**1 MENSURAÇÃO LEGIBILIDADE (LE)**

**1.1 Literatura LE**

A Legibilidade (LE) corresponde a complexidade do texto, concentra-se no texto para medir a dificuldade de compreensão, ou seja, a capacidade que o texto possui de permitir ao leitor de interpretar o seu significado de forma adequada (Telles & Salotti, 2021). Portanto, redigir parágrafos longos, com palavras extensas e de difícil compreensão, acaba desviando o real significado que se deseja transmitir nos relatórios financeiros (Gkikas *et al.*, 2022).

**1.2 Mensuração LE**

**(Moreno & Casasola, 2016)**

A *proxy* LE (variável independente de interesse), será coletada por meio das Notas Explicativas (NE) das Demonstrações Financeiras Padronizadas (DFP) das companhias. A complexidade das demonstrações contábeis será mensurada pela *proxy* LE, obtida por meio da ferramenta mais utilizada para estimar o nível de LE dos relatórios financeiros, o índice de Facilidade de Leitura de Flesch (FLF). Esta que mensura o nível de facilidade linguística de um texto e é considerada como uma das técnicas mais precisas, confiáveis e práticas para essa finalidade (Holtz & Santos, 2020).

O índice FLF será obtido com o auxílio do *software* Microsoft Word 2010, uma vez que este possui uma função de revisão do texto que fornece resultados da FLF. Para a mensuração do índice FLF os arquivos com extensão em *.pdf* serão convertidos para o formato *.doc* do Microsoft Word, sendo as tabelas, cabeçalhos e figuras, excluídas dos arquivos, para não provocar distorções no referido índice (Moreno & Casasola, 2016). A fórmula de facilidade de leitura do índice FLF, considera o comprimento da palavra medido em sílabas e o comprimento da frase medido em número de palavras. O fator palavra mede a dificuldade semântica e a velocidade de reconhecimento, já o fator frase mede a capacidade de memória de curto prazo.

A fórmula do índice FLF é especificada a seguir, conforme a Equação 1.

FLF = 206,835 – (84,6 x CP) – (1,015 x CF) (Equação 1)

Em que:

FLF = é a medida de facilidade de leitura: medido pelo índice FLF;

CP = comprimento da palavra: medido pelo número de sílabas dividido pelo número de palavras; e

CF = comprimento da frase: medido pelo número de palavras dividido pelo número de frases.

Quanto maiores as palavras e as frases - pior a legibilidade (compreensão do texto). O FLF aplicada a fórmula, de 100 em 100 palavras.

Os valores são mensurados numa escala de 0 a 100, classificando o texto, por nível de facilidade de leitura, do nível de legibilidade mais difícil para o mais fácil, conforme a Tabela 1.

Tabela 1

**Níveis de Legibilidade do Índice Flesch**

|  |  |
| --- | --- |
| **Pontuação do FLF** | **Nível de Legibilidade** |
| 00 – 25 | Muito Difícil |
| 25 – 50 | Pouco Difícil |
| 50 – 75 | Fácil |
| 75 – 100 | Muito Fácil |

**Fonte**: Martins *et al.* (1996).

**(Gomes, Ferreira, Martins, 2018)**

As empresas foram analisadas quanto à variação do volume de NEs por meio da diferença do número de páginas, caracteres, palavras, sílabas, frases e parágrafos contidos nas notas explicativas. Para a coleta desses dados, utilizaram-se softwares criados pelos próprios autores combinando funções das seguintes bibliotecas para Python: PyPDF2, Pdf2Txt e regex.

Criou-se um programa para contar as páginas dos arquivos em “.pdf” e exportar os  
resultados para uma tabela com o nome de cada empresa, o ano de referência das notas  
explicativas e o número de páginas de cada um dos arquivos. Enquanto que para a conferência da variação no número de palavras foi necessário extrair o texto de cada arquivo em “.pdf” e salvá-lo em um arquivo “.txt”. Como o tamanho das letras, o espaçamento entre elas ou outro artifício visual podem influenciar a quantidade de páginas, sem necessariamente afetar a quantidade de informações, é importante destacar que não foram encontradas diferenças nos layouts das notas explicativas de uma mesma empresa nos dois anos analisados. Portanto, uma variação positiva (ou negativa) de páginas de fato indica aumento (ou redução) de informação divulgada.

Antes do cálculo do índice de Flesch, utilizou-se a ferramenta Character Count Tool,  
disponível sem custos no sítio eletrônico com o mesmo nome da ferramenta, para coletar os dados quanto ao número de caracteres, sílabas, frases e parágrafos. Ao realizar upload das notas explicativas em formato “.pdf” na ferramenta em questão, ela fornece um relatório completo contendo todos esses dados, entre outros. Essa ferramenta também fornece o tempo estimado de leitura do documento de texto, calculado com base na velocidade média de leitura de um indivíduo que é de aproximadamente 200 palavras por minuto – variável também utilizada neste trabalho.

Este trabalho utiliza como dados o número de páginas, caracteres, palavras, sílabas,  
sentenças e parágrafos e, como indicadores, utiliza-se o tempo estimado de leitura (em  
minutos), a quantidade de palavras sobre o número de parágrafos, a quantidade de palavras sobre o número de frases e a quantidade de sílabas dividida pelo número de palavras. Todos esses dados e indicadores representam o tamanho (rapidez de leitura) das notas explicativas. Já o índice de legibilidade de Flesch é um indicador de facilidade ou qualidade de leitura.

**2 PROPOSTA DE MENSURAÇÃO DA LEGIBILIDADE PARA A TESE**

Mensurar a legibilidade pelo Python:

CP = comprimento da palavra: medido pelo número de sílabas dividido pelo número de palavras; e

CF = comprimento da frase: medido pelo número de palavras dividido pelo número

E após coletar esses dados, inseri-los na fórmula de FLF para mensurar a LE.