

Instituto Politécnico de Coimbra  
Instituto Superior de Contabilidade  
e Administração de Coimbra

Dília Susana Alves Bernardino

Aplicação de métodos analíticos para auditoria  
e deteção de anomalias no ficheiro SAFT-PT:  
O caso de uma empresa comercial

Aplicação de métodos analíticos para Auditoria e deteção de anomalias no ficheiro SAFT-PT

Dília Susana Alves Bernardino

ISCAC | 2017

Coimbra, outubro de 2017



Instituto Politécnico de Coimbra  
Instituto Superior de Contabilidade  
e Administração de Coimbra

Dília Susana Alves Bernardino

Aplicação de métodos analíticos para auditoria  
e deteção de anomalias no ficheiro SAFT-PT:  
o caso de uma empresa comercial

Dissertação submetida ao Instituto Superior de Contabilidade e Administração de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de **Mestre em Auditoria Empresarial e Pública – Ramo Auditoria Financeira** realizada sob a orientação da Doutora Isabel Maria Mendes Pedrosa e coorientação do Doutor Raul Manuel Silva Laureano.

Coimbra, Outubro de 2017

## **TERMO DE RESPONSABILIDADE**

Declaro ser a autora desta dissertação, que constitui um trabalho original e inédito, que nunca foi submetido a outra Instituição de ensino superior para obtenção de um grau acadêmico ou outra habilitação. Atesto ainda que todas as citações estão devidamente identificadas e que tenho consciência de que o plágio constitui uma grave falta de ética, que poderá resultar na anulação da presente dissertação.

## AGRADECIMENTOS

Nesta página dedicada aos agradecimentos, quero deixar uma palavra de reconhecimento a todos aqueles que durante esta etapa me ajudaram a continuar e a vencer. A ordem pela qual agradeço é aleatória, não há nem últimos nem primeiros, porque cada um colaborou à sua maneira, e todos juntos foram muito relevantes, bastava faltar um elemento para que todo este trabalho não se realizasse, é um círculo em que ninguém pode ser excluído, pelo que passo a enumerar:

À Professora Doutora Isabel Pedrosa, minha orientadora que é uma pessoa surpreendente e incansável, que sempre disponibilizou o seu precioso tempo para me auxiliar e ajudar. Depositou muita confiança neste trabalho, e espero ter atingido o mínimo das suas expectativas, sim porque as suas expectativas são do tamanho do universo. Foi uma verdadeira companheira e amiga, pelo seu pragmatismo e percepções consistentes possibilitou a conclusão deste trabalho. Dificilmente se encontra uma pessoa com o seu ânimo e energia, bem-haja.

Ao meu coorientador Professor Doutor Raul Laureano, pela sua inteira objetividade dedicação, disponibilidade e profissionalismo.

À gerente da empresa que teve a gentileza e coragem de me disponibilizar a base de dados, bem como pelo tempo, apoio e dedicação que sempre me dispensou.

Ao Sr. Presidente do ISCAC, Dr. Manuel Castelo Branco pela disponibilidade para solicitar a licença para uso do Caseware IDEA e à Caseware IDEA Inc., na pessoa de Alain Soublière, que gentilmente cedeu uma licença completa deste *software*, fundamental para a realização deste trabalho.

À minha família: à minha mãe que se desdobrou tantas vezes para me ajudar e que “*fez das tripas coração*”, sempre compreendeu o meu *stress* e angústia, ao meu pai pela dedicação, ao meu irmão, à minha cunhada e ao meu “pequeno” querido sobrinho.

Ao meu melhor amigo e companheiro, Filipe e à sua família pelo apoio e dedicação.

Às minhas “sócias” do mestrado que sempre me apoiaram, e nunca me deixaram “morrer na praia”, o vosso companheirismo foi fundamental e decisivo.

Às minhas “melhores” amigas que são especiais (vocês sabem! sim, agora já podemos combinar viagens e petiscos!!!).

## **ACKNOWLEDGMENTS**

This research was possible due to the cooperation of Caseware IDEA Inc and Alain Soublière, Chief Product Strategist, who kindly provided a full license of Caseware IDEA that was fundamental to this project success.

## RESUMO

A temática da fraude que, nos tempos mais recentes, tem vindo ao de cima, é conhecida como uma infração comum, podendo ser realizada por qualquer pessoa ou resultar de conluio, tendo como fim o decréscimo patrimonial das organizações.

Ao nível da auditoria externa o objetivo é obter evidência apropriada e suficiente que forneça uma segurança elevada, embora não absoluta, de que as demonstrações financeiras estão isentas de fraude ou de erros materiais. Já ao nível da auditoria interna, o objetivo passa pela identificação do risco de fraude, através da definição de estratégias com o fim de conceber a capacidade para prevenir e antecipar as necessidades de controlo.

Atualmente, as organizações geram e armazenam mais informação em formato eletrónico. Como existem avultadas quantidades de dados, as transações que são alvo de planos abusivos são “escondidas” e são de difícil deteção pelos meios tradicionais. Assim, cada vez mais os métodos analíticos se tornam fundamentais na identificação de potenciais indícios de fraude.

O presente trabalho de investigação consiste em aferir e clarificar sobre a utilidade da aplicação de métodos analíticos ao ficheiro SAFT - *Standard Audit File for Tax Purposes*, como ferramenta utilizada pelos auditores com o objetivo de deteção de erros/anomalias nas vendas e ofertas de uma empresa comercial, utilizando, para esse efeito Caseware IDEA e Microsoft Excel. Foram analisadas as vendas correspondentes aos 6 primeiros meses de 2017 de uma empresa comercial, num total de mais de 132430 registos. Quanto às vendas foi aplicada a Lei de Benford e o *Even Amount/Números Redondos*, entre outros testes intermédios, tendo-se verificado que o valor das vendas não está em conformidade com a Lei de Benford. Para analisar os produtos oferecidos, usou-se o teste Z-score e o *Same-Same-Same* que permitiram identificar anomalias relativamente a vários artigos.

O principal contributo deste trabalho centra-se na proposta de melhorias ao Sistema de Controlo Interno da empresa, como consequência dos métodos analíticos aplicados e dos resultados obtidos.

**Palavras-chave:** Fraude; Auditoria; Análise de Dados; Métodos analíticos; Deteção de Anomalias

## **ABSTRACT**

The issue of fraud has recently been more frequent and is a common infraction that can be perpetrated by a single person or arise as a result of collusion, causing asset decrement or theft in organizations.

External audits aim to obtain relevant and sufficient evidence that points to high, but not absolute, assurances that the financial statements are free from fraud or material misstatement. On the other hand, the main goal of internal auditing is to identify the risk of fraud by developing strategies that provide the capacity to prevent it and anticipate control needs.

Organizations nowadays generate and store more information in electronic format. Because there are large amounts of data, the transactions that are subject of abuse are concealed and harder to detect by traditional means. Thus, analytical methods become increasingly fundamental in identifying and uncovering potential evidence of fraud.

This paper tries to assess the usefulness of applying analytical methods to the SAFT file (Standard Audit File for Tax Purposes), so as to detect errors/anomalies in the sales and offerings of a commercial company, through the use of Caseware IDEA and Microsoft Excel software. The analyzed data set corresponds to the first two sales quarters of 2017 of a commercial company, with over 132,430 records. Benford's Law and Even Amount/Rounding Numbers were applied to the sales figures, among other intermediate tests, and it was found that the sales value did not comply with Benford's Law. The Z-score test and the Same-Same-Same test were used to identify potential anomalies in the products.

The results obtained from the analytical methods applied led to the main contribution of this work, which is the proposal of several improvements to the Internal Control System of the company in question.

**Keywords:** Fraud; Audit; Data Analysis; Analytical Methods; Anomaly Detection

# ÍNDICE GERAL

INTRODUÇÃO .....	1
1 Revisão de Literatura .....	4
1.1 A Fraude.....	4
1.1.1 Conceito de Fraude .....	4
1.1.2 Fraude Ocupacional .....	5
1.1.3 Árvore da Fraude e Classificação dos Tipos de Fraude.....	6
1.1.4 Modelos Teóricos e Ocorrência da fraude .....	8
1.1.4.1 Modelo do Triângulo da Fraude.....	8
1.1.4.2 Teoria do Diamante da Fraude.....	10
1.1.5 Abordagem da Fraude à Auditoria.....	11
1.1.5.1 Conceito de Auditoria .....	11
1.1.5.2 Auditoria Interna e a Fraude .....	11
1.1.5.3 Controlo Interno, o COSO e a Fraude.....	13
1.1.5.4 Auditoria Externa e Fraude .....	14
1.2 Fraude e Análise de Dados.....	16
1.2.1 Análise de Dados .....	16
1.2.2 Ferramentas de Análise de Dados.....	16
1.2.3 Análise de Dados e Auditoria .....	17
1.2.3.1 Vantagens da Aplicação de Análise Dados.....	17
1.2.4 Anomalias .....	18
1.2.5 Análise de Dados e Fraude .....	18
1.2.6 Software de Análise de Dados Para Auditoria.....	19
1.2.6.1 Técnicas de Auditoria Assistidas Por Computador – CAAT .....	19
1.2.6.2 Programas disponíveis para Análise de Dados.....	20
1.3 Abordagem de Métodos Analíticos na Detecção de Anomalias.....	22
1.3.1 Lei de Benford .....	22
1.3.2 Number Duplication Test ou Teste da Duplicação dos Números.....	24

1.3.3	Z-Score.....	25
1.3.4	Relative Size Factor Test, RSF.....	26
1.3.5	Same-Same-Same Test .....	27
1.3.6	Same Same Different Test .....	28
1.3.7	Even Amount/Números Redondos .....	29
2	Metodologia do Estudo .....	30
2.1	Caracterização da Empresa e dos Procedimentos Relativos às Vendas.....	30
2.2	Objetivos Específicos.....	34
2.3	Recolha e Compreensão dos Dados .....	34
2.3.1	Notas Introdutórias .....	34
2.3.2	Processo de Tratamento dos Dados .....	35
2.3.3	A Distinção Entre Amostra, Bónus e Oferta.....	45
2.4	Testes Para Análise dos Dados .....	48
3	Apresentação do Estudo e Resultados .....	50
3.1	Verificar se Existem Anomalias ao Nível dos Valores Totais das Faturas.....	50
3.2	Identificar Situações Anómalas em Relação aos Produtos Oferecidos.....	56
3.3	Sugestões de Melhoria no Sistema de Controlo Interno .....	73
	CONCLUSÃO .....	74
	REFERÊNCIAS BIBLIOGRÁFICAS .....	76
	APÊNDICES .....	81
	APÊNDICE 1. Ligações entre tabelas extraídas do ficheiro SAFT .....	82
	APÊNDICE 2. Como é que se executa um Sumário no IDEA – Clientes por vendedor .....	84

## ÍNDICE DE FIGURAS

Figura 1 - Árvore da Fraude .....	7
Figura 2 - Triângulo da Fraude de Cressey.....	9
Figura 3- Diamante da Fraude .....	10
Figura 4 - Controlo Interno do Coso - Integrated Framework.....	13
Figura 5 - Lei de Benford distribuição de frequência do primeiro dígito.....	23
Figura 6 - Exemplo de compras semelhantes na mesma data pelo mesmo titular do cartão .....	28
Figura 7 - Recursos Humanos por género na empresa TADIMA LDA .....	32
Figura 8 - Número de Funcionários por Departamento empresa TADIMA LDA .....	32
Figura 9 - Processo de emissão de documentos de faturação na empresa TADIMA LDA .....	33
Figura 10 - Arquivo de Ficheiros.....	36
Figura 11 - Importação de ficheiros do SAFT .....	36
Figura 12 - Formatação de campos na importação .....	37
Figura 13 - Importação de dados - Tipos de campos.....	37
Figura 14 - Ficheiros utilizados na análise de dados da empresa TADIMA LDA.....	39
Figura 15 - Campos do ficheiro Clientes .....	39
Figura 16 - Campos do ficheiro CLIENTES_VENDEDORES.....	40
Figura 17 - Campos do ficheiro Customer 2017.....	41
Figura 18 - Campos do ficheiro Invoice 2017 .....	41
Figura 19 - Número de Faturas (FT) e de Notas de Crédito (NC).....	42
Figura 20 - Campos do ficheiro Line 2017 .....	42
Figura 21 - Campos do ficheiro Product 2017.....	43
Figura 22 - Estatísticas de Campo – Tipo de campo: numérico – Ficheiro: FT_ICV ....	44
Figura 23 - INVOICES FT - FOF INVOICES com GROSSTOTAL 0,00€ .....	44

Figura 24 - Estatísticas de Campo – Tipo de campo: data – Ficheiro: FT_ICV.....	45
Figura 25 - Extração da INVOICE_ID 1038 .....	47
Figura 26 - Extração da INVOICE_ID 3025 .....	47
Figura 27 - Extração das INVOICENO = FOF .....	48
Figura 28 - Detecção de Quebras de Sequência .....	50
Figura 29 - Detecção Quebras Sequência FA.....	51
Figura 30 - Same Invoice Same InvoiceDate Same Customer .....	51
Figura 31 - Extração INVOICES Cliente 30003 .....	52
Figura 32 - Aplicação da Lei de Benford .....	53
Figura 33 - Lei de Benford - Primeiro Dígito aplicado ao GROSSTOTAL.....	53
Figura 34 - Lei de Benford - Dois Primeiros Dígitos aplicada ao GROSSTOTAL .....	54
Figura 35 - Extração dos Altamente Suspeitos para dois primeiros dígitos 15 .....	55
Figura 36 - EVEN AMOUNT para GROSSTOTAL.....	56
Figura 37 - FT Produtos oferecidos .....	57
Figura 38 - Produtos Oferecido sem transporte, paletes e artigos tipo DOC.....	58
Figura 39 - Quantidade de Paletes oferecidas.....	59
Figura 40 - Ficheiro composto pelo total de artigos cujo SETTLEMENTAMOUNT é zero .....	61
Figura 41 - Ficheiro composto pelo total de artigos vendidos.....	61
Figura 42 - Ficheiro ARTIGOS SETTLEMENTAMOUNT E UNITPRICE MAIOR 062	
Figura 43 - Ficheiro com extração do produto 431101 .....	63
Figura 44 - Preços praticados para o PRODUCTCODE 431101 .....	63
Figura 45 - Extração PRODUCTCODE 431101 .....	64
Figura 46 - Soma da QUANTITY e do valor do SETTLEMENTAMOUNT por CLIENTE .....	65
Figura 47 - Soma dos produtos oferecidos por total de linhas e por cliente.....	65

Figura 48 - Z-Score para quantidades oferecidas e para quantidades vendidas por cliente .....	68
Figura 49 - Diferença entre as vendas e os produtos oferecidos ordenado na diferença por ordem crescente .....	69
Figura 50 - Cliente 72453 - Diferença entre as vendas e os produtos oferecidos ordenado pelo S_TOTAL por ordem decrescente .....	70
Figura 51 - Cliente 31049 - Diferença entre as vendas e os produtos oferecidos ordenado na diferença por ordem crescente .....	70
Figura 52 - Teste Same-Same-Same - PRODUCTCODE/A + INVOICE_ID/A + INVOICEDATE/A para o CLIENTE = "31049" .....	71
Figura 53 - INVOICE_ID 5100 análise do campo SETTLEMENTAMOUNT .....	71
Figura 54 - Cliente 71407 diferença entre as vendas e os produtos oferecidos ordenado na diferença por ordem crescente .....	72
Figura 55 - Cliente 71407 - Diferença entre as vendas e os produtos oferecidos ordenado pelo S_TOTAL por ordem decrescente .....	72

## ÍNDICE DE TABELAS

Tabela 1- Princípios e componentes do COSO .....	14
Tabela 2 - Tabela Parcial de Z-Score.....	25
Tabela 3 - Evolução das Vendas da empresa TADIMA LDA .....	31
Tabela 4 - Relação entre os dados do SAFT e os Procedimentos internos da empresa..	38
Tabela 5 - Calculo do Z-Score para as quantidades oferecidas .....	66
Tabela 6 - Z-Score por Cliente por quantidades vendidas.....	67

## **Lista de abreviaturas, acrónimos e siglas**

ACFE - *Association of Certified Fraud Examiners*

ACL - *Audit Command Language*

AICPA - *Association of International Certified Public Accountants*

CAAT - *Técnicas de Auditoria Assistidas por Computador*

COSO – *Committee of Sponsoring Organizations of the Treadway Commission*

CSV – *Comma Separated values*

DA – *Análise de Dados*

DF – *Produtos Danificados*

FT - *Faturas*

IDEA - *Interactive Data Extraction & Analysis*

IFAC – *International Federation of Accountants*

IPAI – *Instituto Português de Auditoria Interna*

NC – *Notas de Crédito*

NDT – *Number Duplication Test*

RSF – *Relative Size Factor*

SAFT- *Standard Audit File for Tax Purposes*

SSD - *Same-Same-Different*

SSS - *Same-Same-Same*

XML - *Extensible Markup Language*

## **INTRODUÇÃO**

### **Enquadramento do Tema**

A fraude é um fenómeno sistémico que provoca avultados prejuízos às organizações, sendo de difícil deteção já que, na maioria das vezes, a fraude é cometida pelos próprios funcionários que são conhecedores da organização e do seu modo de funcionamento.

A prática da auditoria depara-se com grandes e constantes desafios, que envolvem a análise de um elevado e crescente volume de informação.

Atualmente, as organizações geram e armazenam mais informação em formato eletrónico do que nunca, e, embora sejam realizadas mais análises aos dados disponíveis, a fraude persiste. Como existem avultadas quantidades de dados, as transações que são alvo de planos abusivos são escondidas e são de difícil deteção pelos métodos tradicionais. Assim, cada vez mais os métodos analíticos são fundamentais na identificação de erros/anomalias ou potenciais indícios de fraude por parte dos auditores.

Quando se pensa em medidas antifraude, as pessoas tendem a concentrar-se do lado da despesa, mas também há oportunidades para a existência de fraude do lado dos proveitos.

Assim sendo, com este estudo de caso pretende-se dar resposta à seguinte questão “conseguirão as empresas detetar anomalias nas vendas (análise ao ficheiro SAFT), usando ferramentas de *Análise de Dados (DA)*?”

### **Objetivos e Contributos**

Atendendo à questão colocada são definidos os seguintes objetivos:

- analisar se existem situações anómalas para os valores totais das faturas;
- identificar se existem situações anómalas em relação aos produtos oferecidos.

A concretização destes objetivos é efetuada através de um estudo de caso em que se analisam os dados constantes do ficheiro SAFT de uma empresa comercial e, em particular, o valor total das vendas euros, as quantidades vendidas, os descontos das ofertas e preços unitários de venda. São também realizadas análises aos descontos, preços,

reversões de vendas e notas de crédito, que podem identificar anomalias indiciadoras de fraude.

Deste modo, este estudo é uma mais-valia para a empresa, por proporcionar à gestão informação útil sobre se os procedimentos de controlo interno (a nível do processo de vendas), nomeadamente, se estão a ser cumpridos e, caso não estejam, fornece uma ferramenta no sentido de inculcar uma cultura de honestidade que reja os funcionários pelos valores e princípios éticos, devendo ser uma base para a prevenção, deteção e correção de irregularidades. Adicionalmente, proporciona uma estratégia antifraude, para tentar restringir a oportunidade, monitorizar e diminuir a pressão, e limitar a capacidade de racionalização. Assim, a empresa é mais eficiente na gestão dos recursos disponíveis, e com isso incrementa os seus resultados financeiros.

## **Abordagem Metodológica**

A pesquisa consiste numa revisão da literatura e num estudo de caso aplicado a uma empresa comercial pelo que, relativamente ao método de pesquisa utilizada, optou-se por um estudo de caso, considerando-se a disponibilidade de dados para a realização do mesmo.

Atendendo aos objetivos selecionaram-se os métodos analíticos mais adequados, os quais foram aplicados aos dados disponíveis no ficheiro SAFT relativo aos primeiros seis meses de 2017. Esses dados foram convertidos e importados para a ferramenta de análise de dados Caseware IDEA e, também, para o Microsoft Excel.

## **Estrutura**

A dissertação encontra-se estruturada da seguinte forma: no presente capítulo é feito um enquadramento do tema, dos objetivos, é abordada a pertinência e a metodologia do estudo.

No primeiro capítulo, apresenta-se a revisão da literatura em que se efetua um enquadramento da fraude, o que é, em que situações ocorre, por quem pode ser cometida, os tipos de fraude, e como a auditoria pode contribuir para a sua deteção. Neste capítulo, também é abordada a análise de dados, as ferramentas existentes e a sua ligação com a fraude e com a auditoria, bem como as vantagens de se utilizar *software* de análise de dados em auditoria. Para encerrar este capítulo, são abordados alguns métodos analíticos para deteção de anomalias.

Seguidamente, no segundo capítulo, é realizado um enquadramento relativamente à empresa alvo do estudo, uma descrição dos objetivos específicos e a recolha e compreensão dos dados, bem como os métodos utilizados na sua análise.

No terceiro, é apresentado o estudo, os testes e os resultados detalhados de cada temática abordada. Apresenta-se, ainda, um conjunto de recomendações – que poderão ser integradas no Sistema de Controlo Interno da empresa – como consequência das análises efetuadas.

Por fim, são apresentadas as conclusões, contributos, as limitações encontradas no âmbito deste trabalho, bem como um conjunto de sugestões para investigações futuras.

## **1 Revisão de Literatura**

Este capítulo tem por finalidade introduzir os principais conceitos abordados na presente investigação. No primeiro ponto será apresentado um breve enquadramento teórico da fraude, sendo expostos os diversos tipos de fraude, os fatores que a podem propiciar e a relação entre a fraude e a auditoria. No segundo ponto será abordada a análise de dados e a sua relação com a fraude, bem como um estudo do software disponível no mercado para dar suporte à análise de dados e ainda à investigação realizada pelo auditor na deteção de erros/anomalias e fraude. A revisão de literatura será finalizada com a descrição de alguns métodos analíticos que podem ser usados na deteção de anomalias/fraude.

### **1.1 A Fraude**

No mundo do comércio, as organizações incorrem em despesas para produzir e vender os seus produtos ou serviços, estes custos percorrem toda a linha: mão-de-obra, impostos, publicidade, rendas, matérias-primas, investigação e desenvolvimento e, também, a fraude. Este último custo é muito diferente dos restantes, porque se encontra oculto, mesmo que refletido nos números de perdas e lucros (WELLS, 2009).

#### **1.1.1 Conceito de Fraude**

Existem várias definições e classificações de fraude, depende de múltiplos critérios, diferentes perspetivas, finalidades e da própria realidade em que se insere.

Fraude pode ser definida como: “*1-Ato de má-fé praticado com o objetivo de enganar ou prejudicar alguém; burla; engano; logração.*

*2-Ato ou comportamento que é ilícito e punível por lei...*”<sup>1</sup>

É relevante fazer a distinção entre erro e fraude, visto que são conceitos distintos mas, por vezes, confundidos: os erros são ações involuntárias de omissão, distração, desconhecimento ou má interpretação (Bunget, Grigori, & Dumitrescu, 2009), enquanto que a fraude é definida, como “*quaisquer atos ilegais caracterizados pelo engano, encobrimento ou violação da confiança. Tais atos não dependem de ameaça de violência ou de força física. As fraudes são perpetradas por indivíduos e organizações para se*

---

<sup>1</sup> <https://www.infopedia.pt/dicionarios/lingua-portuguesa/fraude> consulta em 17-06-2017

*apropriarem de dinheiros, bens ou serviços; para evitarem o pagamento ou perda de serviços; ou para obterem vantagens pessoais ou comerciais” (IPAI, 2009, p. 37), contudo, a fraude é cometida com vista a prejudicar uma pessoa ou uma organização. De modo análogo o IFAC na ISA 240 - *The Auditor’s Responsibilities Relating to Fraud in an Audit of Financial Statements* define fraude como sendo “*um ato intencional praticado por um ou mais indivíduos entre a gerência, os encarregados da governação, os empregados ou terceiros, envolvendo o uso propositado de falsidades para obter uma vantagem injusta ou ilegal*”(IFAC, 2012, p. 3).*

A fraude está conexas às noções de engano e de intencionalidade. Joseph Wells considera que “*a fraude pode incluir qualquer crime para obtenção de lucro, utilizando como principal modus operandus o logro. Existem apenas três modos de retirar ilegalmente dinheiro a uma vítima: força, logro ou furto*”(WELLS, 2009, p. 18). Precedentemente, a fraude implica uma má conduta intencional, realizada com o intuito de evitar a deteção e é concebida para enganar outros, resultando, por isso, em perdas para esses e ganhos para o prevaricador (Bunget *et al.*, 2009). A fraude é um termo usado para descrever uma grande variedade de comportamentos desonestos, tais como suborno, corrupção, falsificação, declarações falsas, conluio e ocultação de factos relevantes (HM Treasury, 2012).

A ACFE<sup>2</sup> (*Association of Certified Fraud Examiners*) considera que o termo fraude abrange muitas formas de má conduta, embora a definição legal de fraude seja muito específica, para a maioria das pessoas o uso comum é muito mais amplo e geralmente abrange qualquer tentativa de enganar outra parte para obter um benefício (ACFE, 2012).

### **1.1.2 Fraude Ocupacional**

A fraude cometida contra uma organização pode ser realizada tanto internamente pelos funcionários, gerentes, diretores, ou proprietários, como externamente, por clientes, ou

---

<sup>2</sup> A ACFE é a maior organização do mundo no combate à fraude, sendo também o primeiro fornecedor de formação antifraude. Inclui mais de 75.000 membros, com o objetivo de contribuir para a redução mundial da fraude nas organizações e inspirar a confiança pública (in [www.acfe.com](http://www.acfe.com)).

fornecedores, pelo que a ACFE define fraude ocupacional como “*o uso de uma profissão para enriquecimento pessoal por meio do uso deliberado indevido ou da má aplicação dos recursos ou ativos da organização empregadora*” (ACFE - Association of Certified Fraud Examiners, 2016, p. 6). O conceito, em termos genéricos, de fraude ocupacional assume uma natureza económica, e deve sublinhar-se o facto de ter sempre subjacente a intencionalidade de quem a perpetra. Com o objetivo de classificar as fraudes e os abusos ocupacionais foi criada a *Árvore da Fraude*.

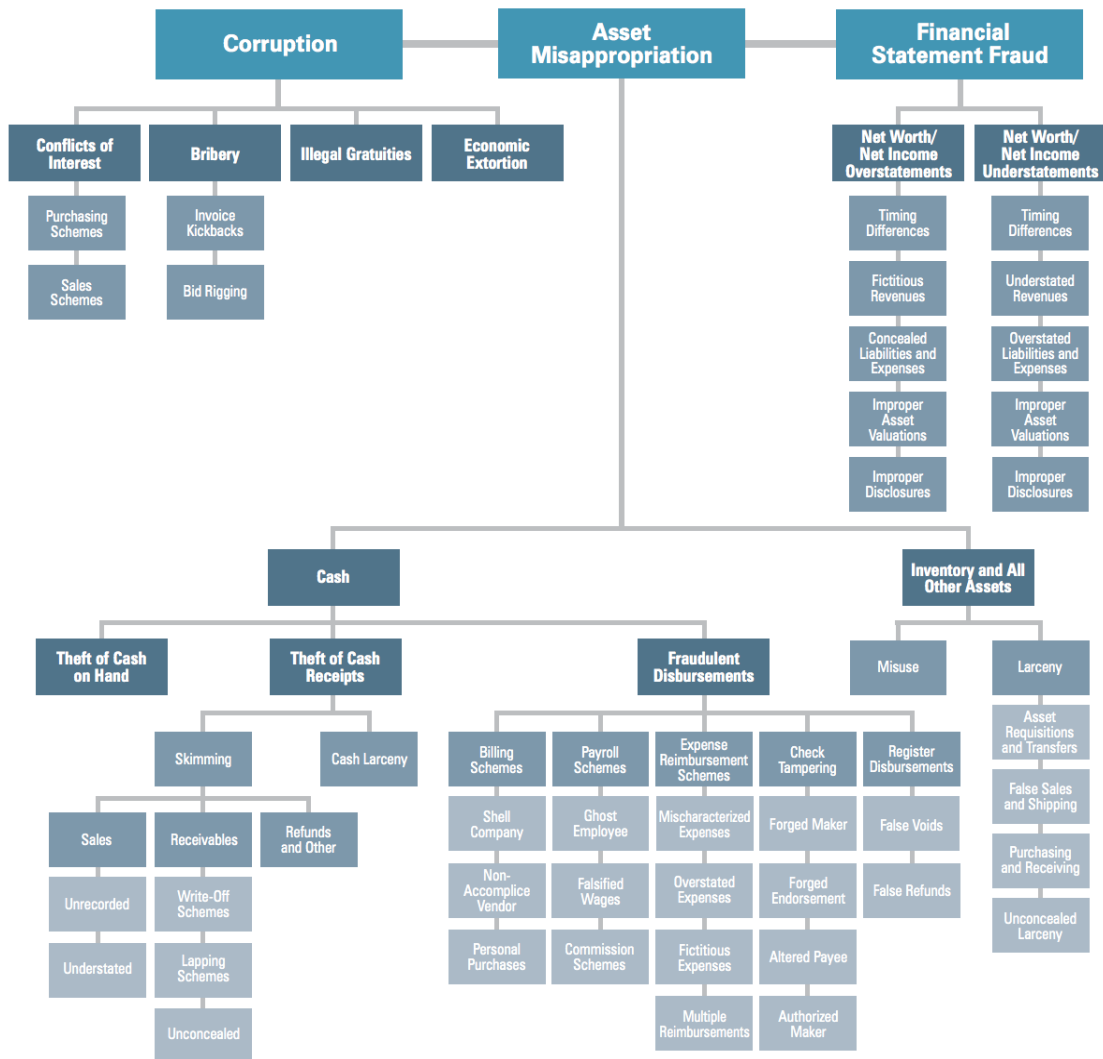
### **1.1.3 *Árvore da Fraude e Classificação dos Tipos de Fraude***

A *Árvore da Fraude* surgiu no primeiro relatório da ACFE sobre a Fraude e Abuso Ocupacional, publicado em 1996 (ACFE, 2016). Ao longo do tempo, a ACFE foi realizando pequenos ajustes na estrutura da árvore de modo a melhorá-la. A estrutura da árvore da fraude, vinte anos após a sua criação, detém, ainda, a mesma consistência, mas os mecanismos e abordagens utilizados pelos profissionais que a estudam e investigam têm-se vindo a reestruturar para ir ao encontro dos defraudadores que adotam novos esquemas.

A *Árvore da Fraude* subdivide-se em corrupção, apropriação indevida de ativos e relatório financeiro fraudulento (ACFE - Association of Certified Fraud Examiners, 2016), conforme pode ser observado na Figura 1 - *Árvore da Fraude*.

Na corrupção é utilizada a influência ou o poder para se conseguir uma vantagem (para o autor e/ou terceiros) contrária ao dever para com a entidade empregadora ou ao direito da contraparte. Neste grupo integram-se formas de fraude como o suborno, as gratificações ilegais, o conflito de interesses e a extorsão económica (ACFE - Association of Certified Fraud Examiners, 2016). O suborno representa a prática de oferecer dinheiro ou benefícios a um indivíduo, em troca de um ato ilícito, diligenciando a obtenção de vantagens. No caso das gratificações ilegais estas constituem uma recompensa paga, de forma ilícita, pela realização de um determinado serviço. O conflito de interesses evidencia o favorecimento pessoal, ou de um terceiro, pelo desenvolvimento de esquemas de compras ou vendas. A extorsão económica compreende a obtenção de vantagens por parte de outrem, com recurso à coação, chantagem ou violência (Association of Certified Fraud Examiners, 2012).

Figura 1 - Árvore da Fraude



Fonte: Report to The Nations 2016 Global Fraud Study (ACFE - Association of Certified Fraud Examiners, 2016, p. 11)

A apropriação indevida de ativos envolve o furto ou utilização indevida dos ativos da empresa. Relativamente ao furto, as situações mais habituais respeitam a dinheiro e a inventários. No furto de dinheiro, há a considerar o furto (por exemplo, de dinheiro de caixa), os pagamentos fraudulentos (nos quais se incluem, por exemplo, pagamentos a empresas ou empregados fantasma, compras de bens e serviços para uso pessoal, despesas fictícias ou sobreavaliadas, falsificação de cheques e fraudes na operação), e a sonegação/encobrimento (que abrange a sonegação de vendas, de contas a receber e de reembolsos). No que respeita a inventários, encontram-se tipificadas duas situações: a utilização indevida e o furto (este é possível por várias formas, como por exemplo através

de falsas vendas e remessas ou através das compras). É nesta categoria (apropriação indevida de ativos) onde se verifica o maior número de fraudes (ACFE - Association of Certified Fraud Examiners, 2016).

Por fim, os relatórios financeiros fraudulentos, em que a informação contida nos relatórios não é verdadeira e tem como objetivo induzir o leitor em erro. A Árvore da Fraude divide os relatórios em financeiros e não-financeiros, especificando que, para os primeiros, se pode considerar quer a sobreavaliação, quer a subavaliação de ativos e receitas. Nesta categoria de fraude inclui-se também a prestação de informação não financeira, como exemplo a divulgação efetuada por uma empresa relativa à invenção de um presumível novo produto (com a finalidade de suscitar vantagens no mercado por parte da empresa) que afinal não está ainda em condições de ser comercializado. Habitualmente este tipo de fraude é cometida ao nível do órgão de gestão, com o intuito de melhorar a aparência económica das organizações (Bishop *et al.*, 2007), ou para poder beneficiar diretamente da fraude através da venda de ações ou receção de prémios de desempenho (ACFE - Association of Certified Fraud Examiners, 2016).

#### **1.1.4 Modelos Teóricos e Ocorrência da fraude**

##### ***1.1.4.1 Modelo do Triângulo da Fraude***

O modelo teórico do triângulo da fraude, desenvolvido por Donald Cressey, refere que, a prática de um ato de natureza fraudulenta é sempre antecedida de um processo de decisão por parte do respetivo autor e cujo sentido parece derivar da avaliação que faz sobre determinados aspetos que contextualizam o seu "*aqui e agora*". Trata-se de uma espécie de equação, cujo resultado depende das três variáveis que a integram, e que são a pressão, a racionalização e a oportunidade (Cressey, 1953), pelo que pode ser visualizado na Figura 2.

A pressão/incentivo, própria da vida particular, nomeadamente da que resulta de necessidades urgentes de liquidez financeira, incluindo-se, a título de exemplo, a existência de dívidas, hábitos de jogo ou de consumos aditivos. A pressão também pode resultar de simples "*necessidades*" relacionadas com a ostentação de determinados símbolos, associados à posse de certos objetos, que façam o sujeito sentir-se integrado no grupo social com o qual se revê. Em qualquer dos casos, a pressão resulta de uma necessidade premente de alcançar dinheiro ou bens com valor de natureza material, a fim

de solucionar um problema (no primeiro caso), ou simplesmente para aumentar as disponibilidades financeiras (os dividendos resultantes da prática do ato adicionam-se às receitas normais da sua atividade profissional) para poder aceder à posse de tais símbolos de integração social (Arens *et al.*, 2007).

A atitude/racionalização entendida como a capacidade possuída pelo sujeito para racionalizar os diversos dados que possui sobre a realidade que o rodeia, e, cujo somatório, entre eventuais perdas e ganhos o levam a decidir ou não pela prática do ato. A racionalização, não raramente, é apresentada apenas como uma justificação, às vezes como uma desculpa criada pelo sujeito desviante para a concretização da sua motivação criminal empresarial (Hencsey, 2016).

Quanto à oportunidade para a prática de um ato de natureza fraudulenta, pode dizer-se que a inexistência de um sistema de controlo interno eficaz facilita a realização de atos fraudulentos (Dal-Ri Murcia, Borba, & Schiehl, 2008).

*Figura 2 - Triângulo da Fraude de Cressey*



Fonte: Adaptado de Arens *et al.* (2007, p. 316)

Hencsey considera que nenhum modelo é capaz de descrever sozinho a complexidade da motivação para a realização de crimes empresariais e que não é pela subjetividade da racionalização que este argumento se faz verdadeiro (Hencsey, 2016). Alguns autores, críticos do triângulo da fraude tais como Kassem e Higson (2012), citaram a obra de Dorminey, Fleming, Kranacher, & Riley (2011) argumentando que esta teoria sozinha, não pode explicar os desvios empresariais, uma vez que dois dos seus elementos a racionalização e a pressão não podem ser observados.

A fraude pode ocorrer quando a intuição dos indivíduos lhes diz que é aceitável cometer fraude por causa de uma racionalização (Murphy & Dacin, 2011).

#### **1.1.4.2 Teoria do Diamante da Fraude**

A Teoria do Diamante da Fraude foi apresentada por Wolfe e Hermanson, em 2004, e trata-se de uma expansão da versão do triângulo da fraude, em que o elemento capacidade foi adicionado às três componentes iniciais (pressão, racionalização e oportunidade) conforme consta da Figura 3. Ainda que a pressão possa coexistir com a oportunidade e racionalização, é improvável que haja fraude a menos que a capacidade para a executar esteja presente. Por outras palavras, o potencial autor deve ter habilidade e capacidade para cometer fraude (Wolfe e Hermanson, 2004).

*Figura 3- Diamante da Fraude*



**Fonte:** Adaptado de Wolfe & Hermanson (2004)

A oportunidade abre a porta para a fraude, o incentivo (isto é, pressão) e a racionalização levam uma pessoa para a porta, já a capacidade permite que a pessoa possa reconhecer a porta aberta como uma oportunidade e dela tirar proveito. Corresponde à situação de ter as características, competências ou aptidões necessárias para o ato de cometer fraude (Wolfe e Hermanson, 2004), ou seja, a pessoa reconheceu uma oportunidade particular de fraude e teve a capacidade para a transformar em realidade. Posição, inteligência, ego, coação, engano e stress são os elementos demonstrativos de capacidade (Wolfe & Hermanson, 2004). Mackevicius & Lukas Giriunas (2013) consideram que todas as pessoas que possuem motivação, oportunidade e racionalização podem cometer fraude, mas têm que ter capacidade para a realizar ou esconder. Já anteriormente, W. S. Albrecht, Williams e Wernz (1995), entenderam que a capacidade é de particular importância

quando se trata de uma fraude em grande escala ou que se prolongue no tempo. Além disso, Albrecht *et al.* (1995) acreditam que apenas a pessoa que tenha capacidade extremamente alta, será capaz de perceber o controlo interno existente (ou a falta dele), para identificar os pontos fracos e usá-los no planeamento e implementação da fraude.

## **1.1.5 Abordagem da Fraude à Auditoria**

### ***1.1.5.1 Conceito de Auditoria***

Batista da Costa considera que não existe uma definição rígida do que é a auditoria, embora determinados termos ou expressões sejam utilizados nas definições dadas por diversos autores e organismos profissionais (Costa, 2010). Arens alude que auditoria é a acumulação e avaliação de informações baseadas em evidências, para determinar e informar sobre o grau de correspondência entre a informação e os critérios estabelecidos, devendo a auditoria ser realizada por uma pessoa independente e competente (Arens *et al.*, 2007).

A Morais & Martins destaca que “*a Auditoria é o processo sistemático de objetivamente obter e avaliar prova acerca da correspondência entre informações, situações ou procedimentos e critérios preestabelecidos, assim como comunicar conclusões aos interessados*”<sup>3</sup> (Morais & Martins, 2013, p. 19).

### ***1.1.5.2 Auditoria Interna e a Fraude***

“*A auditoria interna é uma atividade independente de garantia e de consultoria, destinada a acrescentar valor e a melhorar as operações de uma organização. Ajudar a organização a alcançar os seus objetivos, através de uma abordagem sistemática e disciplinada, na avaliação e melhoria da eficácia dos processos de gestão de risco, de controlo e de governação*” (IPAI, 2009, p. 10).

A função dos auditores internos relativamente à fraude foi contemplada no IPPF – *International Professional Practices Framework* do IIA - *The Institute of Internal*

---

<sup>3</sup> Auditing Concepts Committee, Report of the Committee on Basic Auditing Concepts, The Accounting Review, Vol. 47 1972, p. 18.

*Auditors.* As normas mais relevantes sobre a responsabilidade da auditoria interna no que cinge ao risco de fraude são as seguintes:

- Proficiência (1210.A2) - Os auditores internos devem ter conhecimento suficiente para avaliar o risco de fraude e o modo como este é gerido na organização, mas não é esperado que os auditores internos tenham o mesmo conhecimento de uma pessoa cuja primeira responsabilidade seja a de detetar e investigar fraude.
- Gestão de Risco (2120.A2) - A atividade de auditoria interna deve avaliar a potencial ocorrência de fraude e a forma como a organização faz a gestão do risco de fraude.
- Objetivos do Trabalho de Auditoria (2210.A2) - Os auditores internos devem considerar a probabilidade de erros significativos, fraude, não conformidades e outras exposições ao desenvolver os objetivos do seu trabalho.

Há necessidade de verificação e revisão periódicas do controlo interno, para reduzir o risco de ocorrência de erros ou para que as tentativas de fraude não fiquem encobertas por muito tempo. É responsabilidade da auditoria interna testar o controlo interno e assimilar os possíveis erros e deficiências que possam acontecer e definir soluções.

Os auditores internos são colaboradores da empresa, podem apoiar diretamente o órgão de gestão na implementação de sistemas mais adequados, podendo monitorizar de forma continuada e sistemática o sistema de controlo interno, através da identificação e investigação dos sinais de alerta (*red flags*<sup>4</sup>) suscetíveis de serem indiciadores de fraude (Rezaee, 2002).

---

<sup>4</sup> Uma bandeira vermelha é um conjunto de circunstâncias que são de natureza incomum ou variam da atividade normal. É um sinal de que algo está fora do normal e pode precisar de ser mais investigado. As bandeiras vermelhas não indicam culpa ou inocência, apenas fornecem possíveis sinais de alerta de fraude (Dinapoli & Hancox, n.d.).

### **1.1.5.3 Controlo Interno, o COSO<sup>5</sup> e a Fraude**

O Controlo Interno segundo o COSO (Committee of Sponsoring Organizations of the Treadway Commission) é definido como o processo realizado pelas pessoas da organização (administração, gestão e remanescentes colaboradores), concebido para dar garantia razoável de fiabilidade sobre o atingimento de objetivos relacionados com as operações, relato de informação e conformidade (McNally, 2013), ou seja, é um processo na medida em que é um meio para atingir um fim, e não um fim em si mesmo; é realizado por pessoas, pelo que não é meramente a existência de políticas, procedimentos, manuais, sistemas e formulários; está destinado a dar garantia razoável de fiabilidade, não sendo uma garantia total, à administração, acionistas e demais *stakeholders*; destinado ao atingimento de objetivos em uma ou mais categorias.

*Figura 4 - Controlo Interno do Coso - Integrated Framework*



**Fonte:** McNally (2013, p. 4)

É visível através da Figura 4, que o COSO prevê três categorias de objetivos - operações, relato e conformidade - e tem cinco componentes integradas: ambiente de controlo, avaliação de riscos, atividades de controlo, informação e comunicação, e atividades de monitorização (Tabela 1- Princípios e componentes do COSO). O Framework continua a ser adaptável à estrutura de uma determinada organização, permitindo que se considere controlos internos de uma entidade, divisão, unidade operacional e/ou nível funcional,

---

<sup>5</sup> COSO – Committee of Sponsoring Organizations of the Treadway Commission é uma iniciativa conjunta de cinco entidades: American Accounting Association; American Institute of Certified Public Accountants; Financial Executives International; Institute of Management Accountants; e The Institute of Internal Auditors.

como por exemplo, para um centro de serviços partilhado. Finalmente, é mantido o papel importante do julgamento da administração na conceção, implementação e manutenção do controlo interno, bem como a avaliação da sua eficácia.

*Tabela 1- Princípios e componentes do COSO*

<b>Ambiente de Controlo</b>
1. Demonstra comprometimento com a integridade e os valores éticos
2. Independência e supervisão
3. Estabelece as estruturas, autoridade e responsabilidades
4. Demonstra comprometimento e competência
5. Impõe Responsabilidade
<b>Avaliação de riscos</b>
6. Especificação de objetivos suficientemente claros
7. Identifica e analisa os riscos
8. Avalia o risco de fraude
9. Identifica e avalia as mudanças significativas
<b>Atividades de controlo</b>
10. Seleciona e desenvolve atividades de controlo
11. Seleciona e desenvolve atividades sobre a tecnologia
12. Implementa através de políticas e procedimentos
<b>Informação e comunicação</b>
13. Uso de informações relevantes
14. Transmite internamente
15. Comunica externamente
<b>Atividades de monitorização</b>
16. Realiza avaliações contínuas e/ou independente
17. Avalia e comunica deficiências

**Fonte:** Adaptado McNally ( 2013, p. 5)

É de salientar que, na avaliação dos riscos, faz parte o princípio n.º 8, que é relativo à avaliação do risco de fraude, pelo que os reguladores e os *stakeholders* têm maiores expectativas em relação à supervisão da governança, gestão de risco e deteção e prevenção de fraude (McNally, 2013).

#### **1.1.5.4 Auditoria Externa e Fraude**

Para Marques (1997) o trabalho da auditoria externa incide, por regra ou fundamentalmente, sobre as áreas contabilística e financeira, sendo este tipo de trabalho muitas vezes designado por auditoria financeira. Costa considera que o auditor é responsável por obter uma segurança razoável, de que as demonstrações financeiras

consideradas como um todo, estão isentas de distorções materiais, causadas por fraude ou erro (IFAC, 2012 ; Costa, 2010).

A SAS nº 99<sup>6</sup> exige que os auditores, durante as fases de planeamento e de execução da auditoria, questionem e reconheçam a possível ocorrência de fraude nas demonstrações financeiras. Da aplicação das disposições do SAS nº 99 é esperado melhor desempenho dos auditores na deteção de fraude, identificação dos riscos de fraude e realização de testes de auditoria e execução de procedimentos de auditoria para testar o sistema de controlo interno (AICPA, 2002). Recentemente, o AICPA (*Association of International Certified Public Accountants*) adotou iniciativas no sentido de melhorar a educação antifraude dos investidores e de fornecer instrumentos de auxílio na proteção contra a fraude.

A responsabilidade pela prevenção e deteção de fraude cabe à gestão, mas a pressão exercida sobre os auditores para partilhar os resultados das suas análises de dados com a administração só se tornará mais forte, pelo que dá origem a dilemas éticos. Os auditores consideram a ISA 240 sobre fraude e ISA 520 - *Analytical Procedures* sobre procedimentos analíticos como áreas problemáticas, todavia, a ISA 240 identifica explicitamente os procedimentos analíticos como um método para avaliar o risco de fraude (IFAC, 2012). No entanto, os auditores consideram que os procedimentos estão desatualizados, e não são amplamente utilizados no contexto da fraude, embora possam ser usados para mostrar certos erros básicos. No contexto dos procedimentos analíticos, os requisitos da ISA 520 (IFAC, 2010) para que as expectativas desenvolvidas sejam suficientemente precisas na identificação de uma distorção relevante, não se adequem aos procedimentos muito granulares permitidos pela análise de dados, tornando esses procedimentos inválidos para fins de testes substantivos (ICAEW, 2016).

---

<sup>6</sup> Statement on Auditing Standards No. 99: Consideração da fraude numa auditoria às demonstrações financeiras, abreviada como SAS 99, é uma declaração de auditoria emitida pelo Conselho de Normas de Auditoria do Instituto Americano de Contadores Públicos Certificados (AICPA) em outubro de 2002.

## **1.2 Fraude e Análise de Dados**

### **1.2.1 Análise de Dados<sup>7</sup>**

Análise de Dados é o processo de análise de conjuntos de dados, com o fim de extrair conclusões sobre as informações neles contidos, através do auxílio de sistemas e *software* especializados. As tecnologias e as técnicas de Análise de Dados são utilizadas pelas empresas comerciais e industriais para ajudar na tomada de decisão e por cientistas e investigadores para verificar ou refutar modelos, teorias e hipóteses científicas.<sup>8</sup> Resumindo, é a aplicação de sistemas informáticos na análise de grandes conjuntos de dados para apoiar na tomada de decisão, sendo os dados avaliados, selecionados, tratados, visualizados e analisados, e posteriormente, os resultados são interpretados (Runkler, 2012).

A análise de dados fornece informações sobre o conjunto de dados, identifica as relações e estruturas de dados subjacentes, suposições e hipóteses de teste, reconhece variáveis de relacionamentos causais e deteta anomalias (Gee, 2015).

A Análise de Dados pode ser aplicada na deteção de fraude e nos algoritmos de identificação de exceções, os algoritmos são muitas vezes baseados em esquemas de deteção de anomalias (Deloitte, 2012).

### **1.2.2 Ferramentas de Análise de Dados**

A Análise de Dados envolve extração, usando campos dentro da estrutura básica dos dados em vez do formato dos registos. Quanto melhor for a qualidade dos dados, mais simples será a sua visualização e proceder a análises corretas. Atualmente, existem muitas empresas dedicadas a obter dados e a colocá-los no formato certo, mas nesse ponto não existe evidência para apoiar uma opinião, pelo que o valor da auditoria vem da análise dos dados (ICAEW, 2016).

---

<sup>7</sup> O termo usado em inglês é *Data Analytics*.

<sup>8</sup> Definição adaptada de: <http://searchdatamanagement.techtarget.com/definition/data-analytics>

### **1.2.3 Análise de Dados e Auditoria**

A auditoria contínua é o uso de DA em tempo contínuo ou em tempo real, permitindo que a gestão e a auditoria identifiquem e relatem mais rapidamente a atividade fraudulenta. O acompanhamento contínuo das transações, sujeitas a certos sinais de alerta (*red flags*), pode promover uma análise mais rápida das transações de alto risco (Bishop *et al.*, 2007).

#### **1.2.3.1 Vantagens da Aplicação de Análise Dados**

John W. Tukey, no seu livro *Exploratory Data Analysis* de 1977, distingue entre dois tipos de análise de dados: exploratória e confirmatória. A análise exploratória é ascendente e tácita. No que respeita à análise confirmatória, é uma análise descendente e dedutiva (Hoaglin, 2003).

A análise exploratória começa com os dados e o auditor a colocar um conjunto de questões como: o que pode ter corrido mal; onde é que os riscos aparentam estar; existência de indicadores de fraude; e quais as asserções em que se devem focar. É o tipo de análise que mais se adequa à fase de planeamento da auditoria, nomeadamente no conhecimento da entidade e do seu ambiente, identificar os riscos significativos e posteriormente traçar os procedimentos de auditoria ( Marques, 2016)

A análise confirmatória inicia-se com os objetivos e asserções e o auditor coloca perguntas como: os resultados do modelo estão de acordo com as expetativas definidas? Existem desvios significativos? Formam um padrão não usual? Este tipo de análise é a que possibilita ao auditor avaliar se as asserções definidas estão materialmente corretas, e, no seu objetivo final, garantir se as demonstrações financeiras estão isentas de distorções materialmente relevantes. Fazem parte deste tipo de técnicas os seguintes procedimentos:

- Identificar e analisar padrões ou desvios não usuais;
- Mapear e visualizar a performance da entidade, através da análise de dados a um nível de desagregação superior e através de várias dimensões;
- Construção de modelos estatísticos ou preditivos que expliquem relações entre os dados e permitam identificar desvios não usuais; e
- Combinar informação de várias fontes e de análises distintas de forma a ganhar um melhor entendimento dos processos ou da entidade.

São várias as vantagens da utilização de técnicas de análise de dados, no decurso do processo de auditoria das quais se destacam: maior segurança nas análises efetuadas, porque pode-se trabalhar com a população inteira, sendo possível realizar análises mais eficientes e abrangentes que permitem visualizar e identificar padrões não usuais e desvios que mereçam uma análise mais aprofundada; a aptidão para trabalhar um grande volume de dados; permite transformar a auditoria num processo contínuo; uma maior eficiência na deteção de fraude; e na Redução de “falsos positivos” ( Marques, 2016).

#### **1.2.4 Anomalias**

No processo de análise de dados, "Detetar uma fraude é como encontrar uma agulha no palheiro"<sup>9</sup>. Normalmente, as transações fraudulentas em registos eletrónicos são poucas em relação à grande quantidade de registos em conjuntos de dados. Transações fraudulentas ou com anomalias não são a norma. Outras anomalias, como anomalias nos registos contabilísticos, são devidas a procedimentos inadequados ou a outras deficiências do controlo interno. Essas fraquezas serão repetitivas e ocorrerão com frequência no conjunto de dados, ou seja, acontecem com regularidade e consistentemente em períodos específicos, como, por exemplo, no final do ano. Compreender o negócio, as suas práticas e procedimentos ajudam a explicar a maioria das anomalias (Gee, 2015).

#### **1.2.5 Análise de Dados e Fraude**

A proatividade com a qual alguns esquemas de fraude originam transações identificáveis, amplia a efetiva capacidade de uma organização projetar e implementar análises de dados. A Análise de Dados pode ser usada para assegurar a eficácia no controlo preventivo e detetivo de fraude (Bishop *et al.*, 2007). Existe um amplo leque de ferramentas especializadas em Análise de Dados, capazes de suportar e melhorar a atividade antifraude, mas na prática os gerentes não estão a servir-se delas (Banarescu, 2015).

Os responsáveis pelo controlo interno têm de analisar todas as transações que ocorrem mas tal processo não pode ser realizado manualmente: exige o uso de ferramentas de Análise de Dados. De um modo geral, as organizações trabalham com grandes volumes de dados, pelo que é estritamente necessário implementar processos de contínua

---

<sup>9</sup> Provérbio Popular

monitorização, a fim de identificar anomalias nos fluxos de dados ou padrões comportamentais potencialmente fraudulentos (Banarescu, 2015). A título de exemplo, os criadores dos cartões de crédito conceberam controlos e análise de dados para tentar minimizar o risco de esquemas de fraude (Hess & Cottrell, 2016), a que por norma são muito vulneráveis.

Para se aplicar e interpretar os resultados, usando métodos estatísticos tradicionais ou avançados, o auditor ou o investigador têm de dispor de um bom conhecimento sobre o ambiente do negócio, bem como estar familiarizado com o *software* utilizado para a análise (Gee, 2012).

## **1.2.6 Software de Análise de Dados Para Auditoria**

### **1.2.6.1 Técnicas de Auditoria Assistidas Por Computador – CAAT**

Nos últimos anos, foram desenvolvidos muitos pacotes de *software* de Análise de Dados para auxiliar os investigadores e auditores no seu trabalho normal, bem como de deteção de fraudes e anomalias, que podem ser denominados por CAAT, *Computer-Assisted Audit Techniques* ou Técnicas de Auditoria Assistidas por Computador.

Os objetivos da utilização de CAAT em Análise e Extração de Dados são vários, tais como: aceder e analisar informação de distintas proveniências; garantir integridade dos dados originais; utilizar técnicas de amostragem; registar o histórico dos procedimentos e das análises efetuadas; e automatizar análises (Pedrosa, 2015).

Existem várias vantagens associadas à utilização de CAAT's na análise e extração de dados em auditoria, pelo que as principais são: o facto de o processamento de dados ser mais rápido; a maleabilidade e facilidade de utilização; a possibilidade de análise do universo completo, permitindo a deteção de *outliers*<sup>10</sup> e posterior uso de técnicas de amostragem; a dispensa de conhecimento relativos a SQL (*Structured Query Language*);

---

<sup>10</sup> Um *outlier* é uma observação atipicamente grande ou pequena. *Outliers* podem ter um efeito desproporcional sobre os resultados estatísticos, como a média, o que pode resultar em interpretações equivocadas. Por exemplo, um conjunto de dados inclui os valores: 1, 2, 3, e 34. O valor médio, 10, que é maior do que a maioria dos dados (1, 2, 3), é muito afetado pelo ponto de dados extremo 34. Neste exemplo, o valor médio faz transparecer que os valores dos dados são mais elevados do que realmente são. Devem-se investigar os outliers porque eles podem fornecer informações úteis sobre os dados ou processo em investigação. <https://support.minitab.com/pt-br/minitab/18/help-and-how-to/graphs/supporting-topics/exploring-data-and-revising-graphs/identifying-outliers/> consultado em 20-08-2017

a melhor eficiência e performance das auditorias; e o fornecimento de dados que aumentam a credibilidade da auditoria (Pedrosa, 2015).

A análise de dados em larga escala permite: a contagem de registos, fazer somatórios, identificar os tipos de campos, agrupar e resumir por tipo de dados, classificar as populações em termos estatísticos, analisar registos em função do seu conteúdo, fazer diversos tipos de apuramentos (média, máximos, mínimos, desvio-padrão entre outros), testar a sequência dos registos (verificar a existência de informação duplicada, ou registos que foram anulados ou estão em falta) (Pedrosa, 2015).

### ***1.2.6.2 Programas disponíveis para Análise de Dados***

Dada a amplitude de usos que pode ser dada à Análise de Dados, não é possível enumerar todos os programas disponíveis, pelo que, dos que são mais amplamente usados, incluem-se os seguintes:

- **ACL** (Audit Command Language) é uma ferramenta forte na detecção de fraude através da Análise de Dados que permite melhorias ao nível da eficiência da análise, da realização de testes repetitivos, com uma cobertura mais rápida e alerta precoce, sendo, nesta categoria, a ferramenta mais utilizada pelos auditores em todo o mundo. Embora o foco principal seja a auditoria (incluindo técnicas como amostragem), tem disponível um módulo de fraude. O ACL inclui também uma linguagem de programação chamada *ACLScript* que possibilita a automatização de procedimentos<sup>11</sup>.
- **CaseWare IDEA** tem características muito semelhantes ao ACL, mas a interface é ligeiramente diferente. O IDEA centra-se primordialmente na auditoria, mas as versões recentes incluem um crescente número de técnicas de detecção de anomalias/fraude. A linguagem de programação do IDEA, designada *IDEAScript* é semelhante ao *Visual Basic* o que possibilita a automatização dos procedimentos, ou seja, é um instrumento de Análise de Dados abrangente, com bom interface com o utilizador, fácil de usar porque analisa rapidamente 100% dos dados, garantindo a integridade dos mesmos, abrindo caminho para auditorias mais rápidas e efetivas (Albrecht, Albrecht, Albrecht, & Zimbelman, 2016).

---

<sup>11</sup> [www.acl.com](http://www.acl.com)

- **Microsoft Office e ActiveData**, trata-se de um *plug-in* para o Microsoft Office que fornece procedimentos aperfeiçoados de Análise de Dados. Uma vez que se pode ser executado no Excel e no Access, herdou as capacidades nativas do *Office* permitindo consultas, relatórios, análises numéricas e utilização de linguagem *Visual Basic*. Esta solução é uma boa opção para os utilizadores que desejam uma interface familiar e uma alternativa menos dispendiosa do que o ACL e IDEA e que, acima de tudo, pretendam capitalizar os seus conhecimentos de Excel (Pedrosa, 2015).
- **TopCAATs**, suplemento para Excel que permite realizar a: Extração; *Monetary Unit Sampling* (“MUS”); Detecção de duplicados; Testes associados à Lei de Benford; *Join*, acrescentar, sumários, estatísticas; Evidências de diferenças entre folhas de cálculo; testes de sequências de numeração; extração dos principais registos; extração de duplicados; *Outliers*, entre outros. Este *software* inclui ferramentas de descoberta interativa para ajudá-lo a detetar anomalias e características incomuns nos dados a investigar (Pedrosa, 2015).
- O **Picalo** é semelhante ao IDEA e ACL, trata-se de um *software open source*, mas incorpora “detetores”, ou seja, pequenos *plug-ins* para detetar indicadores de fraude, ou seja, estes *plug-ins* descobrem um indicador específico, como a correspondência de endereços de fornecedores ou comparações de valores de faturas com a Lei de Benford.<sup>12</sup>

Existem ainda outros programas que, embora não sejam ferramentas específicas para análise de dados em auditoria, poderiam possibilitar realizar análises: SAS e SPSS (programas de análise estatística com módulos de fraude disponíveis); linguagens de programação tradicionais como Java, Perl, Python, Ruby, Visual Basic e outros; e uma série de pacotes especializados de *Data Mining* (W. Steve Albrecht et al., 2016).

Da análise do *Survey Report on Data Analysis Audit Software* publicado em 2012 dos questionários realizados a auditores, concluiu-se que quem têm *software* de Análise de Dados 64% tem ACL, 16% IDEA e 47% Microsoft Access (Kaplan, 2012). O mesmo estudo agora relativo a 2015, indica que não houve evolução relevante quanto a ferramentas para análise de dados: 60% das empresas dizem que possuem uma solução

---

<sup>12</sup> <http://www.iacae.org>

para Análise de Dados e, dessas, 39% indica que possui ACL, 18% IDEA, 36% Access e 70% Excel (Kaplan, 2016).

Em Portugal, um estudo realizado junto dos revisores oficiais de contas apresenta um cenário um pouco distinto já os auditores indicam que utilizam o Excel (100%), o IDEA (30%), o ACL (12%) sendo que apenas 7% indica utilizar ferramentas de *Data Mining* (Pedrosa, 2015).

### **1.3 Abordagem de Métodos Analíticos na Detecção de Anomalias**

Para aplicar e interpretar efetivamente os resultados utilizando métodos estatísticos tradicionais ou avançados, o auditor ou o investigador têm que ter um bom conhecimento sobre o ambiente do negócio, e têm que estar familiarizados com o software utilizado para a análise (Gee, 2012).

Para Gee, os métodos analíticos tradicionais de Análise de Dados incluem: extração; estatísticas; ordenação; duplicações; amostras; resumos; estratificação; junção; e comparações. Para realizar estas análises mais simples pode usar-se o Microsoft Excel ou Access (Gee, 2012).

Vários investigadores sugeriram o uso da análise de dígitos como uma ferramenta para que os auditores detetassem dados suspeitos (Wallace, 2002, Nigrini, 2011). A taxa de ocorrência real ou padrão de dígitos dentro dos dados é comparada à taxa de ocorrência ou padrão de dígitos hipotéticos para determinar se os dados podem conter dados suspeitos. Embora as técnicas não identifiquem quais pontos de dados são suspeitos, as técnicas podem alertar o auditor sobre a possível presença de dados suspeitos (Lowe, 2000).

Assim, enumeraram-se alguns dos testes que se podem utilizar para Análise de Dados a nível de ferramentas de auditoria, e que podem ser executados através do software de análise de dados.

#### **1.3.1 Lei de Benford**

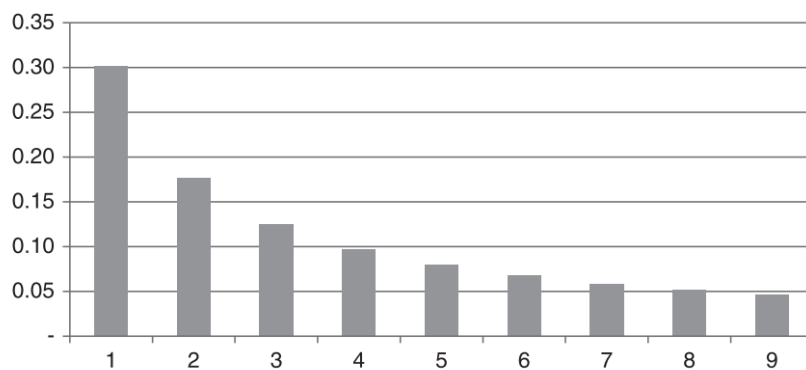
No final do século XIX, Simon Newcomb, astrónomo e matemático, verificou que as primeiras páginas dos livros de logaritmos que começavam com o dígito 1 se encontravam mais gastas. Newcomb concluiu das suas observações que existem mais números que começam com o número um, do que com números maiores, tais como os números oito ou nove (Durtschi, Hillison, & Pacini, 2004). Mais tarde, em 1920, Frank Benford ao

trabalhar como físico nos laboratórios de pesquisa da GE em Schenectady, Nova York, também observou que as primeiras páginas dos livros de tabelas de logaritmo estavam mais gastas do que as últimas, e com isso, Benford extrapolou que os registos de números com os primeiros dígitos mais baixos eram mais frequentes, porque havia mais números com baixos dígitos em primeiro lugar no mundo, ou seja, indicando que os números mais usados começam com o dígito 1 do que com o dígito 9 (Nigrini, 1999).

A Lei de Benford afirma que os dígitos e as sequências de dígitos num conjunto de dados seguem um padrão previsível, pelo que permite realizar análises de dígitos em dados numéricos, que ajudam a identificar anomalias como a manipulação sistemática de dados, fraude potencial e outras irregularidades. A Lei de Benford identifica duplicações incomuns ou excessivas de dígitos (Mark J. Nigrini, 2000).

Conforme observado na Figura 5, os números cujo primeiro dígito mais significativo é 1 devem ocorrer 30 por cento do tempo, enquanto os números que começam com o primeiro dígito de 9 só ocorrem com uma frequência de 4,6 por cento.

*Figura 5 - Lei de Benford distribuição de frequência do primeiro dígito*



Fonte: Baesens, Van Vlasselaer, & Verbeke (2015, p. 50)

M. J. Nigrini & L. I. Mittermaier (1997) descrevem as condições em que a Lei de Benford constitui um conjunto válido de frequências de dígitos esperados. Essas condições são resumidas da seguinte forma:

- A lista de números deve descrever fenômenos de dimensões semelhantes; e
- Os números não devem ter máximos ou mínimos internos; e
- Não devem ser atribuídos números usados para descrever elementos de um conjunto de dados (exemplo: número segurança social ou números de conta corrente).

A Lei de Benford faz parte de muitos planos de auditoria e é frequentemente utilizada pelos auditores, no entanto, nem sempre é bem interpretada, de modo que os resultados da aplicação da Lei de Benford devem ser o ponto de partida para o auditor (Gee, 2015). Uma análise mais complexa aplicando a lei de Benford a um conjunto de transações pode identificar situações de fraude (Hess & Cottrell, 2016). Embora a análise de Benford, por si só, não seja uma maneira "segura" de detetar a fraude, pode ser uma ferramenta útil para auxiliar na identificação de alguns itens para novos testes e, portanto, deve ajudar os auditores na indagação da deteção de fraude nas demonstrações financeiras (Durtschi et al., 2004).

### **1.3.2 Number Duplication Test ou Teste da Duplicação dos Números**

O *Number Duplication Test* (NDT) é um teste que pode ser usado para proporcionar mais informação decorrente dos testes da Lei de Benford. O NDT pode extrair números específicos que instigaram os picos no teste de primeira ordem (teste de dois dígitos é um exemplo) e no teste da soma (teste avançado da Lei de Benford). Os picos no teste de primeira ordem são causados por números que ocorrem com maior frequência do que o esperado, enquanto os picos no teste da soma geralmente devem-se a elevadas quantidades dos mesmos números que se repetem com mais frequência do que o normal (Gee, 2015).

O teste da soma foi desenvolvido por Nigrini (2011) que, ao simular uma distribuição de Benford, descobriu que as somas dos números pertencentes a cada grupo de primeiros dígitos 10, 11, 12,..., 99 resultavam em valores aproximadamente iguais, ou seja, 1/90. Porém, o autor verificou que os dados reais raramente se configuram com tal padrão, uma vez que eles possuem alguns valores muito elevados ou muitas repetições de números de valor médio (Nigrini, 2011).

O resultado do NDT é uma tabela em formato de relatório que exhibe: a classificação, o valor que foi duplicado, e a contagem para cada montante. O relatório mostra-nos quais os números que ocorreram com maior frequência e com que frequência ocorreram. A tabela seria ordenada por contagem decrescente, de modo que a quantidade que ocorreu mais frequentemente apareça primeiro (Nigrini, 2011). Este teste também pode detetar possíveis ineficiências de processamento, ou seja, quando uma pequena transação foi processada mais que uma vez. Permite a verificação da combinação entre os padrões e os acordos de crédito (correções e vazios) e os números positivos. Permite uma estreita

associação entre os números positivos (números de vendas) e anulados, corrigidos ou reembolsados ( Nigrini, 2011).O NDT fornece informações detalhadas para identificar picos do teste de soma e o primeiro teste de dois dígitos, o que ajuda a determinar se é necessário pesquisar mais (Gee, 2015).

### **1.3.3 Z-Score**

O termo Z-score, Z-value, Z-ratio ou Z é uma medida estatística de um número em relação à média do grupo de números. Refere-se a pontos ao longo da curva normal padronizada. O ponto central da curva tem um valor Z de 0. Os Z-value à direita de 0 são positivos e os Z-value à esquerda são valores negativos. Um Z-score está acima da média se for à direita de 0, e está abaixo da média se for à esquerda do ponto central 0. A distância da média é medida pelo desvio padrão. Se o Z-score for 0, é 0 desvios padrão da média e é igual à média. O Z-score é calculado pela diferença entre o número e a média, e dividindo a diferença obtida pelo desvio padrão.

$$Z = \frac{(X - \mu)}{\sigma}$$

X representa o número bruto,  $\mu$  representa a média da população e  $\sigma$  é o desvio padrão.

Um Z-score é padronizado, a unidade pode medir a altura, o peso, os níveis de educação ou as pontuações dos testes. O Z-score é sempre relativo à média que é o centro ou designado como zero. O Z-score diz muito sobre a distribuição dos números no seu conjunto de dados e pode destacar extremos. Com o Z-score, a área sob a curva normal pode ser determinada pelo cálculo do computador ou observando as tabelas conforme exemplificado na Tabela 2 - Tabela Parcial de Z-Score.

*Tabela 2 - Tabela Parcial de Z-Score*

<b>Z</b>	<b>Area between Mean and Z</b>
0.30	.1179
0.50	.1915
0.80	.2881
1.00	.3413
1.50	.4332
1.80	.4649
1.96	.4750
2.58	.4951
3.90	.4999

Fonte: Gee (2015, p. 82)

Uma pontuação Z de 1,50 indica que 43,32 por cento da área sob a curva normal está localizada entre a média e a pontuação Z de 1,50. Quanto maior o valor absoluto do Z-Score, mais o número está longe da média ou da norma. O auditor pode querer examinar transações que são *outliers* extremos.

#### **1.3.4 Relative Size Factor Test, RSF**

Para Nigrini (2011), o *Relative Size Factor Test* (RSF) identifica subconjuntos em que o maior valor está fora de linha com os outros montantes para esse subconjunto. Trata-se de um teste relevante para detetar erros. Ao existirem diferenças podem ficar a dever-se ao facto de o maior registo pertencer a outro subconjunto, ou de pertencer ao subconjunto em análise, mas o valor numérico estar incorretamente registado. Este teste foi desenvolvido em meados da década de 90, após uma empresa em Cleveland, ter transferido por engano para uma instituição de caridade \$ 600,000, que deveriam de ter sido para um fornecedor. Até então nunca tinha sido transferida uma quantia tão avultada para uma instituição de caridade, se a empresa tivesse executado um teste de razoabilidade antes de autorizar a transferência, teria percebido que o valor transferido era muito superior ao que tinha doado até à data, ou seja, o teste RSF identifica subconjuntos onde a maior quantidade é significativamente maior do que os outros itens no subconjunto.

$$RSF = \frac{MAIOR\ REGISTO\ NUM\ SUBCONJUNTO}{SEGUNDO\ MAIOR\ REGISTO\ NUM\ SUBCONJUNTO}$$

A desvantagem deste teste é que é bastante difícil de programar, porque um subconjunto com apenas um registo não pode ter um RSF, uma vez que não há um segundo número maior. Geralmente é uma boa ideia excluir todos os registos menores do que <1 ou <10, para evitar pequenos números que influenciam os resultados, por outro lado incluir números negativos poderia nos dar um RSF negativo, o que não faz sentido ou seja, as etapas na execução deste teste são para excluir os números pequenos e irrelevantes, e de seguida, excluir todos os subconjuntos com apenas um registo. Então, será necessário identificar o maior e o segundo maior número para cada subconjunto e de estabelecer uma regra que se aplique se o maior e o segundo maior número forem ambos iguais (Nigrini, 2011).

*Outliers* podem não ser os maiores montantes em todo o conjunto de dados, mas são elevados em relação aos elementos particulares do subconjunto. Grandes diferenças podem ser atribuídas a erros como o registo pertencente a outro subconjunto ou a quantidades que estão incorretamente atribuídas (por exemplo, ponto decimal deslocado). Grandes diferenças podem ser uma indicação de atividade fraudulenta, como fraude ocupacional de contas a pagar, faturas falsas para créditos de imposto, compensação de receitas de lavagem de dinheiro ou vendas de produtos para empresas relacionadas (preços de transferência offshore).

### **1.3.5 Same-Same-Same Test**

O objetivo do *Same-Same-Same* Test (SSS) é identificar duplicações anormais como indicadores potenciais de erros ou fraude (Gee, 2015). O SSS é o mais básico dos testes de duplicação dentro do subconjunto. O objetivo deste teste é identificar duplicações exatas. O teste é chamado de *Same-Same-Same*, independentemente de quantos campos são usados para determinar se os registos são duplicados (Nigrini, 2011).

Nigrini considera que o investigador forense tem necessariamente de ter alguma criatividade para identificar potenciais situações indiciadoras de fraude. Exemplos onde se pode aplicar este teste: os inventários terem o mesmo número de produto e a mesma quantidade disponível; relativamente a um cartão bancário poder haver casos em que o mesmo cartão tenha o mesmo valor cobrado mais do que uma vez na mesma data (o que geralmente acaba por ser uma compra dividida para manter o limite do controlo). Quem executa os testes de análise forense, deverá ter criatividade para identificar duplicações que são anomalias. O SSS foi usado para encontrar duplicações anormais em tabelas de dados relativas a: contas a pagar; pagamentos relativos a seguros de apólices de saúde, propriedade e acidentes; reembolsos a funcionários; reembolsos a clientes; inventários; ativos e extratos de pagamentos.

Nigrini considera fundamental testar os controlos, porque os pagamentos duplicados podem ser fraude, e é mais provável a sua ocorrência quando se inicia um novo ano fiscal, se altera o sistema de contas a pagar ou se adquire uma entidade, e se assume a responsabilidade pela função das contas a pagar após a data de aquisição (Nigrini, 2011).

Na Figura 6 é apresentado o resultado de um teste em que se identificou: os mesmos cartões, as mesmas datas, o mesmo fornecedor, e as mesmas quantidades. O campo mais à direita é a contagem e a maioria dos casos foi para duas compras idênticas.

Figura 6 - Exemplo de compras semelhantes na mesma data pelo mesmo titular do cartão



CardNumber	Date	MerchantName	Amount	PurchaseCount
5142189945	12/28/2010	GE REUTER-STOKES, INC	\$24,845.00	2
5142189945	9/18/2010	DELL MARKETING L.P.	\$23,130.00	2
5142121593	12/22/2010	DELL MARKETING L.P.	\$6,296.00	2
5142288601	9/11/2010	CHEM-TECH CONSULTING GR	\$3,144.00	2
5142189945	5/10/2010	LYNDE-ORDWAY CO, INC.	\$3,115.00	2
5142117663	8/18/2010	INORGANIC VENTURES INC	\$2,542.68	2
5142123687	10/13/2010	RETAIL DEBIT ADJUSTMENT	\$2,500.00	2
5142131721	5/11/2010	FBM COMPUTER & OFFICE SU	\$2,500.00	2
5142149042	3/13/2010	CROWN PLAZA HOTEL	\$2,500.00	4
5142153646	3/22/2010	INFORMATION MAPPING	\$2,500.00	4
5142251068	4/11/2010	COLE INDUSTRIAL INC	\$2,500.00	2
5142251068	4/24/2010	COLE INDUSTRIAL INC	\$2,500.00	2
5142257356	2/28/2010	BFS SOLUTIONS	\$2,500.00	2
5142122592	8/3/2010	PC *PC CONNECTION FED	\$2,499.00	2
5142205500	4/5/2010	MICRON GOVERNMENT SYSTE	\$2,499.00	2
5142196337	7/22/2010	MWI*MICRO WAREHOUSE	\$2,494.90	2
5142221788	9/12/2010	DELL MARKETING L.P.	\$2,494.00	2
5142289869	6/7/2010	LABSOURCE INC	\$2,492.00	3
5142196337	2/16/2010	DELL MARKETING L.P.	\$2,490.00	2
5142278415	11/15/2010	COMBINED DIGITAL	\$2,485.00	6

Fonte: Mark J Nigrini (2011, p. 447)

Resumindo, o teste SSS permite identificar registos que contêm campos com dados que são duplicações exatas de outros registos e duplicações anormais que sejam indicadores de potenciais erros ou fraude, tais como: requisições de pagamentos em duplicado; pagamentos duplicados a fornecedores; pagamentos duplicados de planos de saúde públicos ou privados.

### 1.3.6 Same-Same-Different Test

O *Same-Same-Different Test* (SSD) é usado para identificar registos duplicados semelhantes nos campos selecionados pelo auditor (Gee, 2015). Este teste é preponderante para detetar erros e fraude, deve ser considerado para os projetos de análise forense. Numa recente investigação de compra de cartões numa empresa de serviços públicos, este teste evidenciou vários casos em que dois funcionários dividiam a mesma compra usando cartões diferentes, o que se manifestou uma forma de dividir a compra e colocar a cobrança no mesmo cartão (Nigrini, 2011).

O teste usual é executado de modo a que o campo diferente seja um campo do subconjunto. Pretendem-se transações que estão ligadas a dois subconjuntos diferentes. O pressuposto é que uma das transações é um erro e não deve ter sido vinculada ao

segundo (subconjunto diferente). O relatório SSD usual é ter cada caso correspondente em duas linhas na tabela de resultados (Gee, 2015).

### **1.3.7 Even Amount/Números Redondos**

Números pares ou arredondados normalmente não ocorrem com uma elevada taxa de frequência. Por conseguinte, os números que são arredondados para dezenas, centenas ou milhares podem ser considerados anomalias, e, como tal, deve ser-lhe ser dada especial atenção. Relativamente a faturas a pagamento, geralmente dá-se mais destaque às que têm valores mais altos, mas as de valores baixos, também são suscetíveis a fraude. A título de exemplo, quando se fixam tetos máximos para a realização de determinadas despesas (almoços, jantares, alojamento), para garantir que não há abuso, deve-se verificar tanto os valores como as quantidades (Gee, 2015).

Quem comete fraude, geralmente cria faturas com valores arredondados, ou seja, faturas sem cêntimos. Um modo fácil de identificar faturas com montantes arredondados é usar a função MOD no Excel ou o operador % no IDEA, que permite calcular o resto da divisão inteira de um número por outro (neste caso, poderá ser por 10 ou por 100 para detetar números redondos nas dezenas ou nas centenas, por exemplo), pois permite identificar os fornecedores que têm uma elevada percentagem de faturas com montantes arredondados. Para tal, basta calcular o número de faturas de cada fornecedor com valores arredondados e dividi-lo pelo número total de faturas para esse fornecedor, obtendo a percentagem, seguidamente, ordena-se por ordem decrescente para destacar os fornecedores mais suspeitos (Warner, 2017).

É necessária cautela pois, normalmente, existem muitos montantes que são arredondados, tais como trabalhos de consultoria ou rendas, sendo anomalias de baixo risco (Gee, 2015).

## **2 Metodologia do Estudo**

O estudo de caso representa um processo de análise de dados reais, para proporcionar uma análise sobre determinado tema compreensiva, sistemática e profunda (PATTON, 2002). Os estudos de casos representam a metodologia preferida quando o foco do estudo se encontra em ocorrências contemporâneas inseridas num contexto real (YIN, 2005). O estudo empírico incidirá sobre os elementos relativos à faturação e tentará determinar a eficácia da aplicação de vários testes como uma ferramenta de investigação na determinação de elementos suspeitos que podem ser indiciadores de anomalias/fraude.

Para abordar a potencial relação entre a anomalias/fraude e os métodos analíticos na análise de dados, é necessário rever a literatura relevante. A literatura consultada neste estudo incluiu: livros, revistas, teses de mestrado e doutoramento, artigos baseados na web e outras publicações.

Para o desenvolvimento deste trabalho, torna-se fundamental fazer uma breve caracterização da empresa que será objeto de aplicação de métodos analíticos, para deteção de anomalias/fraude, nomeadamente dos fatores que podem influenciar os resultados obtidos da utilização dos métodos analíticos descritos. Desde já se ressalva o facto de, ao longo do trabalho, se garantir o anonimato da empresa em estudo, optando por, ainda que ocasionalmente, ocultar ou rasurar elementos que, de alguma forma, possam expor a identidade da mesma.

Posteriormente, será descrito como é que se procedeu à recolha dos dados a qual foi realizada com recurso ao Caseware IDEA e ao Microsoft Excel.

### **2.1 Caracterização da Empresa e dos Procedimentos Relativos às Vendas**

A empresa *TADIMA LDA*<sup>13</sup> foi fundada na década de noventa, tem sede em Portugal continental, tendo instalações compostas por um armazém, um escritório e uma pequena loja aberta ao público em geral. Trata-se de uma empresa de cariz comercial que se dedica ao comércio por grosso de alimentos e de alguns acessórios para animais. Em relação aos

---

<sup>13</sup> Nome fictício para salvaguardar o anonimato da empresa.

acessórios podem ser necessárias intervenções de manutenção, o que corresponde a um serviço prestado pela empresa.

Relativamente à atividade da empresa e aos procedimentos de logística das mercadorias (artigos/produtos), os artigos são rececionados no armazém, e, posteriormente, expedidos para os clientes, sendo que a expedição pode ser de duas formas: por transportadora ou à responsabilidade do cliente quando a venda é executada na loja da empresa (o volume de vendas na loja representa aproximadamente 3% do volume total das vendas nacionais e na sua maioria destinam-se a consumidores finais). Também pode existir a situação em que as mercadorias não passam pelo armazém, ou seja, são expedidas diretamente das instalações do fornecedor para as dos clientes, tal apenas acontece no caso dos clientes que compram em grandes quantidades (volumes acima de 20 000 kg que corresponde a contentores ou camiões completos, em que o custo de transporte da mercadoria é suportado pelo cliente. Este tipo de clientes tem um peso significativo no volume de vendas da empresa e, normalmente, são representantes da marca vendida pela empresa TADIMA LDA.

A empresa tem clientes nacionais e estrangeiros, sendo estes últimos referentes maioritariamente ao mercado espanhol. A gama de clientes a nível Nacional engloba o consumidor final, pequenos e grandes retalhistas. No que concerne ao mercado Espanhol, a empresa trabalha com importadores que representam a marca por regiões, tratando-se de venda a retalho por grosso.

Na Tabela 3 está patente a evolução do volume de negócios da empresa nos últimos 3 anos, pelo que se depreende que as vendas têm tido uma evolução notória, sobretudo no mercado de externo.

*Tabela 3 - Evolução das Vendas da empresa TADIMA LDA*

<b>Vendas + Prestação de Serviços</b>	2016	2015	2014
Mercado Nacional	5 058 086,09	4 770 945,23	4 744 149,52
Mercado Exportação	4 537 917,91	4 195 574,33	3 787 418,38
<b>Total</b>	<b>9 596 004,00</b>	<b>8 966 519,56</b>	<b>8 531 567,90</b>

**Fonte:** Adaptado do Relatório e Contas da empresa TADIMA LDA

No que concerne aos recursos humanos afetos à empresa, a TADIMA tem 21 colaboradores, na faixa etária dos 21 aos 51 anos, sendo 13 do sexo masculino e 8 do feminino, conforme patente na Figura 7.

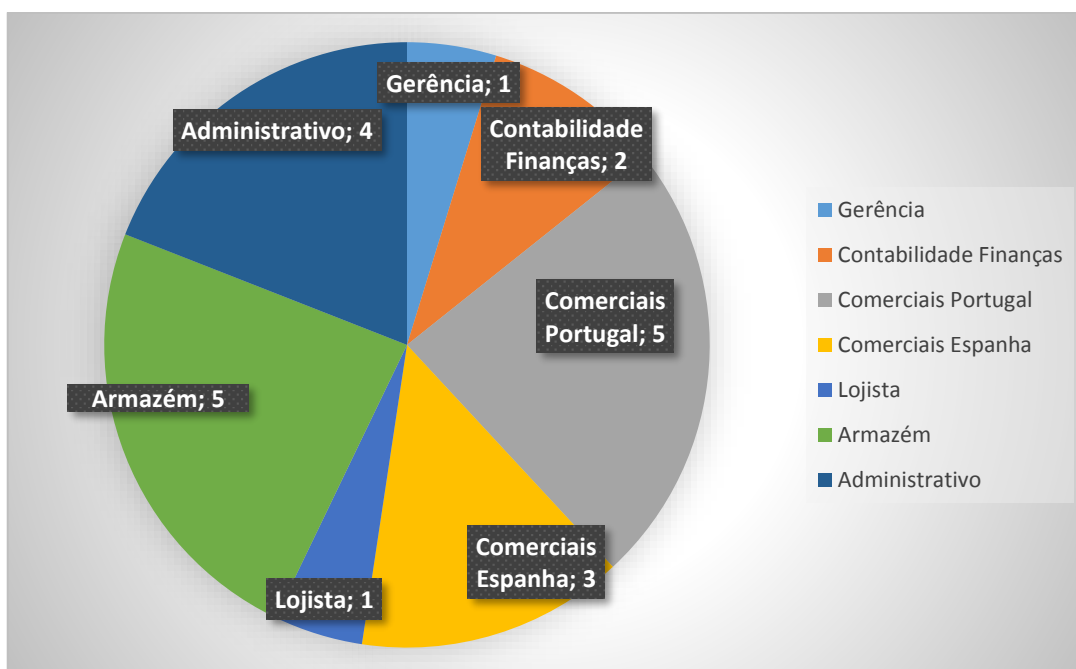
*Figura 7 - Recursos Humanos por género na empresa TADIMA LDA*



Fonte: Elaboração Própria

Na Figura 8 está representada a distribuição dos recursos humanos por categorias: Gerência – 1; Administrativo – 4; Contabilidade/financeira - 2; Vendedores/Comerciais Portugal - 5; Vendedores/Comerciais Estrangeiros – 3; Lojista – 1; Operadores de armazém – 5.

*Figura 8 - Número de Funcionários por Departamento empresa TADIMA LDA*



Fonte: Elaboração Própria

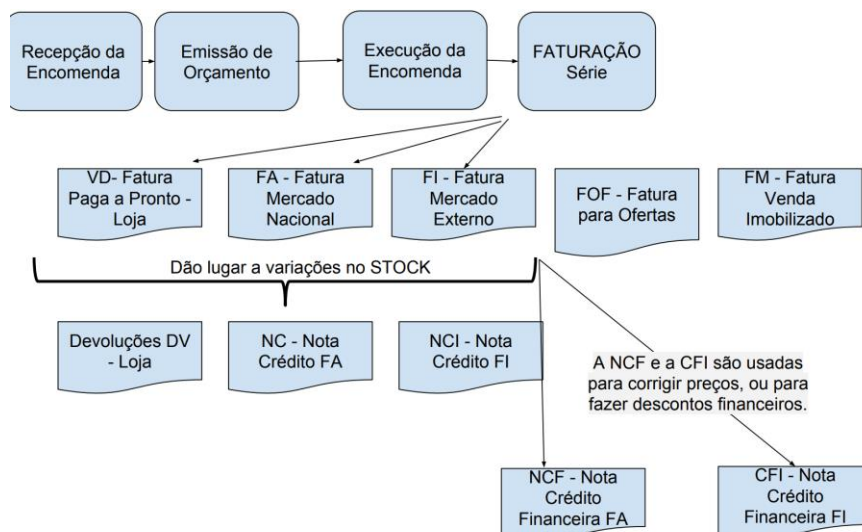
Em termos de dimensão, estamos perante uma pequena empresa, conforme definido nos critérios de atribuição de PME da Comissão Europeia (Comissão Europeia, 2006, p. 14).

A análise de dados do estudo vai incidir sobre os produtos vendidos pela empresa *TADIMA LDA*, e que vão ser extraídos do ficheiro SAFT.

O processo de vendas da empresa encontra-se resumido na Figura 9: o departamento de faturação recebe as encomendas diretamente dos clientes ou por intermédio dos vendedores/comerciais, após receção da encomenda, é elaborado um documento chamado orçamento, que é enviado para o armazém. Os funcionários, à medida que vão elaborando a encomenda, colocam um visto nos artigos disponíveis para entrega. Depois de pronta a encomenda e de conferência dos elementos constantes do orçamento com visto, o orçamento é convertido num documento fiscal, mais concretamente em Fatura (existindo várias séries de faturas: VD; FA; FI; FOF; e FM) sendo no mesmo dia expedida a mercadoria para o cliente.

Posteriormente a este processo, pode acontecer que ou os funcionários que emitem faturas se enganaram a digitar os códigos dos produtos encomendados pelos clientes, ou os funcionários do armazém que enviaram os artigos errados, pelo que se emitem notas de crédito - NC (séries de NC: DV; NC; NCI), para corrigir as possíveis situações anómalas. A empresa emite também notas de crédito financeiras, quando existe necessidade de corrigir preços ou fazer descontos financeiros (séries: NCF e CFI que mais à frente serão referenciadas).

*Figura 9 - Processo de emissão de documentos de faturação na empresa TADIMA LDA*



A nível de devoluções podemos ter duas situações:

1. A mercadoria é devolvida à empresa, isto quando o custo de devolver é inferior ao valor da mercadoria; - Documentos: DV; NC; NCI.
2. A mercadoria não é devolvida à empresa, porque o custo de transporte é superior ao valor da mercadoria: NCF; CFI.

**Fonte:** Elaboração própria

## **2.2 Objetivos Específicos**

Pretende-se que a realização deste estudo seja uma mais-valia para a organização, na tentativa de tentar discernir um conjunto de situações que podem ser consideradas como anomalias/erros ou mesmo fraude, e assim contribuir para o aperfeiçoamento e implementação de melhorias no controlo interno.

Neste contexto, os objetivos da investigação passam por:

1. Verificar se existem anomalias ao nível do montante total das faturas emitidas, através da análise do valor do GROSSTOTAL.
2. Identificar se existem situações anómalas em relação aos produtos oferecidos, através da análise do universo das amostras, bónus e produtos oferecidos: o procedimento associado às ofertas está a ser cumprido? Os produtos que estão a ser oferecidos estão de acordo com o definido? São oferecidos mais do que uma vez os mesmos produtos ao mesmo cliente? A oferta estimula a venda? Ou seja, os clientes a quem os produtos estão a ser oferecidos, compram posteriormente o produto?

## **2.3 Recolha e Compreensão dos Dados**

### **2.3.1 Notas Introdutórias**

A diferença entre população e amostra não é por vezes clara para alguns investigadores, tornando ainda mais difícil esclarecer e decidir os métodos de amostragem a aplicar. A população teórica (ou universo estatístico) corresponde assim ao conjunto de todos os indivíduos com uma característica comum que se submete a um estudo estatístico. A amostra é um subconjunto representativo da população, criteriosamente escolhido, sobre o qual incide a observação<sup>14</sup>.

No caso em estudo, para atingir os objetivos definidos, será analisado o conjunto de dados fornecidos pela empresa, extraídos do ficheiro SAFT e correspondentes ao primeiro semestre de 2017 (janeiro a junho), tendo-se optado por analisar toda a informação

---

<sup>14</sup> <http://analise-estatistica.pt/inicio/2015/03/diferenca-entre-populacao-e-amostra.html>

contida no ficheiro. Numa primeira fase, utilizou-se a versão educação do CaseWare IDEA (a qual suporta um máximo de 10000 registos por ficheiro). Posteriormente, foi solicitada à CaseWare IDEA Inc<sup>15</sup> uma licença completa do IDEA, exclusivamente para este estudo, sem limite de registos, a custo zero e por um período de 2 meses. O pedido foi prontamente acolhido.

O ficheiro SAFT é um ficheiro que, num formato normalizado XML (*Extensible Markup Language*), reúne toda a documentação fiscalmente relevante de uma entidade relativa a um determinado período de tempo (mensal), tendo sido originalmente concebido para efeitos de auditoria fiscal. Este ficheiro engloba: a identificação da empresa, dos clientes, dos produtos vendidos, os documentos de venda fiscalmente relevantes, mesmo que os mesmos tenham sido anulados, diversas séries de faturas, as notas de débito e de crédito. Em Portugal, os sistemas de faturação e de contabilidade passaram a ser desenvolvidos com o propósito de exportar o ficheiro SAFT (PT), o que permite extrair informação normalizada independentemente do *software* de gestão do cliente de auditoria (Correia, 2017).

Como forma de apresentar o presente trabalho e por uma questão de legibilidade, optou-se por colocar sempre os nomes dos campos a negrito, os nomes dos ficheiros a itálico+negrito e os critérios de extração numa linha separada, em itálico e com tipo de letra diferente.

### **2.3.2 Processo de Tratamento dos Dados**

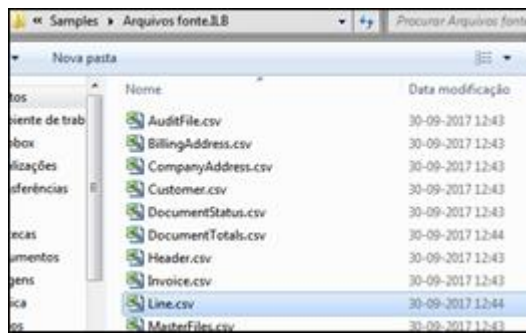
O processo de tratamento dos dados, para posteriormente se poder proceder às análises adequadas, passou pelas seguintes etapas (alguns dos passos são ilustrados para facilitar a compreensão):

- ✓ Exportação do ficheiro SAFT do Programa de faturação;
- ✓ Conversão do ficheiro XML para CSV (através da aplicação XML Converter);
- ✓ Arquivo dos ficheiros na pasta Samples>Arquivos Fonte (Figura 10);

---

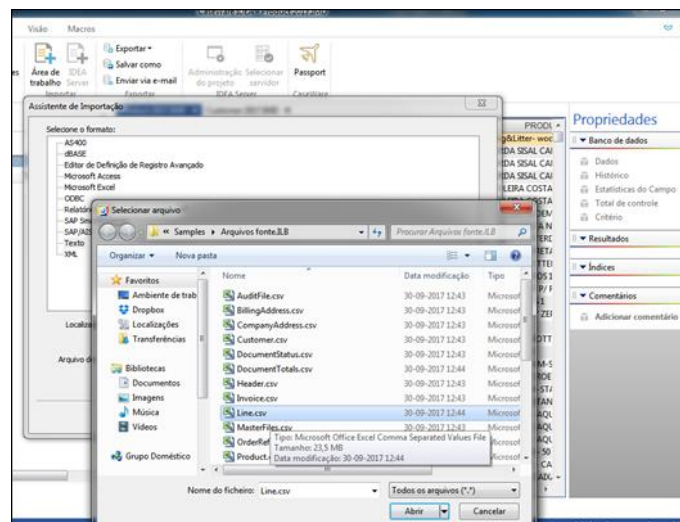
<sup>15</sup> [www.casewareanalytics.com](http://www.casewareanalytics.com)

*Figura 10 - Arquivo de Ficheiros*



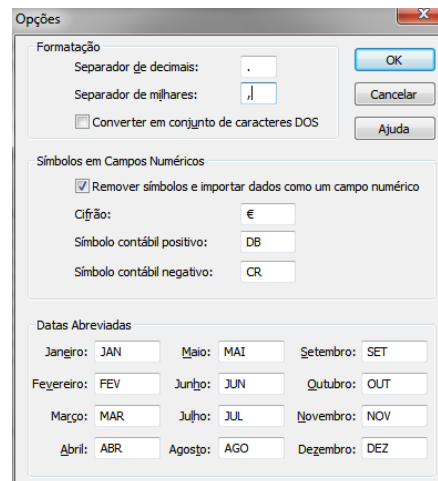
- ✓ Importação dos ficheiros considerados relevantes para o CaseWare IDEA, em que foi dada relevância ao tratamento dos dados integrantes dos campos (Figura 11).

*Figura 11 - Importação de ficheiros do SAFT*



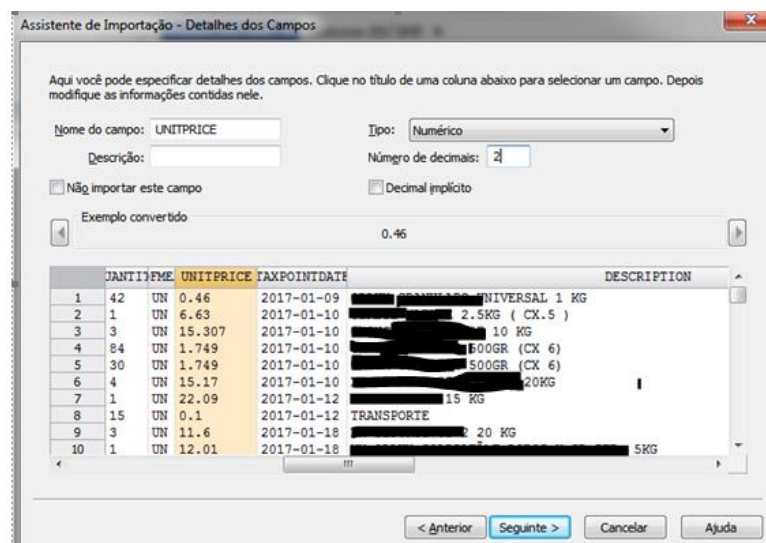
A título exemplificativo, um dos cuidados a ter, ao nível da formatação dos campos numéricos, é a formatação e separação das casas decimais e o separador de milhares, para que os dados sejam importados de modo a poderem ser analisados de forma correta, conforme Figura 12. É necessário verificar se o separador de milhares que se encontra na parametrização da importação é o que consta do ficheiro CSV a importar. O mesmo terá de se verificar para o separador das casas decimais.

*Figura 12 - Formatação de campos na importação*



Na Figura 13 (relativamente a este exemplo foi rasurado a descrição dos artigos, por uma questão de confidencialidade) pode visualizar-se outro exemplo, neste caso relativamente ao campo **UNITPRICE**, que é um campo do tipo Numérico, com 2 casas decimais e que terá de se confirmar se está a ser considerado dessa forma. É também, neste passo que se tem a opção de não importar campos que não sejam relevantes para a análise (fez-se uma exclusão dos que não eram necessários, nomeadamente, no ficheiro invoice: movement\_end\_time; movement\_end\_time; hash\_control; todos os campos que podem ser observados através do ficheiro SAFT podem ser analisados no Apêndice 1).

*Figura 13 - Importação de dados - Tipos de campos*



Na Tabela 1, está patente alguma informação pertinente relativa à informação que está incluída no SAFT no campo **INVOICETYPE** e **INVOICENO**, que dizem respeito ao tipo de documentos fiscais emitidos durante o processo de venda. Estes podem ser desdobrados em FT - Faturas ou NC - Notas de crédito, o **INVOICENO** diz respeito às

*Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro  
SAFT-PT: O caso de uma empresa comercial*

diferentes séries que foram criadas pela empresa para diferenciar as operações. No presente caso, e como já referido anteriormente, o tipo de documento diz respeito às diferentes séries de documentos fiscais que são emitidos pela empresa, e que foram criados pela entidade como critério para diferenciar mais facilmente o tipo de documento que está a ser emitido. O **INVOICENO** identifica a série, o ano e o número de documento. Na referida tabela está uma breve descrição sobre a que se refere o tipo de documento e do tipo de clientes que se pode incluir.

*Tabela 4 - Relação entre os dados do SAFT e os Procedimentos internos da empresa*

<b>INVOICETYPE</b>	<b>Tipo de Documento</b>	<b>INVOICENO</b>	<b>Breve Descrição</b>	<b>Tipo de Clientes</b>
<b>FT</b>	<i>Faturas VD</i>	<b>VD 2017/</b>	<i>Registo de vendas quando o cliente paga a pronto.</i>	<i>Cliente com NIF ou Consumidor Final</i>
<b>FT</b>	<i>Faturas FA</i>	<b>FA 2017/</b>	<i>Registo de vendas quando o cliente é nacional e não paga a pronto.</i>	<i>Cliente com NIF</i>
<b>FT</b>	<i>Faturas FI</i>	<b>FI 2017/</b>	<i>Registo de vendas quando o cliente é estrangeiro e não paga a pronto.</i>	<i>Cliente com NIF</i>
<b>FT</b>	<i>Faturas FOF</i>	<b>FOF 2017/</b>	<i>Registos de vendas em que os artigos são 100% oferta a clientes.</i>	<i>Cliente com NIF</i>
<b>NC</b>	<i>Notas Crédito DV</i>	<b>DV 2017/</b>	<i>Devolução de vendas que são correspondentes com as Faturas VD.</i>	<i>Cliente com NIF ou Consumidor Final</i>
<b>NC</b>	<i>Notas de Crédito NC</i>	<b>NC 2017/</b>	<i>Notas de Crédito NC – Notas de devolução referentes às Faturas FA</i>	<i>Cliente com NIF</i>
<b>NC</b>	<i>Notas de Crédito NCI</i>	<b>NCI 2017/</b>	<i>Notas de Crédito NCI – Notas de devolução relativas às Faturas FI</i>	<i>Cliente com NIF</i>
<b>NC</b>	<i>Notas de Crédito Financeiras NCF</i>	<b>NCF 2017/</b>	<i>Notas de Crédito Financeiras NCF - Clientes Nacionais</i>	<i>Cliente com NIF</i>
<b>NC</b>	<i>Notas de Crédito Financeiras CFI</i>	<b>CFI 2017/</b>	<i>Notas de Crédito Financeiras CFI - Clientes Estrangeiros</i>	<i>Cliente com NIF</i>

Fonte: Elaboração Própria

Do ficheiro SAFT foram importados os seguintes ficheiros: **Customer 2017**; **Invoice 2017**; **Line 2017**; e **Product 2017**. Para além dos elementos extraídos do ficheiro SAFT, acrescentou-se mais dois ficheiros: **Clientes** e **CLIENTES\_VENDEDORES** que contém todos os clientes e vendedores da base de dados da empresa TADIMA LDA, uma vez que o ficheiro SAFT não contém essa informação.

Figura 14 - Ficheiros utilizados na análise de dados da empresa TADIMA LDA

+	Clientes	4.507
+	CLIENTES_VENDEDORES	13
+	Customer 2017	1.517
+	Invoice 2017	11.671
+	Line 2017	132.430
+	Product 2017	2.840

A Figura 14 correspondente ao ambiente de trabalho do IDEA, está patente o conjunto de ficheiros originais do estudo, já que, à medida que se forem realizando testes, esse número irá aumentar já que a maioria das operações geram ficheiros adicionais. À esquerda consta o nome dos ficheiros e à direita o número de linhas que contém cada ficheiro. Assim, de seguida far-se-á uma breve caracterização de cada ficheiro.

O ficheiro **Clientes** tem 4507 linhas que correspondem aos clientes, incluindo todos campos que os caracterizam, os quais são apresentados na Figura 15: CLIENTE é o número de identificação interno do cliente da base de dados da empresa TADIMA LDA, o NOME, a MORADA, o N\_º\_CONTRIBUINTE, o número de VENDEDOR e o NOME\_DO\_VENDEDOR que se encontra associado a cada cliente. Cada cliente tem um vendedor atribuído.

Figura 15 - Campos do ficheiro Clientes

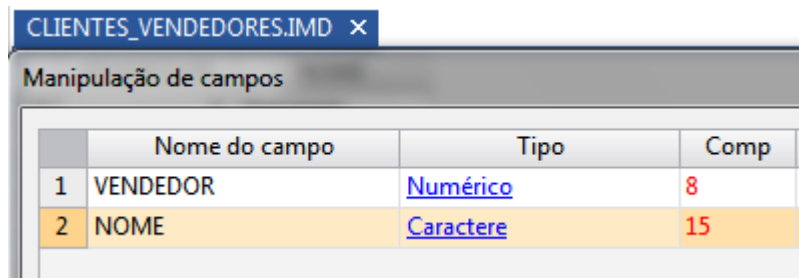


	Nome do campo	Tipo	Comp
1	CLIENTE	Caractere	13
2	NOME	Caractere	50
3	MORADA	Caractere	50
4	N_º_CONTRIBUINTE	Caractere	16
5	VENDEDOR	Caractere	8
6	NOME_DO_VENDEDOR	Caractere	15

O ficheiro **CLIENTES\_VENDEDORES** tem 13 vendedores, ressaltando-se o facto de existir uma diferença entre o n.º de funcionários afetos aos recursos humanos com a

categoria de vendedores/comerciais e o n.º de vendedores mencionado no ficheiro. A situação é elucidada porque a loja e o armazém têm clientes atribuídos, aqui também registados como vendedores, existindo também clientes que foram angariados por outros funcionários que não os vendedores/comerciais. Os campos afetos a este ficheiro estão na Figura 16, nomeadamente o número de **VENDEDOR** e **NOME** do vendedor.

*Figura 16 - Campos do ficheiro CLIENTES\_VENDEDORES*



	Nome do campo	Tipo	Comp	C
1	VENDEDOR	Numérico	8	C
2	NOME	Caractere	15	

O ficheiro *Customer 2017* é composto por 1517 linhas, representando cada linha os diferentes clientes a quem foram emitidas INVOICES no período em estudo. Os campos que integram este ficheiro estão discriminados na Figura 17: o campo **CUSTOMERID** é o campo chave, correspondendo à identificação única do cliente e permite estabelecer união com o ficheiro *Invoice 2017* onde o campo também existe. O campo **ACCOUNTID** contém dados sobre a identificação dos clientes na contabilidade da entidade, sendo que, em alguns casos, se encontra preenchido com o valor “Desconhecido”. Assim, foi necessário criar o campo **ID\_CLIENTE\_N** com o seguinte parâmetro, de modo a ser possível, posteriormente, estabelecer ligação entre a tabela *Customer 2017* e *Clientes* e relacionar com os vendedores da empresa, teve que se criar este novo campo<sup>16</sup>

`@if( ACCOUNTID == "Desconhecido";"00000";@Right( ACCOUNTID ;5))`

---

<sup>16</sup> O ficheiro *Customer 2017*, o campo **ACCOUNTID** é representado sempre nos primeiros cinco dígitos por 21111 e os restantes cinco dígitos são o número de identificação do **CLIENTE** no ficheiro *Clientes*, existindo uma exceção que é o **ACCOUNTID** = desconhecido que é representa os consumidores finais, o **ID\_CLIENTE\_N** será igual a 00000, exemplificando **ACCOUNTID**=2111100001, o **ID\_CLIENTE\_N** será igual a 00001. Os campos que integram este ficheiro são aqueles que permitem identificar o cliente perante a administração fiscal e no cruzamento de informação com a contabilidade da empresa.

*Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro SAFT-PT: O caso de uma empresa comercial*

Figura 17 - Campos do ficheiro Customer 2017

	Nome do campo	Tipo	Comp	Dec	Parâmetro	Descrição
1	CUSTOMERID	Caractere	18			
2	ACCOUNTID	Caractere	14			
3	CUSTOMERTAXID	Caractere	12			
4	COMPANYNAME	Caractere	55			
5	ADDRESSDETAIL	Caractere	55			
6	CITY	Caractere	33			
7	POSTALCODE	Caractere	14			
8	COUNTRY	Caractere	14			
9	REGION	Caractere	32			
10	ID_CLIENTE_N	Caractere virtual	5		@!f(ACCOUNTIC...	

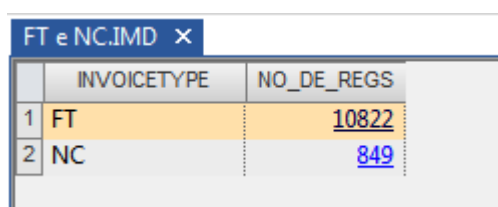
A base de dados **Invoice 2017** tem 11671 linhas que representam o número de FT (Faturas) e de NC (Notas de Crédito) emitidas, os campos que a compõem estão evidenciados na Figura 18. O campo **INVOICE\_ID** é o campo chave que identifica de forma única um documento e que permitirá depois estabelecer a junção com o ficheiro **Line 2017** (que será apresentado em seguida), através do campo **CUSTOMERID** (campo chave para ligar com **Cientes 2017** e se obter, associado a cada documento emitido, toda a informação dos clientes.

Figura 18 - Campos do ficheiro Invoice 2017

	Nome do campo	Tipo	Comp	Dec
1	INVOICE_ID	Caractere	7	
2	INVOICENO	Caractere	14	
3	INVOICESTATUSDATE	Caractere	21	
4	SOURCEID	Caractere	11	
5	SOURCEBILLING	Caractere	3	
6	PERIOD	Caractere	4	
7	INVOICEDATE	Data	8	
8	INVOICETYPE	Caractere	4	
9	CUSTOMERID	Caractere	18	
10	ADDRESSDETAIL	Caractere	55	
11	CITY	Caractere	33	
12	POSTALCODE	Caractere	14	
13	COUNTRY	Caractere	14	
14	LINENUMBER	Caractere	4	
15	PRODUCTCODE	Caractere	15	
16	PRODUCTDESCRIPTION	Caractere	67	
17	QUANTITY	Numérico	8	0
18	UNITPRICE	Numérico	8	2
19	DEBITAMOUNT	Numérico	8	2
20	TAXCOUNTRYREGION	Caractere	7	
21	TAXPERCENTAGE	Numérico	8	2
22	SETTLEMENTAMOUNT	Numérico	8	2
23	TAXPAYABLE	Numérico	8	2
24	NETTOTAL	Numérico	8	2
25	GROSSTOTAL	Numérico	8	2

Através do IDEA e da opção Sumário do campo **INVOICETYPE**, obteve-se um resumo do número de Faturas - FT e de Notas de Crédito – NC que foram emitidas no período em análise, pelo que através da Figura 19 pode-se concluir que foram emitidas 10822 FT e 849 NC, representando um total de 6017170,33€ (conforme será apresentado mais à frente na análise das estatísticas de campo). Este valor poderá ser validado com o que se encontra disponível no SAFT e que corresponde aos dados ao valor que existe no ficheiro **Invoice 2017**.

Figura 19 - Número de Faturas (FT) e de Notas de Crédito (NC)



	INVOICETYPE	NO_DE_REGS
1	FT	10822
2	NC	849

O ficheiro **Line 2017** tem 132430 linhas que correspondem a 132430 itens transacionados. A Figura 20 revela os campos que compõem o ficheiro **Line 2017**. O campo chave é o **INVOICE\_ID** que, par além de permitir identificar de forma única um documento, permite ainda que se possa executar a junção com o ficheiro **Invoice2017** de modo a conhecer todos os dados de Invoice associados a cada uma das linhas do documento. Um cruzamento com o ficheiro **Customer 2017** possibilita, ainda, saber os detalhes do cliente associado ao documento emitido (FT ou NC).

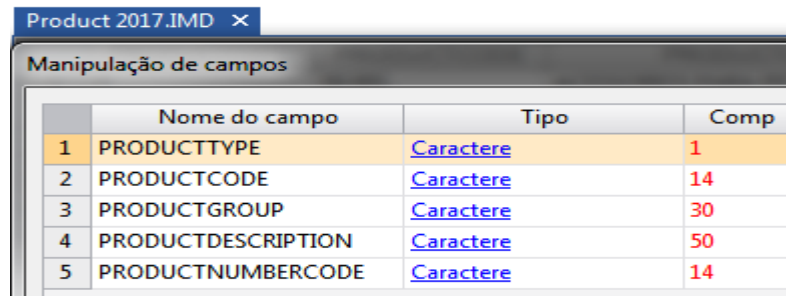
Figura 20 - Campos do ficheiro Line 2017



	Nome do campo	Tipo	Comp	Dec	Parâmetro
1	LINE_ID	Caractere	8		
2	INVOICE_ID	Caractere	7		
3	LINENUMBER	Caractere	5		
4	PRODUCTCODE	Caractere	16		
5	PRODUCTDESCRIPTION	Caractere	94		
6	QUANTITY	Numérico	8	0	
7	UNITPRICE	Numérico	8	2	
8	TAXPOINTDATE	Data	8		AAAAMMDD
9	DEBITAMOUNT	Numérico	8	2	
10	TAXPERCENTAGE	Numérico	8	2	
11	SETTLEMENTAMOUNT	Numérico	8	2	
12	CREDITAMOUNT	Numérico	8	2	

O ficheiro **Product 2017** integra 2840 linhas que representam o número de diferentes artigos que foram transacionados no período em causa. Os campos que compõem este ficheiro são relativos às características e descrição mais pormenorizada dos produtos, conforme se pode deprender pela Figura 21.

Figura 21 - Campos do ficheiro Product 2017



	Nome do campo	Tipo	Comp
1	PRODUCTTYPE	Caractere	1
2	PRODUCTCODE	Caractere	14
3	PRODUCTGROUP	Caractere	30
4	PRODUCTDESCRIPTION	Caractere	50
5	PRODUCTNUMBERCODE	Caractere	14

Por forma a melhorar a interpretação da informação, através do ficheiro **Invoice2017**, executou-se a extração das **INVOICETYPE="FT"** e das **INVOICETYPE="NC"**, criando-se dois novos ficheiros: **FT\_ICV** e **NC\_ICV**. Esta separação foi elaborada para evitar deturpar a análise relativa, por exemplo, às quantidades de produtos oferecidos: o valor das devoluções é positivo, e sempre que se verifica um erro na emissão de uma fatura (FT), esse erro será compensado pela emissão de uma nota de crédito (NC).

Nas propriedades do ficheiro, no banco de , pode seleccionar-se as estatísticas de campo e visualizar informação sobre os dados apresentados. Na Figura 22, e considerando apenas as faturas (FT\_ICV), o campo **GROSSTOTAL** é de 6 017 170,33€, conforme verificado anteriormente, o número de INVOICES tipo FT registadas é de 10822, o valor médio das INVOICES FT foi de 556,01€, o valor máximo de uma INVOICE FT foi de 36346,34€ e o valor mínimo foi 0,00€.

# Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro SAFT-PT: O caso de uma empresa comercial

Figura 22 - Estatísticas de Campo – Tipo de campo: numérico – Ficheiro: FT\_ICV

Estatísticas Numéricas		GROSSTOTAL
▶ Valor líquido		6.017.170,33
Valor absoluto		6.017.170,33
No. de registos		10.822
No. registos zerados		32
Valor positivo		6.017.170,33
Valor negativo		0,00
No. de registos posi...		10.790
No. de registos negat...		0
No. de dados inválidos		0
No. de valores válidos		10.822
Valor médio		556,01
Valor mínimo		0,00
Valor máximo		36.346,34
No. reg do mín.		115
No. reg do máx.		7.776
Desvio padrão da am...		2.017,48
Variância da amostra		4.070.221,76
Desvio padrão da po...		2.017,39
Variância da população		4.069.845,66
Assimetria da populaç...		9.940380
Curtose da população		117.106270

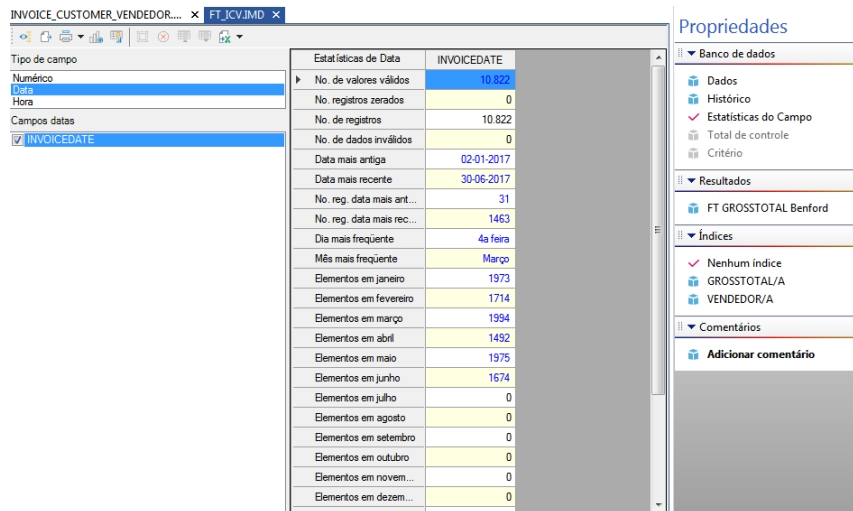
As FT com valor 0,00 podem ser explicadas pelos documentos que são restritos a ofertas ou a patrocínios em produtos. Assim e tal como pode ser visualizado na Figura 23, onde se efetuou um filtro para extrair somente as INVOICES cujo **GROSSTOTAL** tem o valor 0,00€, correspondendo a 32 registos, revelando a extração que as normas da empresa são seguidas e não existem outro tipo de INVOICES com o valor 0,00€.

Figura 23 - INVOICES FT - FOF INVOICES com GROSSTOTAL 0,00€

INVOICE_ID	INVOICENO	PERIOD	INVOICEDATE	INVOICETYPE	LINENUMBER	GROSSTOTAL	CLIENTE	VENDEDOR
1	7775	FOF 2017/12	02	21-02-2017	FT	1	0,00	00000 2
2	7771	FOF 2017/8	02	14-02-2017	FT	1	0,00	00000 2
3	7766	FOF 2017/3	01	12-01-2017	FT	1	0,00	00000 2
4	7776	FOF 2017/13	02	28-02-2017	FT	1	0,00	00000 2
5	7784	FOF 2017/21	03	27-03-2017	FT	1	0,00	00000 2
6	7770	FOF 2017/7	02	10-02-2017	FT	2	0,00	00000 2
7	7780	FOF 2017/17	03	16-03-2017	FT	1	0,00	00726 2
8	7765	FOF 2017/2	01	11-01-2017	FT	1	0,00	30059 24
9	7790	FOF 2017/27	05	31-05-2017	FT	1	0,00	30096 24
10	7782	FOF 2017/19	03	23-03-2017	FT	1	0,00	30185 24
11	7787	FOF 2017/24	04	12-04-2017	FT	1	0,00	30185 24
12	7772	FOF 2017/9	02	15-02-2017	FT	1	0,00	31064 24
13	7781	FOF 2017/18	03	20-03-2017	FT	2	0,00	31064 24
14	7786	FOF 2017/23	04	03-04-2017	FT	1	0,00	40010 5
15	7783	FOF 2017/20	03	27-03-2017	FT	1	0,00	42017 5
16	7778	FOF 2017/15	03	07-03-2017	FT	1	0,00	50189 2
17	7764	FOF 2017/1	01	05-01-2017	FT	1	0,00	72468 18
18	7789	FOF 2017/26	05	26-05-2017	FT	1	0,00	72468 18
19	7795	FOF 2017/32	06	30-06-2017	FT	3	0,00	72468 18
20	7773	FOF 2017/10	02	15-02-2017	FT	1	0,00	72742 18
21	7791	FOF 2017/28	05	31-05-2017	FT	1	0,00	72759 18
22	7774	FOF 2017/11	02	17-02-2017	FT	1	0,00	73180 14
23	7779	FOF 2017/16	03	15-03-2017	FT	1	0,00	73497 24
24	7788	FOF 2017/25	04	17-04-2017	FT	1	0,00	73497 24
25	7767	FOF 2017/4	01	20-01-2017	FT	1	0,00	73620 2
26	7777	FOF 2017/14	03	06-03-2017	FT	1	0,00	73620 2
27	7769	FOF 2017/6	01	26-01-2017	FT	1	0,00	73621 14
28	7793	FOF 2017/30	06	14-06-2017	FT	1	0,00	73655 22
29	7768	FOF 2017/5	01	20-01-2017	FT	1	0,00	73986 2
30	7785	FOF 2017/22	03	27-03-2017	FT	1	0,00	73986 2
31	7792	FOF 2017/29	06	09-06-2017	FT	1	0,00	74196 25
32	7794	FOF 2017/31	06	21-06-2017	FT	1	0,00	74201 1

Continuando a análise das estatísticas do campo do tipo data, relativo as **INVOICESTYPE FT** no ficheiro **FT\_ICV**, na Figura 24 **Erro! A origem da referência não foi encontrada.** podemos conferir que o período das faturas (analisado através do campo **INVOICEDATE**) vai de 02-01-2017 a 30-06-2017, que o dia em que se emitem mais INVOICES é a 4ª Feira, e o mês com maior número de faturas emitidas é março.

*Figura 24 - Estatísticas de Campo – Tipo de campo: data – Ficheiro: FT\_ICV*



### 2.3.3 A Distinção Entre Amostra, Bónus e Oferta

De acordo com o disposto no n.º 7 do artigo 3.º do Código do Imposto sobre o Valor Acrescentado (CIVA), as amostras e ofertas não estão abrangidas pelo conceito de transmissões de bens e, por esse facto, estão excluídas da tributação em sede deste imposto as transmissões efetuadas a título gratuito de bens não destinados a posterior comercialização, quando relativamente a esses bens ou aos elementos que os constituem tenha havido dedução total ou parcial do imposto, e aqueles bens, pelas suas características, ou pelo tamanho ou pelo formato diferentes do produto que constitua a unidade de venda, visem, sob a forma de amostra, apresentar ou promover bens produzidos ou comercializados pelo próprio sujeito passivo, assim como as ofertas de valor unitário igual ou inferior a € 50 e cujo valor global não exceda cinco por mil do volume de negócios do sujeito passivo no ano civil anterior, em conformidade com os usos comerciais.

A Portaria nº 497/2008, de 24 de Junho, do Ministério das Finanças e da Administração Pública, regulamenta as condições delimitadoras do conceito de amostras e de ofertas de

pequeno valor e define os procedimentos e obrigações contabilísticas a cumprir pelos sujeitos passivos do imposto, para efeitos de aplicação do disposto no n.º 7 do artigo 3.º do Código do Imposto sobre o Valor Acrescentado, aprovado pelo Decreto -Lei n.º 394 - B/84, de 26 de Dezembro, pelo que no n.º 1 do Artigo 2.º – Delimitação do conceito de amostras refere que: *“Consideram-se amostras os bens, não destinados a posterior comercialização, de formato ou tamanho diferentes do produto que constitua a unidade de venda ou apresentados em quantidade, capacidade, peso ou medida substancialmente inferiores aos que constituem a unidade de venda que se destinem a apresentar ou promover produtos produzidos ou comercializados pelo sujeito passivo”*, já no Artigo 3.º – Delimitação do conceito de oferta refere que : *“1 – A oferta pode ser constituída por bens comercializados ou produzidos pelo sujeito passivo ou por bens adquiridos a terceiros; 2 – Quando a oferta seja constituída por um conjunto de bens, o valor de € 50, a que se refere o n.º 7 do artigo 3.º do Código do IVA, aplica -se a esse conjunto; 3 – Excluem-se do conceito de oferta os bónus de quantidade concedidos pelo sujeito passivo aos seus clientes.”*

Após esta breve explanação, sobre a distinção entre amostra, ofertas e bónus, e para mais à frente se conseguir uma melhor compreensão das análises e testes a realizar, optou-se por fazer uma descrição de alguns procedimentos internos da TADIMA LDA quanto a ofertas, amostras e bónus.

Nos procedimentos internos da entidade relativamente às ofertas, bónus e amostras, nem sempre há uma distinção efetiva, com exceção da emissão do INVOICENO = FOF, em que, sempre que existe num documento significa que o registo de venda corresponde a uma em que os artigos são 100% oferta a clientes. Relativamente ao bónus estes nem sempre estão bem definidos nas faturas, uma vez que não existe uma regra, pode ser por campanhas, em que num mês se oferece 5+1 (unidades), e no mês seguinte 5+2 (unidades). No caso das amostras, a empresa não dispõe de produtos que possam ser enquadrados na definição anterior (produtos com formato ou tamanho distinto) sendo que apenas se podem considerar amostras artigos publicitários como, por exemplo, calendários, t-shirts, catálogos entre outros.

Pelo que se passa a expor, na empresa TADIMA LDA podem ser encontradas algumas das situações seguintes:

- ✓ Bónus por quantidades vendidas: bónus de uma unidade em função de um determinado número de unidades vendidas de um produto. Tal como pode ser observado na Figura 25, na **INVOICE\_ID** 1038, para o **PRODUCTCODE** 423010 aplicou-se o método “5+1” (esta é a conclusão obtida já que no campo **QUANTITY** existe um “1” e no campo **UNITPRICE** consta 0,00. No caso do **PRODUCTCODE** 461359 o bónus é de uma unidade por 9 unidades vendidas (“9+1”). Na figura estão ainda visíveis situações similares com outros produtos (424182) e outros bónus (“10+5”);

Figura 25 - Extração da INVOICE\_ID 1038

	INVOICE_ID	LINENUMBER	PRODUCTCODE	QUANTITY	UNITPRICE	TAXPOINTDATE	DEBITAMOUNT
1	1038	1	422066	5	2,31	19-01-2017	0,00
2	1038	2	422058	0	0,00	19-01-2017	0,00
3	1038	3	422026	2	4,09	19-01-2017	0,00
4	1038	4	422060	5	4,04	19-01-2017	0,00
5	1038	5	422036	5	4,29	19-01-2017	0,00
6	1038	6	423010	5	1,36	19-01-2017	0,00
7	1038	7	423010	1	0,00	19-01-2017	0,00
8	1038	8	424050	2	5,80	19-01-2017	0,00
9	1038	9	461359	9	2,02	19-01-2017	0,00
10	1038	10	461359	1	0,00	19-01-2017	0,00
11	1038	11	461356	9	1,94	19-01-2017	0,00
12	1038	12	461356	1	0,00	19-01-2017	0,00
13	1038	13	461364	9	2,25	19-01-2017	0,00
14	1038	14	461364	1	0,00	19-01-2017	0,00
15	1038	15	461372	5	3,10	19-01-2017	0,00
16	1038	16	461354	5	1,99	19-01-2017	0,00
17	1038	17	461357	2	5,70	19-01-2017	0,00
18	1038	18	461721	9	1,58	19-01-2017	0,00
19	1038	19	461721	1	0,00	19-01-2017	0,00
20	1038	20	461711	9	1,38	19-01-2017	0,00
21	1038	21	461711	1	0,00	19-01-2017	0,00
22	1038	22	462007	5	1,41	19-01-2017	0,00
23	1038	23	462007	1	0,00	19-01-2017	0,00
24	1038	24	462008	5	1,41	19-01-2017	0,00
25	1038	25	462008	1	0,00	19-01-2017	0,00
26	1038	26	424131	6	1,50	19-01-2017	0,00
27	1038	27	424182	10	2,44	19-01-2017	0,00
28	1038	28	424182	5	0,00	19-01-2017	0,00

- ✓ Ofertas de produtos que por norma, não são vendidos ao cliente: nesta situação, trata-se de uma tentativa de estimular a venda de novos produtos, embora, neste caso, as ofertas sejam pontuais e em unidades mais reduzidas. A Figura 26 revela esta situação em que os artigos com o **PRODUCTCODE** 472688C, 408159, 408157 são oferecidos não existindo nenhum desses produtos faturado nesse mesmo documento, ou seja, não são vendidos ao cliente.

Figura 26 - Extração da INVOICE\_ID 3025

	INVOICE_ID	LINENUMBER	PRODUCTCODE	QUANTITY	UNITPRICE	TAXPOINTDATE	DEBITAMOUNT
1	3025	1	04.04	100	4,88	09-03-2017	0,00
2	3025	2	472442	42	8,60	09-03-2017	0,00
3	3025	3	472688C	15	0,00	09-03-2017	0,00
4	3025	4	408159	2	0,00	09-03-2017	0,00
5	3025	5	408157	2	0,00	09-03-2017	0,00
6	3025	6	499981	2	0,00	09-03-2017	0,00
7	3025	7	499976	4	0,00	09-03-2017	0,00

No entanto, podem existir situações pontuais em que se oferecem muitos produtos aos clientes como bónus para criação de lojas em que os produtos da marca TADIMA LDA sejam vendidos.

Com o objetivo de selecionar as FT que são ofertas (FOF), ou seja, aquelas em que o valor do **GROSSTOTAL**, que é 0,00€, utilizou-se para a extração o critério seguinte sobre o ficheiro **Invoice 2017**, conforme Figura 27 e já referenciado anteriormente:

`@Isini("FOF"; INVOICENO )`

Figura 27 - Extração das INVOICENO = FOF

The screenshot shows a data table with columns: INVOICE\_ID, INVOICENO, PERIOD, INVOICEDATE, INVOICETYPE, QUANTITY, UNITPRICE, SETTLEMENTAMOUNT, and GROSSTOTAL. The table contains 29 rows of data. To the right, a sidebar titled 'Propriedades' shows search criteria: 'Dados' with 'Critério: @Isini("FOF"; INVOICENO )' and 'Resultados' with 'Estratificação' and 'Índices'.

	INVOICE_ID	INVOICENO	PERIOD	INVOICEDATE	INVOICETYPE	QUANTITY	UNITPRICE	SETTLEMENTAMOUNT	GROSSTOTAL
1	7775	FOF 2017/12	02	21-02-2017	FT	1	0,00	0,00	0,00
2	7786	FOF 2017/23	04	03-04-2017	FT	1	0,00	12,50	0,00
3	7783	FOF 2017/20	03	27-03-2017	FT	1	0,00	14,54	0,00
4	7771	FOF 2017/8	02	14-02-2017	FT	1	0,00	5,29	0,00
5	7793	FOF 2017/30	06	14-06-2017	FT	4	0,00	35,60	0,00
6	7780	FOF 2017/17	03	16-03-2017	FT	1	0,00	14,87	0,00
7	7773	FOF 2017/10	02	15-02-2017	FT	1	0,00	0,00	0,00
8	7791	FOF 2017/28	05	31-05-2017	FT	1	0,00	0,00	0,00
9	7778	FOF 2017/15	03	07-03-2017	FT	1	0,00	24,75	0,00
10	7767	FOF 2017/4	01	20-01-2017	FT	2	0,00	85,00	0,00
11	7777	FOF 2017/14	03	06-03-2017	FT	3	0,00	119,40	0,00
12	7766	FOF 2017/3	01	12-01-2017	FT	1	0,00	15,59	0,00
13	7779	FOF 2017/16	03	15-03-2017	FT	1	0,00	10,07	0,00
14	7788	FOF 2017/25	04	17-04-2017	FT	1	0,00	13,77	0,00
15	7776	FOF 2017/13	02	28-02-2017	FT	1	0,00	9,99	0,00
16	7784	FOF 2017/21	03	27-03-2017	FT	1	0,00	16,82	0,00
17	7768	FOF 2017/5	01	20-01-2017	FT	1	0,00	22,63	0,00
18	7785	FOF 2017/22	03	27-03-2017	FT	1	0,00	22,63	0,00
19	7770	FOF 2017/7	02	10-02-2017	FT	1	0,00	9,52	0,00
20	7764	FOF 2017/1	01	05-01-2017	FT	120	0,00	67,20	0,00
21	7789	FOF 2017/26	05	26-05-2017	FT	20	0,00	27,40	0,00
22	7795	FOF 2017/32	06	30-06-2017	FT	210	0,00	111,30	0,00
23	7792	FOF 2017/29	06	09-06-2017	FT	3	0,00	131,70	0,00
24	7790	FOF 2017/27	05	31-05-2017	FT	10	0,00	162,00	0,00
25	7794	FOF 2017/31	06	21-06-2017	FT	2	0,00	32,24	0,00
26	7769	FOF 2017/6	01	26-01-2017	FT	4	0,00	50,04	0,00
27	7774	FOF 2017/11	02	17-02-2017	FT	1	0,00	1,13	0,00
28	7782	FOF 2017/19	03	23-03-2017	FT	1	0,00	8,13	0,00
29	7787	FOF 2017/24	04	12-04-2017	FT	4	0,00	62,36	0,00

## 2.4 Testes Para Análise dos Dados

A análise de dados possui diferentes facetas e abordagens. Neste estudo, serão incorporadas diversas técnicas e testes. Trata-se de dados quantitativos, uma vez que os artigos são expressos em caracteres numéricos e em quantidades.

Dos métodos analíticos descritos na revisão de literatura, e atendendo aos dados contidos na base de dados disponível para estudo, foram aplicados os seguintes métodos:

- Lei de Benford
- Z-Score
- Same-Same-Same Teste
- Even Amount/ Números Redondos

Adicionalmente, utilizaram-se funcionalidades disponíveis no Caseware IDEA (nomeadamente, para importar dados a partir de XML, criar novos campos, aplicar critérios sobre os dados, ordenação de dados, criar tabelas para análises intermédias, juntar tabelas através de uma chave, detetar duplicados, tabelas dinâmicas, importação para IDEA e exportação para Excel, análise de histórico, criação de macros) e no Excel (criação de fórmulas, formatação condicional e ordenação de dados).

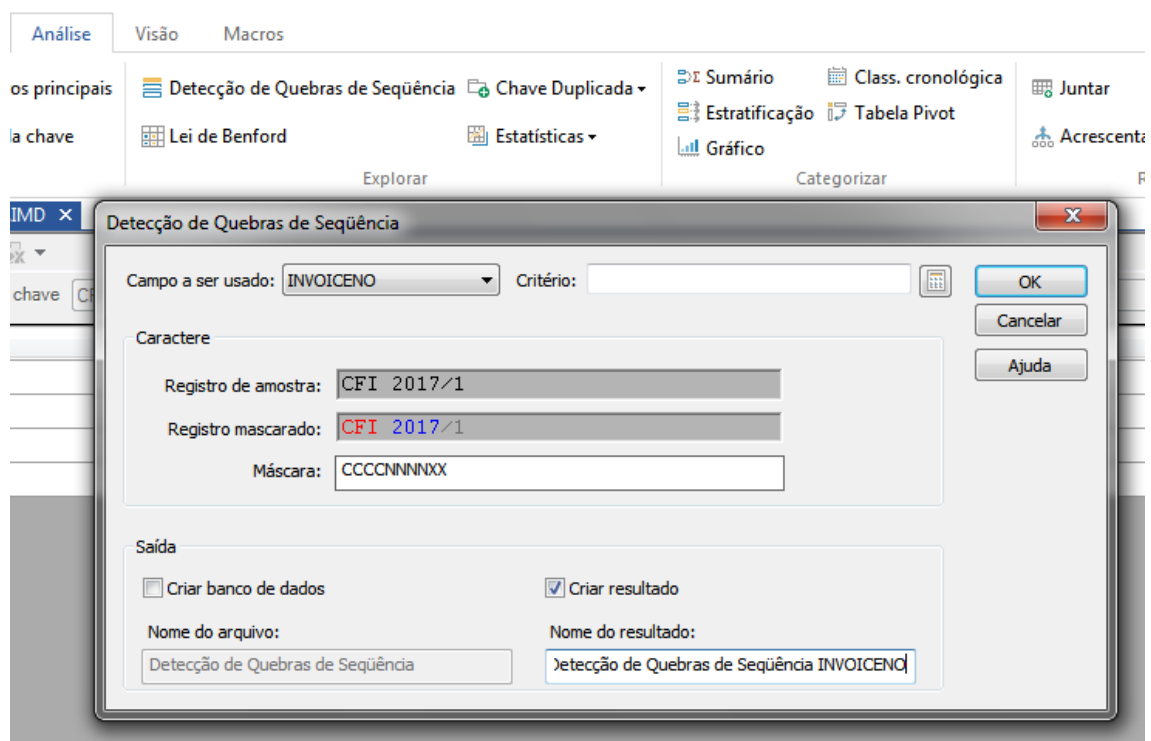
### 3 Apresentação do Estudo e Resultados

Relativamente a este capítulo, para que exista um fio condutor e para facilitar a compreensão, optou-se por colocar o objetivo, a análise e os resultados.

#### 3.1 Verificar se Existem Anomalias ao Nível dos Valores Totais das Faturas

Para responder a este objetivo, a primeira etapa será verificar se existem quebras de sequência ao nível das **INVOICENO**. Para tal realizou-se no IDEA o teste de Deteção de Quebras de Sequência (Figura 28), o campo usado foi **INVOICENO** e os resultados para os diversos tipos de INVOICE (VD, FI; FA; FOF; DV; CFI; NC; NCF;NCD) foram sempre 0 (como são várias séries, apresento apenas a título exemplificativo o resultado para FA, conforme Figura 29).

*Figura 28 - Deteção de Quebras de Sequência*



*Aplicação de métodos analíticos para auditoria e detecção de anomalias no ficheiro  
SAFT-PT: O caso de uma empresa comercial*

Figura 29 - Detecção Quebras Sequência FA

The screenshot shows the IDEA software interface. The main window displays a search for 'FA 2' in the 'Invoice 2017.IMD' file. The results pane shows 'Número total de elementos detectados' and 'Número total de quebras detectadas' both at 0. The 'Propriedades' pane on the right shows search criteria and results.

De seguida, e para verificar se existem situações anómalas a nível dos documentos emitidos pela empresa, em relação ao ficheiro *Invoice 2017* aplicou-se o teste *Same-Same-Same*, no IDEA no Explorar > Chave Duplicada > Extrair Registos Duplicados. As chaves são **INVOICEDATE**; **CUSTOMERID** e **GROSSTOTAL**, todos ordenados por ordem ascendente. Foi gerado um ficheiro de nome **MM INVDATE MMGROSST MM CUST**, cujo resultado é exposto na Figura 30, e onde se verifica que existem 193 registos onde existe a mesma data, o mesmo cliente e o mesmo total da fatura.

Figura 30 - Same Invoice Same InvoiceDate Same Customer

INVOICE_ID	INVOICENO	PERIOD	INVOICEDATE	INVOICETYPE	LINENUMBER	GROSSTOTAL	CLIENTE	VENDEDOR	
16	8788	VD 2017/271	01	13-01-2017	FT	1	31,43	72165	2
17	94	DV 2017/10	01	16-01-2017	NC	3	50,11	73811	2
18	8805	VD 2017/288	01	16-01-2017	FT	1	50,11	73811	2
19	8806	VD 2017/289	01	16-01-2017	FT	1	50,11	73811	2
20	95	DV 2017/11	01	17-01-2017	NC	1	184,32	73403	14
21	8838	VD 2017/321	01	17-01-2017	FT	1	184,32	73403	14
22	8848	VD 2017/331	01	18-01-2017	FT	1	20,76	00000	2
23	8849	VD 2017/332	01	18-01-2017	FT	1	20,76	00000	2
24	97	DV 2017/13	01	19-01-2017	NC	1	31,77	00000	2
25	8879	VD 2017/362	01	19-01-2017	FT	1	31,77	00000	2
26	99	DV 2017/15	01	20-01-2017	NC	1	12,18	00000	2
27	8915	VD 2017/398	01	20-01-2017	FT	1	12,18	00000	2
28	1182	FA 2017/1001	01	24-01-2017	FT	1	1.810,56	30003	24
29	1183	FA 2017/1002	01	24-01-2017	FT	1	1.810,56	30003	24
30	1184	FA 2017/1003	01	24-01-2017	FT	1	1.810,56	30003	24
31	102	DV 2017/18	01	24-01-2017	NC	1	2.150,01	73274	18
32	8960	VD 2017/443	01	24-01-2017	FT	1	2.150,01	73274	18
33	1329	FA 2017/1148	01	26-01-2017	FT	1	1.490,46	73389	14
34	7837	NC 2017/42	01	26-01-2017	NC	1	1.490,46	73389	14
35	105	DV 2017/21	01	27-01-2017	NC	1	18,36	74074	1
36	9027	VD 2017/510	01	27-01-2017	FT	1	18,36	74074	1
37	106	DV 2017/22	01	30-01-2017	NC	1	3,15	30193	24
38	9066	VD 2017/549	01	30-01-2017	FT	1	3,15	30193	24
39	1441	FA 2017/1260	01	31-01-2017	FT	1	57,61	73929	14
40	7846	NC 2017/51	01	31-01-2017	NC	1	57,61	73929	14
41	1517	FA 2017/1336	02	01-02-2017	FT	1	250,28	73453	24
42	7858	NC 2017/63	02	01-02-2017	NC	1	250,28	73453	24

A maioria dos casos tem um **INVOICETYPE** NC e FT, o que quer dizer que foram correções aos documentos, o que significa que um (NC) anula o outro (FT). A situação que se destacou mais foi o cliente 30003, já que o campo **INVOICETYPE** é sempre FT

para os 3 registos sem ter qualquer NC associado. Assim, fez-se, de seguida, uma extração de todas INVOICES (faturas) associadas a esse cliente, para tentar identificar se é uma situação normal. O resultado está exposto na Figura 31, pelo que se entende que este pode ser alvo de análise por parte da empresa de modo a verificar se, de facto, o cliente já liquidou estas faturas, uma vez que no ficheiro em análise não se dispõe dessa informação financeira.

Figura 31 - Extração INVOICES Cliente 30003

PERIOD	INVOICEDATE	INVOICETYPE	CUSTOMERID	LINENUMBER	GROSSTOTAL	CLIENTE	VENDEDOR
1 01	16-01-2017	FT		1	182,32	30003	24
2 01	24-01-2017	FT		1	1.810,56	30003	24
3 01	24-01-2017	FT		1	1.810,56	30003	24
4 01	24-01-2017	FT		1	1.810,56	30003	24
5 02	10-02-2017	FT		1	182,32	30003	24

Após a realização desta análise preliminar, utilizou-se a Lei de Benford para comparar o conjunto de dados da empresa com o resultado esperado da referida lei.

Optou-se por aplicar a Lei de Benford apenas às INVOICE FT, uma vez que a empresa usa as NC para corrigir preços, correção de documentos e de *rappel*<sup>18</sup>, o que poderia originar enviesamentos nos resultados da Lei de Benford. O teste aplicou-se às FT com o **GROSSTOTAL** com valores positivos. A Figura 32, revela o modo como se executou a aplicação da referida Lei, tendo sido aplicados dois testes:

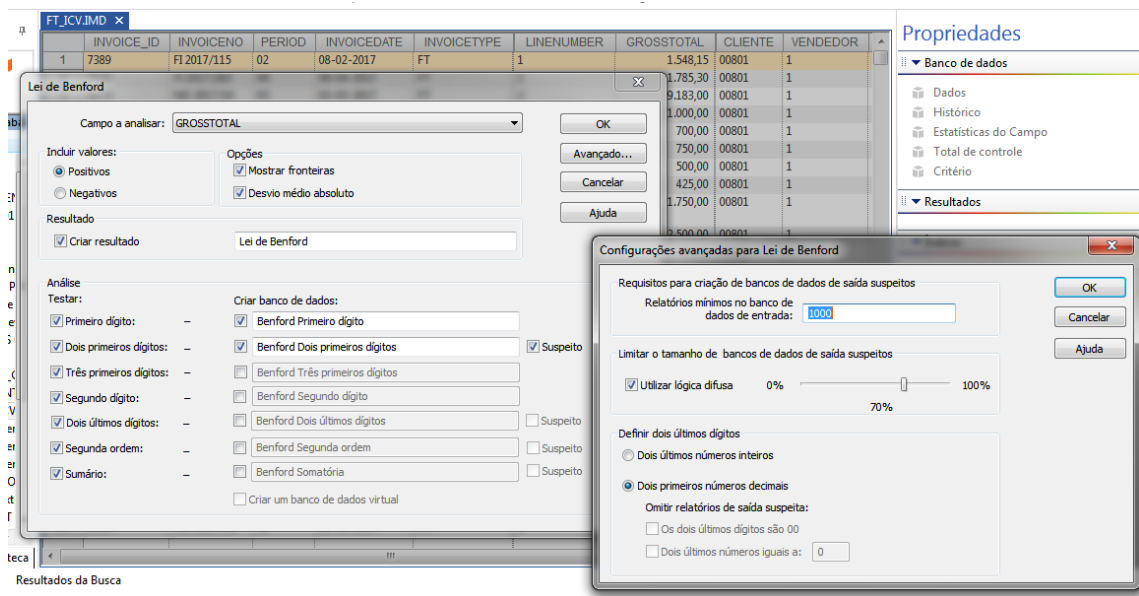
- Primeiro dígito
- Dois primeiros dígitos

---

<sup>18</sup> O "*rappel*" é qualificado como desconto de quantidade porque constitui um abatimento na prestação pecuniária do revendedor estruturalmente ligado ao volume de mercadoria por ele adquirida. Encontra-se, por regra, previsto no contrato geral celebrado entre as partes. O "*rappel*" tanto pode ser fixo, isto é, traduzir-se numa dada percentagem que é abatida ao preço das aquisições, como ser baseado num certo volume de aquisições (em euros, quilos, litros, etc., comportando, por vezes, escalões) e progredir à medida do aumento das aquisições. O que frequentemente se verifica é que o "*rappel*", só se vem a concretizar num momento posterior ao da emissão da fatura. Isto ocorre quando o "*rappel*" é apurado sobre o volume de aquisições realizadas num determinado período de tempo (um trimestre ou um ano, p. ex.). É o chamado "*rappel*" condicional o qual, independentemente do número de transações ocorridas, só se torna líquido no fim do período ([http://bdjur.almedina.net/item.php?field=node\\_id&value=1936065](http://bdjur.almedina.net/item.php?field=node_id&value=1936065), consultado em 27-10-2017).

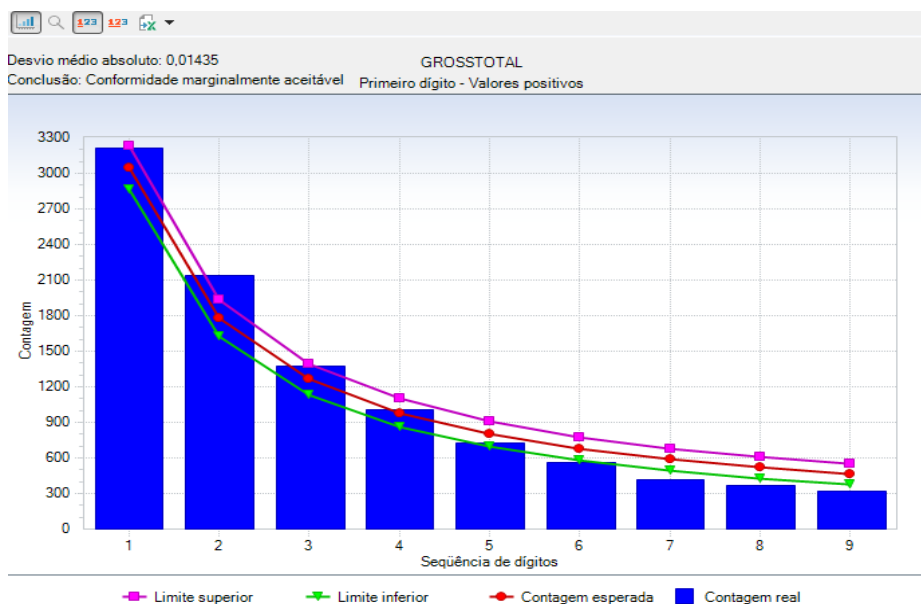
## Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro SAFT-PT: O caso de uma empresa comercial

Figura 32 - Aplicação da Lei de Benford



Por meio da análise ao Figura 33, verifica-se que o conjunto de dados relativo às INVOICES FT está em conformidade marginalmente aceitável com a Lei de Benford quando aplicada ao primeiro dígito, relativamente ao dígito 1, o número de vezes que apareceu na realidade como primeiro dígito foi 3217, o que está ligeiramente acima da contagem esperada pela Lei de Benford que seria 3048, mas não ultrapassou o limite superior. O dígito 2, está na realidade acima do valor esperado e do limite superior. Já os dígitos 6, 7, 8 e 9 estão abaixo do valor esperado e do limite inferior.

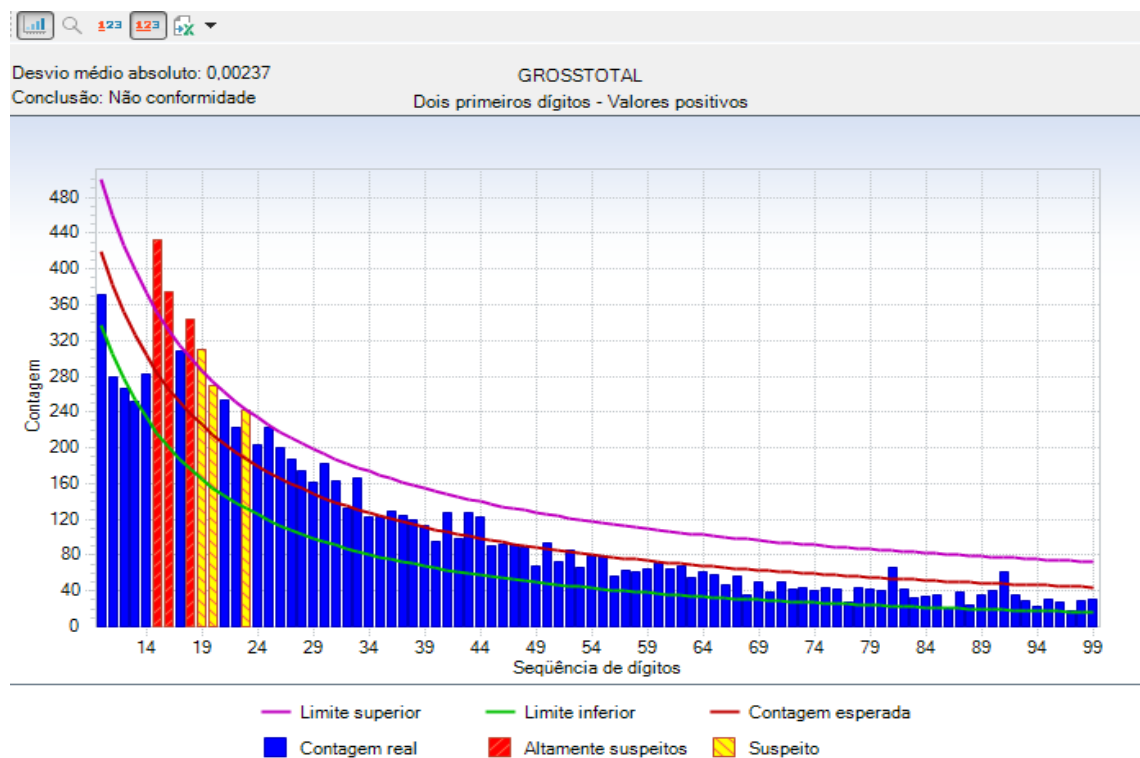
Figura 33 - Lei de Benford - Primeiro Dígito aplicado ao GROSSTOTAL



Para aprofundar um pouco mais a análise e tentar interpretar os diferenciais, é relevante obter a proporção dos dígitos, analisando os primeiros dois dígitos. No Figura 34, está ilustrado o teste dos dois primeiros dígitos para uma melhor análise de precisão. Verifica-se a existência elementos “altamente suspeitos” (a designação atribuída pela Lei de Benford) para os pares de dígitos 15, 16, 18, e “suspeitos” para os pares de dígitos 19, 20, 23, pelo que o conjunto de dados está não conforme com a Lei de Benford.

Será prudente examinar alguns dos conteúdos dos números “altamente suspeitos” e “suspeitos”. No caso dos primeiros estes podem ser explicados, por um lado pelo valor dos preços unitários dos artigos, já que há muitas **INVOICESNO VD** (vendas na loja a consumidores finais) que explicam esses valores, uma vez que os consumidores finais fazem compras ao nível de uma ou poucas unidades, sendo ainda vendidas muitas unidades isoladas. Outro fator que explica os valores “suspeitos” e “altamente suspeitos” corresponde ao facto de a empresa suportar o custo de transporte da mercadoria para vendas superiores a 150€. Assim, na maioria das vezes, os clientes (pequenos retalhistas) ajustam as suas compras no sentido de beneficiarem deste desconto, o que faz oscilar ligeiramente as vendas nos intervalos entre 150,00€ e 250,00€.

Figura 34 - Lei de Benford - Dois Primeiros Dígitos aplicada ao GROSSTOTAL



*Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro  
SAFT-PT: O caso de uma empresa comercial*

O IDEA permite extrair o conjunto de registos a que corresponde cada elemento do gráfico da Lei de Benford, na Figura 35 estão exibidos alguns dos registos cujos dois primeiros dígitos são 15, são os mais destacados de entre os 3 valores considerados “altamente suspeitos”, ou seja, aparecem em 432 registos.

Figura 35 - Extração dos Altamente Suspeitos para dois primeiros dígitos 15

	INVOICE_ID	PERIOD	INVOICEDATE	INVOICETYPE	LINENUMBER	GROSSTOTAL	CLIENTE	VENDEDOR
43	7627	05	12-05-2017	FT	1	1.516,17	40079	4
44	7718	06	19-06-2017	FT	1	1.583,28	49000	1
45	3106	03	13-03-2017	FT	1	153,77	71708	16
46	5564	05	17-05-2017	FT	1	158,45	71708	16
47	4824	04	28-04-2017	FT	1	153,45	70778	22
48	3003	03	09-03-2017	FT	1	151,66	00012	14
49	6082	05	30-05-2017	FT	1	153,75	74166	22
50	870	01	16-01-2017	FT	1	153,86	31028	22
51	3050	03	10-03-2017	FT	1	158,92	71224	16
52	4375	04	17-04-2017	FT	1	152,38	71224	16
53	5647	05	18-05-2017	FT	1	156,81	74093	14
54	10588	04	26-04-2017	FT	1	154,73	00000	2
55	9206	02	07-02-2017	FT	1	15,71	00000	2
56	2728	03	02-03-2017	FT	1	1.515,27	02110	18
57	4808	04	28-04-2017	FT	1	156,55	72693	14
58	6198	06	01-06-2017	FT	1	155,37	74189	16
59	559	01	09-01-2017	FT	1	154,68	72351	14
60	4760	04	27-04-2017	FT	1	151,11	72351	14
61	5239	05	09-05-2017	FT	1	152,48	70207	16
62	6688	06	14-06-2017	FT	1	156,95	70207	16
63	9638	03	02-03-2017	FT	1	155,92	00000	2
64	6157	05	31-05-2017	FT	1	154,46	72959	22
65	963	01	18-01-2017	FT	1	150,10	72776	18
66	11470	06	20-06-2017	FT	1	150,58	73655	22
67	8736	01	11-01-2017	FT	1	155,02	74108	1
68	2910	03	07-03-2017	FT	1	150,53	74017	18
69	10050	03	24-03-2017	FT	1	15,93	00000	2
70	9827	03	13-03-2017	FT	1	155,07	74013	24

Após esta análise, procedeu-se aplicação do teste Even Amount/Números Redondos, para verificar se existem muitas faturas com o GROSSTOTAL arredondado. Para tal, através do ficheiro Invoice 2017, aplicou o seguinte critério de extração:

$$(GROSSTOTAL \% 10) = 0 .AND. GROSSTOTAL < > 0$$

Obteve-se o ficheiro EVEN AMOUNT para GROSSTOTAL, representado pela Figura 36, consegue-se visualizar que não existem nem muitas situações, nem são problemáticas, pois só existem 28 registos com valores arredondados às dezenas e os valores não são significativos, pelo que esta situação está alinhada com o que diz Sunder Gee (referenciado anteriormente no ponto 1.3.7).

Figura 36 - EVEN AMOUNT para GROSSTOTAL

	INVOICE_ID	INVOICENO	PERIOD	NVOICEDATE	INVOICETYPE	GROSSTOTAL
1	8861	VD 2017/344	01	18-01-2017	FT	10,00
2	10374	VD 2017/1857	04	11-04-2017	FT	10,00
3	472	FA 2017/291	01	06-01-2017	FT	40,00
4	3703	FA 2017/3522	03	29-03-2017	FT	40,00
5	9656	VD 2017/1139	03	03-03-2017	FT	40,00
6	10382	VD 2017/1865	04	11-04-2017	FT	40,00
7	8832	VD 2017/315	01	17-01-2017	FT	50,00
8	9760	VD 2017/1243	03	08-03-2017	FT	50,00
9	5544	FA 2017/5363	05	16-05-2017	FT	60,00
10	111105	VD 2017/2588	05	26-05-2017	FT	60,00
11	10601	VD 2017/2084	04	27-04-2017	FT	100,00
12	7459	FI 2017/185	03	07-03-2017	FT	130,00
13	8906	VD 2017/389	01	20-01-2017	FT	150,00
14	10819	VD 2017/2302	05	10-05-2017	FT	150,00
15	3536	FA 2017/3355	03	23-03-2017	FT	160,00
16	5111	FA 2017/4930	05	05-05-2017	FT	160,00
17	675	FA 2017/494	01	11-01-2017	FT	210,00
18	7159	FA 2017/6978	06	28-06-2017	FT	210,00
19	4546	FA 2017/4365	04	20-04-2017	FT	250,00
20	9566	VD 2017/1049	02	27-02-2017	FT	250,00
21	947	FA 2017/766	01	18-01-2017	FT	280,00
22	1038	FA 2017/857	01	19-01-2017	FT	300,00
23	7574	FI 2017/300	04	27-04-2017	FT	330,00
24	7370	FI 2017/96	02	02-02-2017	FT	360,00
25	1708	FA 2017/1527	02	07-02-2017	FT	390,00
26	5125	FA 2017/4944	05	05-05-2017	FT	410,00
27	3059	FA 2017/2878	03	10-03-2017	FT	420,00
28	2484	FA 2017/2303	02	23-02-2017	FT	1.060,00

Propriedades

▼ Banco de dados

- Dados
- Histórico
- Estatísticas do Campo
- Total de controle
- Critério

▼ Resultados

▼ Índices

- Nenhum índice
- INVOICEDATE/A
- GROSSTOTAL/A

▼ Comentários

da Busca

Não está conectado ao IDEA Server Número de Registos: 28

Relativamente, a este objetivo pode-se concluir que as situações anómalas encontradas através da realização dos testes apresentados, são explicados pelo conhecimento da empresa e dos seus procedimentos internos.

### 3.2 Identificar Situações Anómalas em Relação aos Produtos Oferecidos

Para responder a este objetivo tem de se realizar um conjunto de operações de interligação entre ficheiros, de modo a poder abranger, num mesmo ficheiro, o conjunto da informação útil disponível. Assim sendo, inicialmente fez-se a junção do ficheiro *Customer 2017* (sendo este o Banco primário) com o ficheiro *Clientes* (a chave usada foi respetivamente **ID\_CLIENTE\_N** e **CLIENTE**) por forma a ficar com informação relativa aos vendedores. Desta junção surgiu o ficheiro *customer2017\_com vendedor*. De seguida procedeu-se à junção do ficheiro *Invoices 2017* com o anteriormente (*customer2017\_com vendedor*), em que *Invoices 2017* foi o banco primário, a união foi feita através do campo **CUSTOMERID**, assim obteve-se um ficheiro com informação de quem a denominação por extenso do cliente e vendedor associado à fatura (*Invoice*), a que se deu o nome de **INVOICE\_CUSTOMER\_VENDEDOR**. De modo análogo, juntou-se o ficheiro *Line 2017* (banco de dados primário) com o ficheiro **INVOICE\_CUSTOMER\_VENDEDOR**, através do campo **INVOICE\_ID** comum em ambos os ficheiros do qual se obteve o arquivo com o nome

**LINE\_INVOICE\_CUSTOMER\_VENDEDORI**, ou seja, neste último ficheiro ficamos com toda a informação relevante contida num só ficheiro artigos, faturas, clientes e vendedores.

A forma de detetar que o **PRODUCTCODE** foi oferecido é através da visualização do valor patente no campo **QUANTITY** que tem que ser >0 (quando este campo é igual a 0, significa que o cliente encomendou, mas não havia em armazém, trata-se de um procedimento para no futuro se justificar a não expedição de um produto), e **UNITPRICE=0**.

Após a execução destes procedimentos e considerando o último ficheiro criado, procedeu-se à extração dos produtos oferecidos pela empresa, em que o Critério de extração:

*QUANTITY > 0 .AND. UNITPRICE = 0 .AND. INVOICETYPE = "FT".*

Desta extração obtiveram-se 19.516 linhas de registos. Fazendo ordenação do campo quantidade por ordem decrescente, o resultado que se obtêm pode ser visualizado na Figura 37.

*Figura 37 - FT Produtos oferecidos*

	PRODUCTCODE	QUANTITY	UNITPRICE	DEBITAMOUNT	SETTLEMENTAMOUNT	CREDITAMOUNT	INVOICE
1	E461250	1080	0,00	0,00	1.576,80	0,00	7599
2	E461250	1080	0,00	0,00	1.630,80	0,00	7538
3	9995	858	0,00	0,00	17,16	0,00	3910
4	E461251	600	0,00	0,00	870,00	0,00	7599
5	E461251	600	0,00	0,00	930,00	0,00	7538
6	E408359	560	0,00	0,00	0,00	0,00	7374
7	424114	350	0,00	0,00	42,00	0,00	6969
8	424114	350	0,00	0,00	38,50	0,00	5792
9	E461298	240	0,00	0,00	391,20	0,00	7599
10	E461298	240	0,00	0,00	393,60	0,00	7538
11	412481	210	0,00	0,00	111,30	0,00	7795
12	TTR603R	200	0,00	0,00	274,00	0,00	5969
13	424114	200	0,00	0,00	24,00	0,00	488
14	424114	200	0,00	0,00	24,00	0,00	4573
15	424114	200	0,00	0,00	24,00	0,00	2420
16	E461316	180	0,00	0,00	520,20	0,00	7599
17	E461308	180	0,00	0,00	293,40	0,00	7599
18	E461316	180	0,00	0,00	522,00	0,00	7538
19	E461308	180	0,00	0,00	306,00	0,00	7538
20	9995	159	0,00	0,00	0,00	0,00	2045
21	E408539	150	0,00	0,00	0,00	0,00	7523
22	438013	128	0,00	0,00	1.227,52	0,00	5721
23	50037	125	0,00	0,00	18,75	0,00	7354
24	412481	120	0,00	0,00	67,20	0,00	7764
25	408539	120	0,00	0,00	0,00	0,00	4845
26	E461129	117	0,00	0,00	1.544,40	0,00	7344
27	E408539	100	0,00	0,00	0,00	0,00	7649
28	424114	100	0,00	0,00	11,00	0,00	685

Relativamente aos produtos oferecidos, para facilitar a análise optou-se por excluir alguns **PRODUCTCODE**, tais como o código 9995 (relativo ao transporte), os artigos referentes às paletes (em relação a este artigo mais à frente será abordado), e aos elementos que contêm o campo **PRODUCTCODE** igual a DOC (dizem respeito a documentos tipo calendários e t-shirts). Para tal criou-se um novo ficheiro, a partir do **FT Produtos oferecidos**, em que se usou o seguinte critério de extração:

.NOT. @Isini ("PALET"; PRODUCTDESCRIPTION) .AND. .NOT. @Isini ("TRANSP";  
PRODUCTDESCRIPTION) .AND. .NOT. @Isini ("DOC"; PRODUCTCODE)

Obteve-se a tabela que pode ser observada na Figura 38, a qual está ordenada por ordem descendente de quantidade oferecida, tendo-se reduzido o n.º total de registos para análise. Na referida figura pode-se constatar que existem produtos em que o valor do campo **SETTLEMENTAMOUNT** é 0,00€, o que é explicado por esses artigos se referirem a ofertas/amostras de reduzido valor, ou deveria de ser, pelo que também serão analisados posteriormente desses produtos.

*Figura 38 - Produtos Oferecido sem transporte, paletes e artigos tipo DOC*

LINE_INVOICE_CUSTOMER_VENDE...	PRODUCTCODE	QUANTITY	UNITPRICE	DEBITAMOUNT	SETTLEMENTAMOUNT	CREDITAMOUNT	INVOICE_ID
1	E461250	1080	0,00	0,00	1.576,80	0,00	7599
2	E461250	1080	0,00	0,00	1.630,80	0,00	7538
3	E461251	600	0,00	0,00	870,00	0,00	7599
4	E461251	600	0,00	0,00	930,00	0,00	7538
5	E408359	560	0,00	0,00	0,00	0,00	7374
6	E461298	240	0,00	0,00	391,20	0,00	7599
7	E461298	240	0,00	0,00	393,60	0,00	7538
8	TTR603R	200	0,00	0,00	274,00	0,00	5969
9	E461316	180	0,00	0,00	520,20	0,00	7599
10	E461308	180	0,00	0,00	293,40	0,00	7599
11	E461316	180	0,00	0,00	522,00	0,00	7538
12	E461308	180	0,00	0,00	306,00	0,00	7538
13	E408539	150	0,00	0,00	0,00	0,00	7523
14	438013	128	0,00	0,00	1.227,52	0,00	5721
15	50037	125	0,00	0,00	18,75	0,00	7354
16	408539	120	0,00	0,00	0,00	0,00	4845
17	E461129	117	0,00	0,00	1.544,40	0,00	7344
18	E408539	100	0,00	0,00	0,00	0,00	7649
19	50008	100	0,00	0,00	20,00	0,00	3095
20	E461296	96	0,00	0,00	161,28	0,00	7538
21	E461255	96	0,00	0,00	161,28	0,00	7538
22	438013	96	0,00	0,00	920,64	0,00	6281
23	438013	96	0,00	0,00	920,64	0,00	1916
24	E461296	81	0,00	0,00	132,03	0,00	7599
25	E461255	81	0,00	0,00	136,08	0,00	7599
26	E431196	80	0,00	0,00	264,80	0,00	7649
27	E431193	80	0,00	0,00	255,20	0,00	7649
28	431101	78	0,00	0,00	0,00	0,00	4845

Propriet

- ▼ Banco de
- ✓ Dados
- Históric
- Estadísti
- Total de
- Crítéri
- ▼ Resultad
- ▼ Índices
- Nenhur
- QUANT
- ✓ QUANT
- ▼ Comentá
- Adicion

Da análise dos dados é interessante examinar o artigo palete, através da extração dos artigos oferecidos que têm no **PRODUCTDESCRIPTION** a palavra PALET (porque existem vários **PRODUCTCODE** com descrição diferente, mas que se referem ao mesmo produto, como por exemplo: EURO PALETE, CP5 PALET) usou-se o seguinte critério de extração:

*@Isini ("PALET"; PRODUCTDESCRIPTION)*

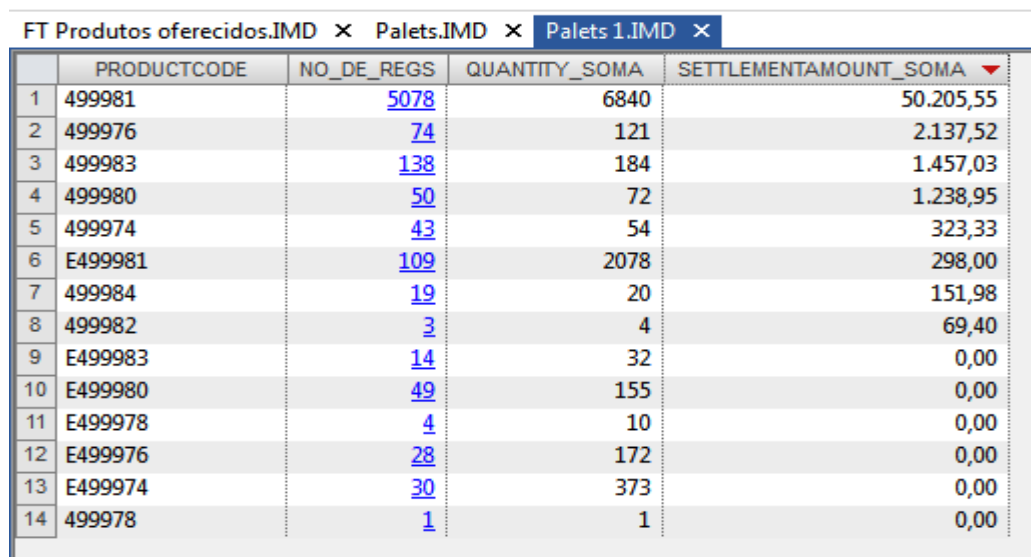
Posteriormente fez-se um sumário (pelo campo **PRODUCTCODE**), que está representado na Figura 39, em que se criou o ficheiro *Palets 1*, composto por:

- **NO\_DE\_REGS** representa o número de linhas em que aparece o produto;
- **QUANTITY\_SOMA** refere-se à soma das quantidades constantes dos referidos registos;
- **SETTLEMENTAMOUNT\_SOMA** é a soma dos descontos realizados por **PRODUCTCODE**.

O ficheiro *Palets 1*, foi ordenado por ordem descendente do campo **SETTLEMENTAMOUNT\_SOMA**, em que relativamente a este campo se pode observar o elevado peso do total de desconto do artigo 499981.

Ao nível do procedimento interno da empresa no que concerne à oferta de paletes, o que se encontra estipulado é que o objetivo é a devolução das mesmas, à exceção dos clientes do mercado externo em que os **PRODUCTCODE** começam por E (exemplos: E499983; E499981).

*Figura 39 - Quantidade de Paletes oferecidas*



	PRODUCTCODE	NO_DE_REGS	QUANTITY_SOMA	SETTLEMENTAMOUNT_SOMA
1	499981	5078	6840	50.205,55
2	499976	74	121	2.137,52
3	499983	138	184	1.457,03
4	499980	50	72	1.238,95
5	499974	43	54	323,33
6	E499981	109	2078	298,00
7	499984	19	20	151,98
8	499982	3	4	69,40
9	E499983	14	32	0,00
10	E499980	49	155	0,00
11	E499978	4	10	0,00
12	E499976	28	172	0,00
13	E499974	30	373	0,00
14	499978	1	1	0,00

Será interessante no futuro que alguém da organização realizasse uma comparação, relativamente aos artigos que têm como **PRODUCTDESCRIPTION** paletes, entre as quantidades que são enviadas para os clientes e as que são efetivamente devolvidas, de modo a avaliar se a empresa está a perder dinheiro com esta operação, uma vez que o peso dos descontos associados a este produto (55 881,76€) no volume de negócios do período em estudo ainda é significativo, ou seja, representa 2% em relação ao volume de vendas.

Com a análise detetou-se que existem produtos que têm valor 0,00 no campo **SETTLEMENTAMOUNT** (refere-se ao valor do desconto, mas no estudo em questão, quando se tem **QUANTITY** > 0.AND. **UNITPRICE** <> 0, o valor do campo **SETTLEMENTAMOUNT** respeita ao preço pelo qual, em caso de venda ao cliente os produtos seriam vendidos), pelo que se vai analisar se, para este tipo de produtos, há valores cujo **UNITPRICE** é diferente de 0. Para tal optou-se por fazer um sumário pelo campo **PRODUCTCODE**, para os artigos com o campo **SETTLEMENTAMOUNT**=0, cujo nome do ficheiro é *Desconto=0* (Figura 40). Posteriormente, realizou-se uma extração com o seguinte Critério:

*QUANTITY > 0.AND. UNITPRICE <> 0*

Neste caso, denominou-se o ficheiro por *Artigos Vendidos* e, por fim, um cruzamento entre estes dois ficheiros (*Artigos Vendidos* e *Desconto=0*) ao que se deu o nome de **ARTIGOS SETTLEMENTAMOUNT E UNITPRICE MAIOR 0**

No ficheiro resultante obtiveram-se 171 registos o que ainda é um número significativo para análise. Assim, utilizou-se o critério:

*.NOT.@ISINI ("PALET"; PRODUCTDESCRIPTION) .AND. .NOT.@ISINI ("TRANS";  
PRODUCTDESCRIPTION) .AND. .NOT. @ISINI ("REPA"; PRODUCTCODE) .AND.  
.NOT. @ISINI ("DOC"; PRODUCTCODE) .AND. .NOT. @ISINI ("FOLHET";  
PRODUCTDESCRIPTION)*

Esta fórmula permite a exclusão de alguns artigos que não são relevantes para a análise, tais como os folhetos ("FOLHET"), reparações ("REPA"), transporte ("TRANS"), as paletes ("PALET") e ainda os artigos cujo **PRODUCTCODE** é "DOC", tendo-se reduzido bastante o número de registos para análise (15 registos finais).

*Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro SAFT-PT: O caso de uma empresa comercial*

Na Figura 41, reflete os arquivos que foram no SAFT com valor de desconto 0,00 €, mas que não se referem a amostras, visto que se encontrou no período vendas com preço unitário acima de 0.

*Figura 40 - Ficheiro composto pelo total de artigos cujo SETTLEMENTAMOUNT é zero*

	PRODUCTCODE	NO_DE_REGS	QUANTITY_SOMA
1	E499981	110	2070
2	E408359	1	560
3	408539	27	536
4	408328	108	495
5	E499974	30	373
6	E408539	2	250
7	431101	9	186
8	E499976	28	172
9	408682	5	161
10	9995	1	159
11	E499980	49	155
12	E408532	3	150
13	E431068	3	120
14	499981	16	101
15	E408179	2	100
16	DOC	5	85
17	DOC	4	72
18	E408328	2	70
19	408160	6	65
20	DOC	1	60
21	408539	1	60
22	408038	7	60
23	DOC	1	50
24	E431192	1	48
25	E431069	1	48
26	DOC	1	40
27	DOC	3	40
28	DOC	6	35
29	E499983	14	32

*Figura 41 - Ficheiro composto pelo total de artigos vendidos*

	PRODUCTCODE	QUANTITY	UNITPRICE	DEBITAMOUNT	SETTLEMENTAMOUNT	CREDITAMOUNT
1	424006	42	0,46	19,32	0,00	0,00
2	E438013	1008	1,00	1.008,00	0,00	0,00
3	438013	1	8,63	8,63	0,27	0,00
4	423002	50	6,24	0,00	0,00	312,00
5	451042	6	17,04	0,00	0,00	102,00
6	421518	6	15,87	0,00	0,00	95,00
7	421379	2	18,90	0,00	0,00	37,00
8	421581	2	22,50	0,00	0,00	45,00
9	422035	4	27,80	0,00	19,62	111,00
10	422003	4	26,27	0,00	18,54	105,00
11	422004	4	27,29	0,00	19,26	109,00
12	460211	6	10,96	0,00	0,00	65,00
13	421992	10	23,80	0,00	0,00	238,00
14	421079	10	18,08	0,00	0,00	180,00
15	421122	10	14,18	0,00	0,00	141,00
16	421123	10	15,11	0,00	0,00	151,00
17	421320	4	45,14	0,00	0,00	180,00
18	9995	2472	0,02	0,00	0,00	49,00
19	411253	1	10,88	0,00	0,00	10,00
20	411091	1	11,88	0,00	0,00	11,00
21	412601	1	4,41	0,00	0,00	4,00
22	423080.	6	3,28	0,00	0,00	19,00
23	441125	10	26,06	0,00	8,06	260,00
24	431119	2	17,07	0,00	9,85	34,00
25	431147	10	33,47	0,00	10,35	334,00
26	431145	6	7,71	0,00	1,43	46,00
27	411546	2	11,50	0,00	0,71	22,00
28	411091	1	13,18	0,00	0,41	13,00

Ao analisar a Figura 42, através do campo **UNITPRICE**, o **PRODUCTCODE** 431101 é que mais se destaca, porque foram vendidas 186 unidades e o seu valor de venda é 15,20€. Em relação aos restantes artigos, os artigos E431069 e E431068 também se destacam, mas não têm um valor assim tão significativo comparando com o artigo 431101.

Figura 42 - Ficheiro ARTIGOS SETTLEMENTAMOUNT E UNITPRICE MAIOR 0

The screenshot shows the IDEA software interface. At the top, there are various toolbars for data manipulation and analysis. The main window displays a table with the following data:

	PRODUCTCODE	NO_DE_REGS	QUANTITY_SOMA	QUANTITY	UNITPRICE	DEBITAMOUNT	CREDITAMOUNT
1	408328	108	495	5	0,02	0,00	0,00
2	408328	1	24	5	0,02	0,00	0,00
3	431101	9	186	1	15,20	0,00	0,00
4	431087	1	4	1	3,98	0,00	0,00
5	412401	1	5	2	5,23	0,00	0,00
6	412331	1	5	2	6,47	0,00	0,00
7	438013	1	3	1	8,63	8,63	0,00
8	431194	1	2	2	3,00	0,00	0,00
9	431193	1	2	1	3,52	0,00	0,00
10	431191	1	2	1	4,91	0,00	0,00
11	431190	1	2	1	3,28	0,00	0,00
12	460111	1	5	1	5,19	0,00	0,00
13	460093	1	5	2	2,93	0,00	0,00
14	E431069	1	48	20	2,45	0,00	0,00
15	E431068	2	120	20	2,60	0,00	0,00

On the right side, the 'Propriedades' panel is visible, showing various data and control options. At the bottom of the window, a status bar indicates: 'Não está conectado ao IDEA Server', 'Número de Registos: 15/171', and 'Espaço em Disco: 17,63 GB'.

Em relação ao **PRODUCTCODE** 431101, fez-se uma extração para visualizar quem são os clientes associados a este produto. No campo **NO\_DE\_REGS**, e clicando no linha correspondente ao referido produto (neste caso é o 9), o IDEA permite extrair esse resultado, pelo que se criou o ficheiro **431101 PRODUCTCODE**, e da visualização se concluiu que foi sempre o mesmo cliente, existindo meses em que foram registadas várias **INVOICEID**, com o mesmo produto, mas variando a quantidade, mas realçando o facto de o campo **SETTLEMENTAMOUNT** ser igual a 0,00€, pelo que esta situação deve ser analisada e investigada pela empresa (Figura 43).

Para perceber que valores poderão estar em causa, isto é, deixar de ser faturados pela empresa, fez-se um sumário relativo ao preço praticado para o referido produto (Figura 44) bem como uma extração relativa **PRODUCTCODE** 431101 (Figura 45).

*Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro SAFT-PT: O caso de uma empresa comercial*

Verifica-se que o preço associado a este produto tem oscilações (valores que variam entre 27,8€ e 3,61€ constatando-se que o valor que mais vezes surge na fatura é o 17,7 € (correspondendo a 45 registos) pelo que esses valores devem ser investigados de modo a perceber se se trata de anomalias ou de campanhas em curso. Caso seja a primeira possibilidade, as situações em causa devem ser investigadas de modo a averiguar se se trata de uma situação de fraude, apurando responsabilidades.

Figura 43 - Ficheiro com extração do produto 431101

LINE_ID	ACCOUNTID	INVOICE_ID	LINENUMBER	PRODUCTCODE	QUANTITY	OFMEAS	UNITPRICE	KPOINTDA	ITAMO	TAXTYPE	TAXCOUR	
1	31029	2111171000	2452	68	431101	21	UN	0,00	22-02-2017	0,00	IVA	PT
2	43886	2111171000	3401	39	431101	6	UN	0,00	21-03-2017	0,00	IVA	PT
3	47243	2111171000	3659	3	431101	6	UN	0,00	28-03-2017	0,00	IVA	PT
4	52127	2111171000	3996	42	431101	6	UN	0,00	05-04-2017	0,00	IVA	PT
5	54606	2111171000	4183	42	431101	12	UN	0,00	11-04-2017	0,00	IVA	PT
6	57913	2111171000	4421	57	431101	6	UN	0,00	18-04-2017	0,00	IVA	PT
7	61790	2111171000	4691	92	431101	12	UN	0,00	26-04-2017	0,00	IVA	PT
8	64025	2111171000	4845	67	431101	78	UN	0,00	28-04-2017	0,00	IVA	PT
9	96217	2111171000	7188	63	431101	39	UN	0,00	29-06-2017	0,00	IVA	PT

o Resultados da Busca

o Samples Não está conectado ao IDEA Server Número de Registos: 9 Esp

Figura 44 - Preços praticados para o PRODUCTCODE 431101

	UNITPRICE	NO_DE_REGS	QUANTITY_SOMA
1	0,00	40	238
2	3,61	1	1
3	4,05	1	2
4	4,18	1	1
5	4,46	1	2
6	13,80	2	25
7	14,19	6	9
8	14,63	4	5
9	15,20	23	25
10	15,21	6	12
11	15,67	3	7
12	15,68	3	6
13	16,22	18	57
14	16,72	18	61
15	17,31	2	2
16	17,70	45	124
17	18,25	1	1
18	18,81	1	1
19	19,23	2	4
20	19,66	2	6
21	19,85	7	7
22	20,23	2	11
23	20,27	21	99
24	20,28	5	10
25	20,90	19	65
26	25,02	8	11
27	27,80	2	10

o Resultados da Busca

o Samples Não está conectado ao IDEA Server Número de Registos: 27 Espaço em Disc

*Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro SAFT-PT: O caso de uma empresa comercial*

Figura 45 - Extração PRODUCTCODE 431101

	PRODUCTCODE	QUANTITY	UNITPRICE	DEBITAMOUNT	SETTLEMENTAMOUNT	CREDITAMOUNT	INVOICE_ID
1	431101	1	15,20	0,00	5,70	15,20	10035
2	431101	1	15,20	0,00	5,70	15,20	10099
3	431101	10	16,72	0,00	41,80	167,20	1011
4	431101	2	15,21	0,00	11,39	30,41	10136
5	431101	2	4,46	0,00	32,88	8,92	10243
6	431101	1	14,19	0,00	6,71	14,19	1047
7	431101	1	19,85	0,00	1,05	19,85	10478
8	431101	1	15,20	0,00	5,70	15,20	10580
9	431101	1	15,20	0,00	5,70	15,20	10634
10	431101	1	16,22	0,00	4,68	16,22	10696
11	431101	1	15,20	0,00	5,70	15,20	10707
12	431101	2	25,02	0,00	5,56	50,04	10749
13	431101	1	15,20	0,00	5,70	15,20	10752
14	431101	1	16,22	0,00	4,68	16,22	10790
15	431101	1	19,85	0,00	1,05	19,85	10819
16	431101	3	16,22	0,00	14,04	48,66	10905
17	431101	1	15,20	0,00	5,70	15,20	10978
18	431101	1	25,02	0,00	2,78	25,02	11071
19	431101	1	15,67	0,00	5,23	15,67	11114
20	431101	1	25,02	0,00	2,78	25,02	11315
21	431101	5	20,90	0,00	0,00	104,50	1138
22	431101	1	0,00	0,00	20,90	0,00	1138
23	431101	1	15,20	0,00	5,70	15,20	11439
24	431101	3	17,70	0,00	9,37	53,09	1146
25	431101	6	17,70	0,00	18,74	106,18	1147
26	431101	4	17,70	0,00	12,49	70,79	1149
27	431101	2	17,70	0,00	6,25	35,39	1155
28	431101	1	19,85	0,00	1,05	19,85	11579

Não está conectado ao IDEA Server Número de Registos: 244

Prosseguindo, passou-se a identificar os clientes com maior número produtos oferecidos, quer em quantidade quer em valor de desconto concedido. Começou-se por fazer uma soma dos produtos oferecidos por cliente, bem como a soma dos valores dos descontos, pelo que daqui se obtiveram 1012 registos, conforme Figura 46, tendo-se utilizado o critério:

*CLIENTE* <> "00000"

que corresponde ao código associado ao nome “Consumidor Final”, o qual se pretendia excluir da análise pelo facto de corresponder a muitos clientes diferentes. Esta extração foi guardada num ficheiro (com o nome *Produtos oferecidos por cliente com valor do desconto*) para posterior utilização.

*Aplicação de métodos analíticos para auditoria e detecção de anomalias no ficheiro  
SAFT-PT: O caso de uma empresa comercial*

Figura 46 - Soma da QUANTITY e do valor do SETTLEMENTAMOUNT por CLIENTE

	CLIENTE	AC_OU	NO_DE_REGS	QUANTITY_SOMA	SETTLEMENTAMOUNT_SOMA
1	40006	ES	63	6562	10.301,35
2	40010	ES	253	3027	15.504,71
3	00554	PT	184	1827	2.773,49
4	70292	PT	217	1625	3.422,72
5	70541	PT	86	1250	1.876,33
6	71328	PT	306	796	4.297,57
7	42295	ES	70	782	4.353,79
8	71407	PT	185	759	3.458,94
9	71000	PT	78	647	778,68
10	40081	ES	46	642	252,39
11	72468	PT	5	554	579,90
12	00008	PT	165	551	2.213,42
13	40035	ES	72	485	1.348,26
14	72583	PT	192	472	3.086,66
15	72453	PT	155	461	8.605,58
16	73274	PT	54	450	1.196,51
17	72371	PT	181	408	1.285,55
18	70159	PT	106	361	1.233,30
19	72944	PT	170	355	1.091,18
20	72173	PT	188	348	746,19
21	73362	PT	172	342	955,21
22	71834	PT	252	331	898,45
23	71787	PT	173	326	1.370,47
24	42017	ES	74	324	1.724,58
25	72870	PT	100	313	1.827,09
26	72657	PT	112	304	991,46
27	74054	PT	105	298	1.712,27
28	73031	PT	10	278	2.489,18
29	73009	PT	125	277	1.118,73

Executou-se um sumário a partir do ficheiro das quantidades oferecidas (*FT Produtos oferecidos*, este ficheiro já foi explicado anteriormente) e selecionou-se a opção de sumarizar para cada cliente (usando o campo **CLIENTE**), para obter o total de artigos que lhes foram oferecidos no período em estudo, conforme Figura 47. Posteriormente e de forma a detetar *outliers* em relação às quantidades oferecidas, fez-se uma exportação dos dados para um ficheiro Excel e aplicou-se o teste do Z-Score.

Figura 47 - Soma dos produtos oferecidos por total de linhas e por cliente

	CLIENTE	AC_OU	NO_DE_REGS	QUANTITY_SOMA	SETTLEMENTAMOUNT_SOMA
1	40006	ES	63	6562	10.301,35
2	40010	ES	253	3027	15.504,71
3	00554	PT	184	1827	2.773,49
4	70292	PT	217	1625	3.422,72
5	70541	PT	86	1250	1.876,33
6	71328	PT	306	796	4.297,57
7	42295	ES	70	782	4.353,79
8	71407	PT	185	759	3.458,94
9	71000	PT	78	647	778,68
10	40081	ES	46	642	252,39

Para se realizar o teste do Z-Score (aplicado conforme descrição na revisão da literatura, secção 1.3.3), usou-se a seguinte fórmula

$$\mathbf{Z-Score} = (C2 - \text{MÉDIA}(\$C\$2:\$C\$1012)) / \text{DESVPAD.P}(\$C\$2:\$C\$1012)$$

Sendo que a fórmula foi aplicada em todas as 1012 células, correspondendo C2 ao valor da observação que se pretende compreender se é um *outlier* ou não e o intervalo C2 a C1012 aquele onde se situam todos os valores das observações. O resultado obtido, ordena do por ordem decrescente e sinalizando (através de formatação condicional) os valores cujo Z-Score é superior a 3, sendo estes os que são considerados *outliers* e que têm que ser alvo de posterior estudo.

Para facilitar a interpretação criou-se um campo que é **ID\_NAC\_OU\_EXP**, em que o valor ES corresponde a clientes estrangeiros e PT a clientes nacionais. No geral, os clientes estrangeiros são grandes armazenistas, pelo que compram grandes quantidades e, conseqüentemente, auferem de maiores descontos, já os clientes nacionais são, na sua maioria pequenos retalhistas, e um número reduzido de cadeias de lojas. Analisando a Tabela 1, pode-se visualizar que os clientes cujos Z-Score são mais elevados são, teoricamente, os clientes que compram em maior quantidade.

*Tabela 5 - Calculo do Z-Score para as quantidades oferecidas*

CLIENTE	Soma QTD	ID_NAC_OU_EXP	Z-SCORE
40010	3027	ES	20,38654
40006	6562	ES	13,43897
72453	461	PT	11,17476
42295	782	ES	5,497728
71328	796	PT	5,422663
71407	759	PT	4,302916
70292	1625	PT	4,254555
72583	472	PT	3,805845
00554	1827	PT	3,387697
31049	70	PT	3,300107
73031	278	PT	3,008084
00008	551	PT	2,639886
72714	164	PT	2,427775
74116	270	PT	2,260861
70541	1250	PT	2,189801
30185	183	PT	2,152548
72870	313	PT	2,124055
42017	324	ES	1,987183
74054	298	PT	1,970746
73354	246	PT	1,884679

O valor médio de quantidades oferecidas por cliente foi de 56 e o desvio padrão de 254.

Para avaliar se os clientes a quem foram oferecidos mais produtos correspondem aos que mais compraram, elaborou-se o mesmo estudo, mas para o apuramento do Z-Score correspondendo às quantidades vendidas por cliente, estudo esse que está patente na Tabela 6.

*Tabela 6 - Z-Score por Cliente por quantidades vendidas*

CLIENTE	QUANTITY_SOMA	Z-Score Q vendidas
40010	100121	11,09841942
42295	90775	10,0305312
40055	86764	9,572228223
40089	84756	9,34279108
40035	74222	8,139160168
71000	72771	7,973366694
42017	71154	7,788605807
40006	68678	7,505694269
40004	66628	7,271458141
00008	65005	7,086011685
49020	61121	6,642219919
73031	51920	5,590899621
40054	50344	5,410823456
40032	49184	5,278280086
00554	47565	5,093290676
01045	44266	4,716341903
71328	41761	4,430116781
40002	34628	3,615089317
70159	27756	2,829884113
70661	26793	2,719850263
01819	26366	2,671060592
70292	26063	2,63643935
74054	25878	2,615300967
40003	25850	2,612101645
40081	25090	2,525262885
40001	25031	2,518521455
70191	24570	2,465846892
72164	24069	2,408601867
40085	24040	2,405288283
71865	23474	2,340616259
40029	23200	2,309308601

O valor médio de quantidades vendidas por cliente foi de 2989 e o desvio padrão de 8752.

Finalmente, procedeu-se à importação para o IDEA das tabelas com os Z-Scores calculados anteriormente em Excel, em dois ficheiros separados denominados **Z-score Q Oferecidas** e **Z-score Q VENDIDAS**, através do qual se fez uma junção, para se poder

comparar os Z-Scores obtidos por cliente, cujo campo chave foi o **CLIENTE**, originando a Figura 48 - Z-Score para quantidades oferecidas e para quantidades vendidas por cliente, em que o campo **Z\_SCORE** diz respeito às ofertas e **Z\_SCORE\_Q\_VENDIDAS** refere-se às quantidades vendidas.

Figura 48 - Z-Score para quantidades oferecidas e para quantidades vendidas por cliente

	CLIENTE	Z_SCORE	Z_SCORE_Q_VENDIDAS
1	40010	20,386540	11,098419
2	40006	13,438966	7,505694
3	72453	11,174758	2,158483
4	42295	5,497728	10,030531
5	71328	5,422663	4,430117
6	71407	4,302916	1,910079
7	70292	4,254555	2,636439
8	72583	3,805845	1,676071
9	00554	3,387697	5,093291
10	31049	3,300107	-0,307280
11	73031	3,008084	5,590900
12	00008	2,639886	7,086012
13	72714	2,427775	0,307789
14	74116	2,260861	0,272825
15	70541	2,189801	2,186706
16	30185	2,152548	-0,168795
17	72870	2,124055	2,238809
18	42017	1,987183	7,788606
19	74054	1,970746	2,615301
20	73354	1,884679	0,580417
21	31064	1,675465	-0,318364
22	72164	1,641818	2,408602
23	02239	1,576045	1,065115
24	71787	1,514372	1,420582
25	30122	1,510406	-0,326933
26	40035	1,484717	8,139160
27	72371	1,400986	-0,036595

Constata-se um diferencial, ou seja, não há uma concordância nos Z-Scores para os CLIENTE 72453, 71407 e 31409. Extraíu-se, para cada um destes 3 clientes, as quantidades oferecidas e as quantidades vendidas por produto, a partir desses ficheiros e através do IDEA, fez-se uma comparação por produto para ver a diferença entre as quantidades vendidas e oferecidas e , assim, detetar anomalias ao nível dos produtos oferecidos para os clientes onde mais se destacam as diferenças. Para facilitar a interpretação das análises posteriores, os campos seguintes são: **P\_NÚMREGS** (corresponde à soma do número registos em que o produto aparece relativamente às quantidades vendidas); **S\_NÚMREGS** (corresponde à soma do número registos em que o produto aparece relativamente às quantidades oferecidas); **P\_TOTAL** (soma das quantidades vendidas por produto); **S\_TOTAL** (soma das quantidades oferecidas por

produto), **DIFERENÇA** corresponde à diferença entre as quantidades vendidas e oferecidas) pelo que se passa a analisar os clientes cujos Z-Scores, estão mais dispares, para os clientes referidos anteriormente.

- **Cliente 72453**

Os resultados relativamente ao cliente 72453 estão evidenciados na Figura 49 e na Figura 50, pelo que na primeira fez-se uma ordenação por ordem ascendente das diferenças entre as quantidades vendidas e as oferecidas. Constata-se que há **PRODUCTCODE** em que foram oferecidos artigos e a empresa não adquiriu os mesmos.

*Figura 49 - Diferença entre as vendas e os produtos oferecidos ordenado na diferença por ordem crescente*

	PRODUCTCODE	P_NÚMREGS	P_TOTAL	S_NÚMREGS	S_TOTAL	DIFERENÇA ▲
1	461606	1	1	6	6	-5
2	421897	0	0	4	4	-4
3	441136	0	0	1	4	-4
4	423002	0	0	2	2	-2
5	421041	1	1	1	1	0
6	401407	1	1	0	0	1
7	423001	1	1	0	0	1
8	441061	1	1	0	0	1
9	460201	1	1	0	0	1
10	460205	1	1	0	0	1
11	461293	1	1	0	0	1
12	461317	1	1	0	0	1
13	461722	1	1	0	0	1
14	RAPPEL	1	1	0	0	1
15	421521	2	2	0	0	2
16	421820	2	2	0	0	2
17	422000	2	2	0	0	2
18	424005	2	2	0	0	2
19	440999	1	2	0	0	2
20	461313	1	2	0	0	2
21	461371	2	2	0	0	2
22	461384	2	2	0	0	2
23	424027	2	3	0	0	3
24	424130	2	3	0	0	3
25	431145	1	5	1	2	3
26	431149	2	10	2	7	3
27	431162	1	5	1	2	3
28	431163	1	5	1	2	3
29	461162	1	3	0	0	3

Posteriormente, fez-se uma ordenação descendente do campo **S\_TOTAL**, para observar quais os artigos que foram oferecidos em maior número ao cliente 72453, e será interessante investigar no futuro se a empresa TADIMA LDA considerando o cliente em causa, incrementou as vendas, por exemplo, para os produtos 431140, 431151, 431143 os quais apresentam um número de ofertas muito elevado em relação às vendas. Este é efetivamente o interesse da investigação: detetar qual das situações se verificou, ou uma anomalia praticada intencionalmente por alguém na empresa ou uma tentativa por parte do vendedor de aumentar as vendas daqueles produtos a este cliente. Pelo que se descreveu, é relevante analisar, no futuro, se as políticas de ofertas de quantidades avultadas em relação ao padrão normal tendo como objetivo o aumento de vendas para um determinado cliente foi atingido, ou seja, o cliente comprou efetivamente esses produtos no futuro ou não.

*Aplicação de métodos analíticos para auditoria e deteção de anomalias no ficheiro  
SAFT-PT: O caso de uma empresa comercial*

Figura 50 - Cliente 72453 - Diferença entre as vendas e os produtos oferecidos ordenado pelo S\_TOTAL por ordem decrescente

	PRODUCTCODE	P_NÚMREGS	P_TOTAL	S_NÚMREGS	S_TOTAL	DIFERENÇA
1	431140	6	93	5	66	27
2	431151	2	40	3	34	6
3	431143	4	45	3	32	13
4	431136	2	32	2	26	6
5	431146	4	32	4	21	11
6	441077	7	62	7	19	43
7	431154	2	25	2	19	6
8	431133	4	40	4	16	24
9	431158	2	20	2	14	6
10	431187	8	68	5	13	55
11	423080	3	27	3	9	18
12	441075	3	26	3	8	18
13	431134	2	20	2	8	12
14	431149	2	10	2	7	3
15	461606	1	1	6	6	-5
16	431153	3	16	2	6	10
17	460375	5	25	5	5	20
18	441065	2	18	2	5	13
19	423075	6	19	5	5	14
20	441136	0	0	1	4	-4
21	441069	2	16	2	4	12
22	431160	3	14	2	4	10
23	431142	1	10	1	4	6
24	431137	1	10	2	4	6
25	431132	1	10	1	4	6
26	424011	10	45	4	4	41
27	421897	0	0	4	4	-4
28	460322	3	15	3	3	12
29	431138	1	9	1	3	6

- **Cliente 31049**

Está evidenciado na Figura 51, pelo que se constata que relativamente a este cliente e, concretamente ao **PRODUCTCODE** 472688C, é possível que existam anomalias, já que foram vendidas 59 unidades e oferecidas 49.

Figura 51 - Cliente 31049 - Diferença entre as vendas e os produtos oferecidos ordenado na diferença por ordem crescente

	PRODUCTCODE	P_NÚMREGS	P_TOTAL	S_NÚMREGS	S_TOTAL	DIFERENÇA ▲
1	408168	0	0	1	1	-1
2	408865	0	0	1	1	-1
3	472563	1	1	0	0	1
4	472535	2	2	0	0	2
5	472478	4	5	1	1	4
6	472646	1	4	0	0	4
7	472688C	7	59	7	49	10
8	472495C	5	34	5	5	29
9	472442	6	195	0	0	195

Para tentar interpretar estes valores anómalos realizou-se o teste *Same-Same-Same*, em que se usou a Chave:

PRODUCTCODE/A + INVOICE\_ID/A + INVOICEDATE/A

O objetivo é encontrar duplicados, usando o Critério:

**CLIENTE** = "31049"

*Aplicação de métodos analíticos para auditoria e detecção de anomalias no ficheiro SAFT-PT: O caso de uma empresa comercial*

O resultado extraído está evidenciado na Figura 52. Aqui importa investigar a razão da empresa vender 8 unidades do produto 472688C e oferecer 7, sucedendo isto todos os meses - excluindo fevereiro – existindo, em janeiro e junho 2 faturas emitidas com o mesmo produto e com a mesma quantidade oferecida.

Figura 52 - Teste Same-Same-Same - PRODUCTCODE/A + INVOICE\_ID/A + INVOICEDATE/A para o CLIENTE = "31049"

PRODUCTCODE	QUANTITY	UNITPRICE	DEBITAMOUNT	SETTLEMENTAMOUNT	CREDITAMOUNT	INVOICE_ID	PERIOD	INVOICEDATE
472478	1	12,67	0,00	0,00	12,67	2962	03	08-03-2017
472478	1	12,08	0,00	0,00	12,08	4220	04	11-04-2017
472478	1	12,08	0,00	0,00	12,08	4220	04	11-04-2017
472495C	6	15,59	0,00	0,00	93,54	2962	03	08-03-2017
472495C	1	0,00	0,00	15,59	0,00	2962	03	08-03-2017
472495C	6	15,59	0,00	0,00	93,54	4220	04	11-04-2017
472495C	1	0,00	0,00	15,59	0,00	4220	04	11-04-2017
472495C	6	15,59	0,00	0,00	93,54	5100	05	05-05-2017
472495C	1	0,00	0,00	15,59	0,00	5100	05	05-05-2017
472495C	6	15,59	0,00	0,00	93,54	6212	06	02-06-2017
472495C	1	0,00	0,00	15,59	0,00	6212	06	02-06-2017
472495C	10	15,59	0,00	0,00	155,90	7240	06	30-06-2017
472495C	1	0,00	0,00	15,59	0,00	7240	06	30-06-2017
472688C	8	14,58	0,00	0,00	116,64	1418	01	30-01-2017
472688C	7	0,00	0,00	102,06	0,00	1418	01	30-01-2017
472688C	8	14,58	0,00	0,00	116,64	2962	03	08-03-2017
472688C	7	0,00	0,00	102,06	0,00	2962	03	08-03-2017
472688C	8	14,58	0,00	0,00	116,64	4220	04	11-04-2017
472688C	7	0,00	0,00	102,06	0,00	4220	04	11-04-2017
472688C	8	14,58	0,00	0,00	116,64	5100	05	05-05-2017
472688C	7	0,00	0,00	1.005,06	0,00	5100	05	05-05-2017
472688C	8	14,58	0,00	0,00	116,64	519	01	09-01-2017
472688C	7	0,00	0,00	102,06	0,00	519	01	09-01-2017
472688C	8	14,58	0,00	0,00	116,64	6212	06	02-06-2017
472688C	7	0,00	0,00	1.005,06	0,00	6212	06	02-06-2017
472688C	11	14,58	0,00	0,00	160,38	7240	06	30-06-2017
472688C	7	0,00	0,00	102,06	0,00	7240	06	30-06-2017

Outra anomalia, que pode ser visualizada é o valor do desconto no valor de 1005,06€, pelo que se extraiu a **INVOICE\_ID** 5100 (também poderia ser a **INVOICE\_ID** 6212, porque apresenta o mesmo valor) conforme se pode constatar na Figura 53, onde terá que analisar qual o critério que foi aplicado para definir o valor a colocar no campo **SETTLEMENTAMOUNT**, atendendo a que, considerando o preço unitário do produto 14,58€ e existindo 7 unidades, o valor do desconto deveria ser 102,06€ e não os 1005,06€ conforme se apresenta na imagem, pode-se suscitar a dúvida se ao nível das quantidades oferecidas se foram faturas as unidades corretas, pelo que se deve tentar investigar, por exemplo através da contagem do stock para o **PRODUCTCODE** 472688C.

Figura 53 - INVOICE\_ID 5100 análise do campo SETTLEMENTAMOUNT

PRODUCTCODE	QUANTITY	UNITPRICE	DEBITAMOUNT	SETTLEMENTAMOUNT	CREDITAMOUNT	INVOICE_ID	PERIOD	INVOICEDATE	INVOIC
472442	42	8,60	0,00	0,00	361,20	5100	05	05-05-2017	FT
472688C	8	14,58	0,00	0,00	116,64	5100	05	05-05-2017	FT
472688C	7	0,00	0,00	1.005,06	0,00	5100	05	05-05-2017	FT
472495C	6	15,59	0,00	0,00	93,54	5100	05	05-05-2017	FT
472495C	1	0,00	0,00	15,59	0,00	5100	05	05-05-2017	FT
472478	2	12,08	0,00	0,00	24,16	5100	05	05-05-2017	FT
499981	2	0,00	0,00	14,90	0,00	5100	05	05-05-2017	FT

• **Cliente 71407**

No que concerne ao CLIENTE 71407, através da análise da Figura 54, que demonstra a diferença entre as vendas e os produtos oferecidos ordenado na diferença por ordem crescente, que foram oferecidos produtos com o PRODUCTCODE a começar por DF, o que é de questionar, uma vez que esse tipo de produtos por norma só são oferecidos aos clientes da loja.

*Figura 54 - Cliente 71407 diferença entre as vendas e os produtos oferecidos ordenado na diferença por ordem crescente*

	PRODUCTCODE	P_NÚMREGS	P_TOTAL	S_NÚMREGS	S_TOTAL	DIFERENÇA ▲
1	408539	0	0	1	50	-50
2	DF441133	0	0	1	48	-48
3	DF441136	0	0	1	48	-48
4	DF441138	0	0	1	48	-48
5	408328	0	0	3	30	-30
6	2012-0000	0	0	1	24	-24
7	431136	1	1	2	5	-4
8	431148	0	0	1	2	-2
9	PCV00041	0	0	1	2	-2
10	431153	1	2	1	3	-1
11	450012	0	0	1	1	-1
12	450028	0	0	1	1	-1
13	411920	1	1	0	0	1

Analogamente, a Figura 55 apresenta a soma dos produtos oferecidos, com ordenação descendente do campo S\_TOTAL. Verifica-se que, para alguns produtos, há uma relação de igualdade entre os campos P\_NÚMREGS e S\_NÚMREGS (uma vez que representam o número de faturas em que apareceu os respetivos produtos – PRODUCTCODE). O que pode ser explicado pela compra de mais quantidade e, como tal, mais ofertas já que o número de registos é o mesmo, de que é exemplo os artigos cujo PRODUCTCODE é 433059, 433055, 433057, 433060.

*Figura 55 - Cliente 71407 - Diferença entre as vendas e os produtos oferecidos ordenado pelo S\_TOTAL por ordem decrescente*

	PRODUCTCODE	P_NÚMREGS	P_TOTAL	S_NÚMREGS	S_TOTAL ▼	DIFERENÇA
1	423080.	12	360	10	60	300
2	408539	0	0	1	50	-50
3	DF441138	0	0	1	48	-48
4	DF441136	0	0	1	48	-48
5	DF441133	0	0	1	48	-48
6	433059	11	242	11	46	196
7	433055	8	180	8	36	144
8	424182	11	191	7	32	159
9	408328	0	0	3	30	-30
10	433057	6	140	6	28	112
11	424185	6	59	4	24	35
12	2012-0000	0	0	1	24	-24
13	424181	13	170	6	23	147
14	433060	10	105	10	21	84
15	433056	10	106	9	20	86
16	431140	6	47	4	19	28
17	431151	5	35	5	15	20
18	431143	8	41	6	15	26

### **3.3 Sugestões de Melhoria no Sistema de Controlo Interno**

Ao longo deste estudo foram apresentados alguns testes ao conjunto de dados disponibilizados pela empresa, nomeadamente: a Lei de Benford, o Z-Score, o *Same-Same* teste e o *Even Amount*/Números Redondos, para aplicar estes testes foi por vezes, necessário acrescentar critérios adicionais, por forma a ser mais fácil a compreensão e interpretação da informação extraída. São exemplo a questão dos artigos referentes às paletes, já que se trata de um produto que é oferecido mas que representa um custo elevado para a entidade e reconhecendo que, por norma, uma encomenda tem pelo menos uma paleta associada.

Relativamente à empresa em estudo, julgo que será de extrema importância que, num futuro próximo e para evitar situações dúbias, se distinga entre o que é de facto “oferta”, “bónus” ou “amostra/publicidade”. As análises permitem detetar situações sistemáticas de ofertas, algo que é perceptível e para se compreender a frequência das ofertas, clientes e produtos associados às mesmas têm que existir algo que as diferencie.

A inexistência desta distinção pode gerar anomalias e dar lugar a situações propícias a fraude. Por outro lado, existem também possíveis implicações fiscais a nível de IVA (Artigo 3º do CIVA) e de IRC (Artigo 23º do CIRC), já que é difícil fazer a distinção no SAFT entre o que é de facto oferta, bónus ou amostra/publicidade. Uma sugestão seria a diferenciação no campo PRODUCTCODE ou na descrição do produto, usando, por exemplo, “O” para ofertas, “B” para bónus e “A” para amostra. Outra possibilidade seria colocar uma linha com a descrição 5+1 (por exemplo), já que seria evidente que, nesta situação, se trataria de um bónus e não de uma oferta.

Como na empresa, existem várias pessoas que emitem faturas, seria prudente, implementar uma política de mensalmente existir um elemento responsável por investigar, verificar e analisar quais os produtos mais vendidos e oferecidos, e depois debruçar-se sobre quem teve mais ofertas e analisar se estavam de acordo com as promoções vigentes no período, e caso hajam falhas/anomalias, tentar corrigi-las ou mitigá-las.

## **CONCLUSÃO**

Com este trabalho pretende-se reforçar a importância do controlo interno nas organizações e da auditoria interna como ferramentas de apoio à gestão na tomada de decisão. Uma das grandes preocupações da profissão de auditor é a capacidade dos auditores de detetarem fraude e irregularidades. Assim, é de extrema relevância a canalização de investimento em formação de recursos humanos na área de análise de dados com o objetivo de estarem preparados para a deteção de anomalias e de erros já que ambos representam custos para a organização: a inexistência de um sistema de controlo interno eficaz facilita a realização de atos fraudulentos.

Um dos objetivos deste estudo foi disponibilizar às empresas soluções de testes analíticos que permitam detetar erros e anomalias, utilizando um ficheiro que todas as empresas têm ao seu dispor e através da realização de testes desenvolvidos no IDEA e em Excel. Ressalva-se que a empresa em causa é uma pequena empresa pelo que, extrapolando estes testes para empresas de maior dimensão, outros erros e anomalias se poderão encontrar.

No que se refere à empresa, relativamente ao primeiro objetivo (de verificar se existem anomalias ao nível do montante das faturas emitidas) foram analisados todos os registos das faturas dos meses de janeiro a junho de 2017. Através da análise do valor do GROSSTOTAL e aplicada a Lei de Benford e o teste *Even Amount*/Números Redondos, entre outros testes intermédios, verificou-se que o valor das vendas não está em conformidade com a Lei de Benford, sendo as anomalias explicadas pela oferta do transporte para vendas superiores a 150€. Da aplicação do teste *Even Amount* conclui-se pela existência de apenas 28 registos de valores arredondados às dezenas, ou seja, os valores arredondados não ocorrem com uma elevada taxa de frequência.

No que concerne ao segundo objetivo (analisar o universo das amostras, bónus e produtos oferecidos) constata-se é muito difícil de determinar se o procedimento associado às ofertas está a ser cumprido, pois a empresa não faz a diferenciação entre amostra, bónus e ofertas ao nível dos códigos dos artigos. Porém, através da aplicação do teste Z-score e o *Same Same Same* teste, entre outros testes intermédios, foi possível identificar anomalias relativamente a produtos, nomeadamente com ofertas consecutivas do mesmo produto em diversos meses e ofertas que não estimularam compras posteriores. No produto “Paletes” verificou-se que deveria ser retirado da análise já que se sabe que

reporta sempre a uma oferta, porém, contabilizados esses totais no final de 6 meses, verifica-se que correspondem a um montante não desprezável.

A análise através da aplicação de métodos analíticos, permitiu detetar anomalias, pelo que estes testes se mostraram eficazes quando aplicados a empresas desta natureza.

### **Limitações do estudo**

As principais limitações do estudo devem-se ao fato da empresa não ter bem definido nem implementado, ao nível dos artigos disponíveis, o que são amostras, ofertas e bónus.

### **Investigação futura**

Relativamente à aplicação da Lei de Benford, seria aliciante realizar este teste a outras empresas do mesmo setor (comércio por grosso de produtos alimentares), com as mesmas características, para verificar se o facto de a empresa suportar o custo de transporte influencia o valor das faturas. Também se poderia fazer a mesma análise mas por vendedor, e investigar quais os produtos que são mais oferecidos por vendedor, bem como avaliar se o GROSSTOTAL das vendas por vendedor segue a Lei de Benford, e quais os que têm um maior número de encomendas a rondar o valor dos 150,00€ (foi neste valor que, através da aplicação da lei de Benford ao GROSSTOTAL, se verificou a existência de mais números suspeitos, porque é o valor mínimo de encomenda para os clientes beneficiarem da oferta do custo de transporte), e tentar investigar se existe potencial para aumentar as vendas para esses mesmos clientes, de modo a incrementar o volume de vendas.

### **Contributos profissionais e científicos**

Ao nível profissional foi alertar e mostrar à organização para a existência de falhas ao nível do controlo interno.

A nível de contribuição científica foi evidenciar a possibilidade de utilizar o IDEA e os métodos analíticos nas auditorias para a deteção de anomalias no contexto português usando um ficheiro que é usado pelas finanças para as suas auditorias fiscais e ser uma forma de as empresas anteciparem as falhas e deficiências do controlo interno.

## **REFERÊNCIAS BIBLIOGRÁFICAS**

- ACFE - Association of Certified Fraud Examiners. (2016). *Report To The Nations 2016, Global Fraud Study*. <https://www.acfe.com/rtnn2016/docs/2016-report-to-the-nations.pdf>
- ACFE - Association of Certified Fraud Examiners. (2012). *Report To The Nations 2012, Global Fraud Study*.  
[https://www.acfe.com/uploadedFiles/ACFE\\_Website/Content/rtnn/2012-report-to-nations.pdf](https://www.acfe.com/uploadedFiles/ACFE_Website/Content/rtnn/2012-report-to-nations.pdf)
- AICPA. (2002). AU Section 316 Consideration of Fraud in a Financial. *October*, (99, 113), 167–218.
- Albrecht, W. S., Albrecht, C. O., Albrecht, C. C., & Zimbelman, M. F. (2016). *Fraud Examination*. (Cengage Learning, Ed.) (5<sup>a</sup>).
- Albrecht, W. S., Williams, T. L., & Wernz, G. W. (1995). *Fraud: Bringing Light to the Dark Side of Business*. In B. Ridge (Ed.).
- Arens, A. A., Elder, R. J., Beasley, M. S., Valladares Franyuti, A. G., Jenkins, J. G., Tech, V., & Dominguez Moran, M. de L. (2007). *Auditoria : un enfoque integral*.
- Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). *Fraud Analytics Using Descriptive, predictive, and social Network Techniques - A Guide to Data Science for Fraud Detection*. Hoboken, New Jersey.: John Wiley & Sons, Inc.
- Banarescu, A. (2015). *Detecting and Preventing Fraud with Data Analytics*, 32(15), 1827–1836. [https://doi.org/10.1016/S2212-5671\(15\)01485-9](https://doi.org/10.1016/S2212-5671(15)01485-9)
- Bishop, T., Bloom, C., Carcello, J., Cotton, D., Daniels, H., Durkin, R., ... Zimbelman, M. (2007). *Managing the Business Risk of Fraud : A Practical Guide*, 1–80. Retrieved from [http://www.acfe.com/uploadedFiles/ACFE\\_Website/Content/documents/managing-business-risk.pdf](http://www.acfe.com/uploadedFiles/ACFE_Website/Content/documents/managing-business-risk.pdf)
- Bunget, O. C., Grigoroï, L., & Dumitrescu, A. C. (2009). *Detecting and Reporting the Frauds and Errors By the Auditor*. *Megatrend Review*, 6(1), 279–291. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=48415019&site=ehost-live>

- Comissão Europeia. (2006). A nova definição de PME - Guia do utilizador e modelo de declaração. *Empresas E Indústria*, 1–50. Retrieved from [http://www.pofc.qren.pt/ResourcesUser/2013/Publicacoes/Guia\\_Definicao\\_PME.pdf](http://www.pofc.qren.pt/ResourcesUser/2013/Publicacoes/Guia_Definicao_PME.pdf)
- Correia, T. (2017). Analisador Saft - Solução de Apoio à Auditoria. *OROC - Ordem Dos Revisores Oficiais de Contas*, pp. 35–37.
- Costa, C. B. (2010). *Auditoria Financeira-Teoria & Prática* (9ª). Letras e Conceitos, Lda.
- Cressey, D. R. (1953). Other People's Money. In Montclair: Patterson Smith (Ed.).
- Dal-Ri Murcia, F., Borba, J. A., & Schiehll, E. (2008). Relevância dos Red Flags na Avaliação do Risco de Fraudes nas Demonstrações Contábeis: a Percepção de Auditores Independentes Brasileiros. *Revista Universo Contábil*, 25–44.
- Deloitte. (2012). Adding insight to audit - Transforming Internal Audit through data analytics. Retrieved July 20, 2017, from <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/audit/us-aers-adding-insight-pov-mobile-061913.pdf>
- Dinapoli, T. P., & Hancox, S. J. (n.d.). Red Flags for Fraud. State of New York Office of the State Comptroller. [http://www.drillie.com/a541/articles/red\\_flags\\_fraud.pdf](http://www.drillie.com/a541/articles/red_flags_fraud.pdf)
- Dorminey, J., Fleming, S., Kranacher, M., & Riley, R. (2011). The evolution of fraud theory. *American Accounting Association Annual Meeting*, 1–58.
- Durtschi, C., Hillison, W., & Pacini, C. (2004). The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting, (January 2004).
- Gee, S. (2012). *Bringing Sophistication to Data Analytics*. [http://www.acfe.com/uploadedFiles/ACFE\\_Website/Content/asiapac/2012/presentations/ppt-7D-Sunder-Gee.pdf](http://www.acfe.com/uploadedFiles/ACFE_Website/Content/asiapac/2012/presentations/ppt-7D-Sunder-Gee.pdf)
- Gee, S. (2015). *Fraud and Fraud Detection A Data Analytics Approach*. (Wiley, Ed.).
- Hencsey, A. C. (2016). A definição, o processo de racionalização no triângulo da fraude e a complexidade de sua construção psicológica. *Observatório de Economia E Gestão de Fraude, Working Pa*.
- Hess, M. F., & Cottrell, J. H. (2016). Fraud risk management: A small business perspective. *Business Horizons*, 59(1), 13–18.

<https://doi.org/10.1016/j.bushor.2015.09.005>

- HM Treasury. (2012). Fraud and the Government Internal Auditor, (January).  
<http://www.kcgaudit.co.uk/wp-content/uploads/2016/01/Fraud-and-the-Government-Internal-Auditor-Treasury-2012.pdf>
- Hoaglin, D. C. (2003). John W. Tukey and Data Analysis. *Statistical Science*, 18(3), 311–318. <https://doi.org/10.1214/ss/1076102418>
- ICAEW. (2016). *Data analytics for auditors*. <https://www.icaew.com/-/media/corporate/files/technical/iaa/tecpln14726-iaae-data-analytics---web-version.ashx>
- IFAC. (2010). International Standard on Auditing 520: Analytical Procedures. *International Federation of Accountants: IFAC*, 433–440. Retrieved from <http://search.proquest.com.ezproxy.is.ed.ac.uk/docview/645249792?accountid=10673%0A>
- IFAC. (2012). ISA 240 - As Responsabilidades do Auditor Relativas a Fraude numa Auditoria de Demonstrações Financeiras, 1–26.
- IPAI. (2009). Enquadramento Internacional de Práticas Profissionais de Auditoria Interna. <https://doi.org/10.1590/S0100-204X2009000200009>
- Kaplan, J. (2012). 2012 Survey Report on Data Analysis Audit Software. *AuditNet*. [www.auditnet.org](http://www.auditnet.org)
- Kaplan, J. (2016). Turning Analytics From “Nice to Have” “Must Have”. [www.auditnet.org](http://www.auditnet.org)
- Kassem, R., & Higson, A. (2012). The New Fraud Triangle Model. *Journal of Emerging Trends in Economics and Management Sciences*, 3(No. 3), 191–195 MAZAR.
- Lowe, R. (2000). Benford’s law and fraud detection. *Chartered Accountants Journal of New Zealand*, 32–36.
- Mackevicius, J., & Lukas Giriunas. (2013). Transformational Research of The Fraud Triangle. *Ekonomika*, 92(4), 150–163.
- Marques, M. (1997). *Auditoria e Gestão*. (Editorial Presença, Ed.) (1ª EDIÇÃO).
- Marques, P. M. B. (2016). Técnicas de Análise de Dados ( Data Analytics ) no contexto de uma auditoria financeira. *OROC - Ordem Dos Revisores Oficiais de Contas*,

(Parte I).

- McNally, J. S. (2013). The 2013 COSO Framework & SOX Compliance: One Approach to an Effective Transition. *The Association of Accountants and Financial Professionals in Business*, 9. <https://doi.org/10.1111/j.1467-8330.1974.tb00606.x>
- Morais, G., & Martins, I. (2013). *Auditoria Interna - Função e Processo*. (Á. Editora, Ed.) (4th ed.).
- Murphy, P. R., & Dacin, M. T. (2011). Psychological Pathways to Fraud : Understanding and Preventing Fraud in Organizations, 99, 601–618. <https://doi.org/10.1007/s10551-011-0741-0>
- Nigrini, M. J. (1999). I've Got Your Number. *Journal of Accountancy*, 187(5), 79–83. Retrieved from <http://ezproxy.library.capella.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=1837306&site=ehost-live&scope=site>
- Nigrini, M. J. (2000). Continuous Auditing. *Ernst & Young Center for Auditing Research and Advanced Technology Accounting Workshop*, 1–23. <https://doi.org/10.1108/02686900110385605>
- Nigrini, M. J. (2011). Forensic Analytics Methods and Techniques for Forensic Accounting Investigations. In Wiley (Ed.).
- Nigrini, M. J., & L. I. Mittermaier. (1997). The Use of Benford's Law as an Aid in Analytical Procedures. *A Journal of Practice and Theory*, 16(2), 1–8.
- PATTON, M. Q. (2002). *Qualitative research & evaluation methods*. (Sage Publications, Ed.) (3rd ed.). London.
- Pedrosa, I. (2015). THESIS: Computer-assisted audit tools and techniques use: determinants for individual acceptance, (May). Retrieved from <https://repositorio.iscte-iul.pt/handle/10071/10017>
- Rezaee, Z. (2002). *Financial Statement Fraud Prevention and Detection*. (Wiley, Ed.). New York.
- Runkler, T. A. (2012). *Data Analytics - Models and Algorithms for Intelligent Data Analysis*. (Springer, Ed.). Germany. Retrieved from <http://www.springer.com/springer+vieweg/it+%26+informatik/datenbanken/book/>

978-3-8348-2588-9

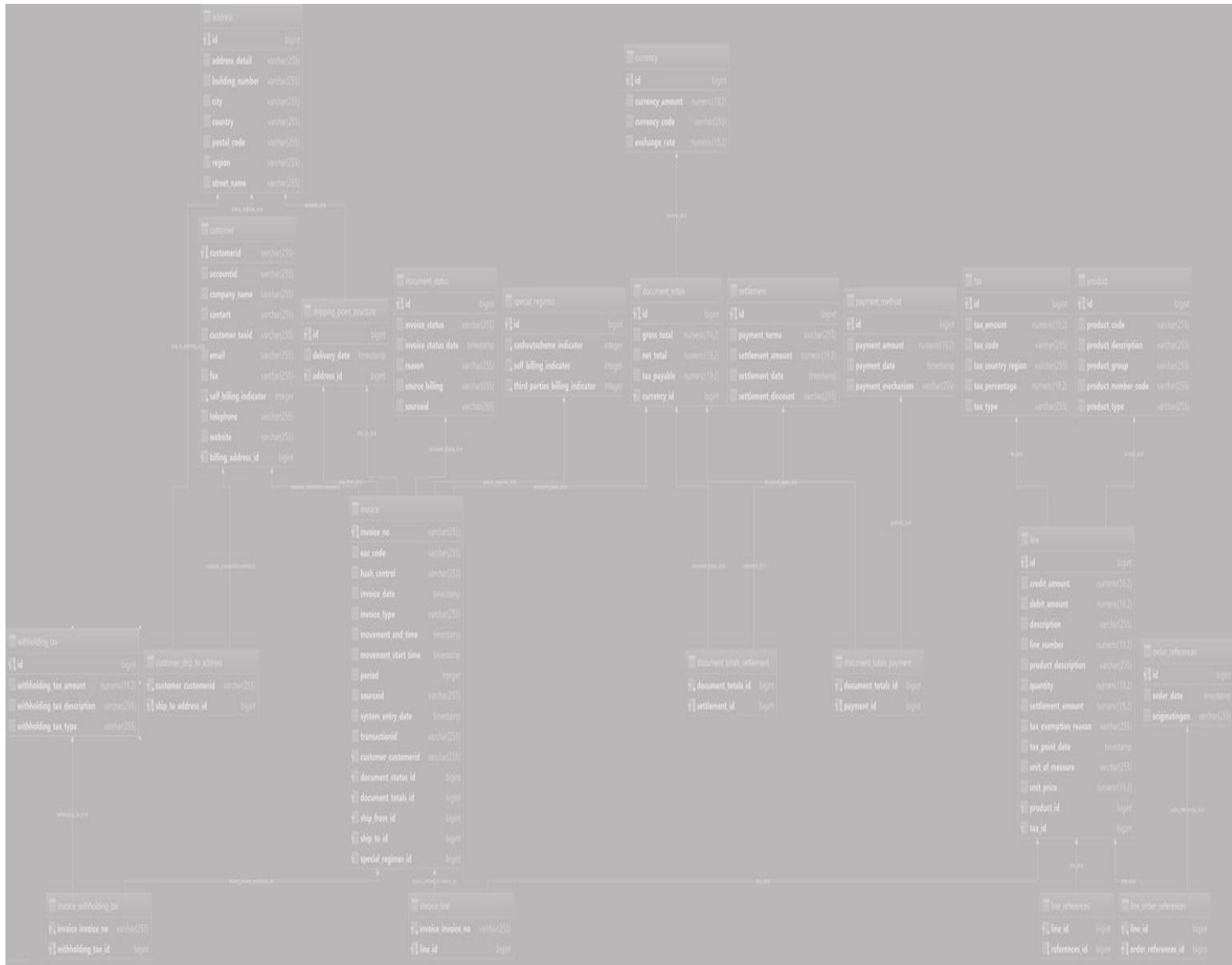
WELLS, J. T. (2009). *Manual da Fraude na Empresa - Prevenção e Detecção* (2a ed.).  
Edições Almedina. SA.

Wolfe, D., & Hermanson, D. R. (2004). The fraud diamond: Considering four elements  
of fraud. *The CPA Journal*, 74, 38–42.

YIN, R. K. (2005). *Estudo de caso: Planeamento e métodos*. (Bookman, Ed.) (3rd ed.).  
Porto Alegre.

## **APÊNDICES**

**APÊNDICE 1. Ligações entre tabelas extraídas do ficheiro SAFT**



## **APÊNDICE 2. Como é que se executa um Sumário no IDEA – Clientes por vendedor**

Neste caso este sumário foi executado a partir da tabela Clientes.

