

**CODERHOUSE - BRASIL  
DATA ANALYTICS**

**ISADORA FERRAZ E FIGUEIREDO  
VINÍCIUS ROCHA CAMARGO**

**WORLD METRICS**

Projeto Final apresentado ao Curso de  
Data Analytics da Instituição CoderHouse  
como parte dos requisitos para a aprovação.

**Professor: Lucas Franco**

**Brasil  
2024**

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>3</b>
<b>2</b>	<b>METODOLOGIA</b>	<b>4</b>
<b>3</b>	<b>CASE</b>	<b>5</b>
3.1	DESAFIO	5
3.2	ANÁLISE DE DADOS	6
3.3	RESULTADOS	6
3.4	RANKING DOS PAÍSES	6
3.5	DASHBOARD	6
<b>4</b>	<b>TRATAMENTO DOS DADOS</b>	<b>9</b>
<b>5</b>	<b>MEDIDAS</b>	<b>12</b>
<b>6</b>	<b>PRÓXIMOS PASSOS</b>	<b>13</b>
	<b>REFERÊNCIA BIBLIOGRÁFICA</b>	<b>14</b>

# 1 INTRODUÇÃO

As migrações humanas são um fenômeno antigo e fundamental na história da humanidade, remontando às primeiras migrações pré-históricas há cerca de 200 mil anos com o *Homo sapiens*. A migração refere-se à deslocação de pessoas de um lugar para outro, comumente associada à busca de melhores condições econômicas, políticas ou sociais. É importante destacar que a migração pode ter diversas causas, como questões ecológicas, políticas, bélicas ou econômicas.

A Metrics World é uma startup que visa simplificar a escolha do país ideal para cada pessoa. Por meio de análise de dados, a empresa traça o perfil individual de cada cliente e identifica países que encaixam com suas características.

Nos capítulos 2 e 3, descreveremos o processo que a empresa utiliza para alcançar os resultados desejados, enquanto os capítulos 4, 5 e 6 abordarão aspectos mais técnicos do estudo.

## 2 METODOLOGIA

A Metrics World, acredita que a escolha do país ideal é uma jornada pessoal e única. Por isso, é utilizada uma metodologia abrangente e personalizada para te guiar nessa importante decisão.

### 1. Coleta de dados:

Primeiramente, são coletados dados demográficos, econômicos, sociais de diversos países, utilizando fontes confiáveis como instituições governamentais e organizações internacionais.

### 2. Desvendando suas preferências:

Em seguida, o cliente responde o formulário. Assim é possível, que sua preferência seja avaliada. E o peso de cada fator que é buscado em um novo país.

### 3. Análise de dados a serviço da sua felicidade:

A análise dos dados e respostas do formulário resulta no desenvolvimento de um modelo matemático, onde o peso de cada variável é determinado pelas respostas do questionário. Então, é elaborado um ranking personalizado dos países que se encaixam ao perfil do cliente.

### 4. Uma parceria para o sucesso:

Para continuar aperfeiçoando o serviço, avaliações periódicas da metodologia é realizada, com base no feedback dos clientes, ajustando o modelo e a plataforma para oferecer um serviço cada vez mais preciso e útil.

## 3 CASE

### 3.1 DESAFIO

Para descrever o processo da Metrics World, inicialmente foram analisadas as preferências do cliente por meio de um formulário. Em seguida, foram utilizados dados do Kaggle que seguem esses interesses, incluindo o Índice de Desenvolvimento Humano (IDH), Produto Interno Bruto (PIB) per capita, Índice de Felicidade e o consumo per capita de bebidas alcoólicas em diversos países. O objetivo foi identificar possíveis relações entre essas variáveis e, se existentes, analisar como se relacionam.

É importante destacar que esses dados do Kaggle foram utilizados apenas para exemplificar o trabalho da nossa empresa.

Todas as conclusões apresentadas serão respaldadas pelas imagens do dashboard, as quais estão disponíveis no final deste capítulo.

**World Metrics - encontre seu lugar no mundo**

A escolha de um país para viver, trabalhar ou estudar é uma decisão importante. Este formulário foi criado para ajudar você a encontrar o país que melhor se encaixa na sua busca por qualidade de vida.

**Instruções:**

Preencha o formulário com **atenção**, indicando o quão importante cada fator é para você. As respostas serão utilizadas para gerar um perfil personalizado e te auxiliar na sua busca pelo país ideal.

---

**Índice de Felicidade \***

☐ 3: Muito importante (positiva)

☐ 2: Importante (positiva)

☐ 1: Pouco importante (positiva)

☐ 0: Neutro

☐ -1: Pouco importante (negativa)

☐ -2: Importante (negativa)

☐ -3: Muito importante (negativa)

Figura 1 – Parte do formulário

Você pode acessar o formulário clicando [aqui](#).

## 3.2 ANÁLISE DE DADOS

A empresa coletou dados sobre o IDH, PIB, Índice de Felicidade e consumo per capita em litros de bebidas alcoólicas em diferentes países. Os dados foram preparados e analisados para identificar padrões e correlações.

## 3.3 RESULTADOS

Após a análise, foi identificado que o IDH, PIB e Índice de Felicidade estão correlacionados positivamente, ou seja, quanto maior o IDH e PIB de um país, maior tende a ser o seu Índice de Felicidade. No entanto, o consumo de bebidas alcoólicas não apresentou uma correlação significativa com essas variáveis.

## 3.4 RANKING DOS PAÍSES

Após analisar a relação entre os dados e obter o formulário preenchido pelo cliente, são criadas medidas para classificar os países de acordo com o perfil do cliente. Isso é feito calculando a média do IDH, Índice de Felicidade e PIB, e subtraindo o valor do consumo de álcool, resultando em um ranking que permite ao cliente analisar de forma quantitativa quais países melhor se encaixam em seu perfil.

## 3.5 DASHBOARD

Por fim é feito um Dashboard, para que o cliente tenha acesso a todos os dados e análises.



Figura 2 – Sumário Dashboard

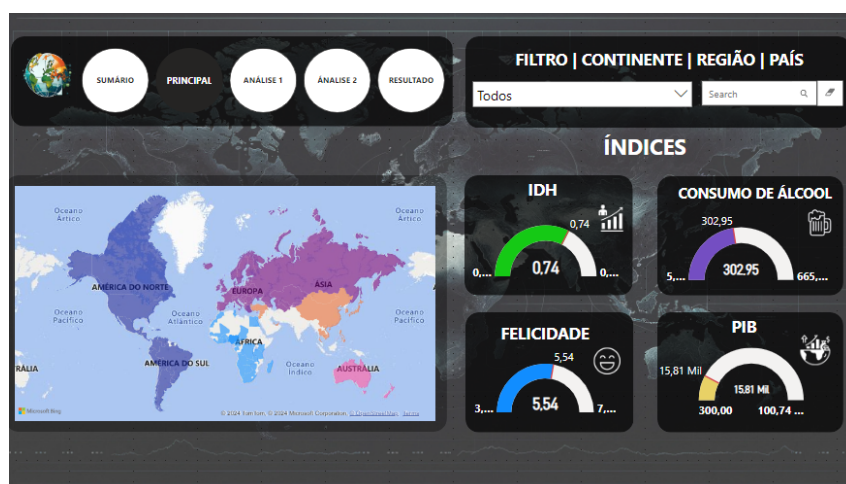


Figura 3 – Principal Dashboard

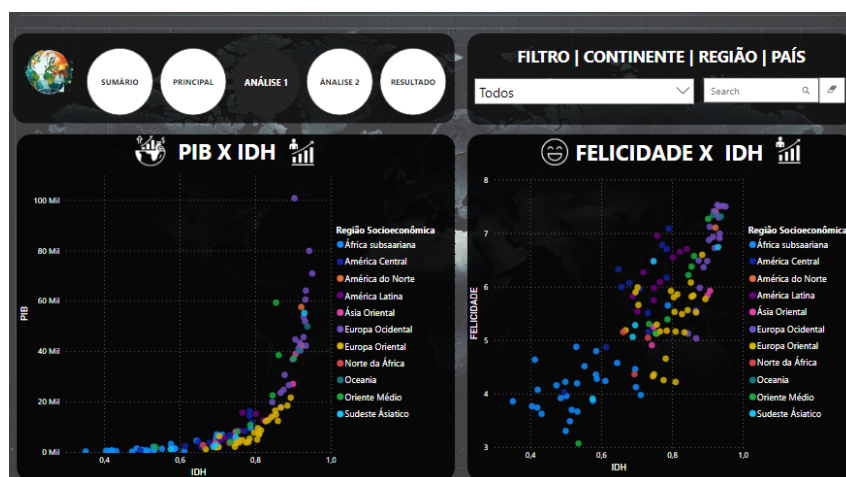


Figura 4 – Análise 1 Dashboard



Figura 5 – Análise 2 Dashboard

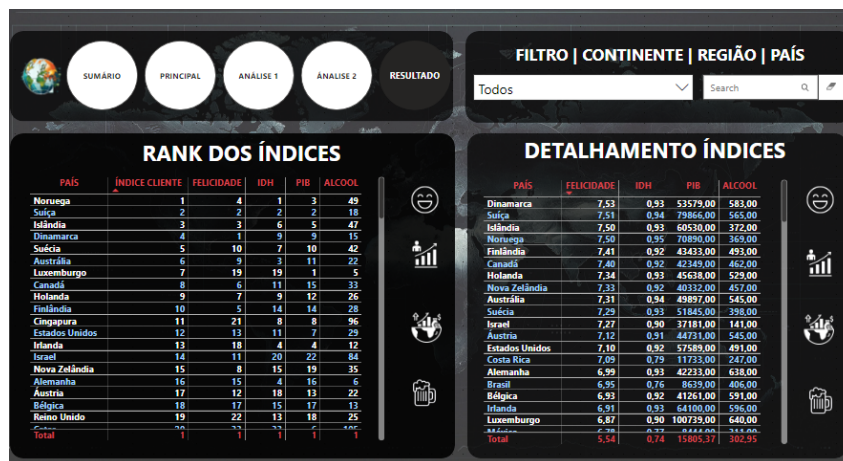


Figura 6 – Resultados Dashboard



## 4 TRATAMENTO DOS DADOS

Ao analisar os dados extraídos da plataforma Kaggle, fora percebido que era necessário realizar alguns ajustes.

- As colunas de índice de felicidade e IDH foram divididas por 1000 para retificar;
- A coluna dos países fora traduzida no google planilhas, porém algumas traduções ficaram literais, por serem poucas traduções foi possível verificar linha por linha;
- A coluna de hemisfério e região socioeconômica fora traduzida, porém utilizando a função substituir valores;
- Para facilitar a análise de consumo de álcool, foram somados o consumo de destilados, vinho e cerveja;
- Para realizar uma análise por continente, fora criada uma coluna e feita substituição de valores;

Índice da Felicidade	IDH
4655	782
6994	934
3866	577
6650	822
4360	749
7313	938
5291	757
6218	846
5956	709
3484	512
5802	805

Figura 7 – Índice de Felicidade e IDH sem ajuste

1.2 Felicidade Ajustada	1.2 IDH ajustado
4,655	0,782
6,994	0,934
3,866	0,577
6,65	0,822
4,36	0,749
7,313	0,938
5,291	0,757
6,218	0,846
5,956	0,709
3,484	0,512

Figura 8 – Índice de Felicidade e IDH pós ajuste

B87     $\text{fx}$  =GOOGLETRANSLATE(A87)

	A	B
87	Turkey	Peru
88	Mali	Eles tinham
89	Togo	Ir
90	Suriname	Suriname
91	Qatar	Catar
92	Kyrgyzstan	Quirguistão
93	Azerbaijan	Azerbaijão
94	United Arab Emirates	Emirados Árabes Unidos
95	Mexico	México
96	Madagascar	Madagáscar
97	Zimbabwe	Zimbábue
98	Cameroon	Camarões

Figura 9 – Tradução no Google planilhas

$A_C^B$ Região	$A_C^B$ Hemisfério
Europa Oriental	norte
Europa Ocidental	norte
África subsaariana	sul
América Latina	sul
Europa Oriental	norte
Oceania	sul
Europa Oriental	norte
Oriente Médio	norte
América Central	norte
África subsaariana	norte
Europa Oriental	norte
América Latina	sul
África subsaariana	sul
América Latina	sul

Figura 10 – Coluna região socioeconômica e hemisfério pós tradução

$t_3^2$ Cerveja_PerCapita	$t_3^2$ Destilado_PerCapita	$t_3^2$ Vinho_PerCapita	1.2 Soma Bebidas
89	132	54	275
346	117	175	638
217	57	45	319
193	25	221	439
21	179	11	211
261	72	212	545
21	46	5	72
42	63	7	112
263	114	8	385
34	4	13	51
142	373	42	557
167	41	8	216
173	35	35	243
245	145	16	406

Figura 11 – Soma Bebidas criada a partir da soma das colunas

## 5 MEDIDAS

Com intuito de analisar individualmente cada dado, foi empregada a função RANKX no DAX.

```
posicao_alcool_1 = RANKX(ALL(Felicidade[País]), Medidas[Soma_Alcool])
```

Figura 12 – Função RANKX para consumo de álcool em litros per capita

```
posicao_pib_1 = RANKX(ALL(Felicidade[País]), Medidas[Soma_PIB])
```

Figura 13 – Função RANKX para PIB per capita

```
posicao_idh_1 = RANKX(ALL(Felicidade[País]), Medidas[IDH_medida])
```

Figura 14 – Função RANKX para IDH

```
posicao_felicidade_1 = RANKX(ALL(Felicidade[País]), Medidas[Felicidade_medida])
```

Figura 15 – Função RANKX para Felicidade

A fim de, efetuar a soma das variáveis de maneira eficaz, todas foram convertidas em porcentagem. Como o IDH e o índice de felicidade têm intervalos definidos, o cálculo da porcentagem foi direto. Entretanto, o consumo de álcool e do PIB de todos os países foram somados e, em seguida, encontrado o percentual do país em relação ao mundo.

```
Soma_PIB = SUM(Felicidade[PIB_PerCapita])
```

Figura 16 – PIB mundial

```
Porcentagem_PIB_1 =  
    DIVIDE(  
        SUM('Felicidade'[PIB_PerCapita]),  
        CALCULATE(SUM('Felicidade'[PIB_PerCapita]), ALL('Felicidade'))  
    )
```

Figura 17 – Porcentagem do PIB de cada país em relação a soma mundial

A mesma lógica fora utilizada para o consumo de álcool.

Então, realizou-se uma média ponderada das variáveis correlacionadas, utilizando como pesos as respostas do formulário, e subtraiu-se o valor das variável não correlacionada.

Por fim, fora feito um rank personalizado com os valores encontrados

```

indice_cliente = DIVIDE(SUMX( Felicidade,
    'Medidas'[porcentagem_felicidade_1]) + 2*SUMX(Felicidade, 'Medidas'[IDH_medida]) + (3*SUMX(Felicidade,
    Medidas[Porcentagem_PIB_1])), 3) - SUMX('Felicidade', 'Medidas'[Porcentagem_alcool_1])

```

Figura 18 – Índice do cliente

```

posicao_indice_cliente_1 = RANKX(ALL(Felicidade[País]), Medidas[indice_cliente])

```

Figura 19 – Rank do cliente

## 6 PRÓXIMOS PASSOS

A World Metrics está constantemente em busca de aprimorar seus serviços e oferecer aos clientes uma experiência cada vez mais completa e personalizada. Nesse sentido, algumas melhorias estão sendo planejadas para aperfeiçoar a análise de dados e as recomendações feitas aos clientes.

Uma das melhorias em curso é a ampliação dos indicadores considerados na análise. Além dos dados atualmente utilizados, a empresa está buscando incluir novos indicadores, como clima, índice de violência e sistema de saúde. Também está será realizada a comparação desses dados com os anos anteriores, a fim de avaliar a estabilidade nos aspectos mencionados.

## REFERÊNCIA BIBLIOGRÁFICA

- [1] Flaticon.  
*<https://www.flaticon.com/br>*
- [2] Kaggle.  
*<https://www.kaggle.com>*
- [3] Microsoft - DAX Overview.  
*<https://learn.microsoft.com/pt-br/dax/dax-overview>*
- [4] Blue Willow AI - Start Here.  
*<https://www.bluewillow.ai/start-here/pt>*