



PUC Minas

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

NÚCLEO DE EDUCAÇÃO A DISTÂNCIA

Pós-graduação *Lato Sensu* em Analytics e Business Intelligence

RELATÓRIO TÉCNICO

LENTES FOTOGRÁFICAS

DECISÃO DE COMPRA BASEADA EM DADOS

Alessandro Lemser

Belo Horizonte

2021

SUMÁRIO

1. Introdução	4
1.1. Contexto	4
1.2. Objetivos.....	5
1.3. Público alvo	5
2. Modelo de Dados.....	5
2.1. Modelo Dimensional.....	5
2.2. Fatos e Dimensões.....	6
3. Integração, Tratamento e Carga de Dados.....	7
3.1. Fontes de Dados	7
3.2. Processos de Integração e Carga (ETL).....	8
4. Camada de Apresentação.....	9
4.1 Dashboard	9
4.2 Análises avançadas.....	9
5. Registros de Homologação	9
5. Conclusões	9
6. Links	9
REFERÊNCIAS.....	10

1. Introdução

1.1. Contexto

Desde a transição do filme fotográfico para o sensor digital, profissionais, entusiastas e amadores na área da fotografia vêm atravessando transformações digitais. A fotografia ficou acessível para mais pessoas tornando-se popular. Muito disso por causa dos smartphones, mas vale ressaltar também a facilidade com que amadores e entusiastas da fotografia encontraram para manusear câmeras profissionais e semi-profissionais. O uso de inteligência artificial em fotografia também tem ajudado a tornar mais simples e acessível o uso de câmeras como a DSLR (Digital Single Lens Reflex) usadas por profissionais. Como exemplos temos o aprimoramento da detecção de cenas e ajuste automático de valores de exposição, tipos específicos de foco para a face, para os olhos entre outras. Isso faz com que mais pessoas se aventurem a usar equipamentos sofisticados e ainda obtenham bons resultados com moderado esforço. Com isso, uma demanda por aquisição de lentes também é esperada, visto que DSLRs são câmeras de lentes intercambiáveis e existe uma grande variedade de lentes, algumas mais indicadas para um determinado motivo fotográfico e outras para motivos em geral.

O mercado de câmeras profissionais e semi-profissionais com lentes intercambiáveis [teve seu auge](#) no ano de 2012 e manteve-se num bom patamar até 2019. Em 2020, em razão da pandemia, as vendas caíram bastante. A indústria fotográfica precisa recuperar o fôlego e uma esperança é uma recente transformação digital: as câmeras *mirrorless*. O mercado vêm demandando, há algum tempo, questões relacionadas com as dimensões do equipamento, buscando câmeras e lentes mais leves e menores, sem comprometer a performance. As câmeras DSLR mantiveram basicamente a mesma estrutura das antigas SLR (Single Lens Reflex) baseadas em filme fotográfico. Desta forma, profissionais e entusiastas puderam continuar utilizando as lentes que já possuíam do sistema antigo. Porém, um novo tipo de câmera chamada *mirrorless* está transformando novamente a área da fotografia. Este tipo de câmera não usa um espelho para refletir a imagem que entra pela lente para um visualizador óptico como nas DSLRs e, com isso, reduz de forma significativa o tamanho do equipamento. Mas só

recentemente, em 2019, que grandes fabricantes como Canon e Nikon, seguindo o pioneirismo da Sony, lançaram câmeras *mirrorless* para profissionais e entusiastas.

Este novo formato tem um grande potencial de alavancar venda de lentes, pois as novas cameras exigem um novo encaixe para lentes onde não será possível encaixar as lentes já existentes sem que um adaptador seja usado. A tendência é que os profissionais e entusiastas comprem novas lentes para o novo sistema, pois existem fatores que melhoram o desempenho e as tornam menores.

O ano de 2020 foi desafiador para a indústria que viu perdas perto da casa dos 50%. As cameras DSLR estão cada vez vendendo menos e existe uma [demanda crescente](#), e mais rentável, no formato *mirrorless*. É esperado um aumento nas vendas para o novo formato. Tanto profissionais como entusiastas terão de analisar as melhores alternativas de acordo com o tipo de fotografia que fazem.

1.2. Objetivos

Este projeto visa auxiliar profissionais e entusiastas da área da fotografia a tomar uma decisão de compra de lentes mais informada. Com base em dados de fotografias existentes em diferentes categorias, tais como esportes casamentos e retratos, poderá se chegar a conclusão da melhor compra, àquela que trará o melhor custo benefício e aplicabilidade.

1.3. Público alvo

Esta solução será usada por profissionais da fotografia. Ainda, a solução visa atender amadores e entusiastas que consomem equipamentos fotográficos.

Os profissionais precisam levar muito a sério o seu equipamento e, diante de uma nova tecnologia, poderão ter a chance de atualizar o seu equipamento tomando uma decisão de compra baseada em dados.

2. Modelo de Dados

2.1. Modelo Dimensional

A seguir estão descritas as perguntas que este trabalho busca responder, levando em consideração o público alvo supramencionado:

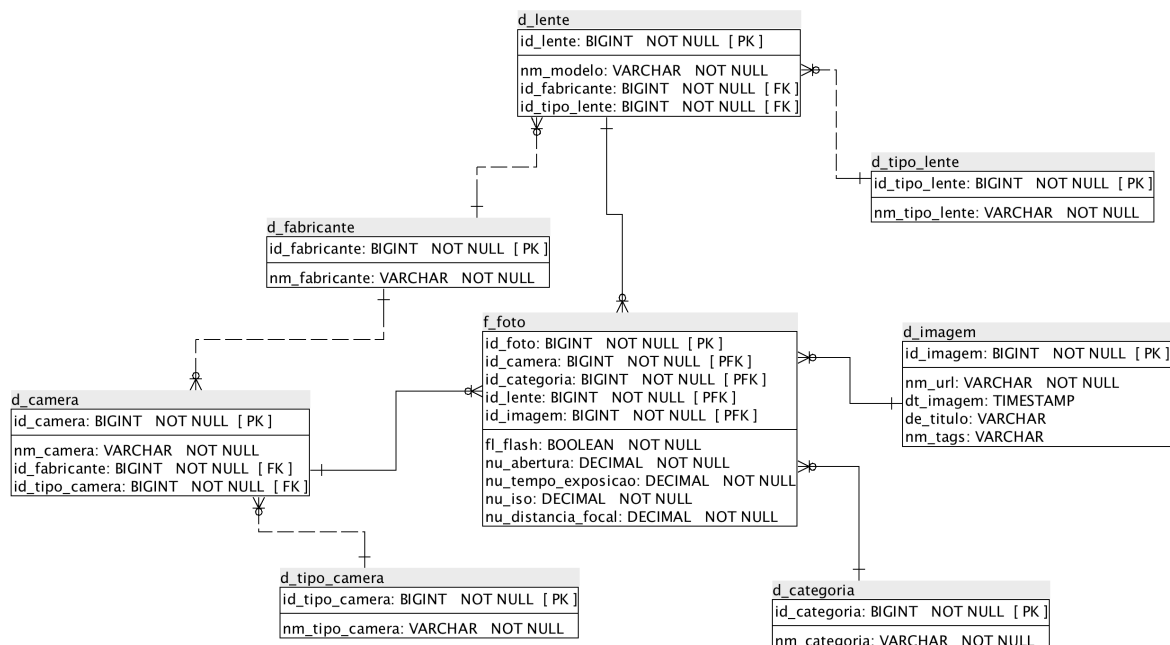
1. Qual a distância focal mais utilizada?

2. Qual tipo de lente é mais utilizada? (distância focal variável (zoom) ou fixa) no total e por categoria?
3. Qual a lente e a distância focal mais utilizada para esportes?
4. Qual a lente e a distância focal mais utilizada para retratos?
5. Qual a lente e a distância focal mais utilizada para natureza?
6. Qual a lente e a distância focal mais utilizada para casamentos?

Com base nas respostas acima, um profissional ou entusiasta poderá pensar a respeito de qual a lente mais utilizada por categoria qual a distância focal mais utilizada por categoria. Dessa forma, poderá realizar uma compra mais assertiva de equipamento. A escolha do equipamento que trará o maior custo benefício é importante tendo em vista os custos envolvidos e, no caso de profissionais, nos lucros obtidos com as fotos.

2.2. Fatos e Dimensões

Aqui são descritas a tabelas de fato e as dimensões definidas para o projeto.

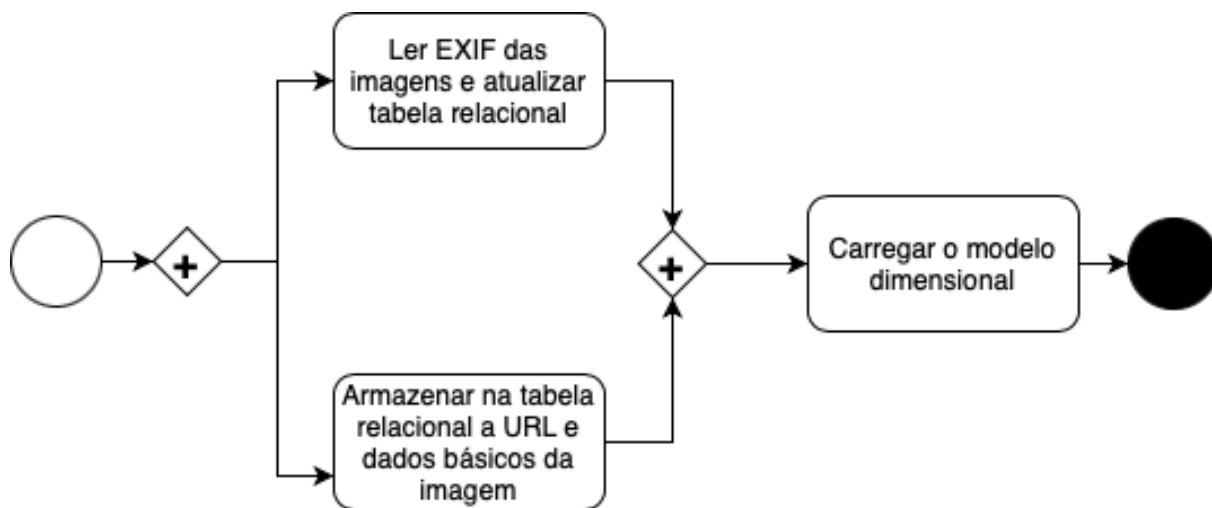


3. Integração, Tratamento e Carga de Dados

3.1. Fontes de Dados

A fonte de dados para obtenção das informações de EXIF das imagens, de onde serão obtidos os dados chave para a tomada de decisão, são as imagens marcadas como públicas no Flickr. Após a leitura das informações via API do Flickr, a seguinte tabela será povoada no banco de dados relacional.

t_fotografias	
id_fotografia:	BIGINT NOT NULL [PK]
de_titulo:	VARCHAR
nm_url:	VARCHAR NOT NULL
dt_coleta:	TIMESTAMP NOT NULL
tp_lente:	VARCHAR
nm_lente:	VARCHAR
nm_fabric_lente:	VARCHAR
nu_dist_focal:	DECIMAL
tp_camera:	VARCHAR
nm_camera:	VARCHAR
nm_fabric_camera:	VARCHAR
nm_categoria_foto:	VARCHAR
nu_abertura:	DECIMAL
nu_tempo_exposicao:	DECIMAL
nu_iso:	DECIMAL
fl_flash:	BOOLEAN
nm_tags:	VARCHAR NOT NULL
nu_dist_focal_35mmEq:	DECIMAL
fl_falha_exif:	BOOLEAN



3.2. Processos de Integração e Carga (ETL)

O processo de integração e carga de dados é realizado por meio de scripts Python ([ver GitHub](#)). O objetivo é fazer com que seja fácil a preparação do ambiente e execução do processo pelos avaliadores deste projeto ou pessoas interessadas em rodar por conta própria sem a necessidade de adquirir produtos ou lidar com ferramentas de integração.

Acima é possível ver as atividades relacionadas com o processo de ETL. As duas primeiras executam em paralelo. Ao final, o modelo dimensional é atualizado.

Todas as informações necessárias para a execução dos scripts de ETL estão disponibilizadas no arquivo **README.md** no [repositório GitHub](#) do projeto.

Serão lidas ao total vinte mil imagens (20,000) distribuídas em quatro categorias: Esporte, casamento, Natureza e retratos. Este limite pode ser aumentado ou o script pode ser executado diariamente obtendo novas fotos públicas disponíveis no Flickr.

Identificação de categorias

O processo categoriza as imagens de acordo com rótulos (*tags*) adicionada pelos próprios fotógrafos (dimensão *d_categoria*). Por exemplo, uma foto de casamento pode conter rótulos como *wedding* e *bride*. Nesta primeira versão, o processo se limitará a utilizar estes rótulos para categorizar a imagem. Entretanto, é possível também expandir o escopo numa próxima versão e utilizar visão computacional para identificar elementos na imagem a fim de reconhecer a cena e categorizar a imagem com maior precisão, visto que os rótulos adicionados pelos fotógrafos podem não representar exatamente a categoria.

4. Camada de Apresentação

4.1 Dashboard

A ser definido.

4.2 Análises avançadas

A ser definido.

5. Registros de Homologação

A ser definido.

5. Conclusões

A ser definido.

6. Links

- GitHub: <https://github.com/alemser/tcc-puc-bia>

REFERÊNCIAS