

# Impacto do consumo de álcool em Diagnostico de Diabetes Mellitus

**Felipe Barbosa Hollerbach**

Pontificia Universidade Catolica de  
Minas Gerais  
Belo Horizonte, Brasil  
fbhollerbach@sga.pucminas.br

**Felipe Viana Carvalhais**

Pontificia Universidade Catolica de  
Minas Gerais  
Belo Horizonte, Brasil  
fcarvalhais@sga.pucminas.br

**Matheus Augusto Moreira**

Pontificia Universidade Catolica de  
Minas Gerais  
Belo Horizonte, Brasil  
matheus.moreira.1001498@sga.pucminas.br

**Rithie Natan Carvalhaes Prado**

Pontificia Universidade Catolica de  
Minas Gerais  
Belo Horizonte, Brasil  
rithie.prado@sga.pucmias.br

**Thiago Soares Mota**

Pontificia Universidade Catolica de  
Minas Gerais  
Belo Horizonte, Brasil  
thiago.mota@sga.pucminas.br

## ABSTRACT

Trabalho feito na matéria de Introdução a Sistemas Inteligentes com o intuito de aprender e aprimorar os conhecimentos referente a Ciência de Dados. Para desenvolvimento do trabalho foram necessários vários passos, dentre eles, o primeiro a coleta do dataset. Muitos dos dados coletados eram desnecessários, portanto houve um pré-processamento, limpeza e redução, para evitar consultas desnecessárias e aumentar a acurácia na correlação entre álcool e diabetes mellitus.

## KEYWORDS

Diabetes Mellitus, Ciência de Dados

## 1 INTRODUÇÃO

A Diabetes Mellitus é uma doença crônica caracterizada por níveis elevados de glicose no sangue. Existem 4 tipos de diabetes: Tipo 1 está relacionado a pouca ou nenhuma produção de insulina pelo pâncreas, Tipo 2 afeta a forma como o corpo processa o açúcar, Pre-diabetes consiste no excesso de açúcar no sangue e se não tratado, tornará do Tipo 2, Diabetes Gestacional ligado a resistência a insulina no corpo devido aos hormônios da placenta.

A diabetes pode levar a vários problemas de saúde, tem maiores riscos a infecções, assim como pode causar: cegueira, insuficiência renal, membros amputados entre outros. Por

ser tratável através da manutenção de um estilo de vida saudável complementado pela aplicação de insulina sintética quando possível, a realização de diagnóstico de diabetes no estágio mais precoce possível é essencial para garantir a sobrevivência daqueles acometidos pela doença. O uso de Ciência de Dados têm se provado cada vez mais eficazes em diagnósticos de doenças como a diabetes devido à capacidade de análise de padrões e massas de dados de grandes tamanhos, sendo uma importante aliada dos médicos na prevenção das complicações da diabetes.

Sendo assim, para o desenvolvimento do trabalho foram necessários coletar a base de dados do IBGE onde há varios dados sobre diversas doenças e hábitos, muitos dos dados coletados eram desnecessários, portanto houve um processamento, limpeza, redução e substituição dos valores faltantes e portanto a relação de diagnósticos de Diabetes Mellitus impactada pelo consumo de álcool.

## 2 METODOLOGIA

### Atributos da base de dados

A base inicial tinha mais de 900 atributos para serem escolhidos, foi feita uma préseleção de algumas variáveis e após a eliminação de variáveis com dados ausentes, dentre as existentes na base de dados, utilizamos as seguintes variáveis:

- C6 - Sexo
- C8 - Idade
- I1 - tem algum plano de saúde (médico ou odontológico), particular, de empresa ou órgão público?
- P1 - O(A) sr(a) sabe seu peso? (mesmo que seja valor aproximado)
- P2 - Quanto tempo faz que o(a) sr(a) se pesou da última vez?
- P9 - Em quantos dias da semana, o(a) sr(a) costuma comer verdura ou legume cozido, como couve, cenoura,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*Inteligência Artificial '21, 19 de Maio de 2021, Belo Horizonte, MG*

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-9999-9/18/06...\$15.00

<https://doi.org/10.1145/1122445.1122456>

chuchu, berinjela, abobrinha? (sem contar batata, mandioca ou inhame)

- P16 - Em quantos dias da semana o(a) sr(a) costuma tomar suco de frutas natural?
- P18 - Em quantos dias da semana o(a) sr(a) costuma comer frutas?
- P20 - Em quantos dias da semana o(a) sr(a) costuma tomar refrigerante ou suco artificial?
- P25 - Em quantos dias da semana o(a) sr(a) come alimentos doces, tais como pedaços de bolo ou torta, doces, chocolates, balas, biscoitos ou bolachas doces?
- P26 - Em quantos dias da semana o(a) sr(a) substitui a refeição do almoço ou jantar por sanduíches, salgados ou pizzas?
- P27 - Com que frequência o(a) sr(a) costuma consumir alguma bebida alcoólica?
- P28 - Quantos dias por semana o(a) sr(a) costuma tomar alguma bebida alcoólica?
- P29 - Em geral, no dia que o(a) sr(a) bebe, quantas doses de bebida alcoólica o(a) sr(a) consome? (1 dose de bebida alcoólica equivale a 1 lata de cerveja, 1 taça de vinho ou 1 dose de cachaça, whisky ou qualquer outra bebida alcoólica destilada)
- P31 - Quantos anos o(a) sr(a) tinha quando começou a consumir bebidas alcoólicas?
- P34 - Nos últimos três meses, o(a) sr(a) praticou algum tipo de exercício físico ou esporte? (não considere fisioterapia)
- P35 - Quantos dias por semana o(a) sr(a) costuma praticar exercício físico ou esporte?
- P37 - Em geral, no dia que o(a) sr(a), pratica exercício ou esporte, quanto tempo dura esta atividade?
- P41 - Quanto tempo o(a) sr(a) gasta, por dia para percorrer este trajeto a pé ou de bicicleta, considerando a ida e a volta do trabalho?
- P45 - Em média, quantas horas por dia o(a) sr(a) costuma ficar assistindo televisão?
- P46 - Perto do seu domicílio, existe algum lugar público (praça, parque, rua fechada, praia) para fazer caminhada, realizar exercício ou praticar esporte?
- P47 - O(A) sr(a) conhece algum programa público no seu município de estímulo à prática de atividade física?
- Q29 - Quando foi a última vez que o(a) sr(a) fez exame de sangue para medir a glicemia, isto é, o açúcar no sangue?
- Q30 - Algum médico já lhe deu o diagnóstico de diabetes?

Todos os atributos são numéricos, os possíveis valores para os atributos são de 1 a 7, os valores mínimos e máximos para os atributos são 1 a 7. A base é supervisionada, possui a classe

Q30 que indica se algum médico já lhe deu o diagnóstico de diabetes.

Porém, após escolhermos essas variáveis, um pré processamento se tornou necessário. Apesar de extensa, a base continha vários dados ausentes e colunas com poucos dados. Primeiramente removemos as colunas com mais de 40% de dados ausentes. Após a remoção das colunas pouco populadas, sobraram três variáveis nas quais também não são relevantes para análise que queremos, logo essas também foram descartadas.

## Ferramentas

Para a construção de um argumento para a relação proposta, foram utilizadas Python para o tratamento da base de dados e Excel para construção de gráficos e cálculos estatísticos.

Em relação a parte estatísticas foram utilizados:

- Gráficos de barra.
- Mediana
- Média
- Moda
- Desvio Padrão
- Percentis

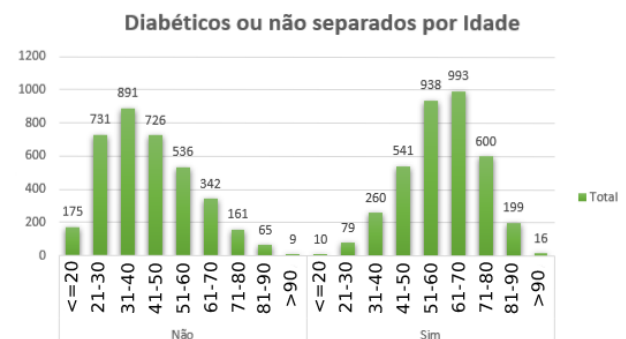
## 3 ANÁLISE

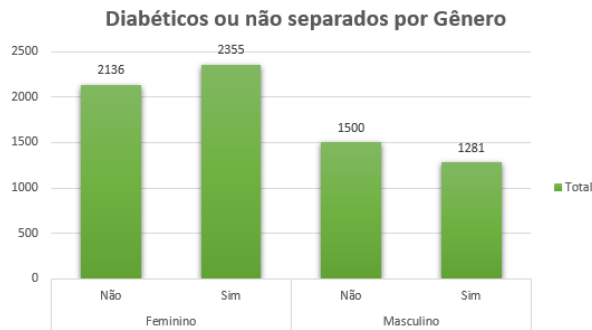
Nesta seção iremo abordar a distribuição gráfica e estatísticas sobre o tema abordado.

### Gráficos

Em primeira instância, temos abaixo duas distribuições:

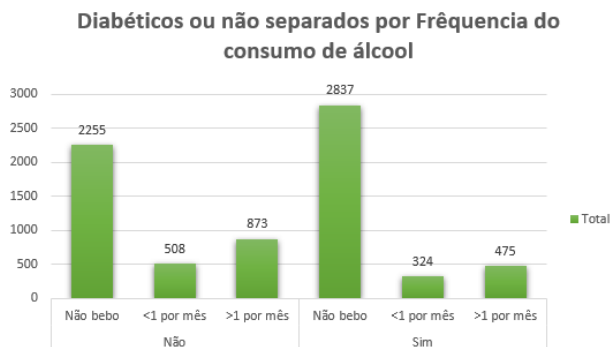
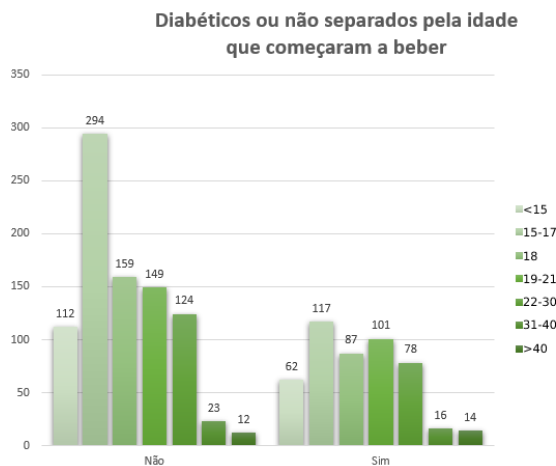
- Diagnóstico de Diabetes separado por Idade - separação por positivo e negativo.
- Diagnóstico de Diabetes separado por Sexo - separação por positivo e negativo.



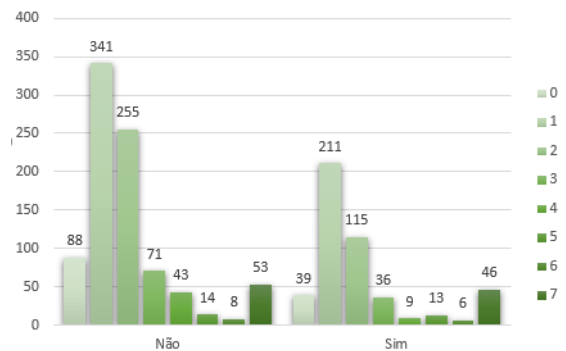


Em segunda instância, temos a separação gráfica com base no consumo de álcool.

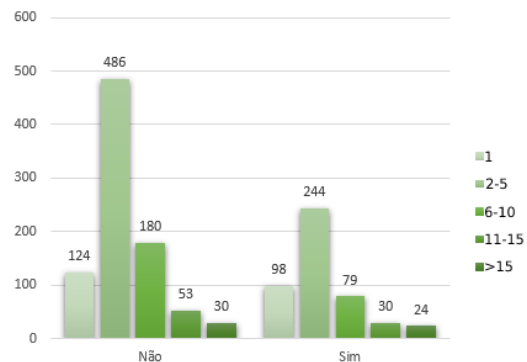
- Diabéticos separados pela idade que começaram a beber - separação por positivo e negativo.
- Diabéticos separados pela frequência do consumo de álcool - separação por positivo e negativo.
- Diabéticos separados pela frequência em dias do consumo de álcool - separação por positivo e negativo.
- Diabéticos separados pela quantidade de doses de álcool - separação por positivo e negativo.



**Diabéticos ou não que bebem separados pelo número de dias da semana que bebem**



**Diabéticos ou não separados pelo número de doses quando bebem**



## Estatística

| Estadística                                   | Homens      | Homens diabéticos | Homens Álcool | Homens D. Álcool | Mulheres    | Mulheres diabéticas | Mulheres Álcool | Mulheres D. Álcool |
|---|-------------|-------------------|---------------|------------------|-------------|---------------------|-----------------|--------------------|
| Número  | 870         | 542               | 153           | 89               | 476         | 331                 | 39              | 29                 |
| Mediana                                       | 47          | 39                | 53            | 44               | 41          | 37                  | 47              | 42                 |
| Média idade                                   | 46,04107916 | 41,19188192       | 39,94771242   | 45,47191011      | 42,56901769 | 38,61329305         | 46,43589744     | 41,41179311        |
| Moda  | 46          | 36                | 46            | 48               | 31          | 30                  | 44              | 44                 |
| Desvio padrão                                 | 14,54481791 | 11,76127913       | 14,67646559   | 14,38677072      | 15,50941001 | 11,47348794         | 16,29141665     | 15,16461198        |
| Percentil 25%/Q1                              | 35          | 31                | 39            | 34               | 31          | 29                  | 32              | 30                 |
| Percentil 50%/Q2                              | 37          | 36,75             | 42            | 38               | 34          | 37                  | 36,5            | 31                 |
| Percentil 75%/Q3                              | 47          | 42,25             | 53            | 44               | 41          | 39                  | 47              | 42                 |
| Percentual de diabéticos em relação ao total  | X           | 62,30%            | X             | 36,23%           | X           | 69,25%              | X               | 6,07%              |
| Percentual de diabéticos em relação ao álcool | X           | X                 | X             | 58,17%           | X           | X                   | X               | 74,36%             |
| Percentual que consume A. além de 4           | X           | X                 | 17,58%        | X                | X           | 9,16%               | X               | X                  |

Distribuição estatística de diabéticos em relação ao consumo de álcool.

## 4 CONCLUSÃO

No mundo atual a diabetes é uma doença na qual atinge uma grande parte da população e um dos diferentes desafios na área médica é o diagnóstico precoce desta doença, na qual quanto mais cedo o diagnóstico mais eficaz será o tratamento. Visto esta situação, na área da computação temos as técnicas de Ciência de Dados na qual podemos ajudar com a análise dos dados e aplicações com o intuito de prever uma possível diabetes já que estas técnicas aprendam com os dados inseridos e consegue processar uma quantidade de informações na qual um ser humano levaria anos para processar/entender.

Em primeira instância, o pico de pessoas diagnosticadas com Diabetes está na faixa entre 30 e 80 anos. Na qual corrobora com a média de idade de mulheres e homens(38,6 e 41,1) que obtiveram diagnóstico positivo.

Em seguida, temos a porcentagem do consumo de álcool por pessoas homens e mulheres que consomem álcool por mais de 4 dias(17,59 e 8,16 por cento), dentre essas pessoas, a maioria consome mais que duas doses de álcool.

Por último, com base nos gráficos e distribuição estatística, a relação de pessoas que começaram a beber a partir dos 15 anos de idade, pode existir em algum momento dessa passagem de tempo em vida dessas pessoas em que ela pararam de beber independente do diagnóstico de diabetes. O que pode corroborar com o fato de a maioria das pessoas nessa base dados, não consome álcool.

Logo, de acordo com a base de dados, 50 por cento das pessoas são diagnosticadas com diabetes mellitus, mesmo assim, a quantidade de pessoas que bebem representam apenas 12 por cento da base como um todo. Como parte desta distribuição, também estão inclusas as pessoas que bebem mas não foram diagnosticadas com diabetes, chegamos a conclusão que a amostragem geral sobre o assunto não é o suficiente para definir se o consumo de álcool por si só realmente impacta no desenvolvimento de diabetes.

## 5 REFERÊNCIAS

**IBGE PNS - Pesquisa Nacional de Saúde**

**Available in:** <https://www.ibge.gov.br/estatisticas/sociais/saude/29540-2013-pesquisa-nacional-de-saude.html?=&t=o-que-e>