Sílvia Gama  
Faculdade de Ciências da Universidade do Porto

## **Coorientador**

Prof. Dr. Margarida Brito  
Faculdade de Ciências da Universidade do Porto



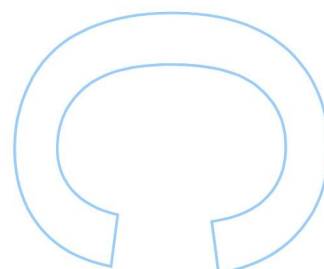
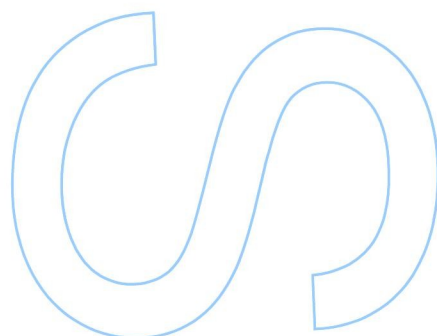
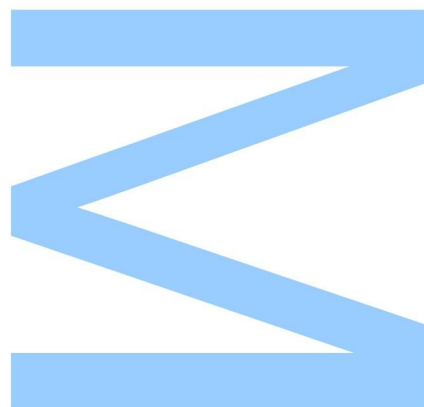




Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto,        /        /





**Francisco Valente Pereira**

**Desenvolvimento de módulo de Análise de  
Dados para identificação de situações de fraude  
e corrupção em processos de contratação  
pública**



***Orientador:** Prof. Dr. Sílvio Gama*

***Coorientador:** Prof. Dr. Margarida Brito*

Departamento de Matemática  
Faculdade de Ciências da Universidade do Porto  
2024



# **Agradecimientos**





## **Resumo**



# **Abstract**



# Motivação



# Table of Contents

<b>Agradecimentos</b>	<b>vii</b>
<b>Resumo</b>	<b>ix</b>
<b>Abstract</b>	<b>xi</b>
<b>Motivação</b>	<b>xiii</b>
<b>Table of Contents</b>	<b>xvi</b>
<b>List of Figures</b>	<b>xvii</b>
<b>List of Tables</b>	<b>xix</b>
<b>Lista de Abreviaturas</b>	<b>xxi</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Estrutura da tese e Etapas da Resolução do Problema . . . . .	1
1.2 Estágio . . . . .	2
1.3 Fraude e Corrupção na Contratação Pública . . . . .	2
1.4 Contratação Pública em Portugal . . . . .	2
1.4.1 Tipos de Procedimento de Contratação Pública . . . . .	3
1.5 Objetivo do estágio . . . . .	4
1.6 Estrutura da tese . . . . .	4
<b>2 DataSet</b>	<b>5</b>
<b>3 Funções</b>	<b>7</b>
3.1 Exploração do Dataset . . . . .	7
3.2 Mini-funções . . . . .	7
3.3 Funções para a primeira flag . . . . .	8
3.4 Flag 1 . . . . .	9
3.5 Flag 2 . . . . .	9

3.6	Análise dos Concursos Públicos . . . . .	10
3.7	Análise dos Ajustes Diretos em Regime Geral . . . . .	10
<b>A</b>		<b>13</b>
<b>Bibliography</b>		<b>13</b>
<b>APPENDICES</b>		<b>13</b>



## **Lista de Figuras**



# Lista de Tabelas

2.1	Número de contratos celebrados para os diferentes tipos de procedimento . . .	5
-----	---	---



# Lista de Abreviaturas

**CPV** Common Procurement Vocabulary

**OCDS** Open Contracting Data Standard

**FCUP** Faculdade Ciências da Universidade do Porto



# CAPÍTULO 1

## Introdução

### 1.1 Estrutura da tese e Etapas da Resolução do Problema

1. Analisar conjunto de dados. Ver nr de contratos, ver nr de procedimentos e nr de tipo de contratos, etc
2. Categorizar as flags de acordo com a facilidade de implementação e valor
3. Focar, numa fase inicial, num conjunto de contratos. Seleccionaram-se apenas ajustes diretos e contratos públicos referentes a CPV's que dizem respeito a serviços de consultoria de IT - 72
4. Construiu-se a primeira flag que compara preço base com preço contratual
5. Construiu-se função que verifica se os preços contratuais dentro dos ajustes diretos caem dentro de um intervalo em torno do preço base
6. Nesta mesma função, inclui-se a presença de um parametro de rácio para comparar com precobase/precocontratual. Nos casos em que este quociente é mt alto vai ser verificado se ha ou nao presença de vários lotes no contrato
7. Construção de uma função que vai analisar ajustes diretos. Primeiro verifica quais os ids dos contratos q ultrapassam o valor maximo permitido por lei q é 20k€. Se ultrapassar dispara a flag. Se num ajuste direto nao houver fundamentação é disparada a flag tambem
8. Os ajustes diretos foram ordenados por ordem crescente de nt de celebracao. Comparou-se o nr de ajustes diretos celebrados com o nt de ajustes suspeitos, calculou-se o racio entre os 2.
9. No caso dos concursos publicos : qnd o preço base é mt maior que o preço contratual verificar se existem varios id's associados a um mesmo nr de anuncio, somar os preços base e contratual e verificar se o rácio ainda é muito grande
10. Atribuiu-se um valor de flag contínuo (entre 0 e 1) ao contratos públicos onde é disparada um flag
11. Construiu-se uma função que dispara um flag caso o valor do prazo de candidatura de um concurso público seja inferior a 6 dias

## 1.2 Estágio

INSERIR TEXTO FORCERA

## 1.3 Fraude e Corrupção na Contratação Pública

FALAR SOBRE O QUE É FEITO A NÍVEL EUROPEU PARA COMBATER FRAUDE E CORRUPÇÃO ( OCD, RED FLAGS, ETC)

## 1.4 Contratação Pública em Portugal

A Contratação Pública em Portugal pode ser classificada de duas formas : aberta e fechada. As regras presentes no Código dos Contratos Públicos (CCP) dizem respeito aos contratos públicos celebrados entre uma entidade adjudicante pública e uma entidade adjudicatária.

O ato de adjudicar consiste em conferir o direito de algo a alguém, conceder algo ao maior licitante ou atribuir algo a alguém por concurso ou por ajuste.

O CCP é aplicado a entidades adjudicantes públicas, tais como o Estado, Regiões Autónomas, Autarquias locais, Institutos públicos, Entidades Administrativas Independentes, Banco de Portugal, Fundações Públicas, Associações Públicas, Associações de que façam parte uma ou várias pessoas coletivas referidas anteriormente e que sejam maioritariamente financiadas por estas. Além destas, são consideradas entidades adjudicantes organismos de direito público, pessoas coletivas e associações <sup>1</sup>. São consideradas, também, entidades adjudicantes organismos com atuação nos setores especiais da água, energia, transportes e serviços postais <sup>2</sup>. Existe, também, a possibilidade de aplicar o CCP a entidades não adjudicantes que pretendem celebrar determinados contratos de empreitadas de obras públicas ou de serviços associados a obras <sup>3</sup>.

Existem duas fases principais no processo de contratação pública. A primeira fase é a **fase preparatória** em que é feita a decisão de realizar um contrato e inclui uma fase preparatória do procedimento e uma fase instrutória que terminará no ato de adjudicação. A segunda fase é a **fase conclusiva** em que é concluído e celebrado o contrato. Existe também uma **fase complementar** que pode ser necessária na eventualidade do contrato público depender de atos posteriores à sua celebração tais como a aprovação, visto e publicidade.

<sup>1</sup> nos termos do artigo 2.º n.º 2, alíneas a), b) e d)

<sup>2</sup> artigo 7.º n.º 1.º

<sup>3</sup> artigo 275.º



### **1.4.1 Tipos de Procedimento de Contratação Pública**

Aquando da formação dos contratos, as entidades adjudicantes devem adotar um dos seguintes tipos de procedimentos :

1. Ajuste Direto
  - (a) Regime Geral
  - (b) Regime Geral ao abrigo do artigo 7º da Lei n.º 30/2021, de 21.05
  - (c) Simplificado
2. Consulta Prévia
  - (a) ao abrigo do artigo 7º da Lei n.º 30/2021, de 21.05
3. Concurso Público
  - (a) Regime Geral
  - (b) Simplificado
4. Concurso limitado por prévia qualificação
  - (a) Regime Geral
  - (b) Simplificado
5. Procedimento de negociação
6. Diálogo concorrencial
7. Parceria para a inovação
8. Ao abrigo de acordo-quadro (art.º 258.º)
9. Ao abrigo de acordo-quadro (art.º 259.º)
10. Disponibilização de bens móveis
11. Serviços sociais e outros serviços específicos
12. Concurso de conceção simplificado

A natureza e designação do tipo de contrato que é possível realizar para cada um dos procedimentos anteriormente enumerados é :

1. Empreitadas de obras públicas
2. Concessão de obras públicas
3. Concessão de serviços públicos
4. Locação ou aquisição de bens móveis
5. Aquisição de serviços
6. Sociedade

## **1.5 Objetivo do estágio**

FALAR DO OBJETIVO GERAL

## **1.6 Estrutura da tese**

FALAR DA ESTRUTURA DA TESE

# CAPÍTULO 2

## DataSet

O conjunto de dados utilizado consiste numa matriz com dimensão  $m \times n$ , sendo  $m = 187198$  o número de contratos celebrados e  $n = 61$  o número de colunas. De entre o universo de contratos celebrados, podemos atentar na seguinte tabela como se distribuem consoante os diferentes tipos de procedimento :

Ajuste Direto Regime Geral	89882
Consulta Prévia	40331
Concurso público	28438
Ao abrigo de acordo-quadro (art.º 259.º)	20624
Ao abrigo de acordo-quadro (art.º 258.º)	5205
Ajuste direto simplificado	1011
Ajuste direto simplificado ao abrigo da Lei n.º 30/2021, de 21.05	919
Consulta Prévia Simplificada	435
Concurso limitado por prévia qualificação	297
Procedimento de negociação	25
Concurso público simplificado	24
Consulta prévia ao abrigo do artigo 7º da Lei n.º 30/2021, de 21.05	16
Ajuste Direto Regime Geral ao abrigo do artigo 7º da Lei n.º 30/2021, de 21.05	9
Serviços sociais e outros serviços específicos	8
Concurso de conceção simplificado	2
Setores especiais – isenção parte II	1
Concurso de ideias simplificado	1

**Tabela 2.1:** Número de contratos celebrados para os diferentes tipos de procedimento

De todas as colunas, as que revelaram maior utilidade foram as colunas referentes ao ID do anúncio, número de anúncio, preço base, preço contratual, data de publicação de concurso, entidades concorrentes, CPV, **PREENCHER COM RESTANTES**.



# CAPÍTULO 3

## Funções

### 3.1 Exploração do Dataset

Numa primeira fase foram criadas funções para explorar o conjunto de dados.

A função **col.names** retorna-nos uma dataframe com todos os nomes das colunas. A função **n\_contracts** retorna-nos o número de linhas, ou seja, o número de contratos totais, da base de dados.

Foi desenvolvida, também, uma função **h** que permite ver uma dataframe de forma mais apresentável.

*É preciso criar uma função com o código que já tenho para categorizar os contratos de acordo com o tipo de procedimento. Idem para o tipo de contrato.*

### 3.2 Mini-funções

Para a construção das *flags* são necessárias várias funções pequenas. Cada função irá ter um único propósito e os outputs de certas funções irão ser usados como inputs de outras funções.

Para identificar um contrato necessitamos do seu id. Assim, criou-se a função **all\_ids** que retorna os ids de todos os contratos da dataframe. Contudo, como não vamos trabalhar com todos os tipos de contratos e procedimentos, é necessário filtrar estes ids. Para isso, fui por partes.

Criei as funções **ajustes\_dir**, **consulta\_prev** e **concurso\_pub** que retornam os ids de todos os ajustes diretos, consultas prévias e concursos públicos, respetivamente.

De seguida, filtrei os contratos por tipo de procedimento e por CPV. Para isso, criei a função **cpv\_direto** que retorna todos os ids de ajustes diretos para serviços de consultoria em IT ( todos os CPV's começados por 72). De seguida, fiz o mesmo para contratos públicos na função **cpv\_cpub**.

De forma a generalizar, criei a função *cpv* que retorna todos os ids de contratos de um tipo de procedimento e para um determinado CPV. Tanto o tipo de procedimento como os primeiros 2 algarismos do CPV são parâmetros de entrada da função.

O primeiro passo, que passa por obter os ids dos contratos desejados está feito. Os id's dos contratos vão ser o input da maioria das funções que se seguem.

Para começar, criaram-se duas funções que permitem ver qual é o contrato associado a um determinado id. A função *contrato* tem como input um único id de um contrato (todas estas funções dizem respeito à tabela *contratos* da DB) e retorna uma dataframe com uma única linha referente ao contrato associado a esse mesmo id. A função *contratos* faz exatamente a mesma coisa mas para um conjunto de id's.

Quando as flags tiverem construídas e quisermos ver o anúncio de um contrato suspeito no site do basegov, podemos fazê-lo usando a função *url* que, novamente, tem como parâmetro de entrada o id de um contrato. Esta função pode ter bastante utilidade nos casos que chamam bastante a atenção, como por exemplo, quando se viu um ajuste direto de 3 milhões de euros.

### 3.3 Funções para a primeira flag

A primeira flag construída tem como objetivo comparar o preço base com o preço contratual. Esta flag não vai funcionar para ajustes diretos visto que não existe preço base.

Primeiro, obteve-se o preço contratual e foi-se, novamente, por passos. A função *preco\_contratual1* devolve, a partir do id de um anúncio, o valor do preço contratual desse mesmo anúncio. A função *preco\_contratual2* faz exatamente a mesma coisa mas para uma tabela genérica. Esta função não vai ter grande utilidade porque só há uma tabela na database. A função *preco\_contrato3* generaliza a primeira função pois retorna um conjunto de preços contratuais referentes a um conjunto de id's de anúncios que leva como input. A função *preco\_contratual4* generaliza a função anterior para qualquer tabela mas, no fundo, também não vai ter grande importância. A função *preco\_contratual3* é a função que importa. Os preços contratuais são retornados no formato de array pois dessa forma torna-se mais fácil manipular os valores, fazer comparações, etc.

De seguida, foi feito o mesmo para o preço base. Foi utilizada a mesma abordagem e chegou-se à função *preco\_base3* que retorna os valores dos preços base para um conjunto de id's. Foi necessário formatar os valores antes de os retornar visto que, no momento da aquisição, estes vinham no formato —.—,—€.

Neste momento já podemos fazer o seguinte : utilizando a função *cpv*, obter os id's de contratos de um dos tipos desejados e para um determinado cpv. Nesta fase inicial, obtiveram-se os ids de ajustes diretos e concursos públicos para CPV's começados por 72.

Com esses id's podemos ver os preços contratuais. No caso dos contratos públicos podemos ver o valor dos preços base. Assim, estamos em condições de construir a primeira flag.

### 3.4 Flag 1

A primeira flag construída compara o preço base com o preço contratual para concursos públicos. Para que ela funcione temos de dar como input o conjunto de preços bases e contratuais associados a um determinado conjunto de contratos que já conseguimos determinar usando as funções criadas anteriormente. É preciso fornecer uma tolerância que vai definir o intervalo em torno do preço base, os ids dos contratos associados e um rácio máximo aceitável entre o preço base e contratual.

O objetivo passa por verificar se o preço contratual cai num intervalo em torno do preço base. Se cair é ativada uma flag. O preço base é conhecido à priori, por isso esta flag não tem muito valor por si só. Quando for acoplada com outras flags pode sugerir um comportamento suspeito. Também é preciso ter em conta como é que o preço base é calculado. Se o preço base for calculado por excesso e o preço contratual for próximo do preço base, pode haver corrupção tanto por parte da entidade adjudicante como da adjudicatária.

Esta função também dispara uma flag nos casos em que o rácio preço base a dividir por preço contratual é muito alto, sendo este limite um parâmetro de entrada da função. Se o preço base for centenas de vezes superior ao preço contratual, é muito provável que o contrato esteja dividido por lotes e, nesse caso, é preciso verificar quais são os lotes associados ao contrato - a partir do número do anúncio pois todos os lotes partilham o mesmo. Tendo o preço contratual de todos os lotes, somam-se e compara-se o resultado ao preço base. Contudo, esta análise dos lotes terá que ser feita numa função externa.

Por fim, o resultado retornado é um conjunto de id's na forma de tuplo. Estes id's podem ser usados posteriormente noutras funções para imprimir os contratos ou para abrir o url da pagina do basegov associado - isto tudo usando as funções definidas anteriormente.

Esta função devia ser adaptada para fazer o seguinte : comparar na mesma o preço contratual com o intervalo em torno do preço base, calcular o rácio e ver se quando estamos a falar de preços bases acima de 100x superior o preço contratual existem ou não lotes e ver os casos em que o preço base é umas 2-10x superior ao preço contratual. Isto porque, para um obra que custe, por ex., 100.000€ e o preço contratual é 50.000€ pode ser estranho. Pode ser o preço base que é mal calculado, pode ser uma forma de afastar outras empresas ( mas isto mais no caso em que o preço base é, por ex., 10x superior ao preço contratual - podem existir muitas empresas capazes de efetuar o serviço A mas se virem no concurso que o preço base é muito grande podem não se candidatar por serem uma empresa mais pequena, por ex. )

### 3.5 Flag 2

Esta função foi criada apenas para ajustes diretos. Não está definida nas flags da OCDS.

Os ajustes diretos tem diferentes limites máximos impostos por lei, mediante o tipo de obra/serviço.

Esta função tem como input uma dataframe correspondente a todos os ajustes diretos realizados com CPV 72, sendo esta obtida fazendo uso das funções *contratos* e *cpv* definidas antes,

e os ids dos contratos associados aos ajustes diretos.

A função vai comparar todos os preços contratuais com o valor de 20.000€ pois é o limite máximo de aquisição de serviços. Esta parte tem de ser atualizada no futuro. Contudo, a maior parte dos ajustes diretos é Aquisição de Serviços. Se os valores forem superiores a este preço, é ativada uma flag.

A segunda parte da função verifica se foi feita uma fundamentação para o ajuste direto. Esta vai ser sempre uma alínea qq do CCP. Mas tem de estar preenchida. Se o valor ultrapassar os 20000€ e não estiver fundamentado é claramente comportamento suspeito. Se o contrato só tiver a parte da fundamentação por preencher, pode ser má prática.

### 3.6 Análise dos Concursos Públicos

Aqui só foram usadas funções já definidas. Criaram-se variáveis que guardam o valor do preço base e contratual dos concursos. Deu-se isso como input à função *redflag*, além dos outros que são definidos por mim, e obtiveram-se as redflags. Representou-se um barplot dos preços contratuais em cima dos preços base para verificar a diferença entre os mesmos.

### 3.7 Análise dos Ajustes Diretos em Regime Geral

Aqui, novamente, foram utilizadas as funções já criadas anteriormente para obter uma dataframe com os contratos referentes a ajustes diretos e os ids associados. Foi feito um sumário estatístico dos valores dos preços contratuais e um histograma e um boxplot. Mas como tem outliers, vê-se mal.

Ordenei os ajustes diretos por fundamentação. Não nenhum campo vazio.

Com isto tudo já se podem calcular as flags usando a função *redflag2*.

É interessante saber qual a proporção de atividades suspeitas por entidade adjudicante e entidade adjudicatária. Para isso, calculou-se primeiro o número de contratos suspeitos e respetiva percentagem.

Para conseguir analisar foi preciso separar as colunas das entidades adjudicantes e adjudicatárias - que estavam no formato Entidade(NIF)(URL) - em três colunas distintas para cada uma das duas entidades. Isto é necessário para filtrar os contratos pelo NIF porque diferentes entidades podem ter o mesmo NIF.

Depois disso, ordenou-se por ordem decrescente de ajustes diretos realizados os NIFS das empresas. Uma das listas ordenadas só para as entidades adjudicantes, outra para as entidades adjudicatárias.

Para esta análise foi necessário criar duas funções que retornem os ids dos ajustes diretos a partir do NIF da empresa. Queremos todos os ajustes diretos celebrados para uma determinada entidade a partir do seu NIF. Isso foi feito a partir das funções *e\_adjudicante* e *e\_adjudicataria*.

De seguida, foi então calculado para cada NIF o número de ajustes diretos realizados usando uma das funções anteriores e os respetivos ID's. Para cada NIF, foi comparado cada um dos



nos id's com a lista de id's dado pela função *redflags2* e feito o rácio para ver a percentagem de contratos suspeitos.

Agora seria interessante verificar se existem, dentro destas empresas, subgrupos entre entidades adjudicantes e adjudicatárias.



# APÊNDICE A

Add appendices.



