

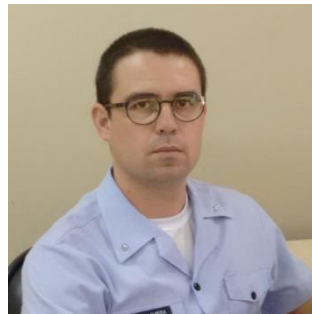
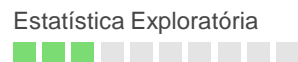
Verão 2017 - IME

Introdução a Análise Exploratória de Dados e Métodos estatísticos



Thiago Costa

Personal Trainer



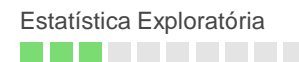
Cleibson Almeida

Militar



Reginaldo Florentino

Arquiteto





Avaliação do Findrisk como instrumento de detecção de diabetes de tipo 2 em um estudo transversal

Análise Descritiva de Dados e Aplicação da Regressão Logística

PESQUISADORA: Dra. Denise Beheregaray

INSTITUIÇÃO: ADJ – Associação de Diabetes Juvenil

FINALIDADE DO PROJETO: Publicação

RESPONSÁVEIS PELA ANÁLISE:

Professora Gisela Tunes

Professora Viviana Giampaoli

Eduardo Carvalho da Silva

William Nilson de Amorim

REFERÊNCIA DESTE TRABALHO:

TUNES, G.; GIAMPAOLI, V; SILVA, E. C.; AMORIM, W. N. Relatório de análise estatística sobre o projeto: “Avaliação do Findrisk como instrumento de detecção de diabetes de tipo 2 em um estudo transversal”. São Paulo, IME-USP, 2012. (ERA-CEA -12P34).



Tópicos Abordados

Estrutura adotada nos slides



Apresentação do Problema



Variáveis em Estudo



Análise Descritiva



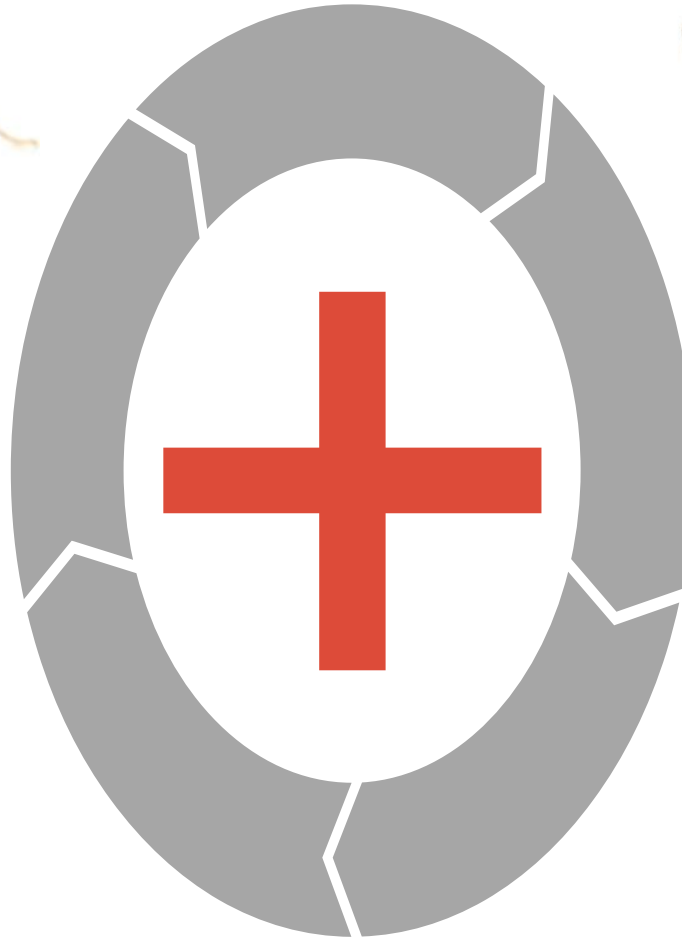
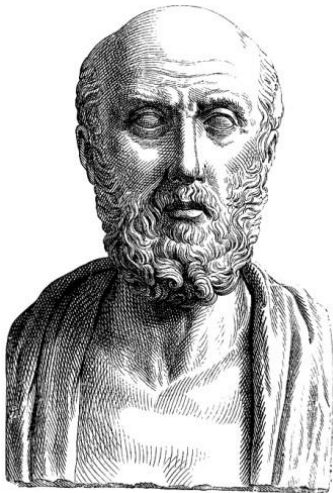
Análise Inferencial



Conclusões

Sobre a Diabetes

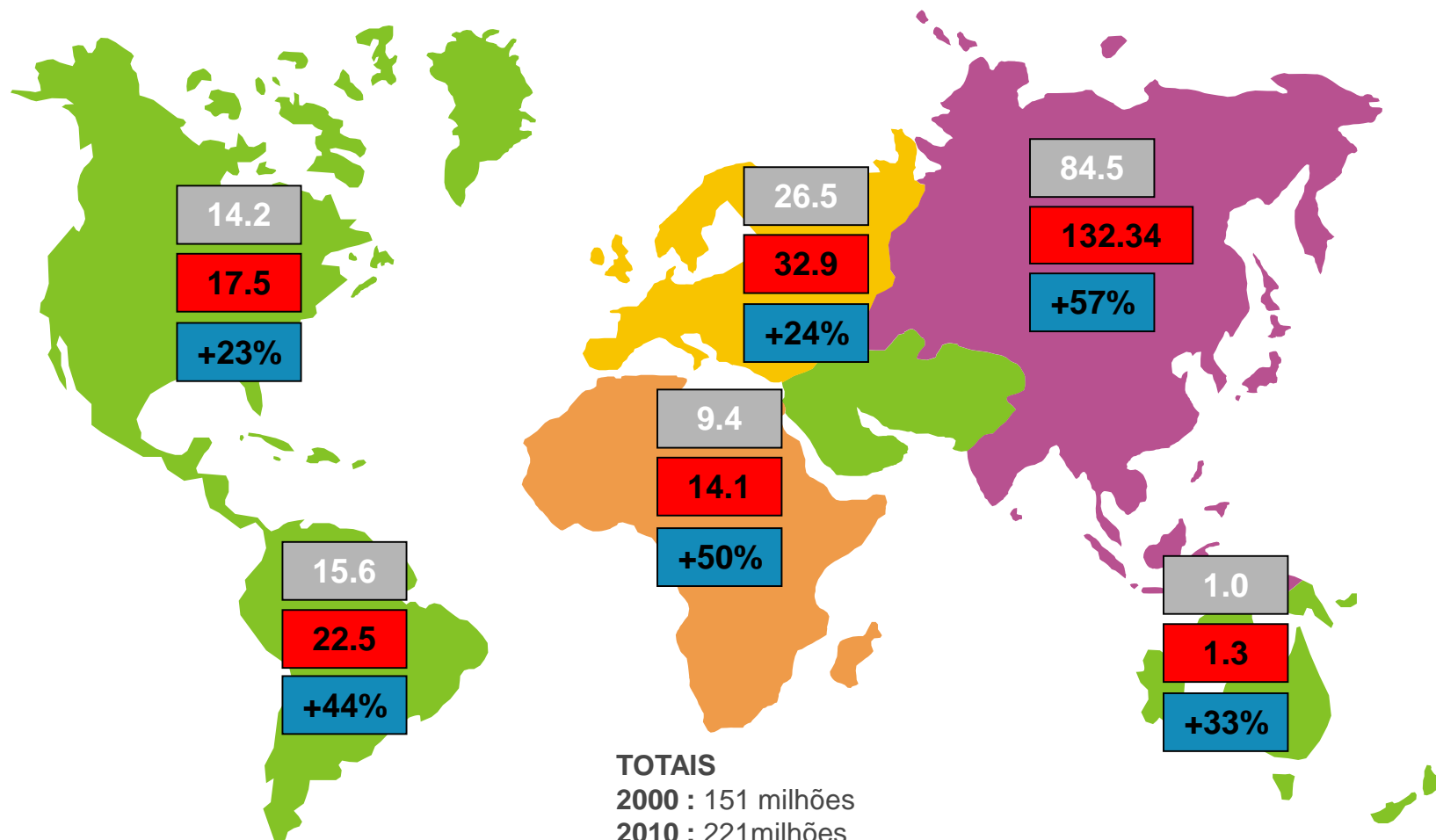
Contextualização



Sobre a Diabetes

Qual é o tamanho disso?

2000 (milhões) 2010 (milhões)



TOTAIS
2000 : 151 milhões
2010 : 221 milhões
Aumento de 46%

Sobre a Diabetes

A classificação

Valores de glicose plasmática (em mg/dl) para diagnóstico de diabetes *mellitus* e seus estágios pré-clínicos.

Categoria	Jejum*	2 h após 75 g de glicose	Casual**
Glicemia normal	< 100	< 140	
Tolerância à glicose diminuída	≥ 100 a < 126	≥ 140 a < 200	
Diabetes <i>mellitus</i>	≥ 126	≥ 200	≥ 200 (com sintomas clássicos)***

*O jejum é definido como a falta de ingestão calórica por no mínimo 8 h.

**Glicemia plasmática casual é aquela realizada a qualquer hora do dia, sem se observar o intervalo desde a última refeição.

***Os sintomas clássicos do DM incluem poliúria, polidipsia e perda não explicada de peso.

Nota: o diagnóstico do DM deve sempre ser confirmado pela repetição do teste em outro dia, a menos que haja hiperglicemia inequívoca com descompensação metabólica aguda ou sintomas óbvios de DM.

Tipo I

- Doença Auto-Imune
- Deficiência Insulina
- Jovens, Adolescentes
- Auto-Anticorpos
- Indivíduos Magros

Tipo II

- 90 % S. Metabólica
- Resistência Insulínica e Deficiência
- Indivíduos Meia Idade
- Depósito Amilóide nas Ilhotas
- Obesidade (85 %)
- Início Geralmente Lento

Sobre o Findrisk

A metodologia adotada para coleta dos dados

TEST DE FINDRISK

1. ¿QUE EDAD TIENES?

Menos de 35 años	→ 0 puntos
De 35-44 años	→ 1 punto
De 45-54 años	→ 2 puntos
De 55-64 años	→ 3 puntos
Mayor de 64 años	→ 4 puntos

2. REALIZAS ALGUNA ACTIVIDAD FÍSICA DURANTE AL MENOS 30 MINUTOS AL DÍA??

SI	→ 0 puntos
NO	→ 2 puntos

3. ¿QUÉ PERÍMETRO DE CINTURA TIENES?? (MEDIDO A LA ALTURA DEL OMBLIGO)

Mujeres		Hombres
Menos de 80 cm	→ 0 puntos	Menos de 94 cm
80 hasta 88 cm	→ 3 puntos	94 hasta 102 cm
Más de 88 cm	→ 4 puntos	Más de 102 cm

4. ¿LE HAN RECETADO ALGUNA VEZ MEDICAMENTOS CONTRA LA HIPERTENSIÓN?

SI	→ 2 puntos
NO	→ 0 puntos

5. ¿CON QUÉ FRECUENCIA COME FRUTA, VERDURAS?

Diario	→ 0 puntos
No Diario	→ 1 punto

6. ¿LE HAN DETECTADO ALGUNA VEZ, EN UN CONTROL MÉDICO, UN NIVEL ALTO DE AZÚCAR EN SANGRE?

SI	→ 5 puntos
NO	→ 0 puntos

7. ¿ÍNDICE DE MASA CORPORAL:

Menor de 25 kg/m ²	→ 0 puntos
Entre 25-30 kg/m ²	→ 1 punto
Mayor de 30 kg/m ²	→ 3 puntos

8. ¿HA HABIDO UN DIAGNÓSTICO DE DIABETES EN ALGÚN MIEMBRO DE SU FAMILIA?

NO	→ 0 puntos
Si: abuelos, tíos, primos	→ 3 puntos
Si: padres, hermanos o hijos	→ 5 puntos

Resultados:

- + de 15 puntos: Riesgo muy alto
- 12-14 puntos: Riesgo moderado
- 7-11 puntos: Riesgo bajo

WWW.MEDICADOO.ES

Questionamentos

O que é? **Questionário**

Quantas perguntas? **Oito**

Quem responde? **Qualquer cidadão**

Como é medida cada resposta? **Em pontos**

O que há na pergunta 3? **Perímetros dif.**

Qual o objetivo?

detectar o risco de desenvolvimento da Diabetes do tipo II em um período de 10 anos

Sobre o Findrisk

Características das perguntas [variáveis] [features]

TEST DE FINDRISK

1 **¿QUE EDAD TIENES?**

Menos de 35 años → 0 puntos
 De 35-44 años → 1 punto
 De 45-54 años → 2 puntos
 De 55-64 años → 3 puntos
 Mayor de 64 años → 4 puntos

2 **REALIZAS ALGUNA ACTIVIDAD FÍSICA DURANTE AL MENOS 30 MINUTOS AL DÍA??**

SI → 0 puntos
 NO → 2 puntos

3 **¿QUÉ PERÍMETRO DE CINTURA TIENES?? (MEDIDO A LA ALTURA DEL OMBLIGO)**

Mujeres
 Menos de 80 cm → 0 puntos
 80 hasta 88 cm → 3 puntos
 Más de 88 cm → 4 puntos

Hombres
 Menos de 94 cm → 0 puntos
 94 hasta 102 cm → 3 puntos
 Más de 102 cm → 4 puntos

4 **¿LE HAN RECETADO ALGUNA VEZ MEDICAMENTOS CONTRA LA HIPERTENSIÓN?**

SI → 2 puntos
 NO → 0 puntos

5 **¿CON QUÉ FRECUENCIA COME FRUTA, VERDURAS?**

Diario → 0 puntos
 No Diario → 1 punto

6 **¿LE HAN DETECTADO ALGUNA VEZ, EN UN CONTROL MÉDICO, UN NIVEL ALTO DE AZÚCAR EN SANGRE?**

SI → 5 puntos
 NO → 0 puntos

7 **ÍNDICE DE MASA CORPORAL:**

Menor de 25 kg/m² → 0 puntos
 Entre 25-30 kg/m² → 1 punto
 Mayor de 30 kg/m² → 3 puntos

8 **¿HA HABIDO UN DIAGNÓSTICO DE DIABETES EN ALGÚN MIEMBRO DE SU FAMILIA?**

NO → 0 puntos
 Sí: abuelos, tíos, primos → 3 puntos
 Sí: padres, hermanos o hijos → 5 puntos

Resumen de Riesgo:

- + de 15 puntos: Riesgo muy alto
- 12-14 puntos: Riesgo moderado
- 7-11 puntos: Riesgo bajo

WWW.MEDICADOO.ES

1

IDADE
 > Quantitativa Contínua

2

IMC
 > Quantitativa Contínua

3

CIRCUNFERÊNCIA DA CINTURA
 > Quantitativa Contínua

4

ATIVIDADE FÍSICA (> 30min)
 > Qualitativa Nominal

5

CONSUMO DE LEGUMES /FRUTAS
 > Qualitativa Nominal

6

USO DE MEDICAMENTO (ANTI-HIPERTENÇÃO)
 > Qualitativa Nominal

7

GLICEMIA ELEVADA (EXAME SAÚDE)
 > Qualitativa Nominal

8

FATOR HEREDITARIEDADE (DIABETES I ou II)
 > Qualitativa Nominal

Sobre o Findrisk

Discretização das variáveis [categorização]

1

IDADE

> Quantitativa Contínua -> Discreta

2

IMC

> Quantitativa Contínua -> Discreta

3

CIRCUNFERÊNCIA DA CINTURA

> Quantitativa Contínua -> Discreta

4

ATIVIDADE FÍSICA (> 30min)

> Qualitativa Nominal

5

CONSUMO DE LEGUMES /FRUTAS

> Qualitativa Nominal

6

USO DE MEDICAMENTO (ANTI-HIPERTENÇÃO)

> Qualitativa Nominal

7

GLICEMIA ELEVADA (EXAME SAÚDE)

> Qualitativa Nominal

8

FATOR HEREDITARIEDADE (DIABETES I ou II)

> Qualitativa Nominal

Idade, classificada em:

- o menos de 45 anos (0 ponto);
- o entre 45 e 54 anos (2 pontos)
- o entre 55 e 64 anos (3 pontos)
- o mais de 64 anos (4 pontos).

**Categorização em faixas
(medida em pontos)**

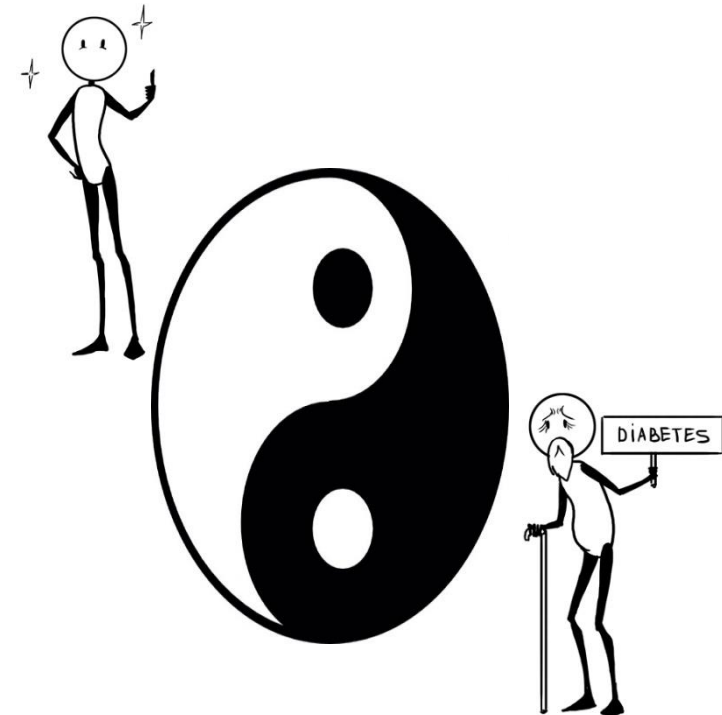
Sobre o Findrisk

Preparação e entendimento das variáveis para a modelagem

1	IDADE > Quantitativa Contínua -> Discreta	
2	IMC > Quantitativa Contínua -> Discreta	+
3	CIRCUNFERÊNCIA DA CINTURA > Quantitativa Contínua -> Discreta	+
4	ATIVIDADE FÍSICA (> 30min) > Qualitativa Nominal	+
5	CONSUMO DE LEGUMES /FRUTAS > Qualitativa Nominal	+
6	USO DE MEDICAMENTO (ANTI-HIPERTENÇÃO) > Qualitativa Nominal	+
7	GLICEMIA ELEVADA (EXAME SAÚDE) > Qualitativa Nominal	+
8	FATOR HEREDITARIEDADE (DIABETES I ou II) > Qualitativa Nominal	+



Score do Teste



- Risco baixo (1 em 100) – pontuação menor que 7;
- Risco levemente elevado (1 em 25) – pontuação de 7 a 11 (inclusive);
- Risco moderado (1 em 6) – pontuação de 12 a 14 (inclusive);
- Risco alto (1 em 3) – pontuação de 15 a 20 (inclusive);
- Risco muito alto (1 em 2) – pontuação maior que 20.

Coleta dos dados

Como foi a aplicação do findrisk



$n = 1062$ (adultos)

$n\alpha = 43$ diabéticos tipo II

Tabulação dos dados

Alguns valores omissos foram encontrados

– Número de valores omissos por variável

Variável	Número de valores omissos
Idade	3
IMC	11
Circunferência da cintura (mulheres)	6
Circunferência da cintura (homens)	3
Prática diária de atividades físicas	15
Consumo de frutas e legumes	15
Uso de medicamentos anti-hipertensivos	15
Histórico de exame com glicemia elevada	17
Histórico de familiar com diabetes tipo I ou II	15



$n = 1062$ (adultos)
 $n\alpha = 43$ diabéticos tipo II

$n - n\alpha = 1019$ amostras

8 variáveis



	A	B	C	D	E	F
1	Amostra	idade	imc	cintura_H	cintura_M	...
2	n0001	0	0	0
3	n0002	4	1	0
4	n0003	3	DM	NA	3	...
5	n0004	DM	1	DM	DM	...
6	n0005	DM	1	1	NA	...
7	n0006	2	2	...	4	...
8	n0007	DM	0	DM	DM	...
9	n0008	2	1	4	NA	...
10	n0009	3	2	NA	0	...
11	n0010	2	0	3	NA	...
12	n0011	0	1	NA	0	...
13

Análise Descritiva

Risco por Gênero

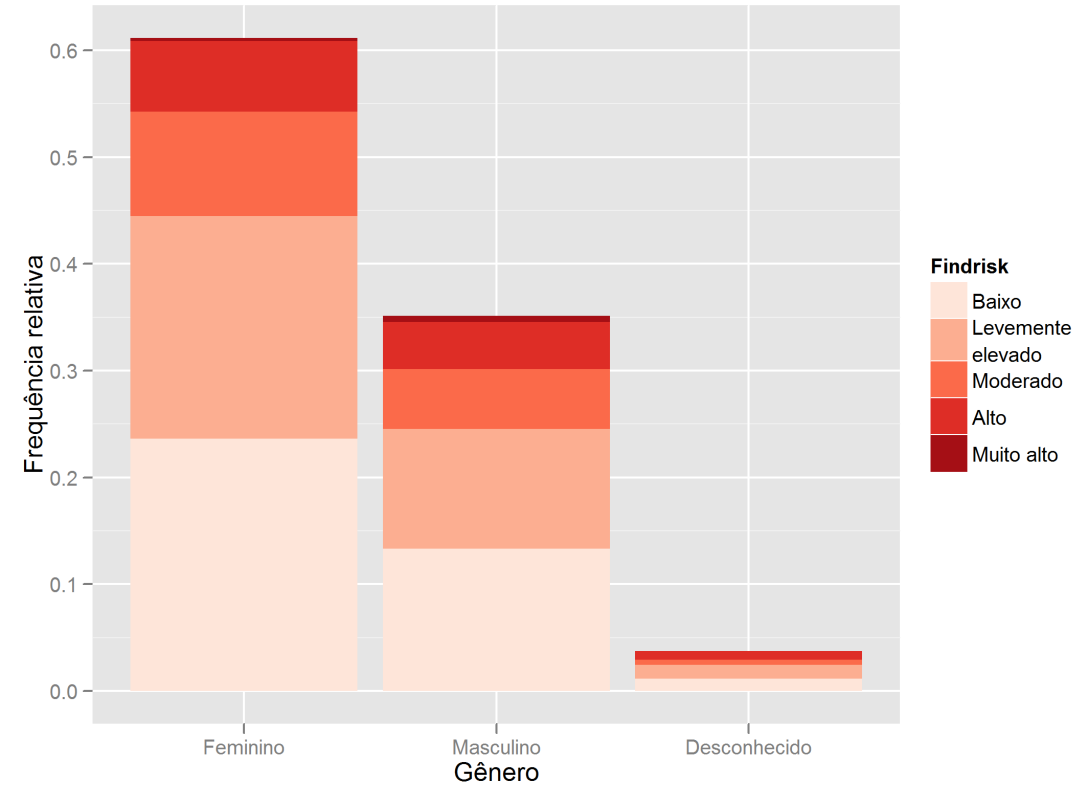


Gráfico B1 – Gráfico de barras da frequência relativa da variável Gênero particionado segundo as categorias do Findrisk.

Análise Descritiva

Risco por Idade

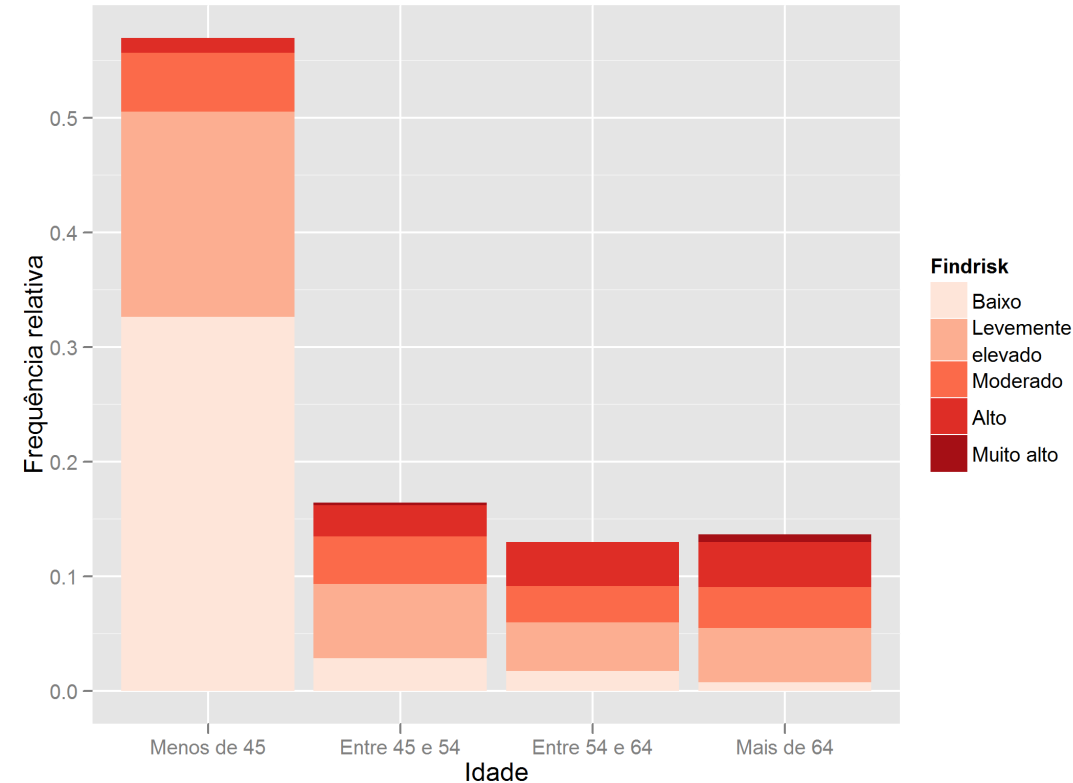


Gráfico B2 – Gráfico de barras da frequência relativa da variável Idade particionado segundo as categorias do Findrisk.

Análise Descritiva

Risco por IMC

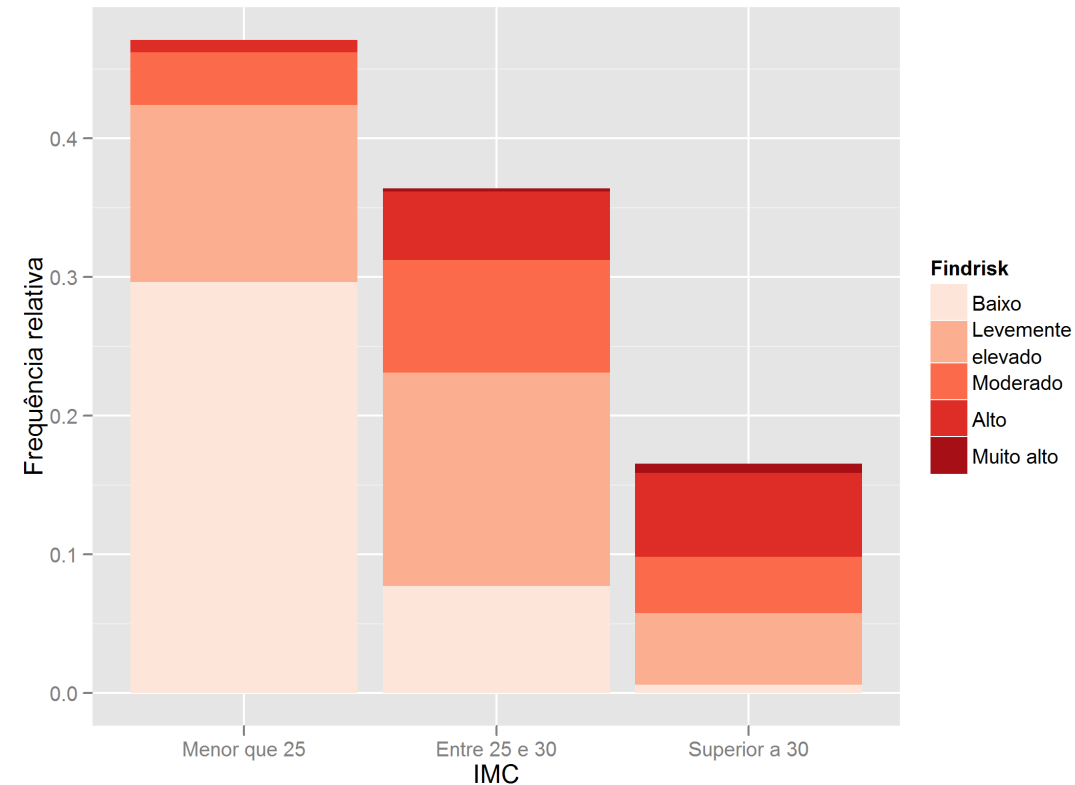


Gráfico B3 – Gráfico de barras da frequência relativa da variável IMC particionado segundo as categorias do Findrisk.

Análise Descritiva

Risco por IMC

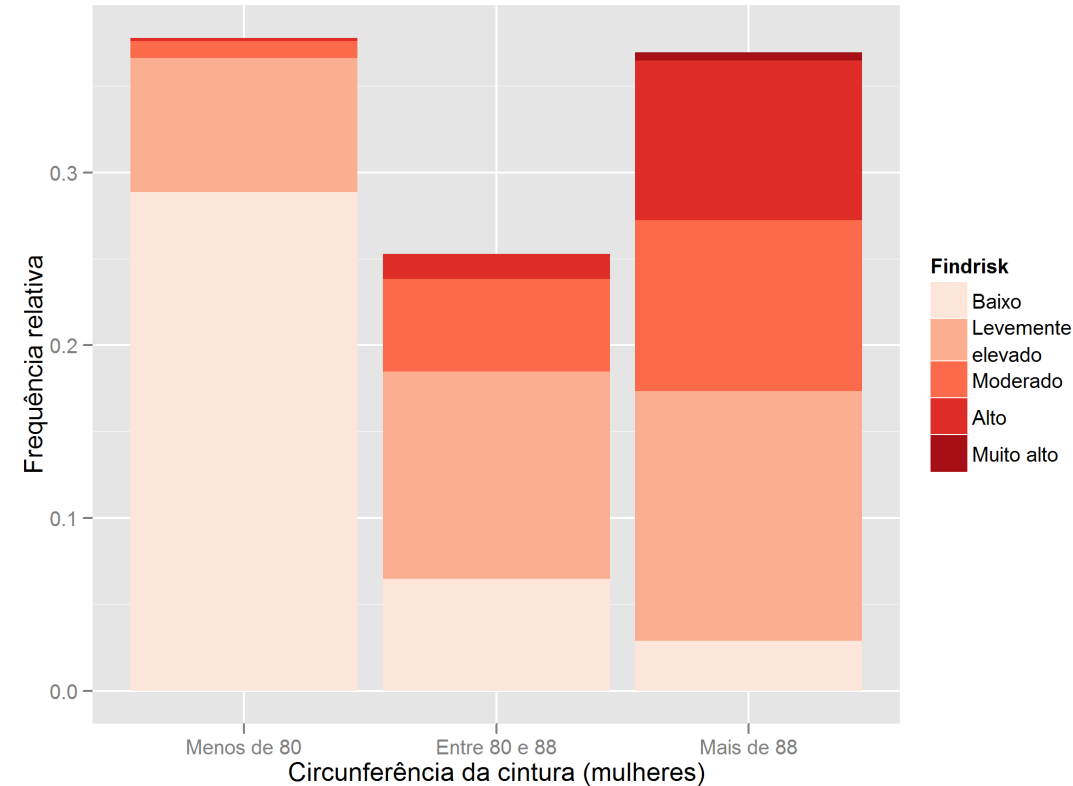


Gráfico B4 – Gráfico de barras da frequência relativa da variável Circunferência da cintura (mulheres) particionado segundo as categorias do Findrisk.

ANÁLISE DOS DADOS

- ANÁLISE DESCRITIVA
 - CARACTERIZAÇÃO DOS PERFIL'S
 - ANÁLISE EXPLORATORIA DAS VARIÁVEIS
 - RELAÇÕES ENCONTRADAS ENTRE AS VARIÁVEIS
- ANÁLISE INFERENCIAL
 - QUAIS VARIÁVEIS ESTÃO MAIS ASSOCIADAS COM DIAGNÓSTICO DE DM 2
 - ESTUDAR O PODER DA PREDIÇÃO DO ESCORE DO FINDRISK
 - AJUSTE DO MODELO

ANALISE DESCRITIVA

- CARACTERIZAÇÃO DOS PERFIL'S

ANALISE DESCRITIVA

- ANALISE EXPLORATORIA DAS VARIÁVEIS

ANALISE DESCRITIVA

- RELAÇÕES ENCONTRADAS ENTRE AS VARIÁVEIS

ANALISE INFERENCIAL

- QUAIS VARIÁVEIS ESTÃO MAIS ASSOCIADAS COM DIAGNÓSTICO DE DM 2

ANALISE INFERENCIAL

- ESTUDAR O PODER DA PREDIÇÃO DO ESCORE DO FINDRISK

ANALISE INFERENCIAL

- AJUSTE DO MODELO
 - MODELO 1
 - MODELO 2

CONCLUSÃO