

Aula 11: Que factores explicam a insatisfação com as condições de trabalho?

Estimação de Modelos de Regressão Logística

Docente: Amílcar Moreira

Data & Hora: 17/12/2020, 18:00-20:00

Local: Edifício F2, Sala 102

E se a nossa variável dependente não é uma variável contínua?

Existe um conjunto de opções ...

Sendo que a escolha do modelo de regressão mais adequado dependerá da natureza da variável dependente e (em alguns casos) da natureza das variáveis dependentes



ANÁLISE DE DADOS EM GRH

Aula 11: Estimação de Modelos de Regressão Logística

Variável Dependente	Modelo	Variáveis Independentes
Contínua	Regressão Linear Múltipla (OLS)	Todos
Nominal (2 Categorias)	Regressão Logística Regressão Probit	Todos
Nominal (2+ Categorias)	Regressão Multinomial	Todos
Nominal (2 Categorias)	Regressão Loglinear	Nominais
Ordinal	Ordinal Logit Ordered Probit	Todos
Contagens (contém apenas número inteiros, número finito de valores)	Poisson Regression	Contínuas
Tempo até um evento (Análise de Sobrevivência)	Cox Regression	Todos

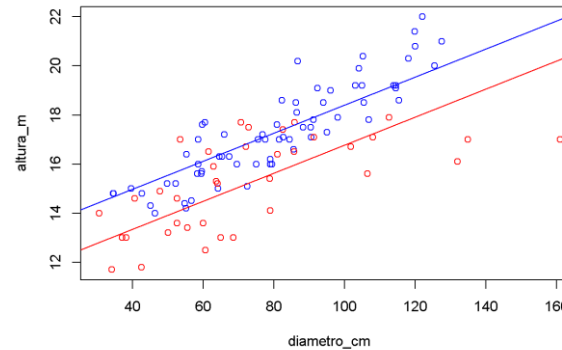
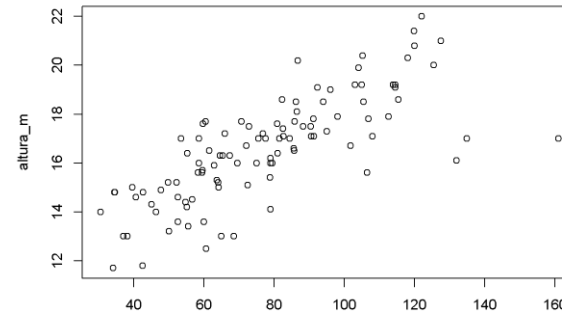
- **Nesta aula vamos focar-nos num dos modelos de regressão mais comum:**

O Modelo de Regressão Logística

Porque é não podemos usar o modelo de regressão linear quando temos uma variável dependente binomial?

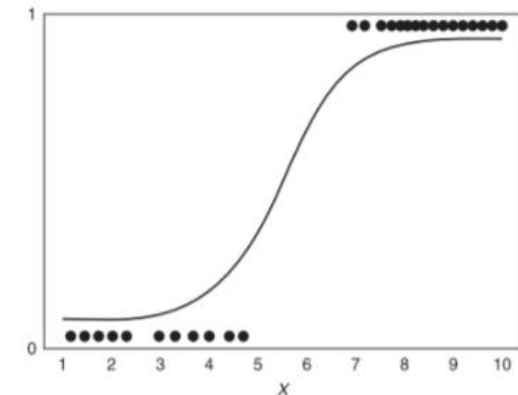
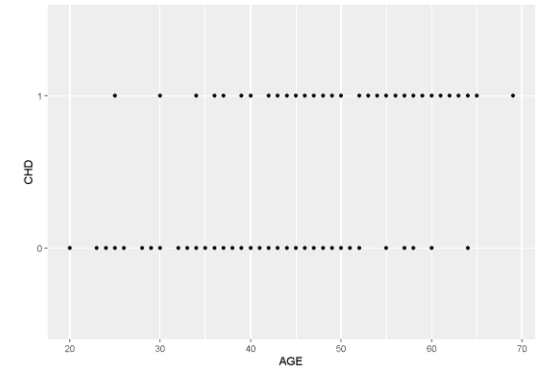
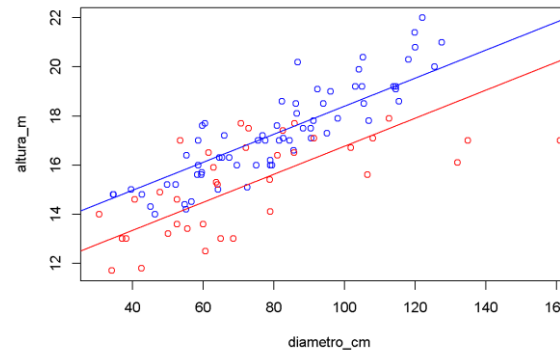
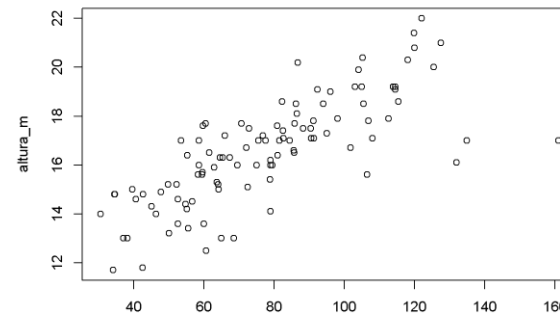
Porque é não podemos usar o modelo de regressão linear quando temos uma variável dependente binomial?

- Numa variável dependente contínua os valores distribuem-se à volta de uma linha (a linha de regressão).



Porque é não podemos usar o modelo de regressão linear quando temos uma variável dependente binomial?

- Numa variável dependente contínua os valores distribuem-se à volta de uma linha (a linha de regressão).
- Ora não é isso que acontece com uma variável dependente binária...
- A regressão logística utiliza a curva logística para assim representar a relação entre a variável dependente e as independentes.
- Esta curva assume que a partir de um dado valor da IV, o valor da DV não se altera



Modelo de regressão linear (simples)

$$Y = \beta_0 + \beta_1 X + E$$

Y	Variável Dependente
β_0	Constante (i.e., valor de Y quando X = 0)
$\beta_1 X$	Coefficiente Beta da variável X (mede o efeito de uma alteração unitária de X sobre o valor médio da variável Y, quando todas as outras variáveis estão fixas)
E	Erro aleatório ou estocástico (reflete a influência de outros factores no no comportamento da variável Y que não podem ser explicadas linearmente pelo comportamento da variável X)

Modelo de regressão logística

$$\ln\left(\frac{p}{1-p}\right) = b_0 + b_1 X_1 + \dots + b_n X_n$$

$\ln\left(\frac{p}{1-p}\right)$	Logaritmo natural das chances (odds) de p acontecer
p	Probabilidade de um evento acontecer
$1-p$	Probabilidade de um evento não acontecer
b_0	Constante
$b_1 X_1$	Coefficiente Beta da variável X (mede o efeito de uma alteração unitária de X sobre o logaritmo das odds da variável dependente)

Implementação do Modelo de Regressão Logística

- 1. Criar a variável dependente*
- 2. Inspeccionar a variável dependente*
- 3. Explorar as relações entre variáveis*
- 4. Definir o modelo de regressão logística*
- 5. Interpretar os resultados do modelo de regressão logística*
- 6. Validar pressupostos do modelo de regressão logística*

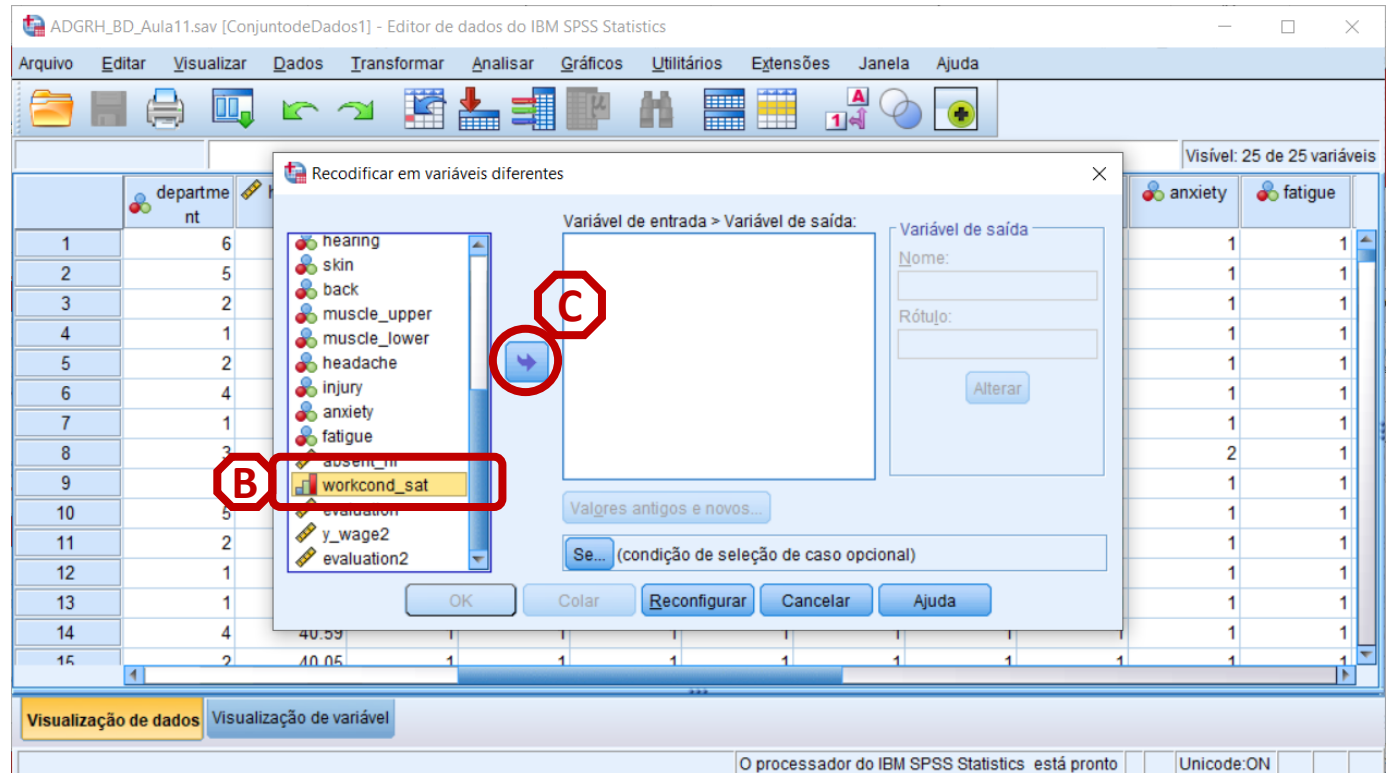
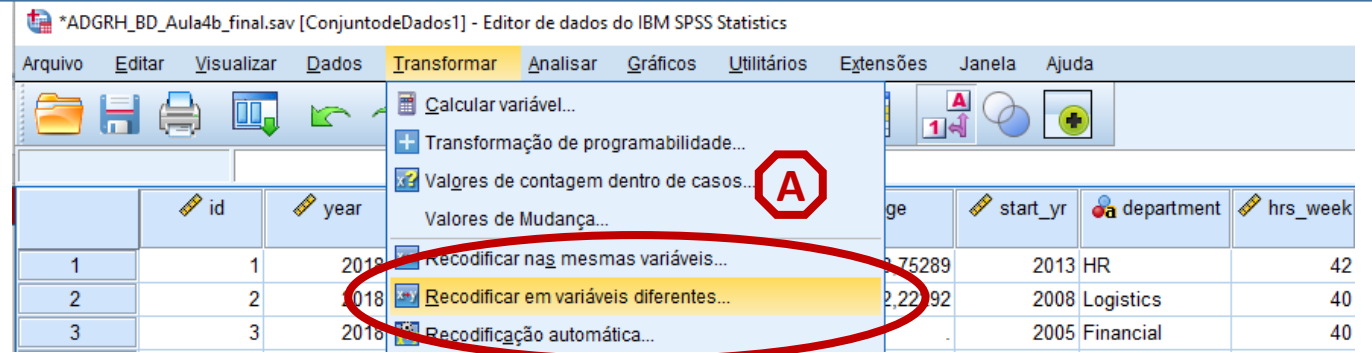
- **Objectivo:**

- Recodificar a variável que mede a satisfação com as condições de trabalho ('workcond_sat'), numa variável dummy em que 1 identifica pessoas que responderam 'Very satisfied' (1) ou 'Satisfied' (2).



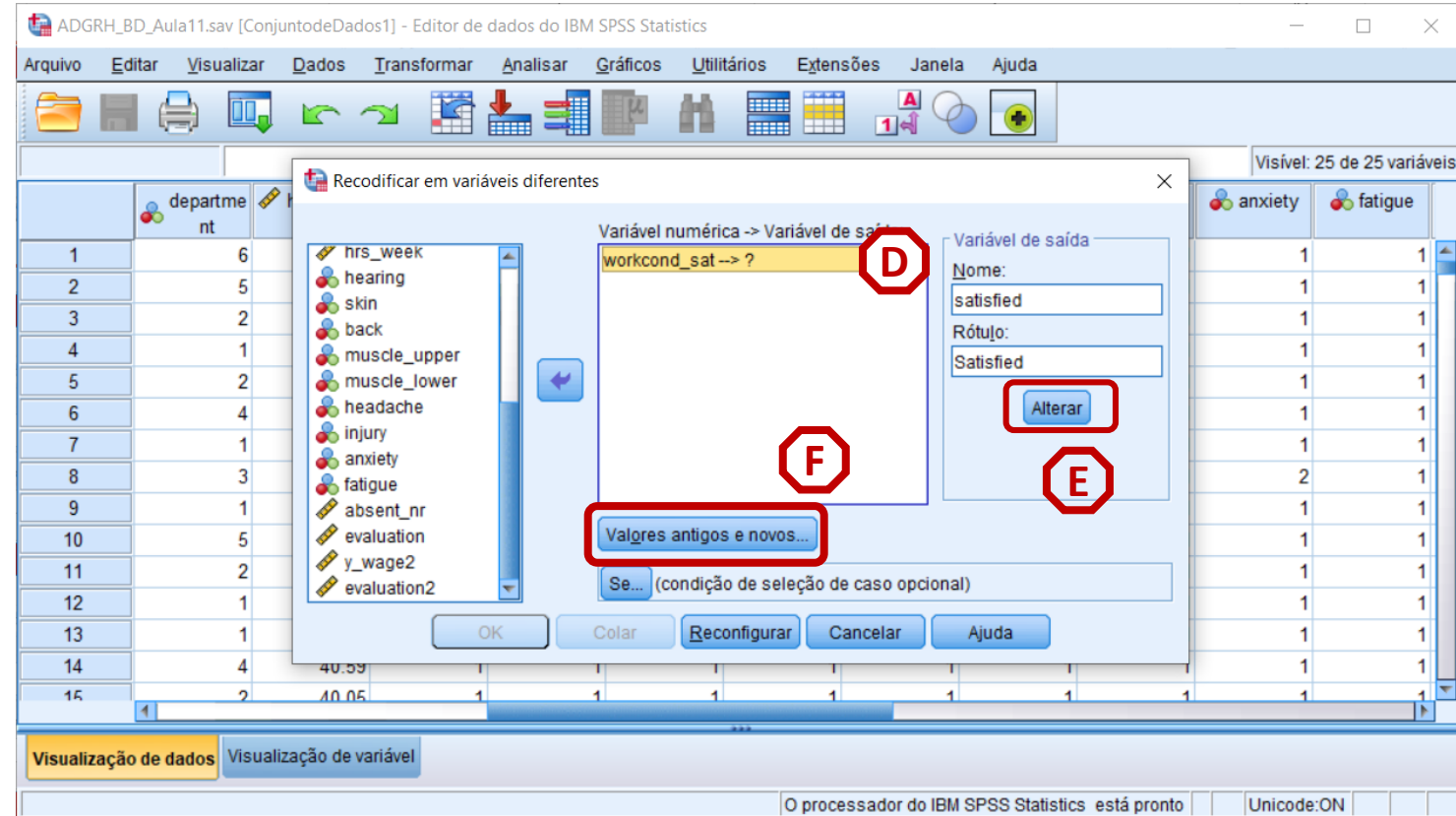
Criar a variável dependente

- Seleccionar 'Transformar'/'Recodificar em variáveis diferentes' **A**
- Seleccionar a variável 'workcond_sat'... **B**
- ... e colocar na caixa da 'variável de entrada'



Criar a variável dependente

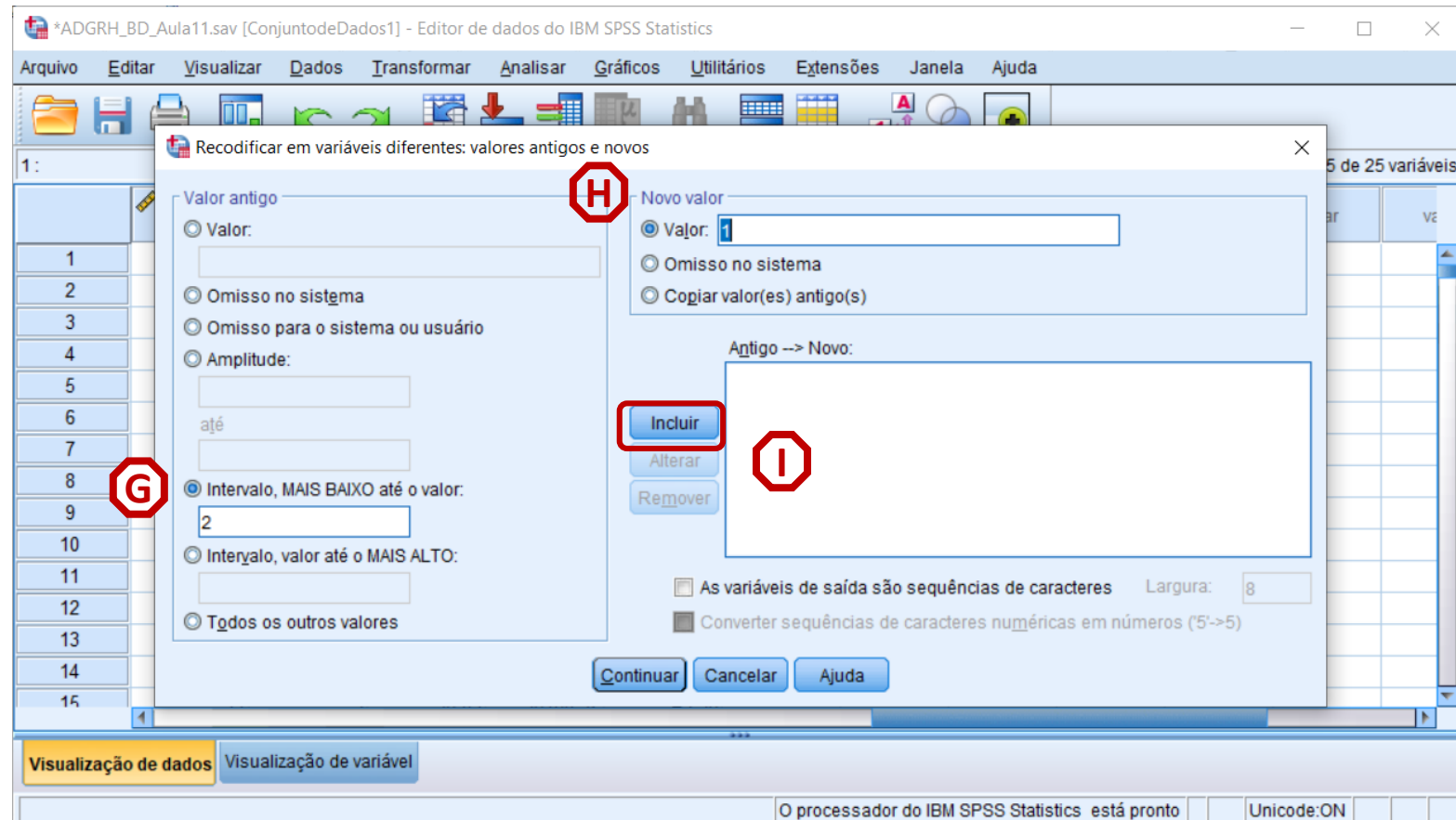
- Seleccionar ‘Transformar’/’Recodificar em variáveis diferentes’ **A**
- Seleccionar a variável ‘workcond_sat’... **B**
- ... e colocar na caixa da ‘variável de entrada’ **C**
- Definir o nome da nova variável (‘satisfied’) e o rótulo da variável (‘Satisfied’) **D**
- Seleccionar o botão ‘Alterar’ **E**
- Seleccionar ‘Valores antigo e novo’ **F**



The screenshot shows the IBM SPSS Statistics interface. The main window is titled 'ADGRH_BD_Aula11.sav [ConjuntodeDados1] - Editor de dados do IBM SPSS Statistics'. The 'Recodificar em variáveis diferentes' dialog box is open, showing the process of creating a new dependent variable from an existing numerical variable. The 'workcond_sat' variable is selected as the input, and the new variable is named 'satisfied' with the label 'Satisfied'. The 'Alterar' button is highlighted, indicating the next step in the process. The 'Valores antigos e novos...' button is also highlighted, indicating the step to define the mapping between the old and new values.

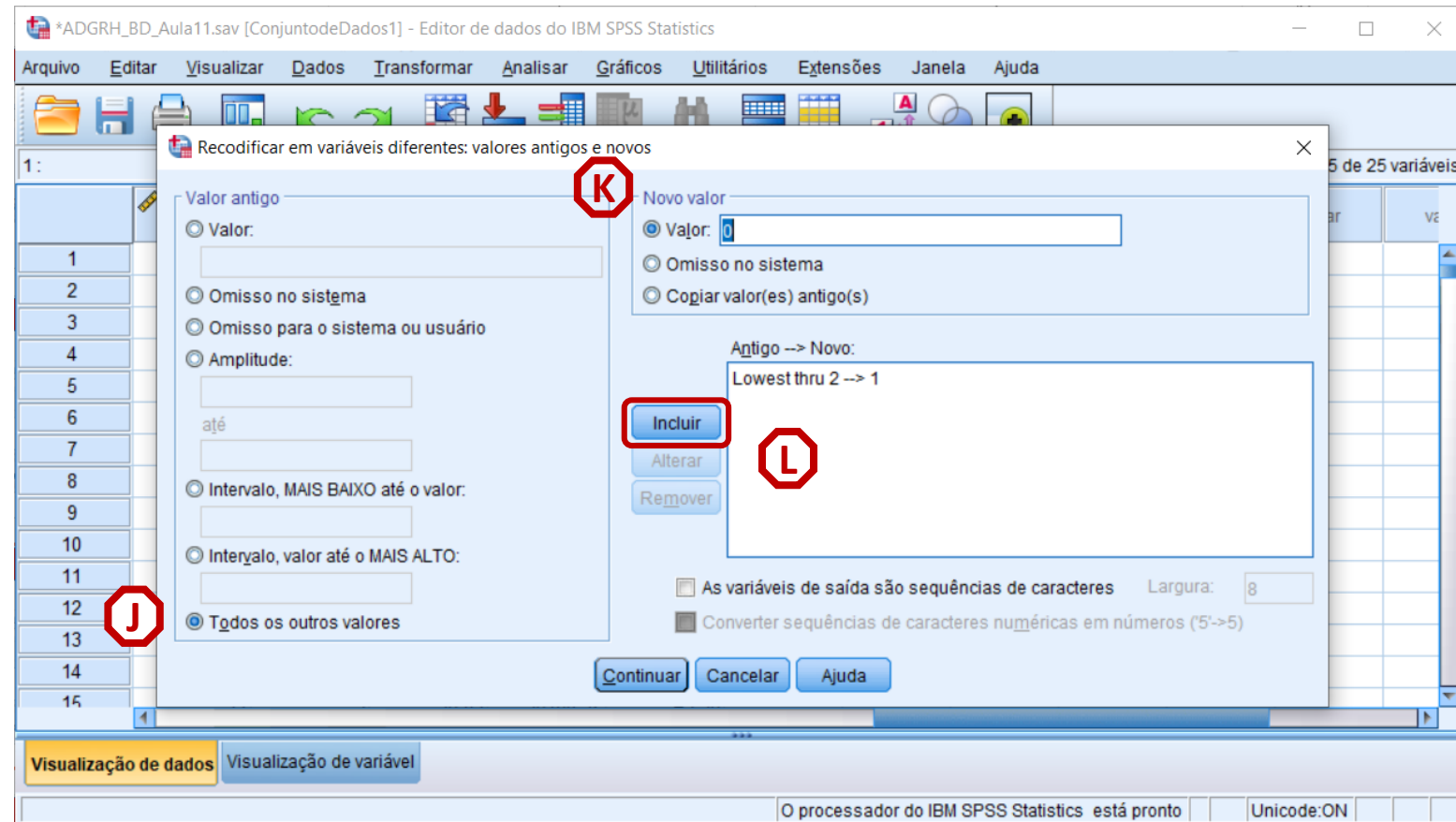
Criar a variável dependente

- Vamos definir que ao valor '1' e '2' na variável 'workcond_sat'
- Corresponde o valor '1', na nova variável ('satisfied')
- Selecionar 'Incluir'



Criar a variável dependente

- Vamos definir que ao valor '1' e '2' na variável 'workcond_sat' **G**
- Corresponde o valor '1', na nova variável ('satisfied') **H**
- Selecionar 'Incluir' **I**
- Vamos definir que 'Todos os outros valores' na variável 'workcond_sat' **J**
- Corresponde o valor '0', na nova variável ('satisfied') **K**
- Selecionar 'Incluir' **L**



Criar a variável dependente

- Vamos definir que ao valor '1' e '2' na variável 'workcond_sat'
- Corresponde o valor '1', na nova variável ('satisfied')
- Selecionar 'Incluir'
- Vamos definir que 'Todos os outros valores' na variável 'workcond_sat'
- Corresponde o valor '0', na nova variável ('satisfied')
- Selecionar 'Incluir'
- Selecionar 'Continuar' / 'OK'

G

H

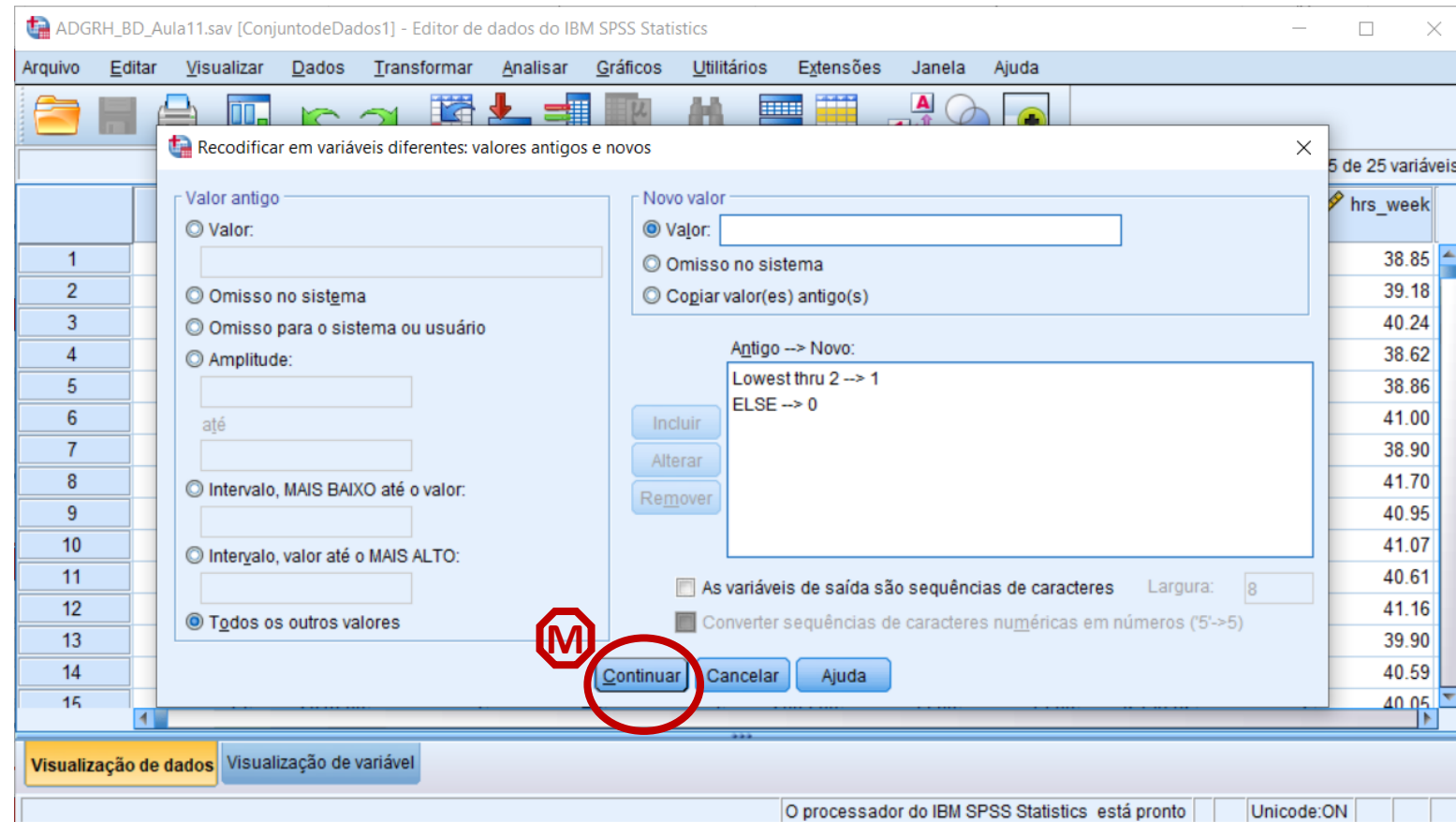
I

J

K

L

M



Criar a variável dependente

- Vamos verificar se a recodificação foi bem sucedida

*ADGRH_BD_Aula11.sav [ConjuntodeDados1] - Editor de dados do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Analisar Gráficos Utilitários Extensões Janela Ajuda

1 : satisfied .0 Visível: 26 de 26 variáveis

	absent_nr	workcond_sat	evaluation	y_wage2	evaluation 2	satisfied	var	var	var	var	var
1	9	3	49.08	17529.60	55.17	.00					
2	10	2	52.58	24843.10	51.75	1.00					
3	15	2	53.80	29874.72	54.79	1.00					
4	10	2	50.42	20994.46	58.30	1.00					
5	11	4	48.43	10044.01	49.89	.00					
6	10	2	54.34	32321.53	52.74	1.00					
7	6	3	49.65	18757.44	47.60	.00					
8	10	4	46.84	9106.36	52.65	.00					
9	8	2	50.76	20432.69	48.48	1.00					
10	17	3	53.33	29327.90	52.40	.00					
11	14	2	50.59	18223.72	47.21	1.00					
12	7	2	54.64	29668.47	48.39	1.00					
13	13	3	56.64	35658.63	50.19	.00					
14	7	3	55.36	35733.01	40.86	.00					
15	11	3	56.09	36160.52	45.36	.00					

Visualização de dados Visualização de variável

O processador do IBM SPSS Statistics está pronto Unicode:ON

Implementação do Modelo de Regressão Logística

2. Inspeccionar a variável dependente

- **Objectivo:**
 - Perceber como se distribuem as observações na variável dependente ('satisfied').
 - Neste caso, através de um Gráfico de Barras...



Gráfico de Barras

- Seleccionar 'Gráficos' / 'Construtor de Gráfico'

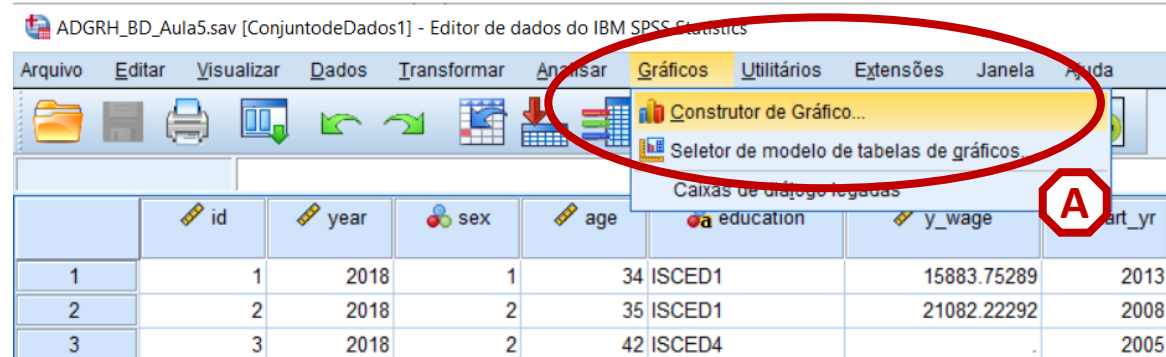


Gráfico de Barras

- Selecione 'Barras'
- Selecionar (com duplo-clique) o Gráfico de Barras (simples)



Construtor de gráfico

Variáveis: A visualização do gráfico usa dados de exemplo

id
year
sex
age
education
y_wage
start_yr
department
hrs_week
hearing

Nenhuma categoria (variável de escala)

Arraste um gráfico da Galeria aqui para usá-lo como seu ponto de início

OU

Clique na guia Elementos Básicos para construir um elemento de gráfico por elemento

Galeria Elementos básicos ID de ponto Títulos/Notas de rodapé

Escolha entre:

Favoritos
Barra
Linha
Área
Setor/Polar
Dispersão/Ponto
Histograma
Alto-Baixo:
Boxplot
Eixos duplos

OK Colar Reconfigurar Cancelar Ajuda

Gráfico de Barras

- Selecione 'Barras' C
- Selecionar (com duplo-clique) o Gráfico de Barras (simples) D
- Selecionar a variável 'satisfied' E
- Colocar a variável 'satisfied' no 'Eixo X' F

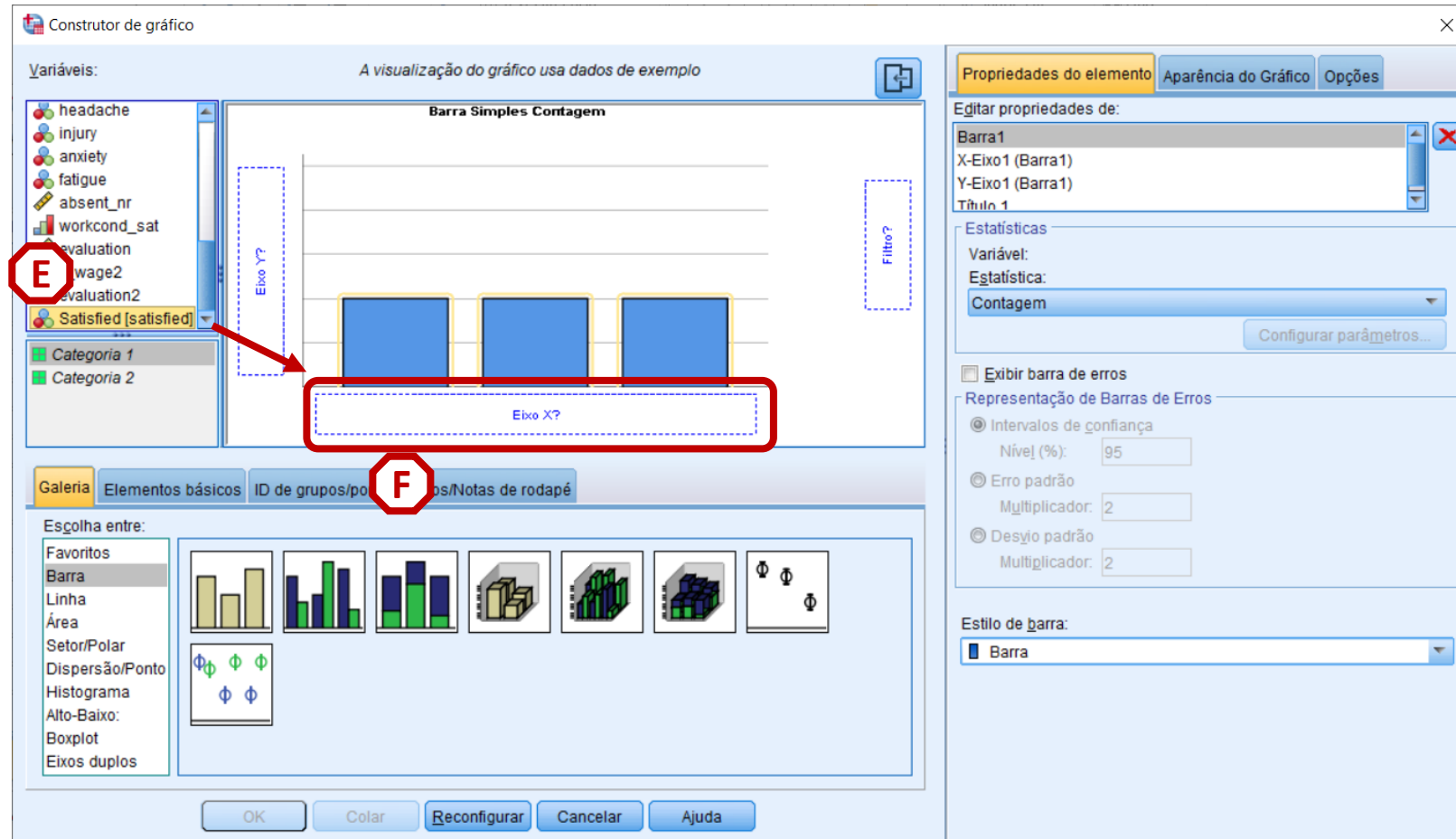


Gráfico de Barras

- Selecione 'Barras' C
- Selecionar (com duplo-clique) o Gráfico de Barras (simples) D
- Selecionar a variável 'satisfied' E
- Colocar a variável 'satisfied' no 'Eixo X' F
- Selecionar 'Porcentagem' G
- Selecionar 'OK' H

Construtor de gráfico

Variáveis: *A visualização do gráfico usa dados de exemplo*

id, year, sex, age, education, start_yr, start_age, experience, y_wage, department

Nenhuma categoria (variável de escala)

Barra Simples Contagem de Satisfied

Contagem

Filtro?

Satisfied

Galeria | Elementos básicos | ID de grupos/ponto | Títulos/Notas de rodapé

Escolha entre:

Favoritos

Barra, Linha, Área, Setor/Polar, Dispersão/Ponto, Histograma, Alto-Baixo, Boxplot, Eixos duplos

OK, Colar, Reconfigurar, Cancelar, Ajuda

Propriedades do elemento | Aparência do Gráfico | Opções

Editar propriedades de:

Barra1

X-Eixo1 (Barra1), Y-Eixo1 (Barra1), Título

Estadísticas

Variável: Contagem

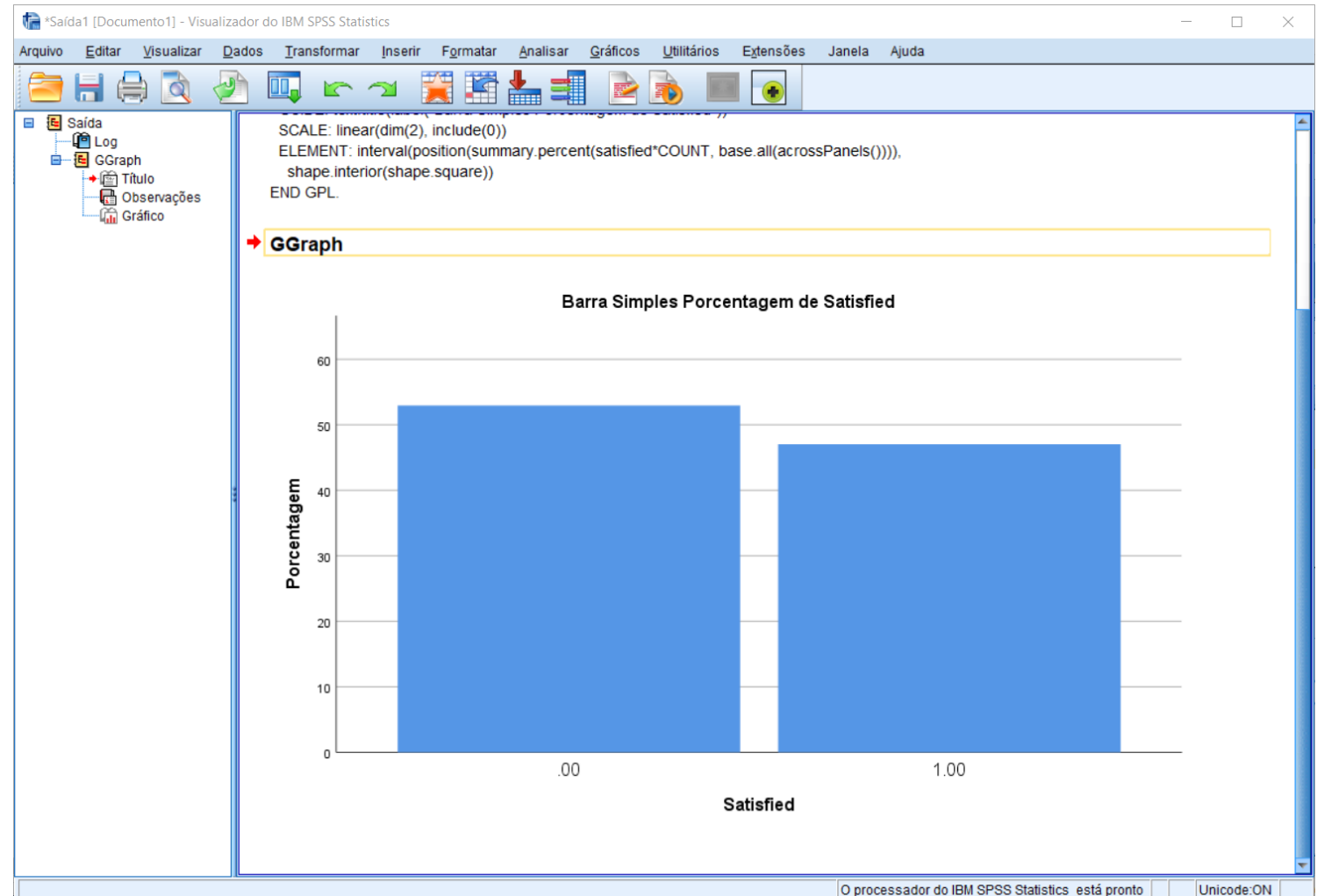
Estadística: Contagem, Contagem cumulativa, **Porcentagem (?)**, Porcentagem acumulativa, Valor, Média, Mediana, Mediana do grupo

Desvio padrão, Multiplicador: 2

Estilo de barra: Barra

Gráfico de Barras

- O gráfico é publicado no 'Visualizador de Resultados'
- Menos de metade dos trabalhadores está satisfeito com as condições de trabalho
- Gráfico dever ser complementado com tabela de frequências (ver Aula 5)



Implementação do Modelo de Regressão Logística

3. Explorar as relações entre variáveis

- **Objectivo:**

- **Perceber como se relaciona a variável dependente ('satisfied') com a variável independente 'sex' que queremos incluir no nosso modelo de regressão**

... neste caso, através de um Gráfico de Barras Sobrepostas

Gráfico de Barras Sobrepostas

- Selecionar 'Gráficos' / 'Construtor de Gráfico'

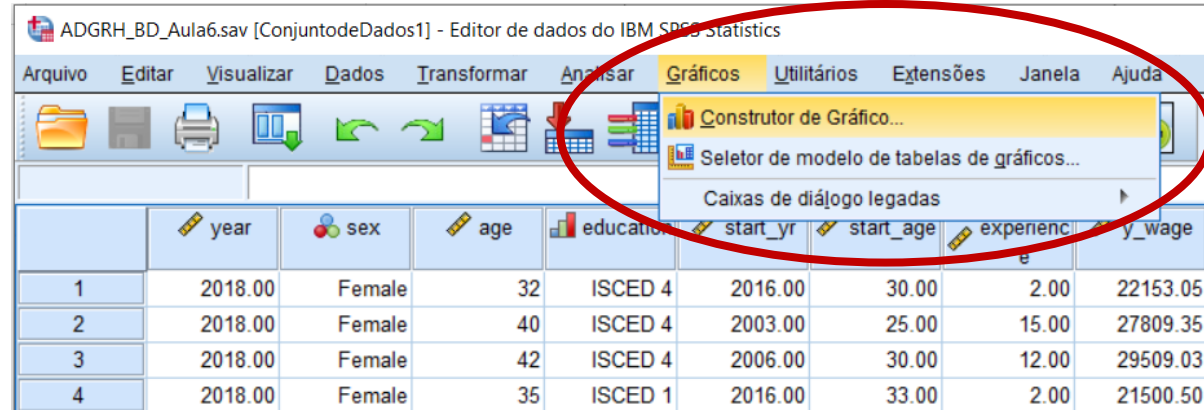


Gráfico de Barras Sobrepostas

- Seleccionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Sobrepostas'

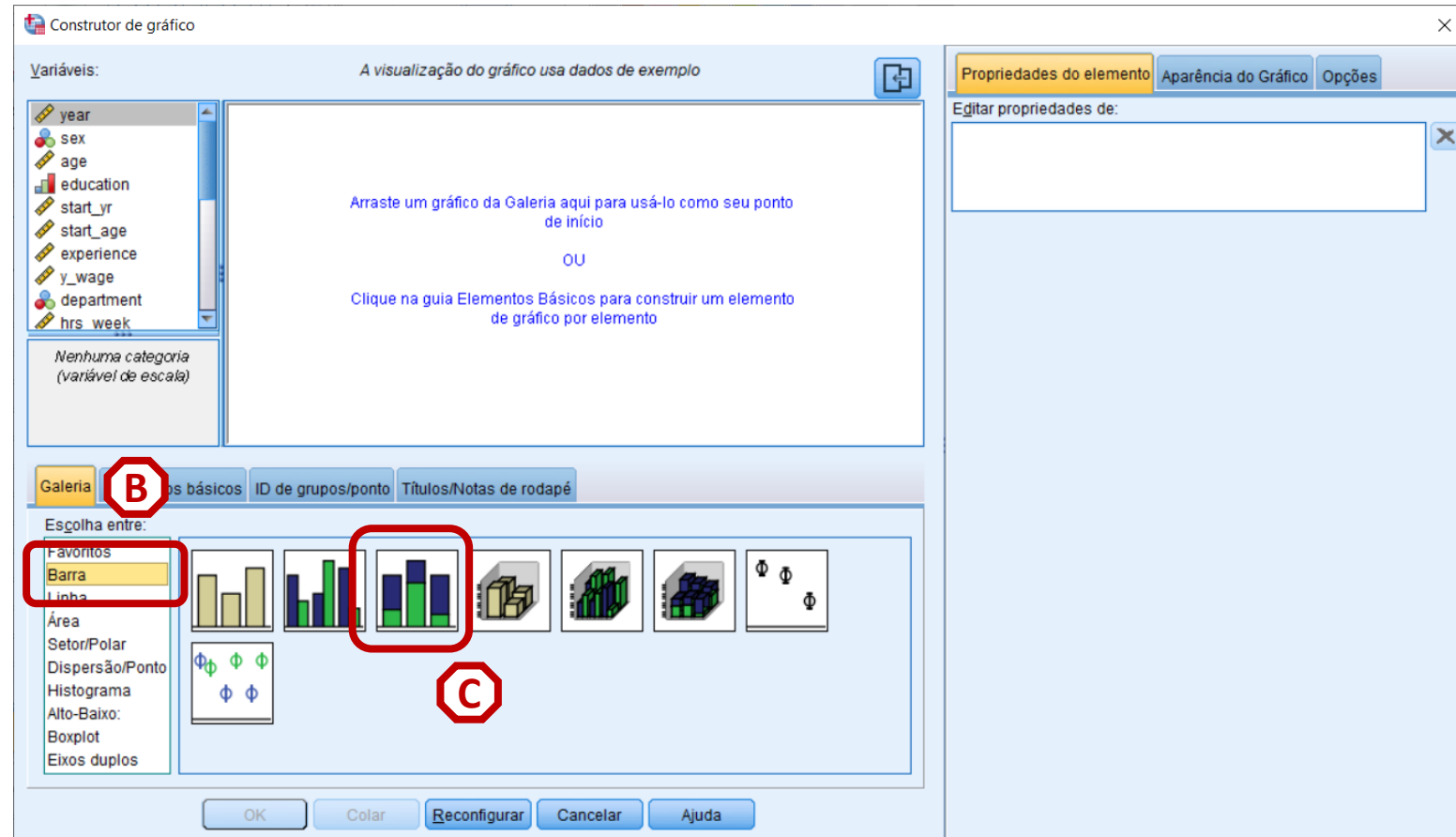


Gráfico de Barras Sobrepostas

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Sobrepostas'
- Selecionar a variável 'satisfied'...

... e colocar na caixa do 'Cluster em X'

Exercício: Colocar a variável 'sex' no 'Eixo X'

A

B

C

D

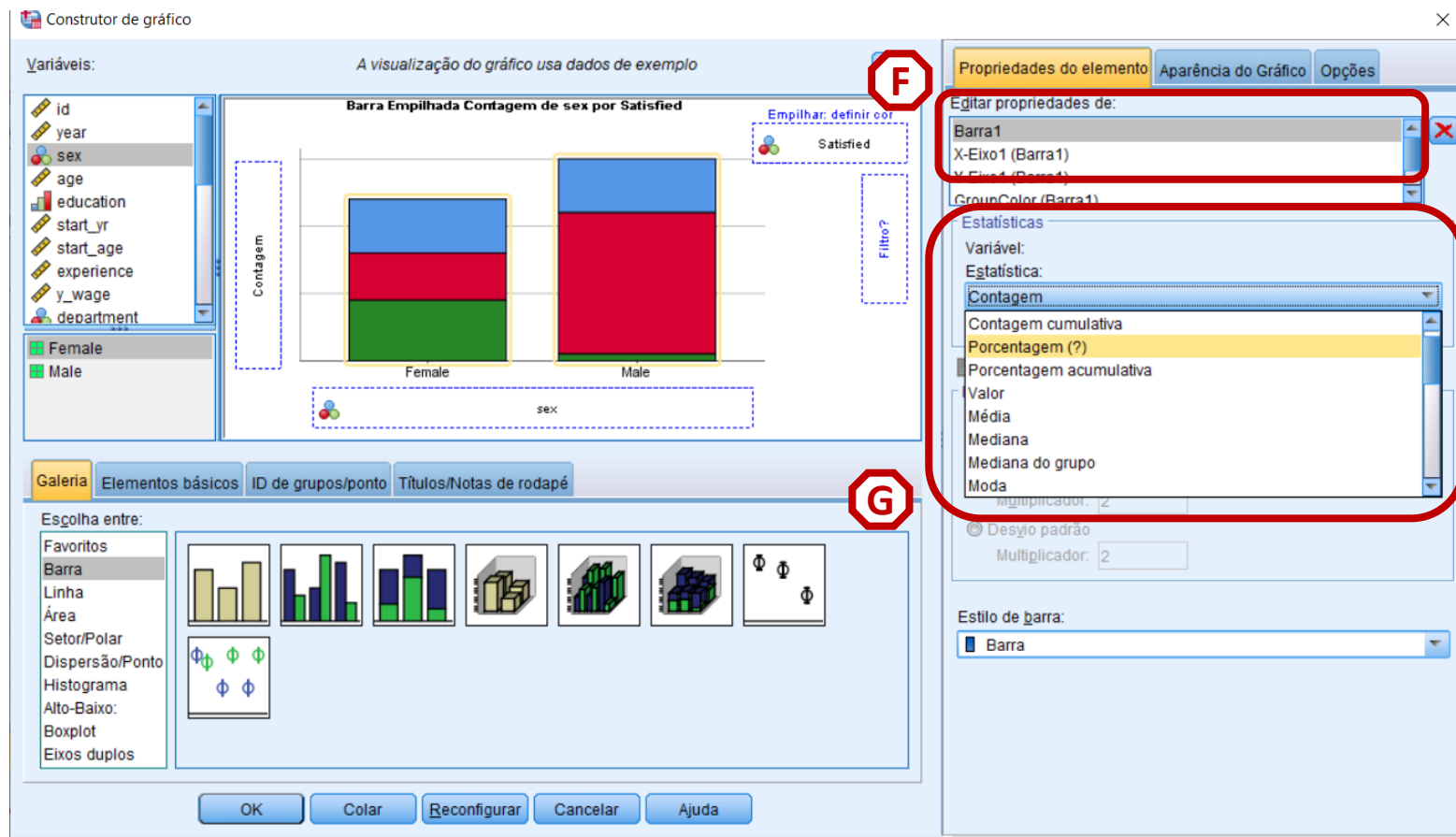
E

Gráfico de Barras Sobrepostas

- Seleccionar 'Barra 1'
- Definir 'Percentagens' como parâmetro de comparação

F

G



Construtor de gráfico

A visualização do gráfico usa dados de exemplo

Variáveis:

- id
- year
- sex
- age
- education
- start_yr
- start_age
- experience
- y_wage
- department

Female

Male

Barra Empilhada Contagem de sex por Satisfied

Empilhar: definir cor

Satisfied

Contagem

Female

Male

sex

Galeria

Escolha entre:

- Favoritos
- Barra
- Linha
- Área
- Setor/Polar
- Dispersão/Ponto
- Histograma
- Alto-Baixo:
- Boxplot
- Eixos duplos

OK

Color

Reconfigurar

Cancelar

Ajuda

Propriedades do elemento

Aparência do Gráfico

Opções

Editar propriedades de:

Barra1

X-Eixo1 (Barra1)

Y-Eixo1 (Barra1)

GroupColor (Barra1)

Estadísticas

Variável:

Estadística:

Contagem

Contagem cumulativa

Percentagem (?)

Percentagem acumulativa

Valor

Média

Mediana

Mediana do grupo

Moda

Multiplicador: 2

Desvio padrão

Multiplicador: 2

Estilo de barra:

Barra

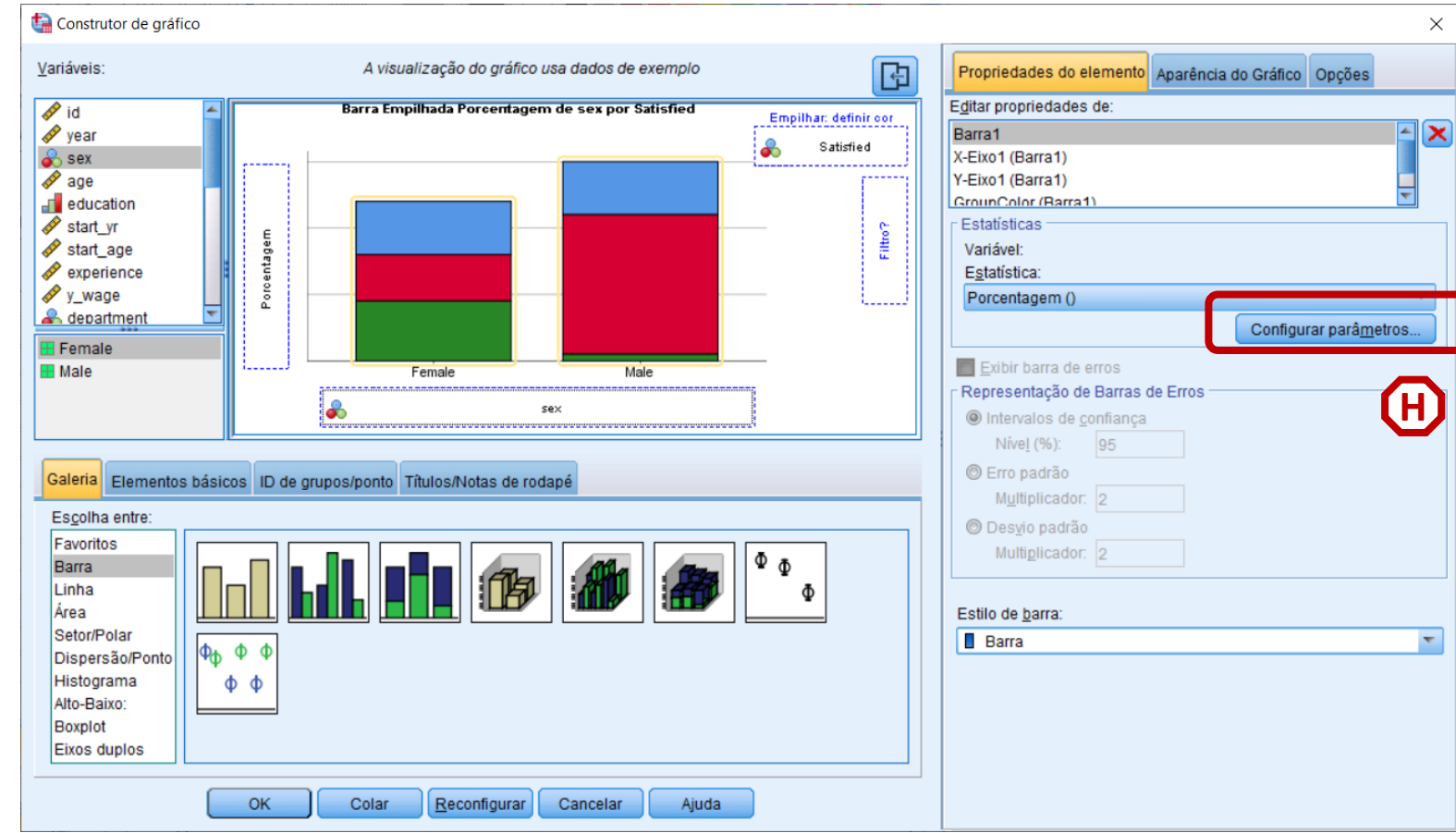
Gráfico de Barras Sobrepostas

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Seleccionar 'Configurar Parâmetros'

F

G

H



H

Gráfico de Barras Sobrepostas

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Seleccionar 'Configurar Parâmetros'
- Seleccionar o denominador 'Total para cada categoria do eixo X'
- Seleccionar 'Continuar' / 'OK'

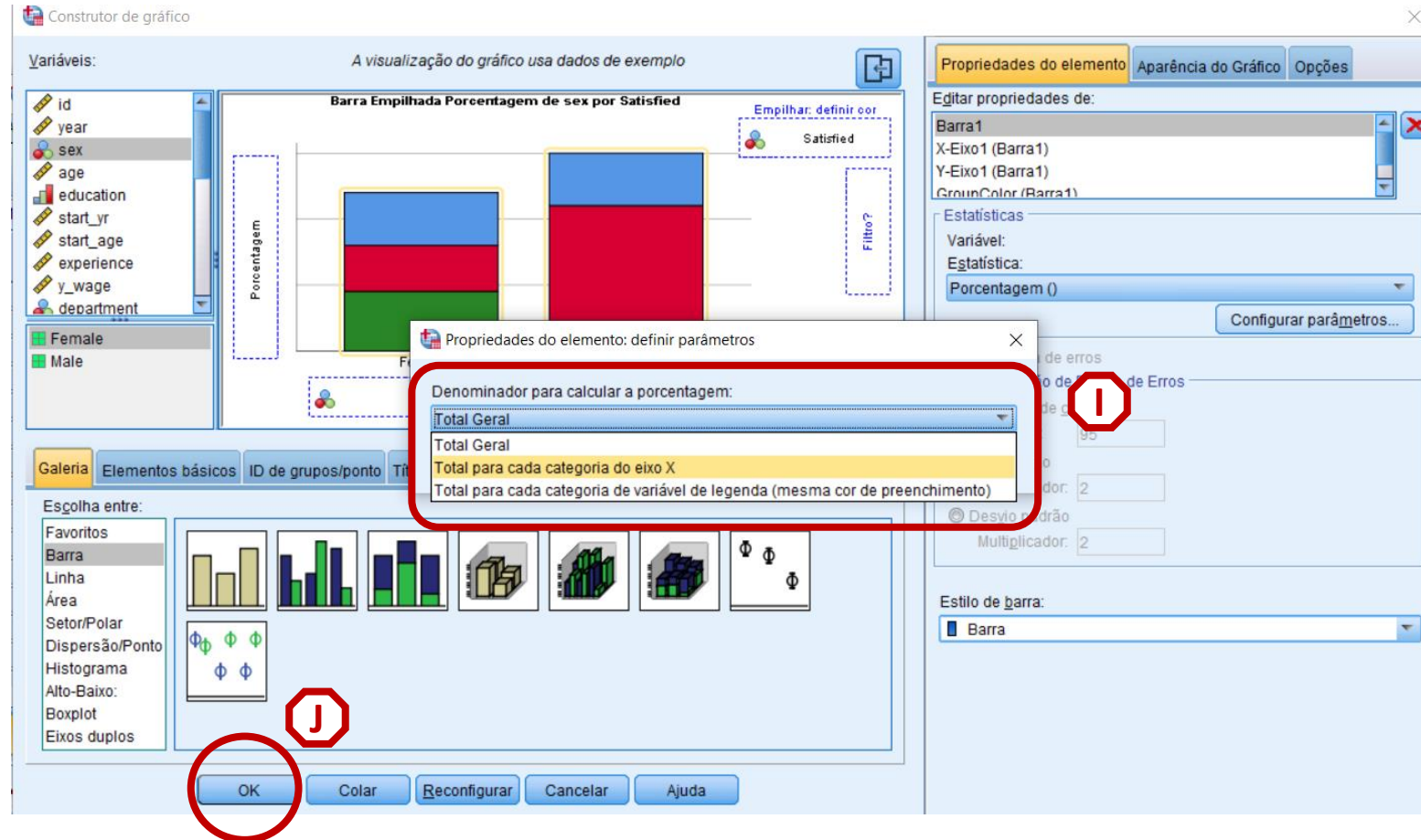
F

G

H

I

J



Construtor de gráfico

A visualização do gráfico usa dados de exemplo

Variáveis:

- id
- year
- sex
- age
- education
- start_yr
- start_age
- experience
- y_wage
- department

Female

Male

Barra Empilhada Porcentagem de sex por Satisfied

Empilhar: definir cor

Satisfied

Porcentagem

Filtro?

Propriedades do elemento: definir parâmetros

Denominador para calcular a porcentagem:

- Total Geral
- Total para cada categoria do eixo X
- Total para cada categoria de variável de legenda (mesma cor de preenchimento)

OK

Color

Reconfigurar

Cancelar

Ajuda

Propriedades do elemento

Aparência do Gráfico

Opções

Editar propriedades de:

Barra1

X-Eixo1 (Barra1)

Y-Eixo1 (Barra1)

GroupColor (Barra1)

Estadísticas

Variável:

Estatística:

Porcentagem ()

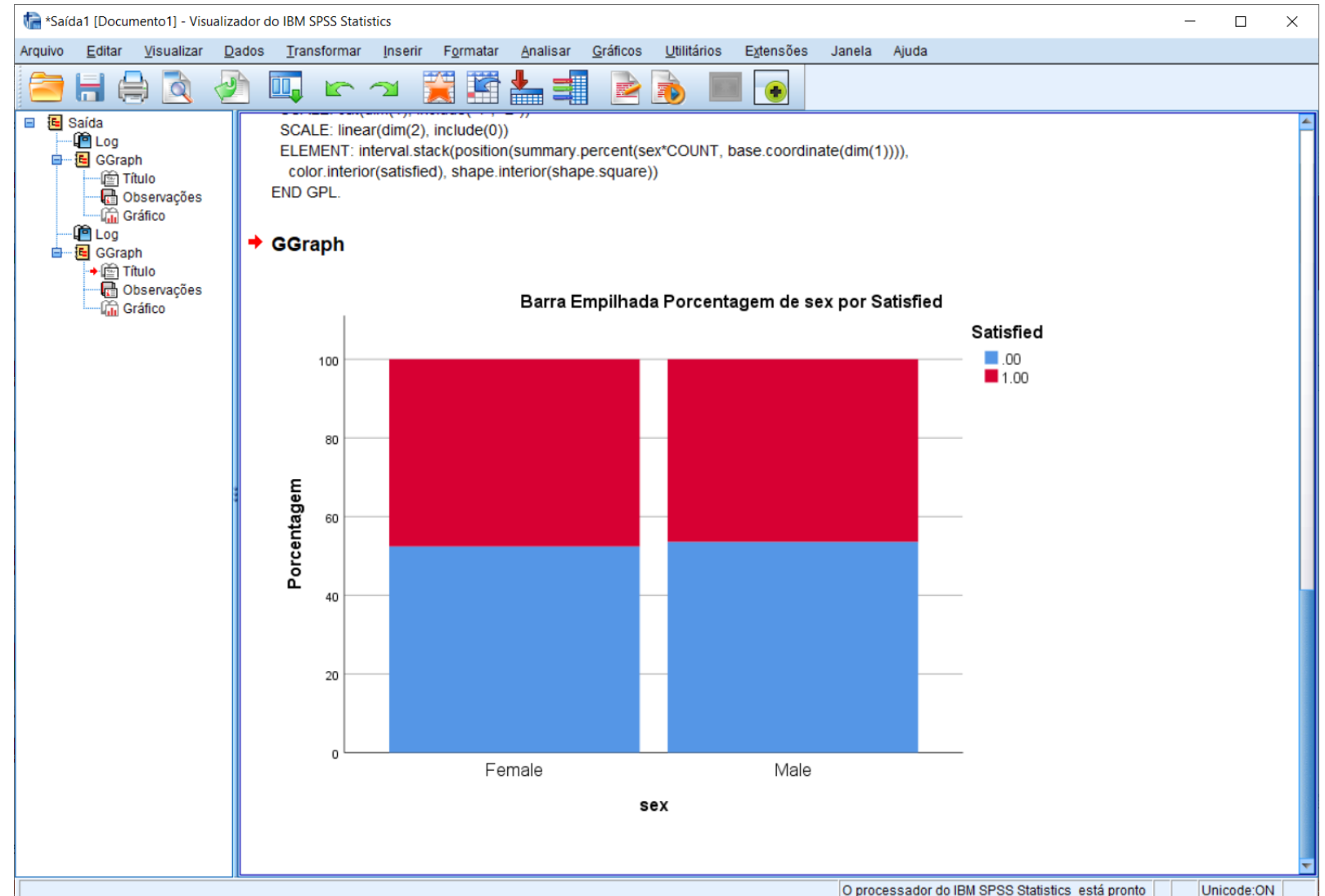
Configurar parâmetros...

Estilo de barra:

Barra

Gráfico de Barras Sobrepostas

- O resultado é publicado no 'Visualizador de Resultados'
- Não se notam diferenças significativas na satisfação com as condições de trabalho em função do sexo...
- Gráfico deve ser complementado com tabela de dupla entrada para estas variáveis (ver Aula 6)



- **Objectivo:**

- **Comparar a idade média dos trabalhadores que estão satisfeitos com as condições de trabalho com a idade média dos trabalhadores que não estão satisfeitos.**

... neste caso, através de um Gráfico de 'Barra de Erro Simples'



Gráfico de Barra de Erro Simples

- Selecionar 'Gráficos' / 'Construtor de Gráfico'

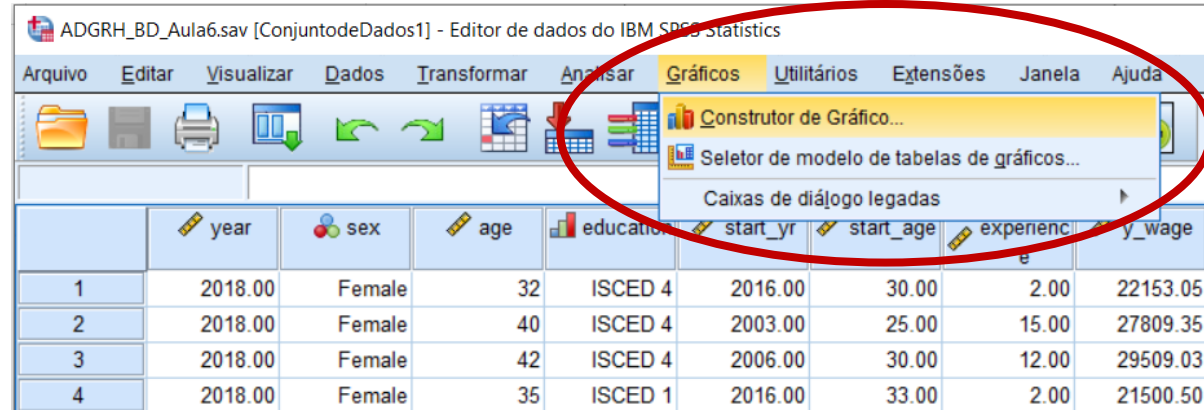


Gráfico de Barra de Erro Simples

- Seleccionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico Barra de Erro Simples'

A

B

C

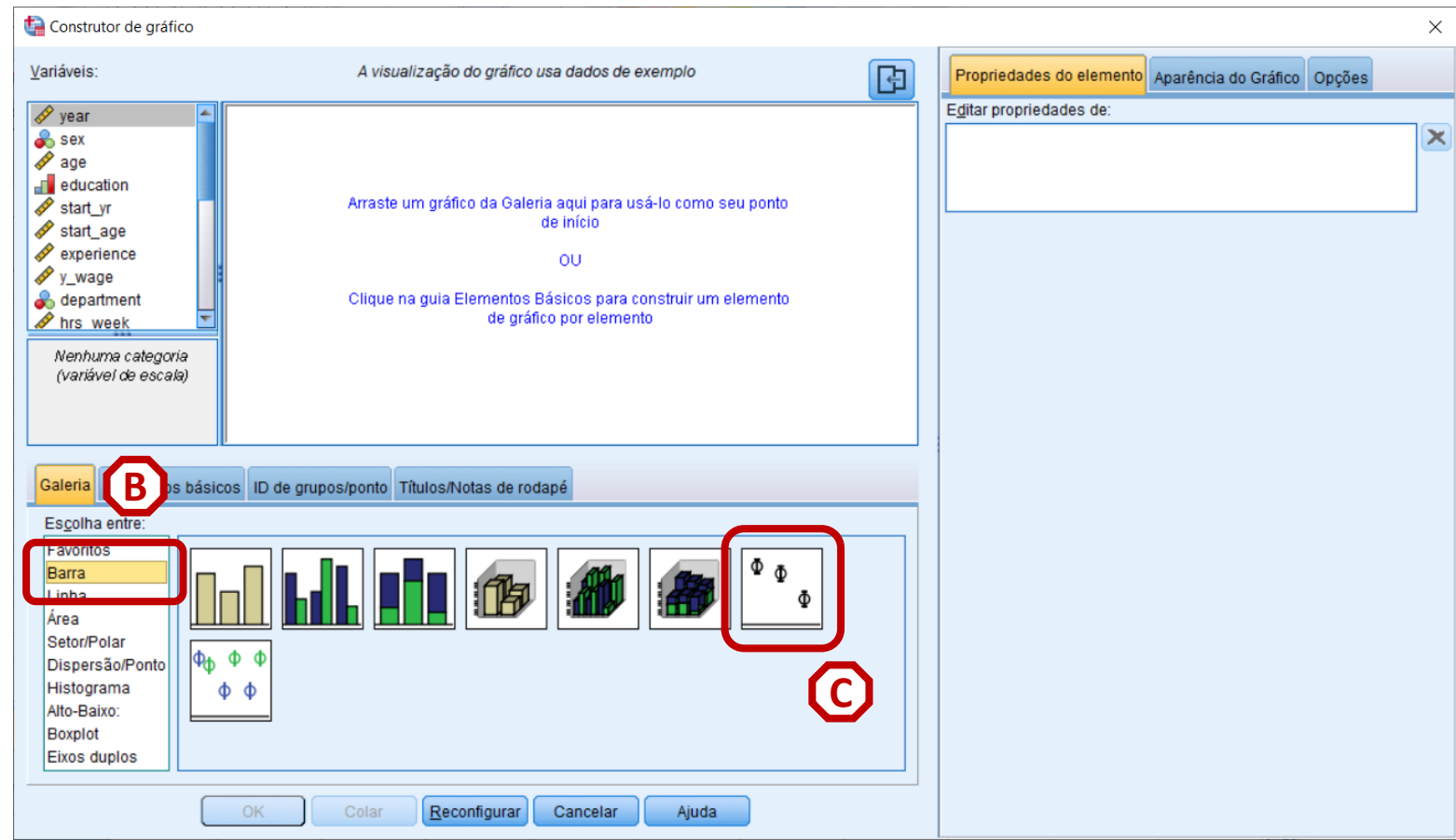


Gráfico de Barra de Erro Simples

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
 - Selecione 'Barra'
 - Duplo-Clique na opção 'Gráfico Barra de Erro Simples'
 - Selecionar a variável 'satisfied'...
- ... e colocar na caixa do 'Eixo X'

Exercício: Colocar a variável 'age' no 'Eixo Y'

A

B

C

D

E

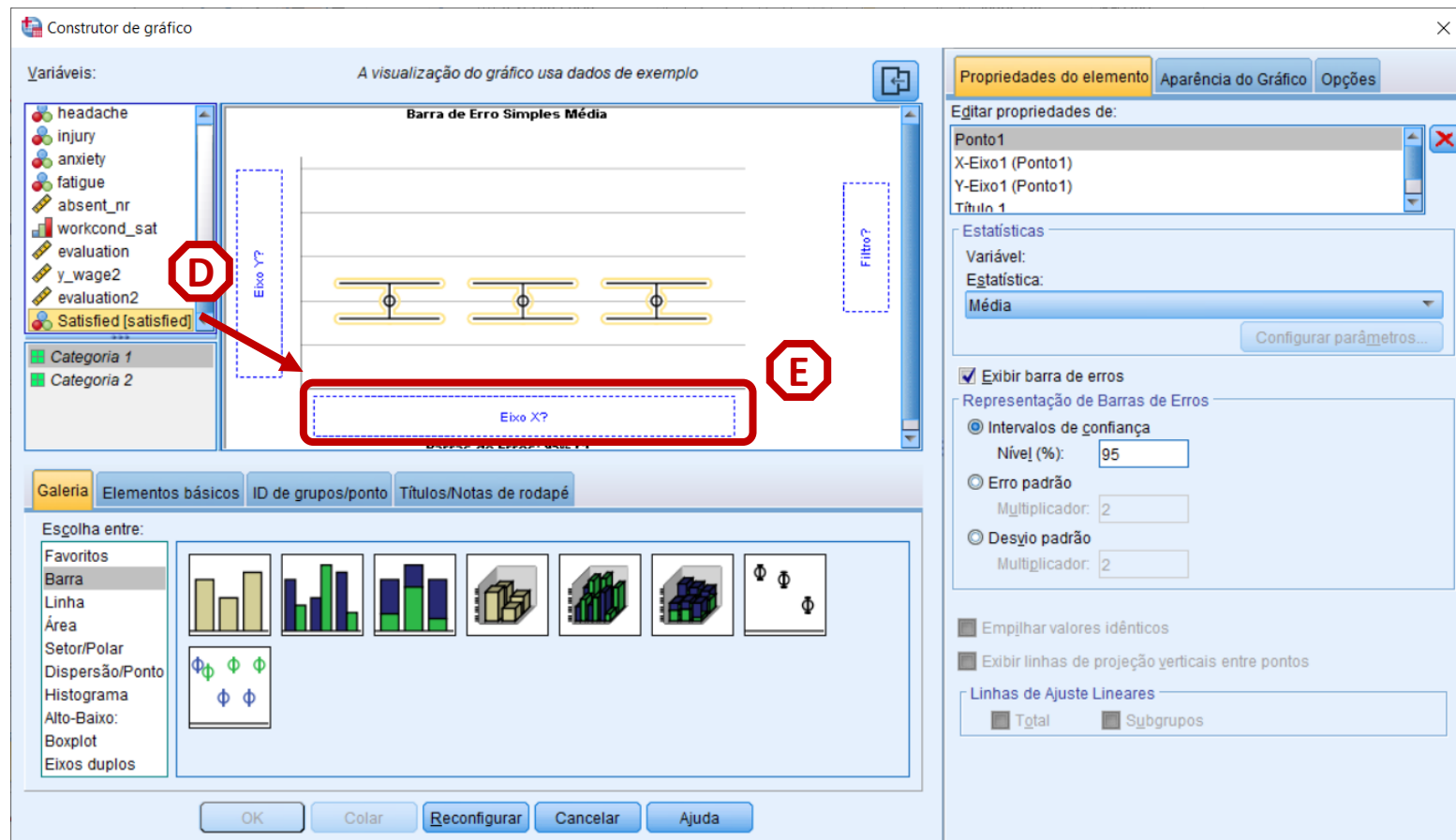


Gráfico de Barra de Erro Simples

- Escolher a 'Média' como estatística a ser representada no gráfico



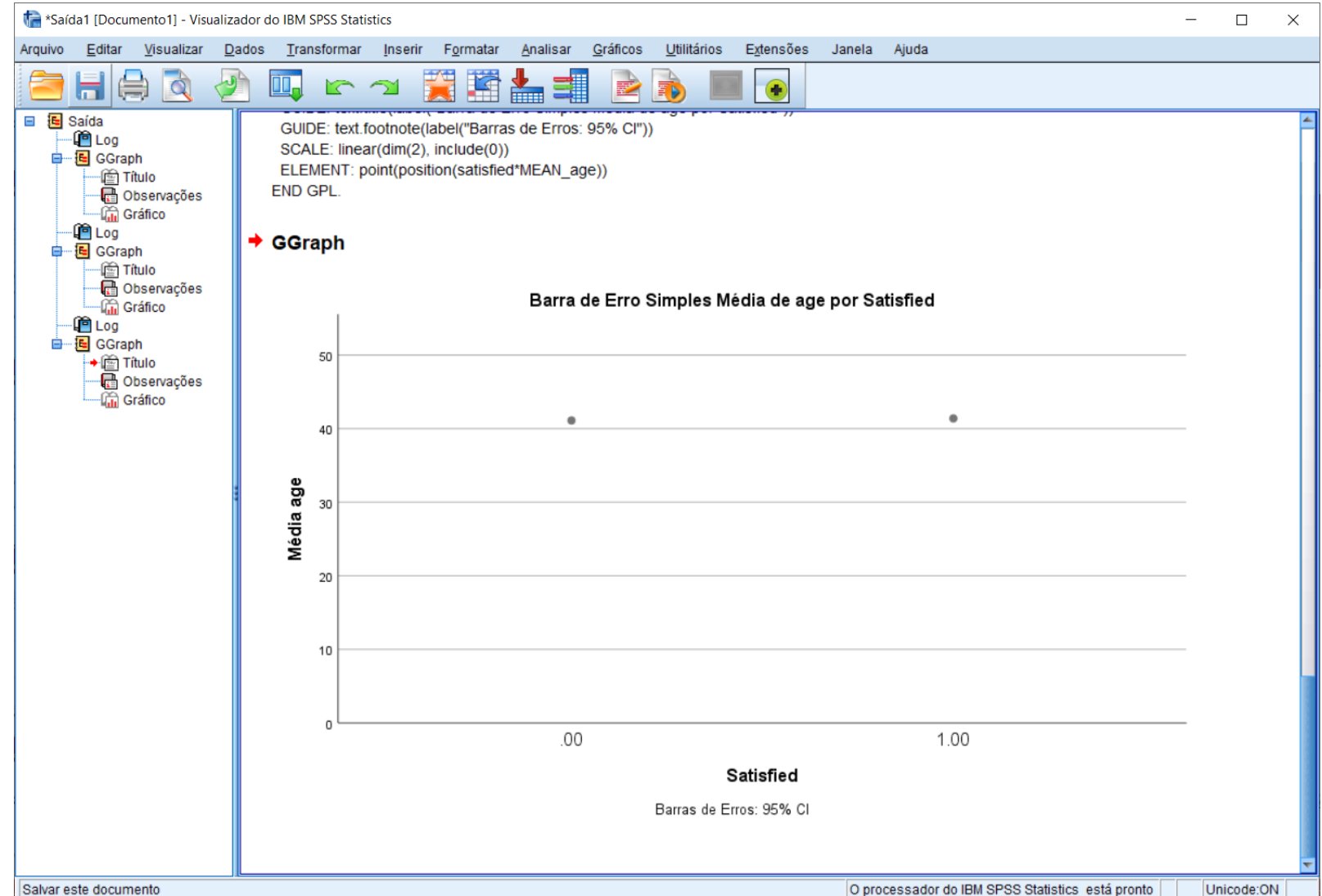
Gráfico de Barra de Erro Simples

- Escolher a 'Média' como estatística a ser representada no gráfico
- Desligar a opção 'Exibir Barra de Erros'
- Clicar 'OK'



Gráfico de Barra de Erro Simples

- O resultado é publicado no 'Visualizador de Resultados'
- Não se notam diferenças significativas em termos da idade média relativamente à sua satisfação com as condições de trabalho ...
- Gráfico deve ser complementado com tabela de comparação de médias (ver Aula 6)



- **Objectivo:**
 - **Perceber como se relaciona a variável dependente ('satisfied') com a variável independente 'education' que queremos incluir no nosso modelo de regressão**

... neste caso, através de um Gráfico de Barras Agrupadas

Gráfico de Barras Agrupadas

- Selecionar 'Gráficos' / 'Construtor de Gráfico'

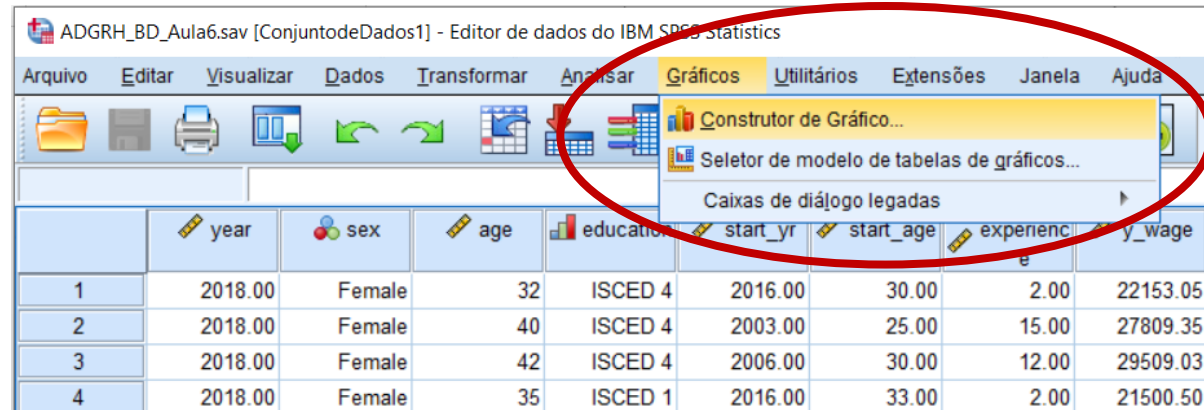


Gráfico de Barras Agrupadas

- Seleccionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Agrupadas'

A

B

C

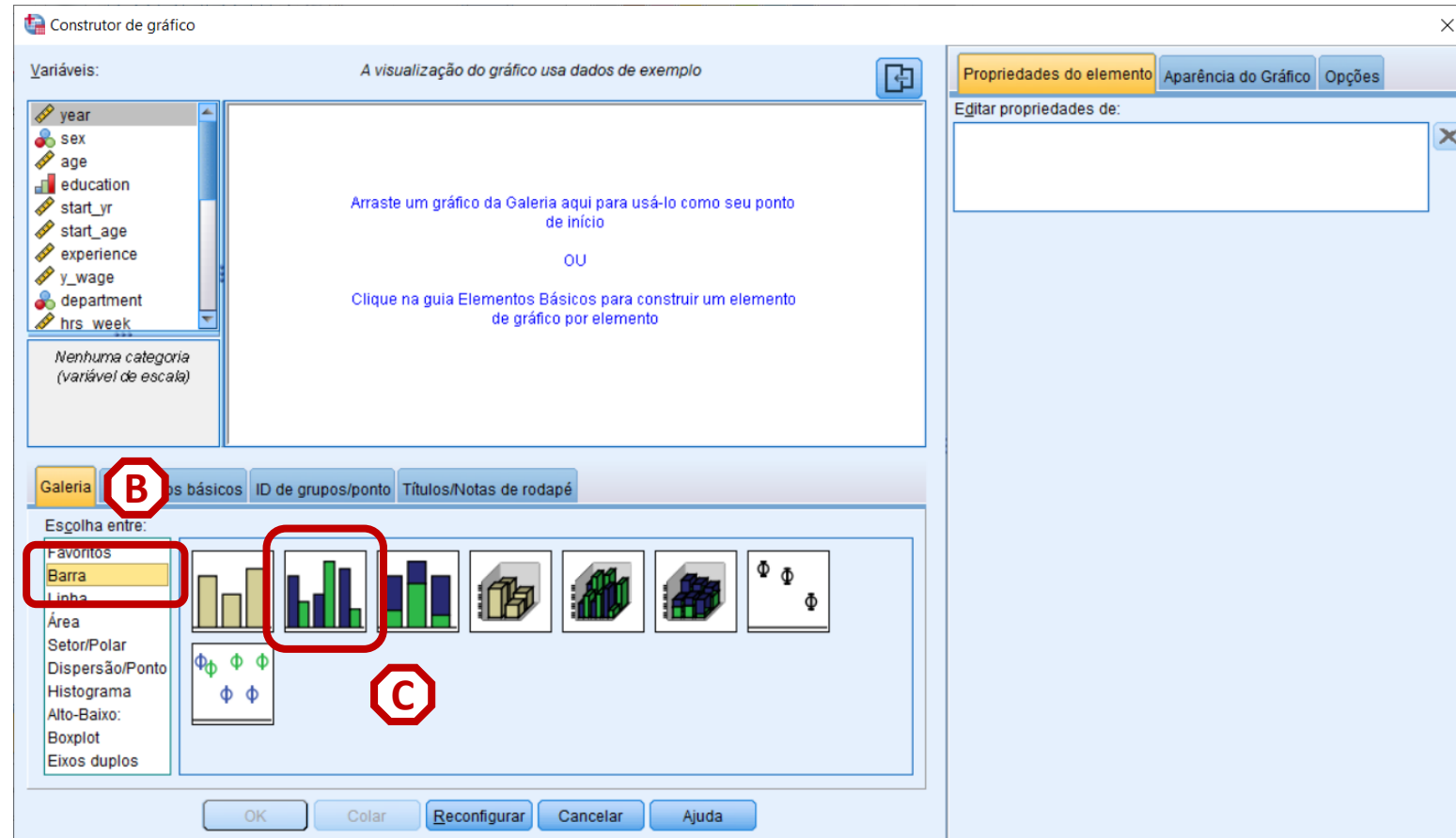


Gráfico de Barras Agrupadas

- Selecionar 'Gráficos' / 'Construtor de Gráfico'
- Selecione 'Barra'
- Duplo-Clique na opção 'Gráfico de Barras Agrupadas'
- Selecionar a variável 'satisfied'...

... e colocar na caixa do 'Cluster em X'

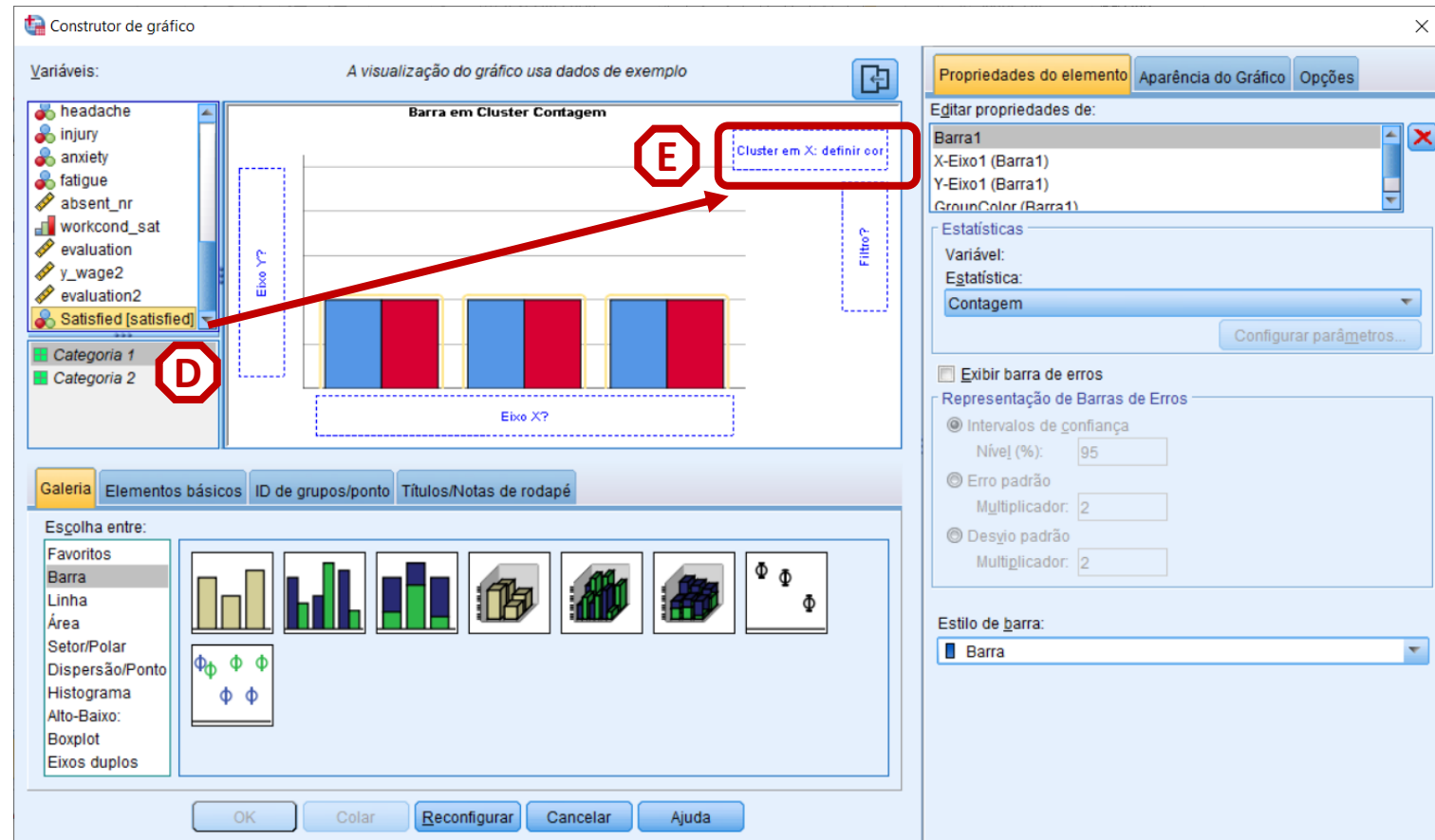
A

B

C

D

E



Construtor de gráfico

A visualização do gráfico usa dados de exemplo

Variáveis:

- headache
- injury
- anxiety
- fatigue
- absent_nr
- workcond_sat
- evaluation
- y_wage2
- evaluation2
- Satisfied [satisfied]
- Categoria 1
- Categoria 2

Barra em Cluster Contagem

Cluster em X: definir cor

Eixo Y?

Eixo X?

Filtro?

Galeria

Elementos básicos ID de grupos/ponto Títulos/Notas de rodapé

Escolha entre:

- Favoritos
- Barra
- Linha
- Área
- Setor/Polar
- Dispersão/Ponto
- Histograma
- Alto-Baixo:
- Boxplot
- Eixos duplos

Propriedades do elemento Aparência do Gráfico Opções

Editar propriedades de:

Barra1

X-Eixo1 (Barra1)

Y-Eixo1 (Barra1)

GroupColor (Barra1)

Estadísticas

Variável:

Estatística:

Contagem

Configurar parâmetros...

Exibir barra de erros

Representação de Barras de Erros

- Intervalos de confiança
- Nível (%): 95
- Erro padrão
- Multiplicador: 2
- Desvio padrão
- Multiplicador: 2

Estilo de barra:

Barra

OK Colar Reconfigurar Cancelar Ajuda

Exercício: Colocar a variável 'education' no 'Eixo X'

Gráfico de Barras Agrupadas

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação

F

G

Construtor de gráfico

Variáveis: A visualização do gráfico usa dados de exemplo

Barra em Cluster Porcentagem de education por Satisfied

Cluster em X: definir cor Satisfied

Porcentagem

education

Propriedades do elemento Aparência do Gráfico Opções

Editar propriedades de: Barra1

X-Eixo1 (Barra1)

Y-Eixo1 (Barra1)

Estadísticas

Variável:

Estadística: Porcentagem ()

Contagem

Contagem cumulativa

Porcentagem ()

Porcentagem acumulativa

Valor

Média

Mediana

Mediana do grupo

Desvio padrão

Multiplicador: 2

Estilo de barra: Barra

Galeria Elementos básicos ID de grupos/ponto Títulos/Notas de rodapé

Escolha entre:

Favoritos

Barra

Linha

Área

Setor/Polar

Dispersão/Ponto

Histograma

Alto-Baixo:

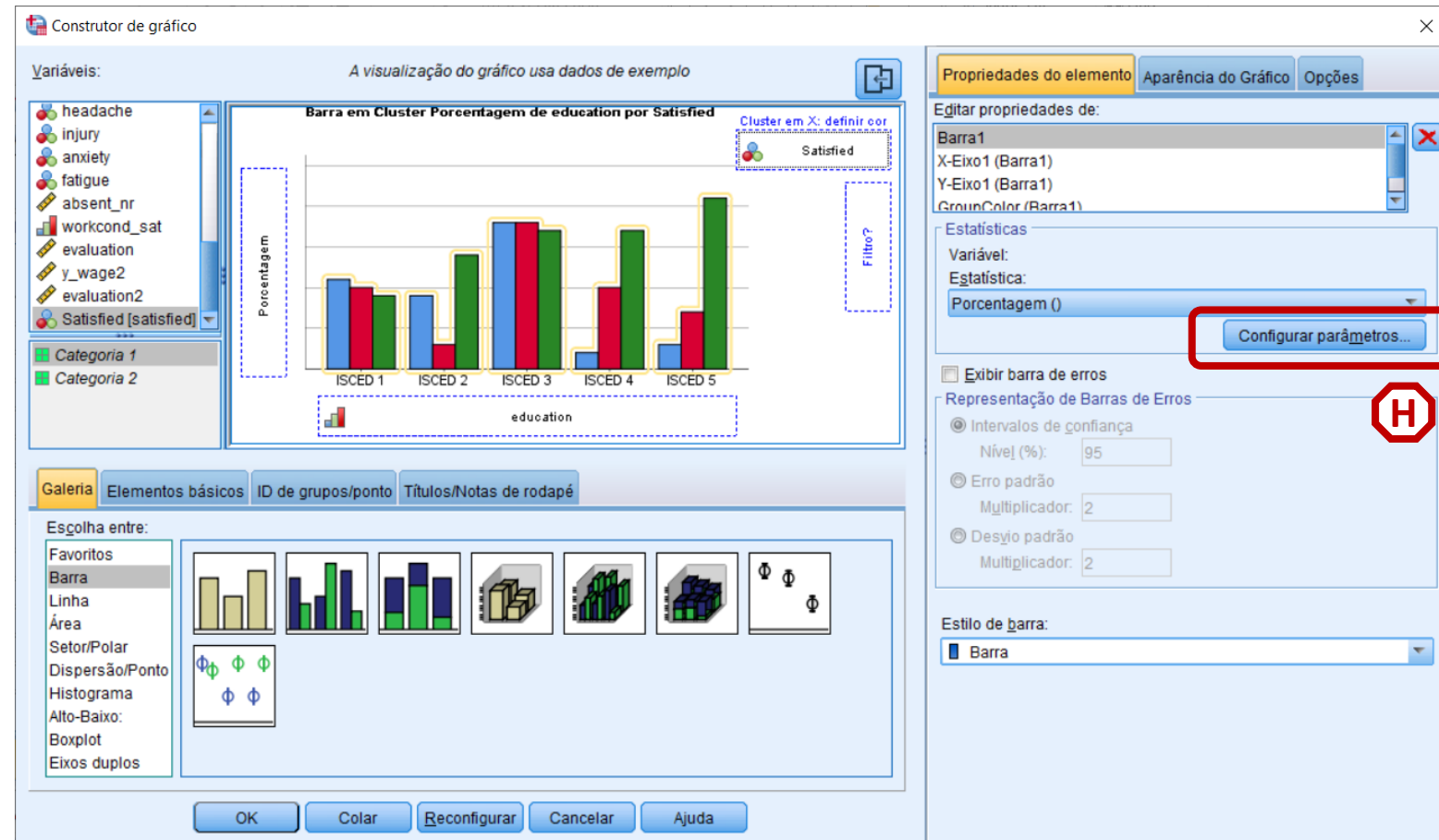
Boxplot

Eixos duplos

OK Colar Reconfigurar Cancelar Ajuda

Gráfico de Barras Agrupadas

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Seleccionar 'Configurar parâmetros'

Construtor de gráfico

Variáveis: A visualização do gráfico usa dados de exemplo

Barra em Cluster Porcentagem de education por Satisfied

Porcentagem

education

ISCED 1 ISCED 2 ISCED 3 ISCED 4 ISCED 5

Cluster em X: definir cor Satisfied

Propriedades do elemento Aparência do Gráfico Opções

Editar propriedades de:

Barra1

X-Eixo1 (Barra1)

Y-Eixo1 (Barra1)

GroupColor (Barra1)

Estatísticas

Variável:

Estatística:

Porcentagem ()

Configurar parâmetros...

Exibir barra de erros

Representação de Barras de Erros

Intervalos de confiança

Nível (%): 95

Erro padrão

Multiplicador: 2

Desvio padrão

Multiplicador: 2

Estilo de barra:

Barra

OK Colar Reconfigurar Cancelar Ajuda

Gráfico de Barras Agrupadas

- Seleccionar 'Barra 1'
- Definir 'Porcentagem' como parâmetro de comparação
- Seleccionar 'Configurar parâmetros'
- Seleccionar o denominador 'Total para cada categoria do eixo X'
- Seleccionar 'Continuar' / 'OK'

F

G

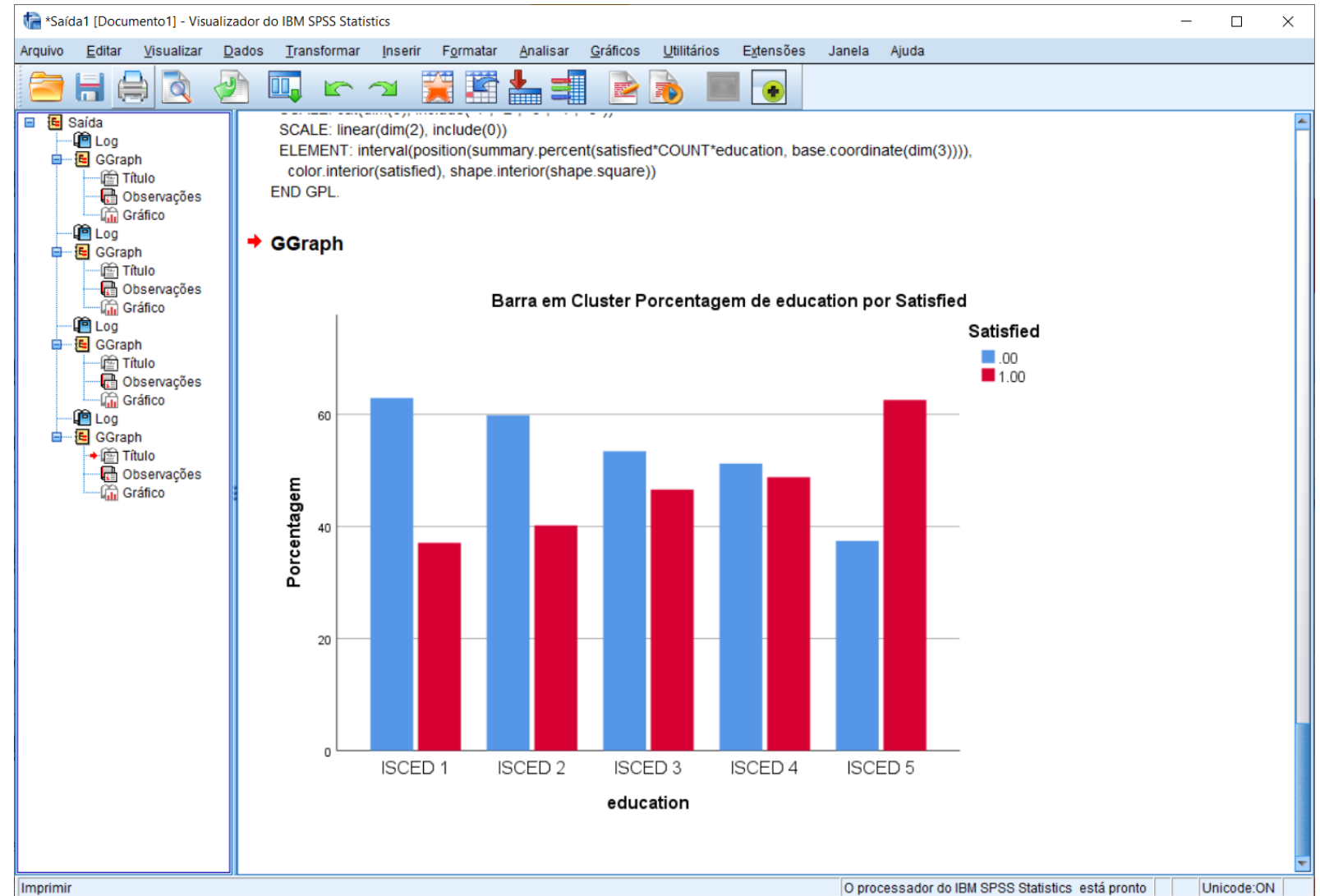
H

I

J

Gráfico de Barras Agrupadas

- O resultado é publicado no 'Visualizador de Resultados'
- Podemos perceber que a percentagem de trabalhadores satisfeitos com as condições do trabalho aumenta em função da educação
- Deve ser complementado com uma tabela de dupla entrada (ver Aula 6)



Implementação do Modelo de Regressão Logística

4. Definir o modelo de regressão logística

- **Objectivo:**

- **Estimar um modelo de regressão logística que explique em que medida a probabilidade de se sentir satisfeito com as condições de trabalho varia em função do sexo, idade e educação.**



Regressão Logística

- Selecionar 'Analisar' / 'Regressão' / 'Logística binária'



*ADGRH_BD_Aula11.sav [ConjuntodeDados] - Editor de dados do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar **Analisar** Gráficos Utilitários Extensões Janelas Ajuda

12 : y_wage2 29668.4665929268 8

	absent_nr	workcond_sat	evaluation
1	9	3	49.08
2	10	2	52.58
3	15	2	53.80
4	10	2	50.42
5	11	4	48.43
6	10	2	54.34
7	6	3	49.65
8	10	4	46.84
9	8	2	50.76
10	17	3	53.33
11	14	2	50.59

Relatórios
Estatística Descritiva
Estatísticas Bayesianas
Tabelas
Comparar Médias
Modelo Linear Geral
Modelos lineares generalizados
Modelos mistos
Correlacional
Regressão
Log linear
Redes neurais
Classificar
Redução de dimensão
Escala

Modelagem Linear Automática...
Linear...
Curva de estimação...
Quadrados Mínimos Parciais...
Logística binária...

Regressão Logística

- Selecionar 'Analisar' / 'Regressão' / 'Logística binária'
- Selecionar a variável 'satisfied'
- Colocar na caixa 'Dependente'

Exercício: Colocar as variáveis 'sex', 'age' e 'education' na caixa 'Covariáveis'

A

B

C

B

The screenshot shows the IBM SPSS Statistics interface with the 'Regressão logística' dialog box open. The 'Dependente:' field is empty, and the 'Covariáveis:' field is also empty. The 'Método:' is set to 'Inserir'. The 'Satisfied [satisfied]' variable is selected in the list on the left. Red circles and boxes highlight the 'Dependente:' field, the 'Satisfied [satisfied]' variable, and the 'Covariáveis:' field.

Regressão Logística

- Selecionar 'Analisar' / 'Regressão' / 'Logística binária'



- Selecionar a variável 'satisfied'

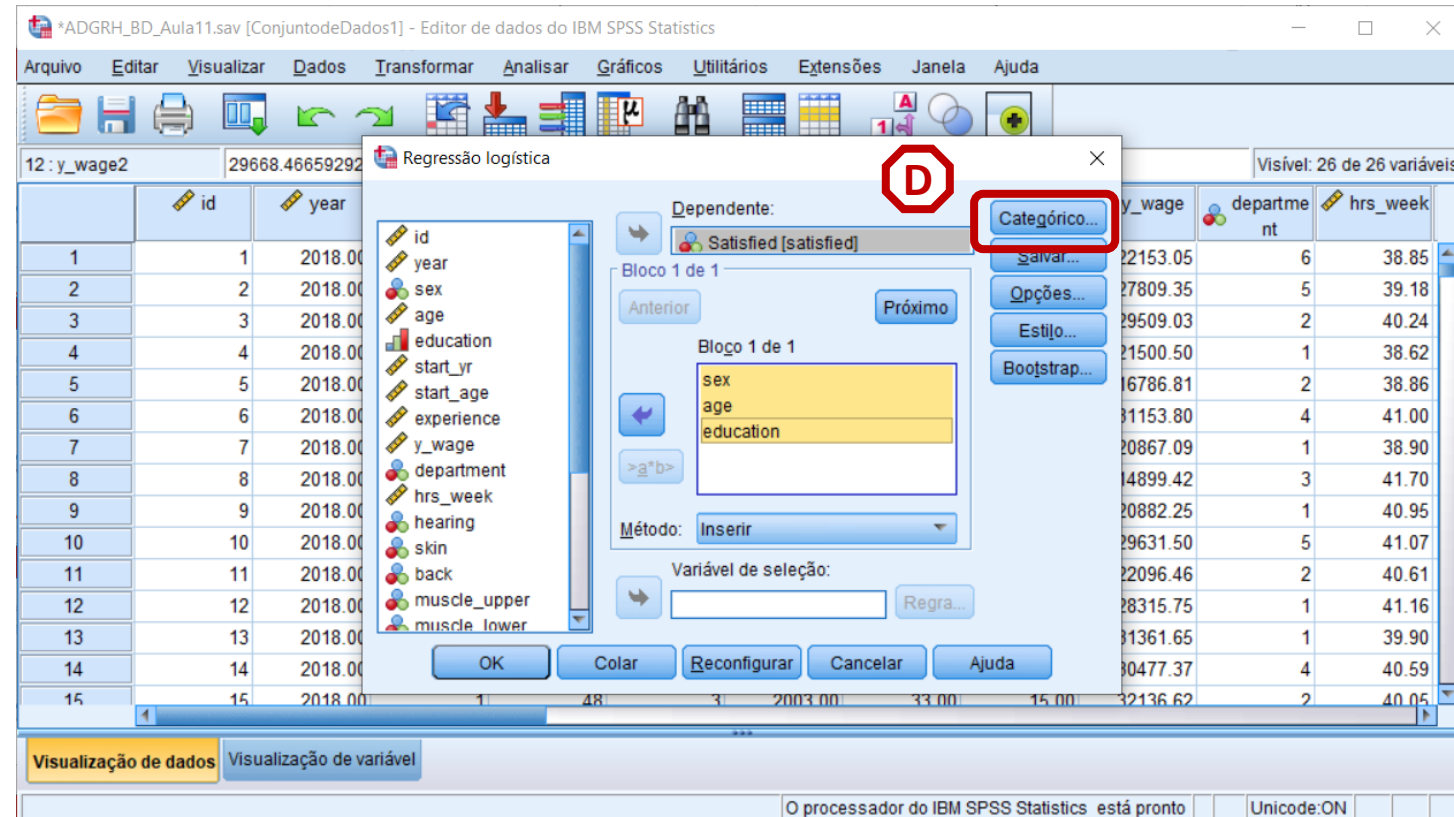


- Colocar na caixa 'Dependente'



Exercício: Colocar as variáveis 'sex', 'age' e 'education' na caixa 'Covariáveis'

- Clicar no botão 'Categórico'



The screenshot shows the IBM SPSS Statistics interface with the 'Regressão logística' dialog box open. The dependent variable is 'Satisfied [satisfied]'. The independent variables are 'sex', 'age', and 'education'. The 'Categorical...' button is highlighted with a red box and a red octagon marker D. The background shows a data table with columns 'id', 'year', 'y_wage', 'department', and 'hrs_week'.

id	year	y_wage	department	hrs_week
1	2018.00	22153.05	6	38.85
2	2018.00	27809.35	5	39.18
3	2018.00	29509.03	2	40.24
4	2018.00	21500.50	1	38.62
5	2018.00	16786.81	2	38.86
6	2018.00	31153.80	4	41.00
7	2018.00	20867.09	1	38.90
8	2018.00	14899.42	3	41.70
9	2018.00	20882.25	1	40.95
10	2018.00	29631.50	5	41.07
11	2018.00	22096.46	2	40.61
12	2018.00	28315.75	1	41.16
13	2018.00	31361.65	1	39.90
14	2018.00	30477.37	4	40.59
15	2018.00	32136.62	2	40.05

Regressão Logística

- Selecionar a variável 'sex'
- Colocar na caixa 'Covariáveis categóricas'



The screenshot displays the IBM SPSS Statistics interface. A dialog box titled 'Regressão logística: definir variáveis de categoria' is open. In the 'Covariáveis' list on the left, the variable 'sex' is highlighted with a red box and a red arrow pointing to the 'Covariáveis categóricas' box on the right. Below this, the 'Alterar Contraste' section is visible, with 'Indicador' selected in the 'Contraste' dropdown and 'Último' selected as the 'Categoria de referência'. The background shows a data table with columns 'wage', 'department', and 'hrs_week'. The status bar at the bottom indicates 'O processador do IBM SPSS Statistics está pronto' and 'Unicode:ON'.

Regressão Logística

- Selecionar a variável 'sex'
- Colocar na caixa 'Covariáveis categóricas'
- Definir a categoria 'male' (valor 1 *) como categoria de referência na variável 'sex'

* neste caso, escolhemos o valor 1 que é o mais baixo (i.e., o primeiro)

- Carregar no botão 'Alterar'



*ADGRH_BD_Aula11.sav [ConjuntodeDados1] - Editor de dados do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Analisar Gráficos Utilitários Extensões Janela Ajuda

12 : y_wage2 29668.46659292

15

1 1 2018.0 2153.05 6 38.85

2 2 2018.0 7809.35 5 39.18

3 3 2018.0 9509.03 2 40.24

4 4 2018.0 1500.50 1 38.62

5 5 2018.0 6786.81 2 38.86

6 6 2018.0 1153.80 4 41.00

7 7 2018.0 2867.09 1 38.90

8 8 2018.0 19.42 3 41.70

9 9 2018.0 882.25 1 40.95

10 10 2018.0 9631.50 5 41.07

11 11 2018.0 2096.46 2 40.61

12 12 2018.0 8315.75 1 41.16

13 13 2018.0 1361.65 1 39.90

14 14 2018.0 30477.37 4 40.59

15 15 2018.0 32136.62 2 40.05

Regressão logística

Regressão logística: definir variáveis de categoria

Covariáveis:

- age
- education

Covariáveis categóricas:

- sex(Indicador)

Alterar Contraste

Contraste: Indicador

Categoria de referência: Último Primeiro

Alterar

Continuar Cancelar Ajuda

OK Colar Reconfigurar Cancelar Ajuda

Visualização de dados Visualização de variável

O processador do IBM SPSS Statistics está pronto Unicode:ON

Regressão Logística

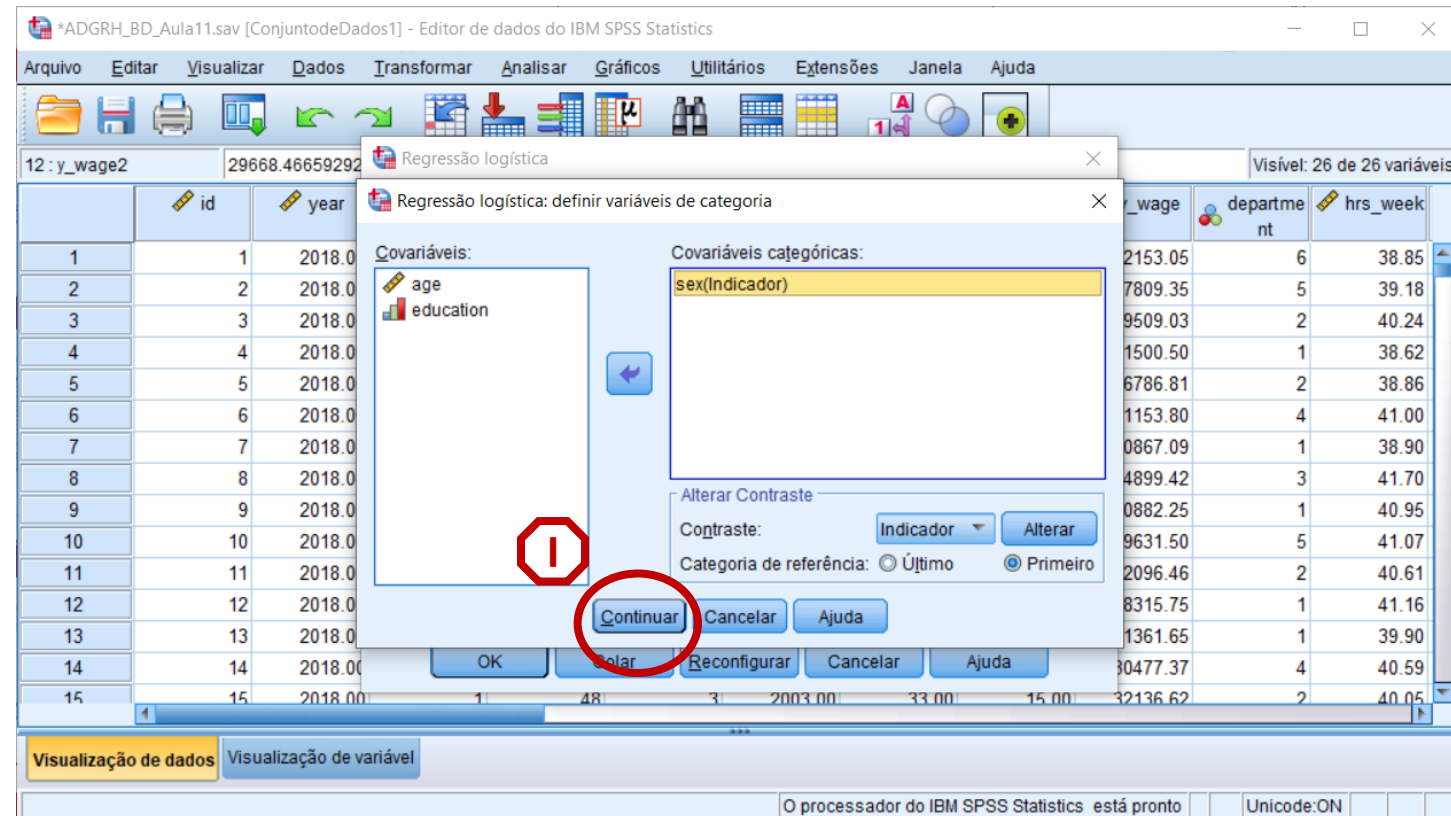
- Selecionar a variável 'sex'
- Colocar na caixa 'Covariáveis categóricas'
- Definir a categoria 'male' (valor 1 *) como categoria de referência na variável 'sex'

* neste caso, escolhemos o valor 1 que é o mais baixo (i.e., o primeiro)

- Carregar no botão 'Alterar'

Exercício: Repetir esta operação para a variável 'education'

- Selecionar 'Continuar'



	id	year	wage	department	hrs_week
1	1	2018.0	2153.05	6	38.85
2	2	2018.0	7809.35	5	39.18
3	3	2018.0	9509.03	2	40.24
4	4	2018.0	1500.50	1	38.62
5	5	2018.0	6786.81	2	38.86
6	6	2018.0	1153.80	4	41.00
7	7	2018.0	0867.09	1	38.90
8	8	2018.0	4899.42	3	41.70
9	9	2018.0	0882.25	1	40.95
10	10	2018.0	9631.50	5	41.07
11	11	2018.0	2096.46	2	40.61
12	12	2018.0	8315.75	1	41.16
13	13	2018.0	1361.65	1	39.90
14	14	2018.0	30477.37	4	40.59
15	15	2018.0	32136.62	2	40.05

Regressão Logística

- Clicar no botão 'Opções'



*ADGRH_BD_Aula11.sav [ConjuntodeDados1] - Editor de dados do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Analisar Gráficos Utilitários Extensões Janela Ajuda

12: y_wage2 29668.46659292

absent_nr workcond_sat

	absent_nr	workcond_sat
1	9	
2	10	
3	15	
4	10	
5	11	
6	10	
7	6	
8	10	
9	8	
10	17	
11	14	
12	7	
13	13	
14	7	
15	11	

Regressão logística

Dependente: Satisfied [satisfied]

Bloco 1 de 1

Anterior Próximo

Bloco 1 de 1

sex(Cat)
age
department(Cat)

Método: Inserir

Variável de seleção:

OK Colar Reconfigurar Cancelar Ajuda

Visualização de dados Visualização de variável

O processador do IBM SPSS Statistics está pronto Unicode:ON

Regressão Logística

- Clicar no botão 'Opções'
- Selecionar a opção 'Qualidade do ajuste de Hosmer-Lemeshow'
- Selecionar a opç 'satisfied'
- Selecionar 'Continuar' / 'OK'

J

K

L

M

*ADGRH_BD_Aula11.sav [ConjuntodeDados1] - Editor de dados do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Analisar Gráficos Utilitários Extensões Janela Ajuda

12: y_wage2 296

Regressão logística: opções

Estadísticas e gráficos

- Gráficos de classificação
- Qualidade do ajuste de Hosmer-Lemeshow
- Listagem por caso de resíduos
- Valgres discrepantes no lado de fora 2 std. dev.
- Todos os casos

Exibir

- Em cada passo
- Na última passo

Probabilidade para Stepwise

Entrada: 0.05 Remoção: 0.10

Corte de classificação: 0.5

Máximo de iterações: 20

- Preservar memória para análises complexas ou grandes conjuntos de dados
- Incluir constante no modelo

Continuar Cancelar Ajuda

departme	hrs_week
6	38.85
5	39.18
2	40.24
1	38.62
2	38.86
4	41.00
1	38.90
3	41.70
1	40.95
5	41.07
2	40.61
1	41.16
1	39.90
4	40.59
2	40.05

Visualização de dados Visualização de variável

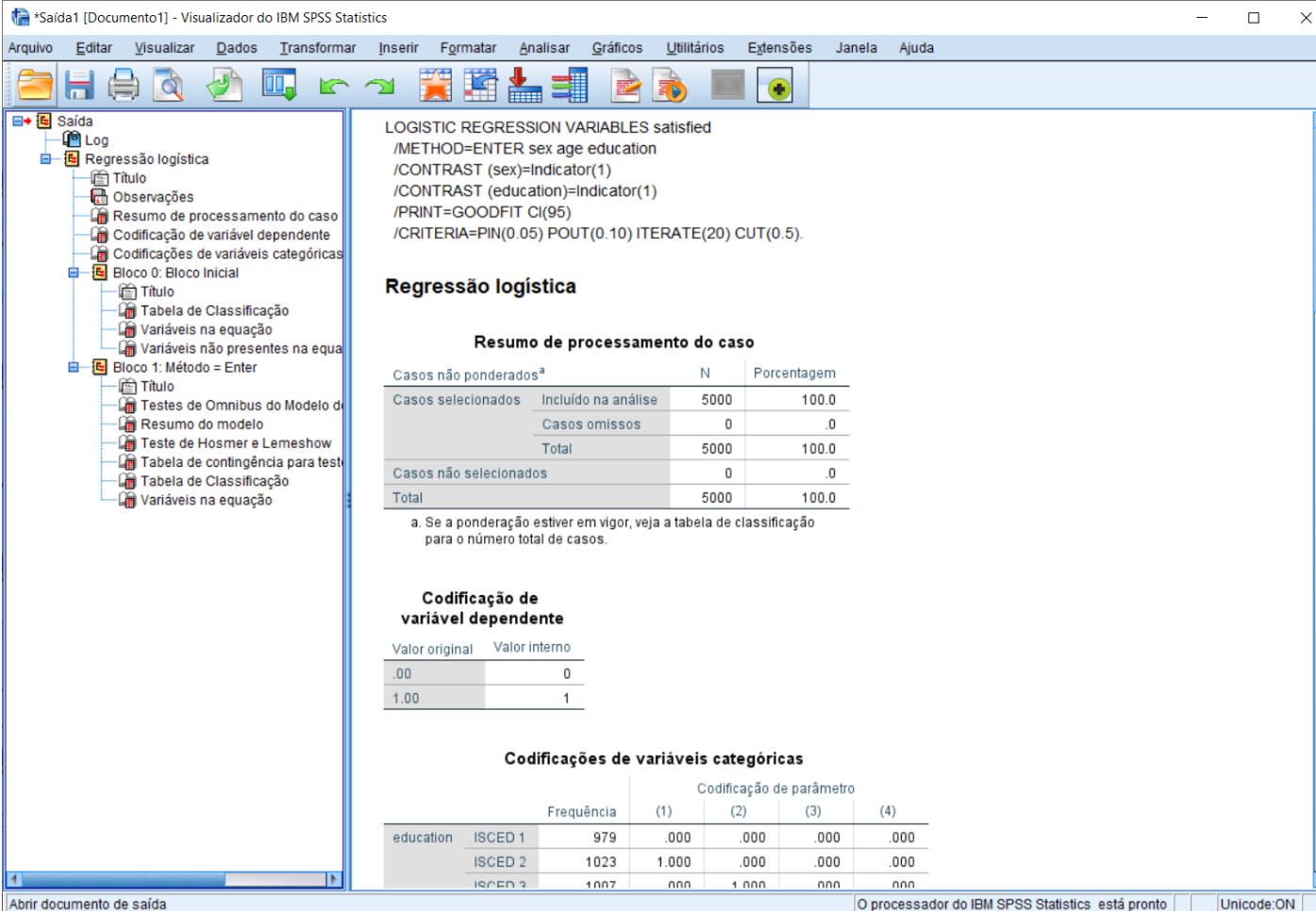
O processador do IBM SPSS Statistics está pronto Unicode:ON

Implementação do Modelo de Regressão Logística

5. Interpretar os resultados do modelo de regressão logística

Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Temos três grupos de resultados
 - Tabelas com informações sobre as variáveis incluídas no modelo
 - Tabelas com os resultados de um modelo de regressão sem variáveis independentes (Bloco 0)
 - Tabelas com os resultados do modelo de regressão que estimamos (Bloco 1)



*Saída1 [Documento1] - Visualizador do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Inserir Formatar Analisar Gráficos Utilitários Extensões Janela Ajuda

LOGISTIC REGRESSION VARIABLES satisfied
/METHOD=ENTER sex age education
/CONTRAST (sex)=Indicator(1)
/CONTRAST (education)=indicator(1)
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).

Regressão logística

Resumo de processamento do caso

Casos não ponderados ^a		N	Porcentagem
Casos selecionados	Incluído na análise	5000	100.0
	Casos omissos	0	.0
Total		5000	100.0
Casos não selecionados		0	.0
Total		5000	100.0

a. Se a ponderação estiver em vigor, veja a tabela de classificação para o número total de casos.

Codificação de variável dependente

Valor original	Valor interno
.00	0
1.00	1

Codificações de variáveis categóricas

		Frequência	Codificação de parâmetro			
			(1)	(2)	(3)	(4)
education	ISCED 1	979	.000	.000	.000	.000
	ISCED 2	1023	1.000	.000	.000	.000
	ISCED 3	1007	.000	1.000	.000	.000

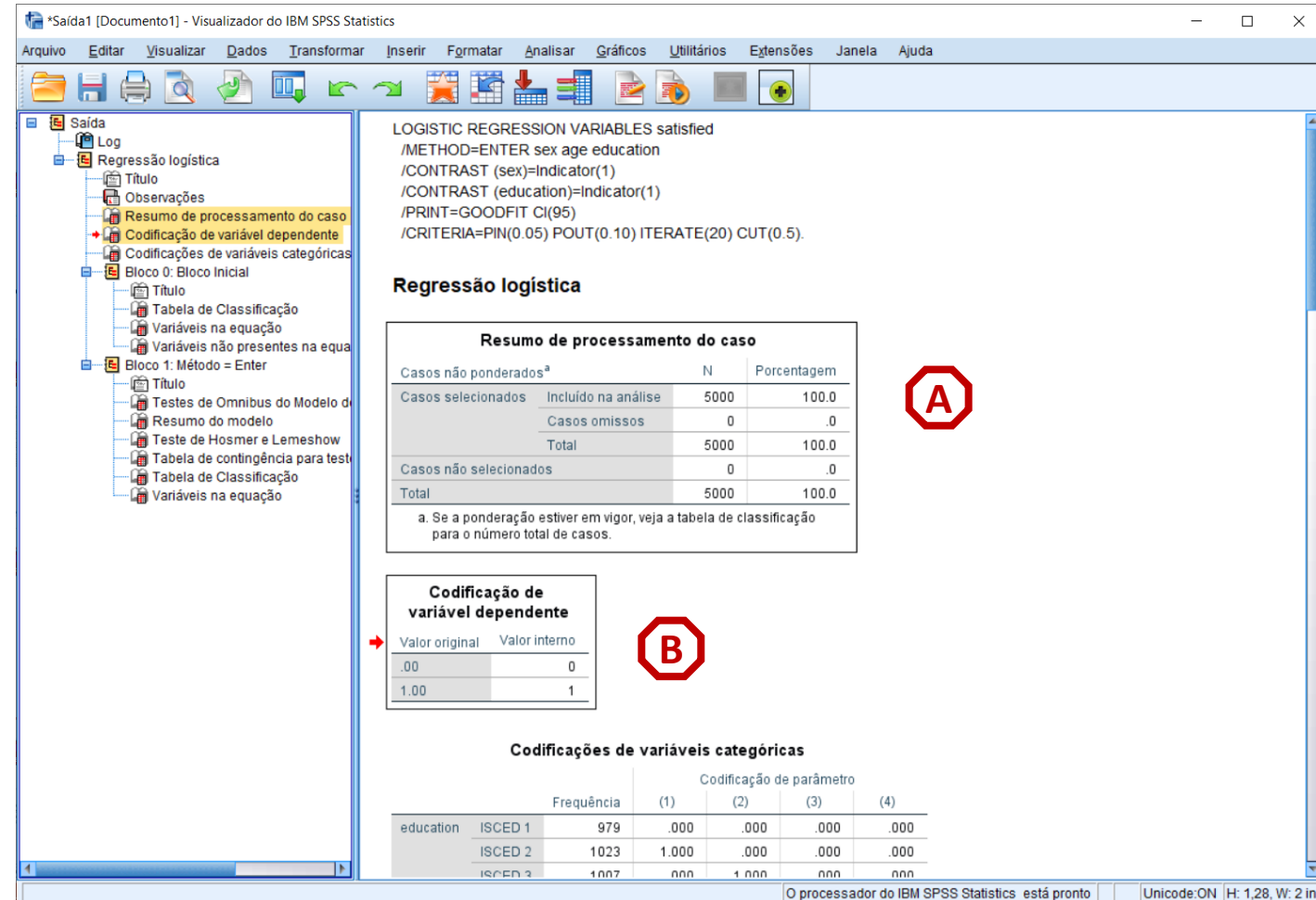
Abriu documento de saída

O processador do IBM SPSS Statistics está pronto

Unicode:ON

Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Estas tabelas dizem-nos:
 - O número de observações que foi integrado no modelo
 - Como está operacionalizada a variável dependente

LOGISTIC REGRESSION VARIABLES satisfied
/METHOD=ENTER sex age education
/CONTRAST (sex)=Indicator(1)
/CONTRAST (education)=indicator(1)
/PRINT=GOODFIT CI(95)
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).

Regressão logística

Resumo de processamento do caso			
Casos não ponderados ^a		N	Porcentagem
Casos selecionados	Incluído na análise	5000	100.0
	Casos omissos	0	.0
	Total	5000	100.0
Casos não selecionados		0	.0
Total		5000	100.0

a. Se a ponderação estiver em vigor, veja a tabela de classificação para o número total de casos.

Codificação de variável dependente		
Valor original	Valor interno	
.00	0	
1.00	1	

		Frequência	Codificação de parâmetro			
			(1)	(2)	(3)	(4)
education	ISCED 1	979	.000	.000	.000	.000
	ISCED 2	1023	1.000	.000	.000	.000
	ISCED 3	1007	.000	1.000	.000	.000

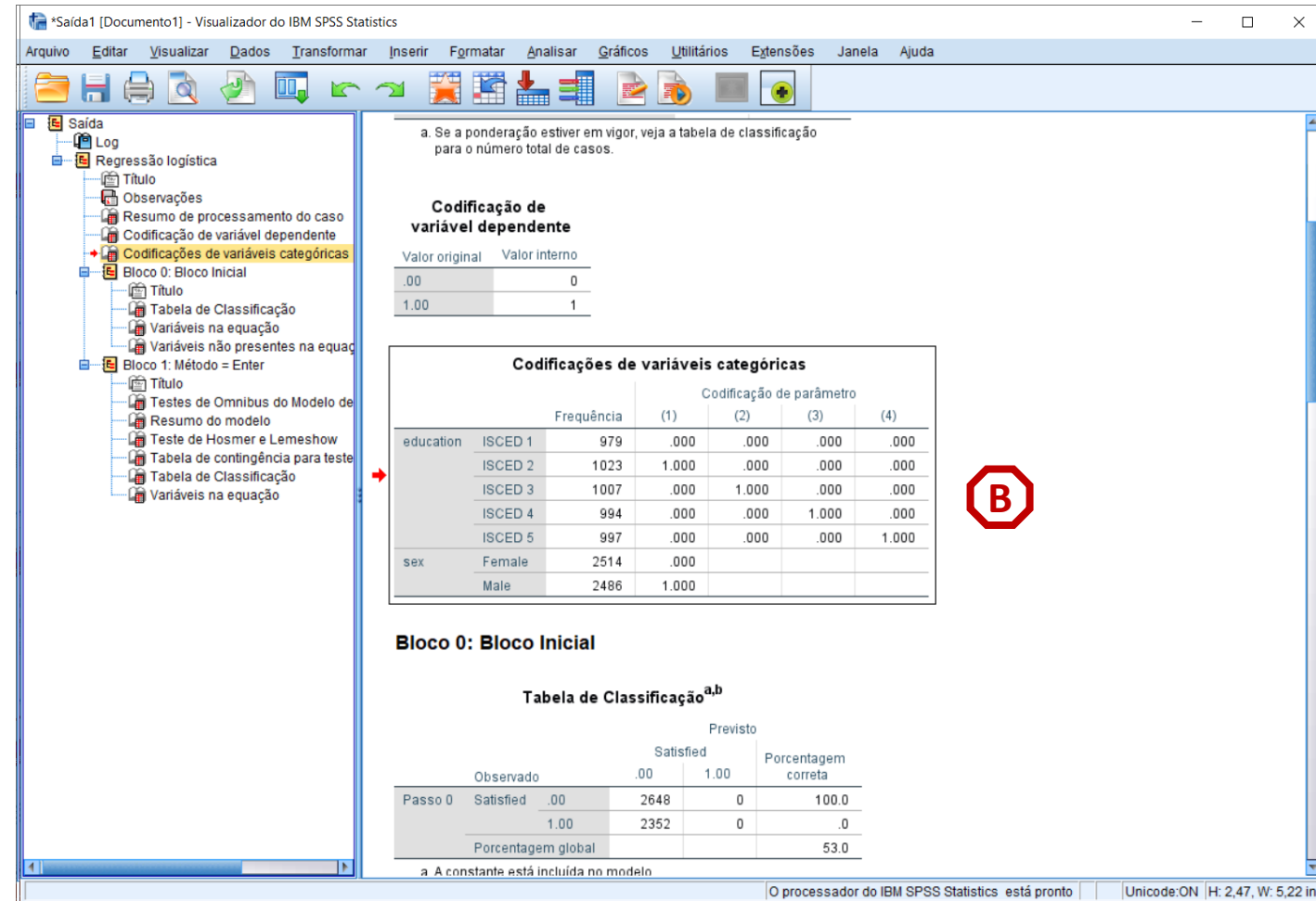
Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
 - Estas tabelas dizem-nos:
 - O número de observações que foi integrado no modelo
 - Como está codificada a variável dependente
 - Como estão codificadas a variáveis independentes *
- * Permite-nos confirmar se as categorias de referência estão bem definidas

A

B

C



Visualizador do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Inserir Formatar Analisar Gráficos Utilitários Extensões Janela Ajuda

Saída

- Log
- Regressão logística
 - Título
 - Observações
 - Resumo de processamento do caso
 - Codificação de variável dependente
 - Codificações de variáveis categóricas**
 - Bloco 0: Bloco Inicial
 - Título
 - Tabela de Classificação
 - Variáveis na equação
 - Variáveis não presentes na equação
 - Bloco 1: Método = Enter
 - Título
 - Testes de Omnibus do Modelo de
 - Resumo do modelo
 - Teste de Hosmer e Lemeshow
 - Tabela de contingência para teste
 - Tabela de Classificação
 - Variáveis na equação

a. Se a ponderação estiver em vigor, veja a tabela de classificação para o número total de casos.

Codificação de variável dependente

Valor original	Valor interno
.00	0
1.00	1

Codificações de variáveis categóricas

		Frequência	Codificação de parâmetro			
			(1)	(2)	(3)	(4)
education	ISCED 1	979	.000	.000	.000	.000
	ISCED 2	1023	1.000	.000	.000	.000
	ISCED 3	1007	.000	1.000	.000	.000
	ISCED 4	994	.000	.000	1.000	.000
	ISCED 5	997	.000	.000	.000	1.000
sex	Female	2514	.000			
	Male	2486	1.000			

Bloco 0: Bloco Inicial

Tabela de Classificação^{a,b}

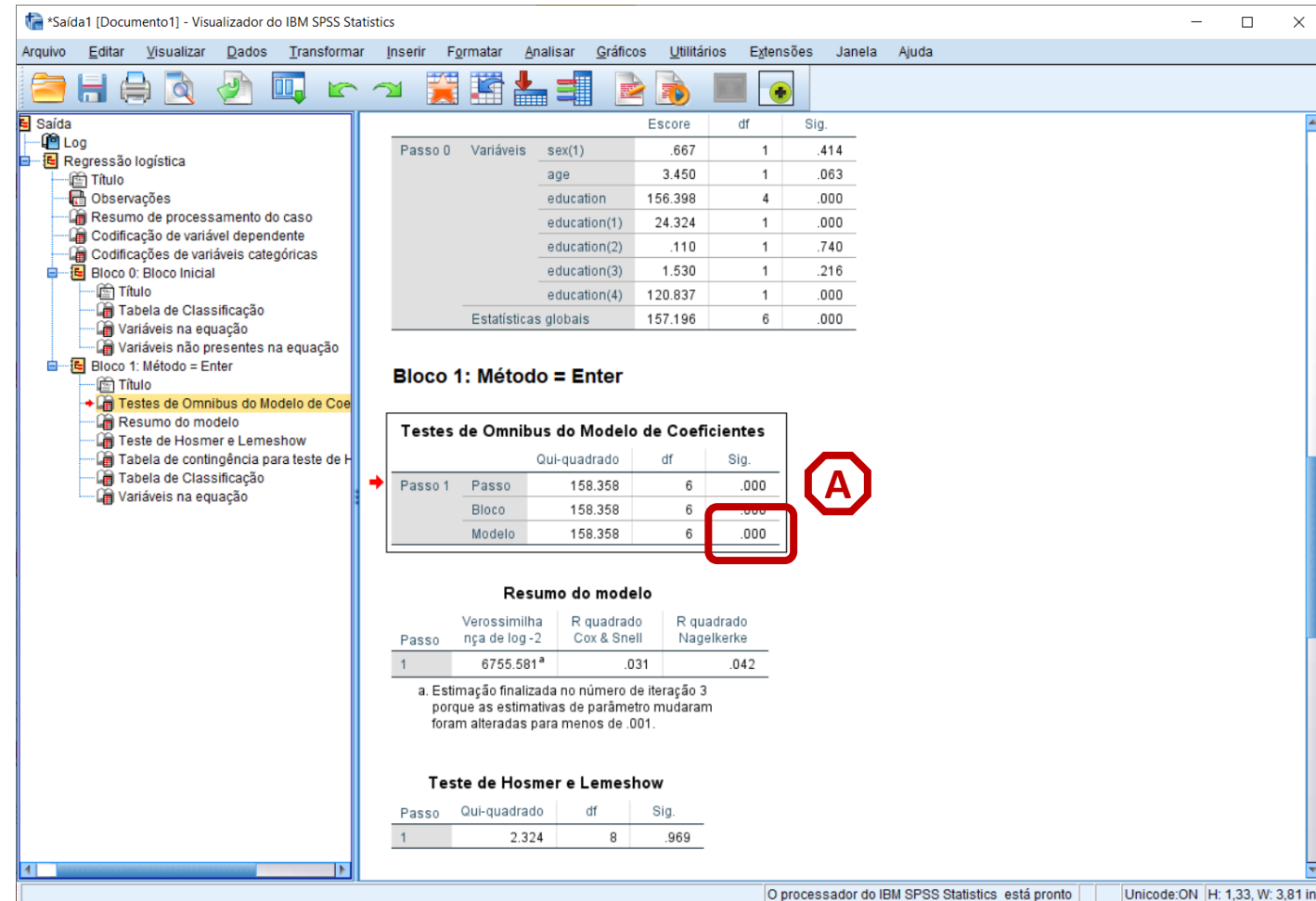
Observado		Previsto		Porcentagem correta
		Satisfied	1.00	
Passo 0 Satisfied	.00	2648	0	100.0
	1.00	2352	0	.0
Porcentagem global				53.0

a. A constante está incluída no modelo

O processador do IBM SPSS Statistics está pronto | Unicode:ON | H: 2,47, W: 5,22 in

Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Esta tabela apresenta o resultado de um teste de Chi-Quadrado que determina se, comparado com um modelo sem variáveis independentes, este modelo é estatisticamente significativo.
- Interpretação:
 - 'Sig.' < 0.01 / 0.05 / 0.1 o modelo é estatisticamente significativo
- Neste caso, o nosso modelo é estatisticamente significativo

*Saída1 [Documento1] - Visualizador do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Inserir Formatar Analisar Gráficos Utilitários Extensões Janela Ajuda

Saída

- Log
 - Regressão logística
 - Título
 - Observações
 - Resumo de processamento do caso
 - Codificação de variável dependente
 - Codificações de variáveis categóricas
 - Bloco 0: Bloco Inicial
 - Título
 - Tabela de Classificação
 - Variáveis na equação
 - Variáveis não presentes na equação
 - Bloco 1: Método = Enter
 - Título
 - Testes de Omnibus do Modelo de Coeficientes
 - Resumo do modelo
 - Teste de Hosmer e Lemeshow
 - Tabela de contingência para teste de H
 - Tabela de Classificação
 - Variáveis na equação

		Escore	df	Sig.
Passo 0	Variáveis			
	sex(1)	.667	1	.414
	age	3.450	1	.063
	education	156.398	4	.000
	education(1)	24.324	1	.000
	education(2)	.110	1	.740
	education(3)	1.530	1	.216
	education(4)	120.837	1	.000
	Estatísticas globais	157.196	6	.000

Bloco 1: Método = Enter

Testes de Omnibus do Modelo de Coeficientes				
		Qui-quadrado	df	Sig.
Passo 1	Passo	158.358	6	.000
	Bloco	158.358	6	.000
	Modelo	158.358	6	.000

Resumo do modelo

Passo	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
1	6755.581 ^a	.031	.042

a. Estimação finalizada no número de iteração 3 porque as estimativas de parâmetro mudaram foram alteradas para menos de .001.

Teste de Hosmer e Lemeshow

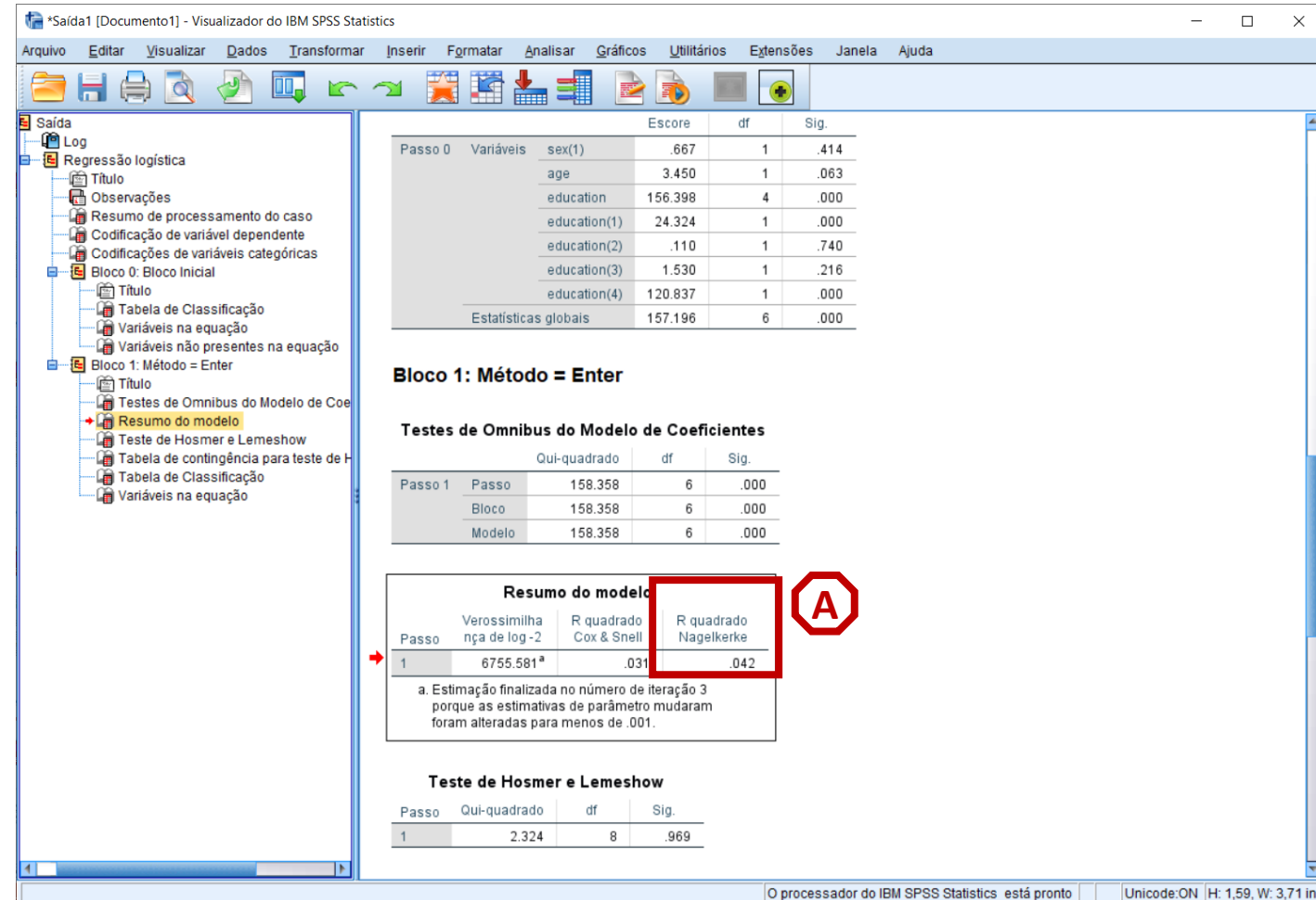
Passo	Qui-quadrado	df	Sig.
1	2.324	8	.969

O processador do IBM SPSS Statistics está pronto | Unicode:ON | H: 1,33, W: 3,81 in

Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Esta tabela apresenta os resultados de um conjunto de testes, nomeadamente o teste R^2 Nagelkerke, que é um pseudo- R^2 .
- Não pode ser interpretado de forma literal como o R^2
- Interpretação:

O model explica aproximadamente 40% da variação na variável dependente



*Saída1 [Documento1] - Visualizador do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Inserir Formatar Analisar Gráficos Utilitários Extensões Janela Ajuda

Passo 0	Variáveis	Escore	df	Sig.
	sex(1)	.667	1	.414
	age	3.450	1	.063
	education	156.398	4	.000
	education(1)	24.324	1	.000
	education(2)	.110	1	.740
	education(3)	1.530	1	.216
	education(4)	120.837	1	.000
Estatísticas globais		157.196	6	.000

Bloco 1: Método = Enter

Testes de Omnibus do Modelo de Coeficientes

Passo 1	Passo	Qui-quadrado	df	Sig.
	Bloco	158.358	6	.000
	Modelo	158.358	6	.000

Resumo do modelo

Passo	Verossimilhança de log -2	R quadrado Cox & Snell	R quadrado Nagelkerke
1	6755.581 ^a	.031	.042

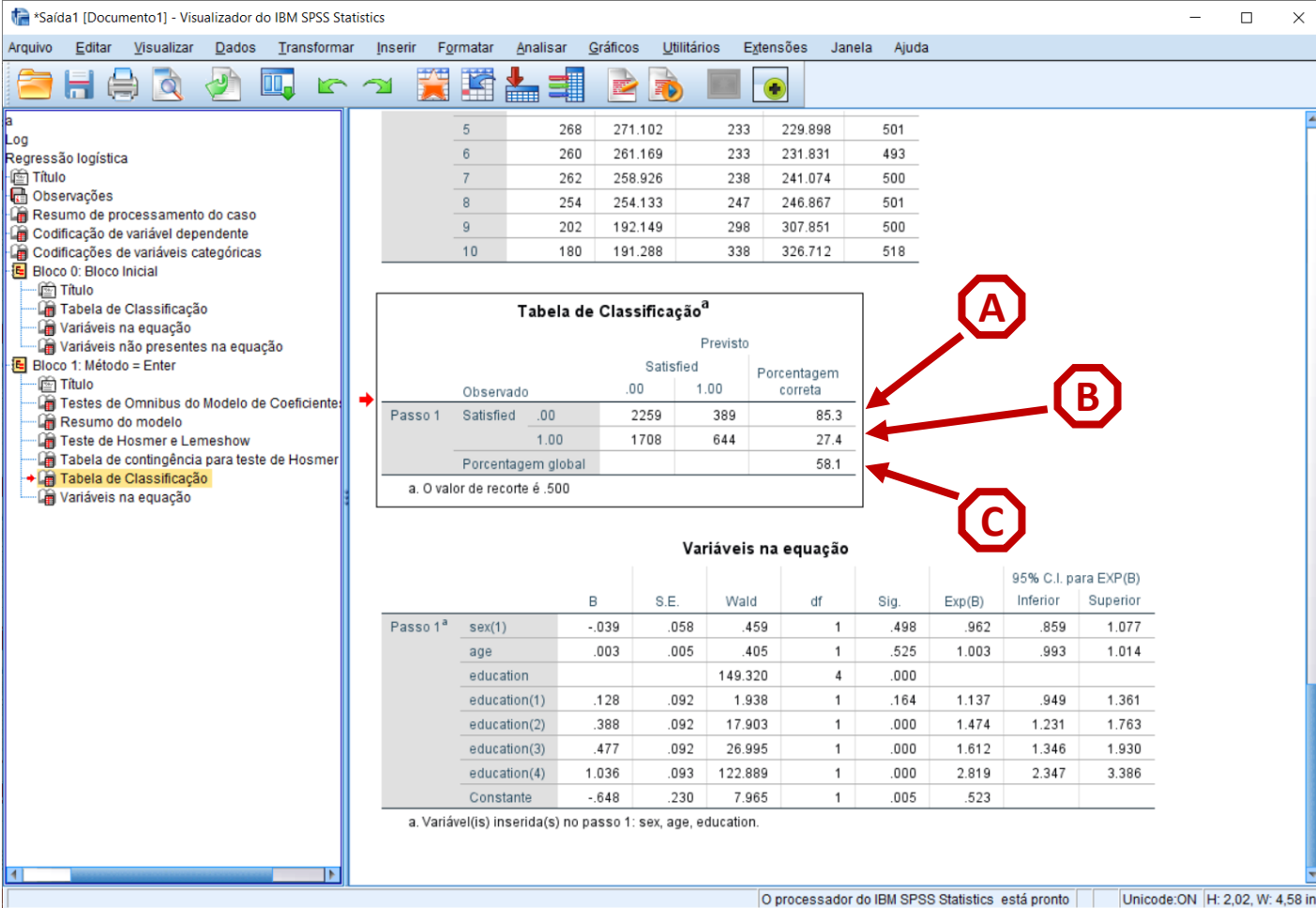
a. Estimação finalizada no número de iteração 3 porque as estimativas de parâmetro mudaram foram alteradas para menos de .001.

Teste de Hosmer e Lemeshow

Passo	Qui-quadrado	df	Sig.
1	2.324	8	.969

Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- Esta tabela dois indicadores da capacidade preditiva do modelo.
 - **Especificidade:** A proporção de casos negativos correctamente identificados (verdadeiros negativos)
 - **Sensibilidade:** A proporção de casos positivos correctamente identificados (verdadeiros positivos)
 - **Precisão:** A proporção de casos correctamente identificados



The screenshot shows the IBM SPSS Statistics Results Viewer window. The left pane displays the navigation tree with 'Tabela de Classificação' selected. The main area shows the following tables:

Table 1: Data Rows

5	268	271.102	233	229.898	501
6	260	261.169	233	231.831	493
7	262	258.926	238	241.074	500
8	254	254.133	247	246.867	501
9	202	192.149	298	307.851	500
10	180	191.288	338	326.712	518

Table 2: Tabela de Classificação^a

Observado	Previsto	Previsto		Porcentagem correta
		Satisfied	1.00	
Satisfied	.00	2259	389	85.3
1.00	1.00	1708	644	27.4
Porcentagem global				58.1

a. O valor de recorte é .500

Table 3: Variáveis na equação

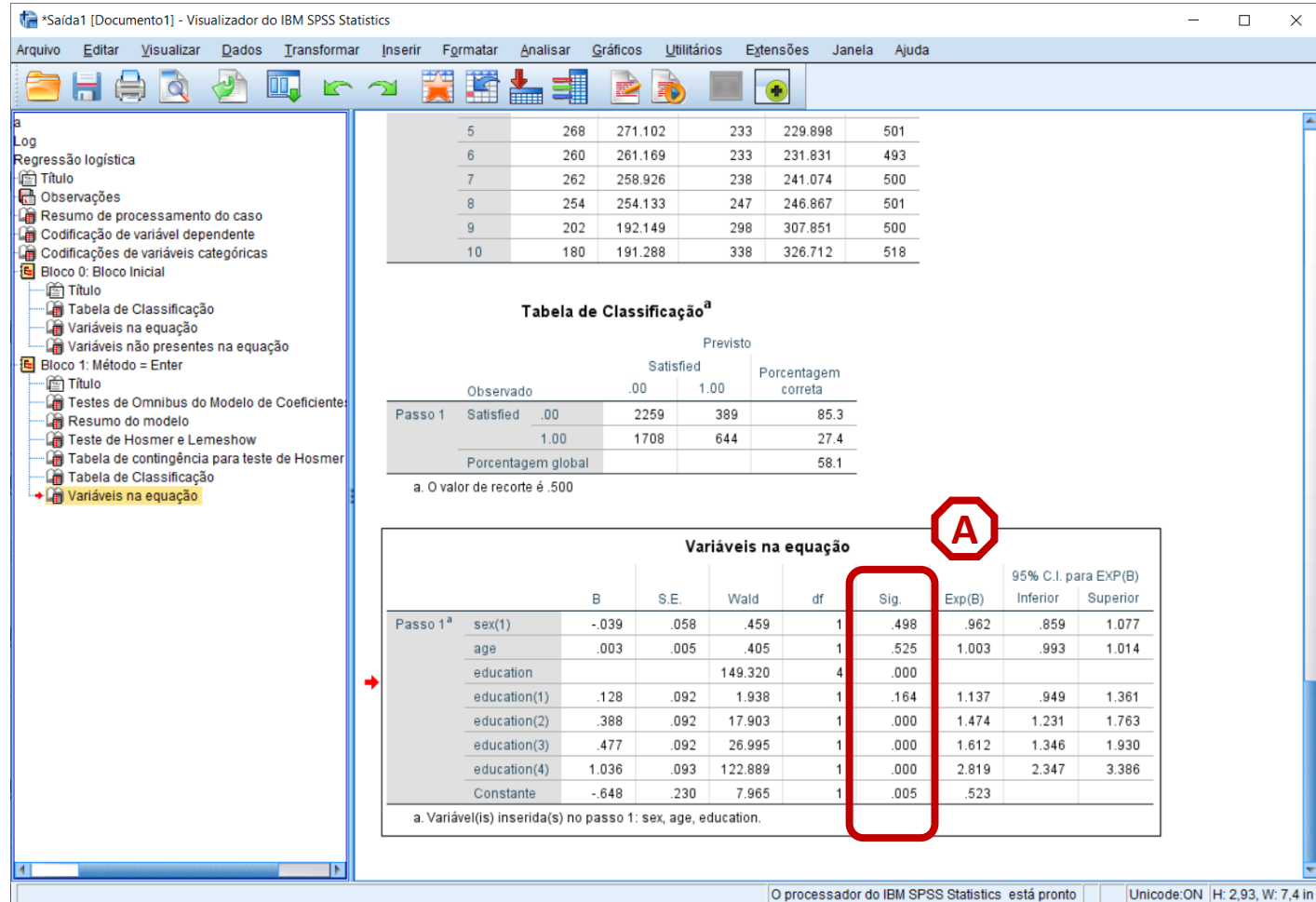
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Inferior	Superior
Passo 1 ^a sex(1)	-.039	.058	.459	1	.498	.962	.859	1.077
age	.003	.005	.405	1	.525	1.003	.993	1.014
education			149.320	4	.000			
education(1)	.128	.092	1.938	1	.164	1.137	.949	1.361
education(2)	.388	.092	17.903	1	.000	1.474	1.231	1.763
education(3)	.477	.092	26.995	1	.000	1.612	1.346	1.930
education(4)	1.036	.093	122.889	1	.000	2.819	2.347	3.386
Constante	-.648	.230	7.965	1	.005	.523		

a. Variável(is) inserida(s) no passo 1: sex, age, education.

Annotations: Three red octagons labeled A, B, and C are placed to the right of the classification table. Red arrows point from octagon A to the 'Satisfied' row, from octagon B to the '1.00' row, and from octagon C to the 'Porcentagem global' row.

Regressão Logística

- O resultado é publicado no ‘Visualizador de Resultados’
- Esta tabela apresenta a informação sobre o efeito das variáveis independentes sobre a variável dependent.
- A primeira coisa a olhar é se a relação das IV’s com a DV é estatisticamente significativa, i.e.
 - ‘Sig.’ < 0.01 / 0.05 / 0.1
- Neste caso, observa-se que as variáveis ‘sex’ e ‘age’ não são significativas.



*Saída1 [Documento1] - Visualizador do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Inserir Formatar Analisar Gráficos Utilitários Extensões Janela Ajuda

Log
Regressão logística
Título
Observações
Resumo de processamento do caso
Codificação de variável dependente
Codificações de variáveis categóricas
Bloco 0: Bloco Inicial
Título
Tabela de Classificação
Variáveis na equação
Variáveis não presentes na equação
Bloco 1: Método = Enter
Título
Testes de Omnibus do Modelo de Coeficientes
Resumo do modelo
Teste de Hosmer e Lemeshow
Tabela de contingência para teste de Hosmer
Variáveis na equação

5	268	271.102	233	229.898	501
6	260	261.169	233	231.831	493
7	262	258.926	238	241.074	500
8	254	254.133	247	246.867	501
9	202	192.149	298	307.851	500
10	180	191.288	338	326.712	518

Tabela de Classificação^a

	Observado	Previsto		Porcentagem correta	
		Satisfied	1.00		
Passo 1	Satisfied	.00	2259	389	85.3
	1.00	1708	644	27.4	
Porcentagem global					58.1

a. O valor de recorte é .500

Variáveis na equação

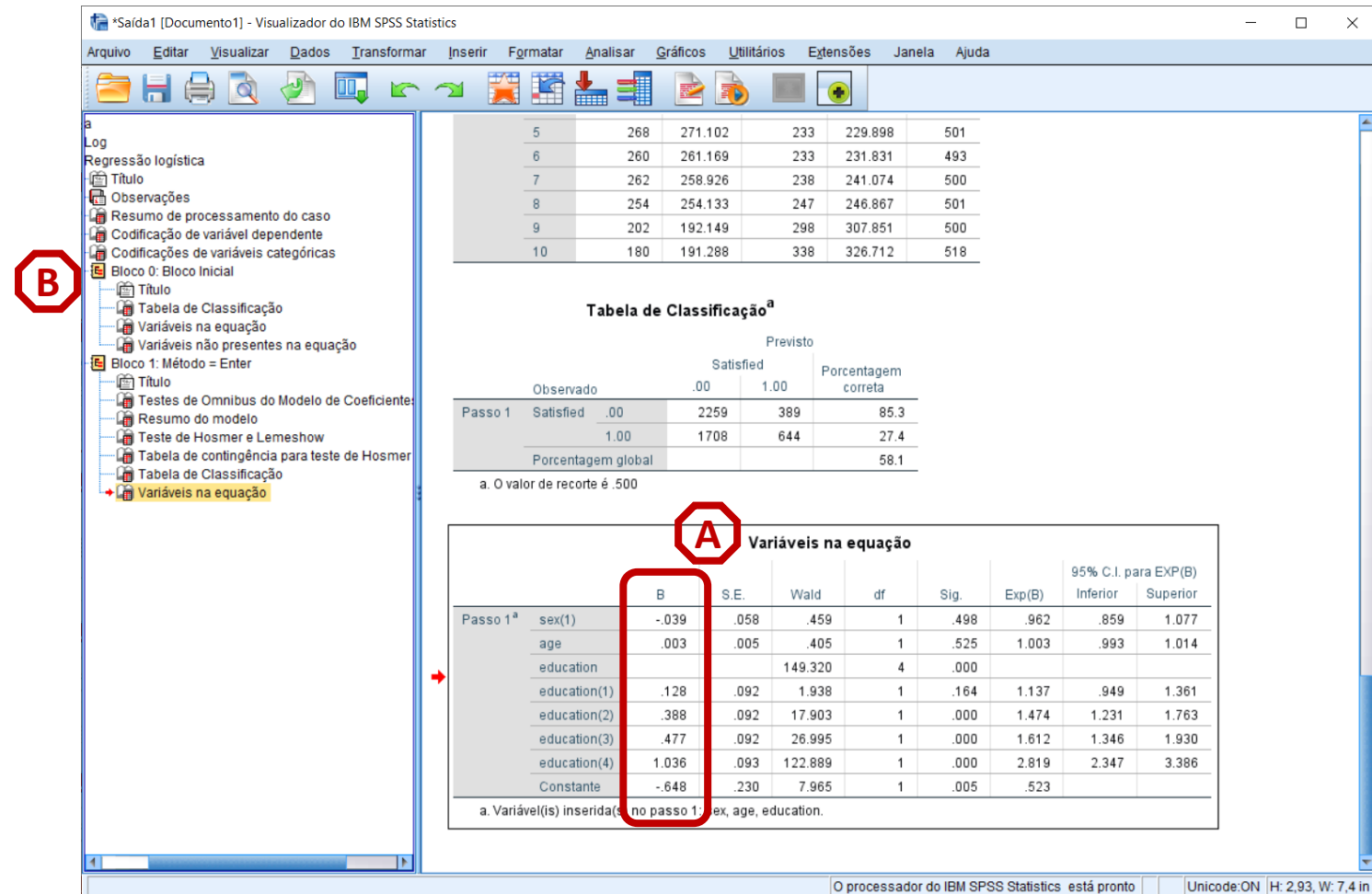
	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)		
							Inferior	Superior	
Passo 1 ^a	sex(1)	-.039	.058	.459	1	.498	.962	.859	1.077
	age	.003	.005	.405	1	.525	1.003	.993	1.014
	education			149.320	4	.000			
	education(1)	.128	.092	1.938	1	.164	1.137	.949	1.361
	education(2)	.388	.092	17.903	1	.000	1.474	1.231	1.763
	education(3)	.477	.092	26.995	1	.000	1.612	1.346	1.930
	education(4)	1.036	.093	122.889	1	.000	2.819	2.347	3.386
	Constante	-.648	.230	7.965	1	.005	.523		

a. Variável(is) inserida(s) no passo 1: sex, age, education.

O processador do IBM SPSS Statistics está pronto | Unicode:ON | H: 2,93, W: 7,4 in

Regressão Logística

- O resultado é publicado no 'Visualizador de Resultados'
- De seguida, queremos ir ver os coeficientes do modelo:
 - Ser homem diminui o logaritmo da chance (log odds) de estar satisfeito com as condições de trabalho em 0.39...
 - Por cada ano adicional de vida, o logaritmo da chance (log odds) de estar satisfeito com as condições de trabalho aumenta em 0.03.
 - Comparado com ter apenas o ensino básico (ISCED1), ter uma licenciatura aumenta o logaritmo da chance (log odds) de estar satisfeito com as condições de trabalho em 1.036.



Visualizador do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Inserir Formatar Analisar Gráficos Utilitários Extensões Janela Ajuda

Log

Regressão logística

- Log
- Regressão logística
- Título
- Observações
- Resumo de processamento do caso
- Codificação de variável dependente
- Codificações de variáveis categóricas
- Bloco 0: Bloco Inicial
 - Título
 - Tabela de Classificação
 - Variáveis na equação
 - Variáveis não presentes na equação
- Bloco 1: Método = Enter
 - Título
 - Testes de Omnibus do Modelo de Coeficientes
 - Resumo do modelo
 - Teste de Hosmer e Lemeshow
 - Tabela de contingência para teste de Hosmer
 - Tabela de Classificação
 - Variáveis na equação

Tabela de Classificação^a

Observado	Previsto	Satisfied		Porcentagem correta
		.00	1.00	
Passo 1 Satisfied .00		2259	389	85.3
1.00		1708	644	27.4
Porcentagem global				58.1

a. O valor de recorte é .500

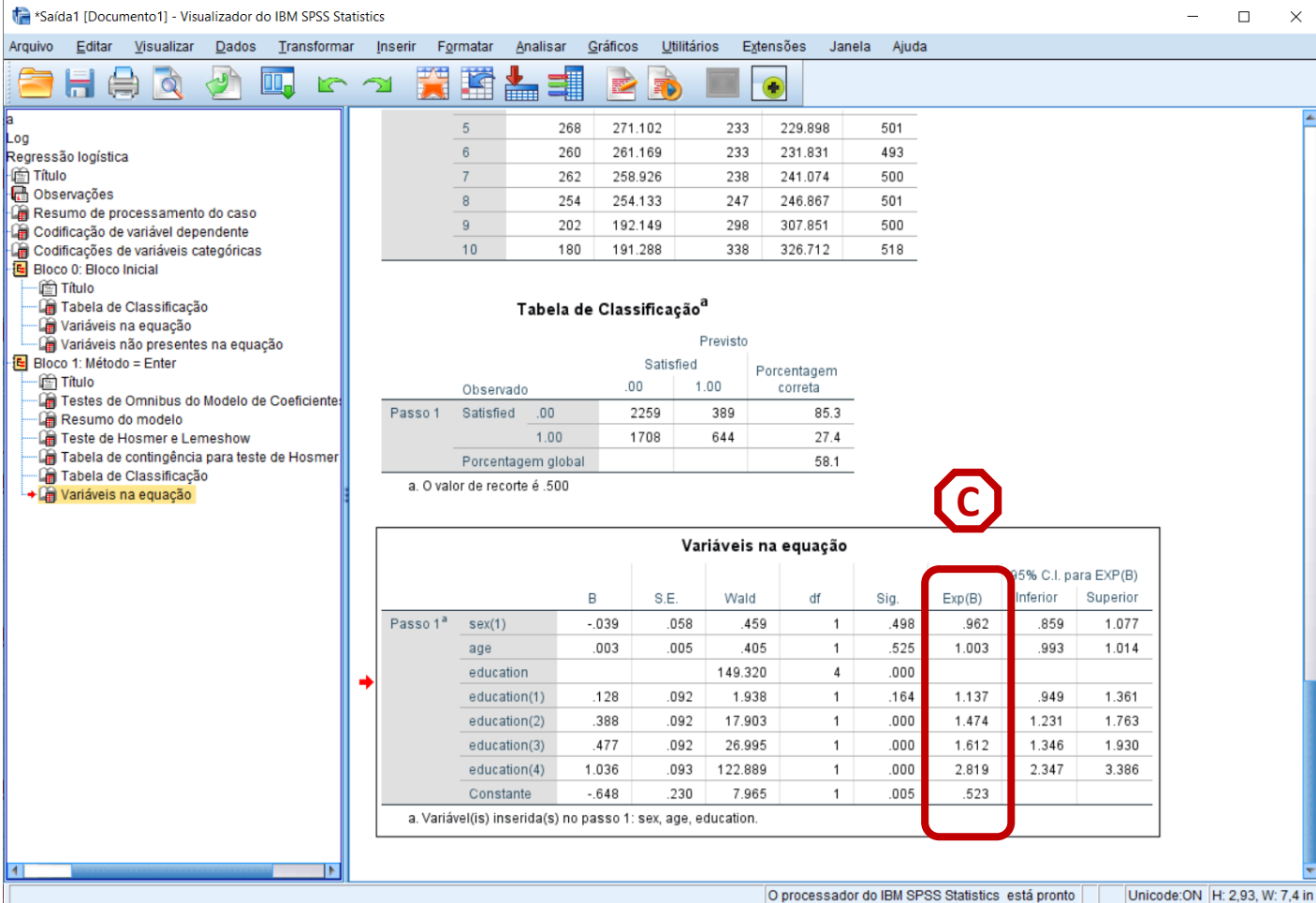
Variáveis na equação

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
							Inferior	Superior
Passo 1 ^a sex(1)	-.039	.058	.459	1	.498	.962	.859	1.077
age	.003	.005	.405	1	.525	1.003	.993	1.014
education			149.320	4	.000			
education(1)	.128	.092	1.938	1	.164	1.137	.949	1.361
education(2)	.388	.092	17.903	1	.000	1.474	1.231	1.763
education(3)	.477	.092	26.995	1	.000	1.612	1.346	1.930
education(4)	1.036	.093	122.889	1	.000	2.819	2.347	3.386
Constante	-.648	.230	7.965	1	.005	.523		

a. Variável(is) inserida(s) no passo 1: sex, age, education.

Regressão Logística

- ISTO NÃO É NADA INTUITIVO!!
- Por isso é que nos interessa olhar para os odd-ratios (razão de chances)
- Interpretação:
 - Um odds ratio maior que 1 significa que um aumento em x leva a um aumento nas chances de $y = 1$.
 - Um odds ratio menor que 1 significa que um aumento em x leva a uma diminuição nas chances de $y = 1$.

*Saída1 [Documento1] - Visualizador do IBM SPSS Statistics

Arquivo Editar Visualizar Dados Transformar Inserir Formatar Analisar Gráficos Utilitários Extensões Janela Ajuda

Log

Regressão logística

Título

Observações

Resumo de processamento do caso

Codificação de variável dependente

Codificações de variáveis categóricas

Bloco 0: Bloco Inicial

Título

Tabela de Classificação

Variáveis na equação

Variáveis não presentes na equação

Bloco 1: Método = Enter

Título

Testes de Omnibus do Modelo de Coeficientes

Resumo do modelo

Teste de Hosmer e Lemeshow

Tabela de contingência para teste de Hosmer

Tabela de Classificação

Variáveis na equação

5	268	271.102	233	229.898	501
6	260	261.169	233	231.831	493
7	262	258.926	238	241.074	500
8	254	254.133	247	246.867	501
9	202	192.149	298	307.851	500
10	180	191.288	338	326.712	518

Tabela de Classificação^a

	Observado	Previsto		Porcentagem correta	
		Satisfied	1.00		
Passo 1	Satisfied	.00	2259	389	85.3
	1.00	1708	644		27.4
Porcentagem global					58.1

a. O valor de recorte é .500

Variáveis na equação

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
								Inferior	Superior
Passo 1 ^a	sex(1)	-.039	.058	.459	1	.498	.962	.859	1.077
	age	.003	.005	.405	1	.525	1.003	.993	1.014
	education			149.320	4	.000			
	education(1)	.128	.092	1.938	1	.164	1.137	.949	1.361
	education(2)	.388	.092	17.903	1	.000	1.474	1.231	1.763
	education(3)	.477	.092	26.995	1	.000	1.612	1.346	1.930
	education(4)	1.036	.093	122.889	1	.000	2.819	2.347	3.386
	Constante	-.648	.230	7.965	1	.005	.523		

a. Variável(is) inserida(s) no passo 1: sex, age, education.

O processador do IBM SPSS Statistics está pronto | Unicode:ON | H: 2,93, W: 7,4 in

Regressão Logística

- Melhor ainda, podemos calcular a mudança percentual nas odds por cada unidade adicional na IV:

$$\Delta \% \text{ das Odds} = 100 * (\text{Odds Ratio} - 1)$$

- Exemplo:

- As chances de estar satisfeito/a com as condições de trabalho são 3.8% - $100 * (0.962 - 1)$ - mais baixas para homens, em comparação com mulheres.
- Cada aumento de um ano na idade leva a um aumento de 1,3% - $100 * (1,013-1)$ das chances de estar satisfeito/a com as condições de trabalho .

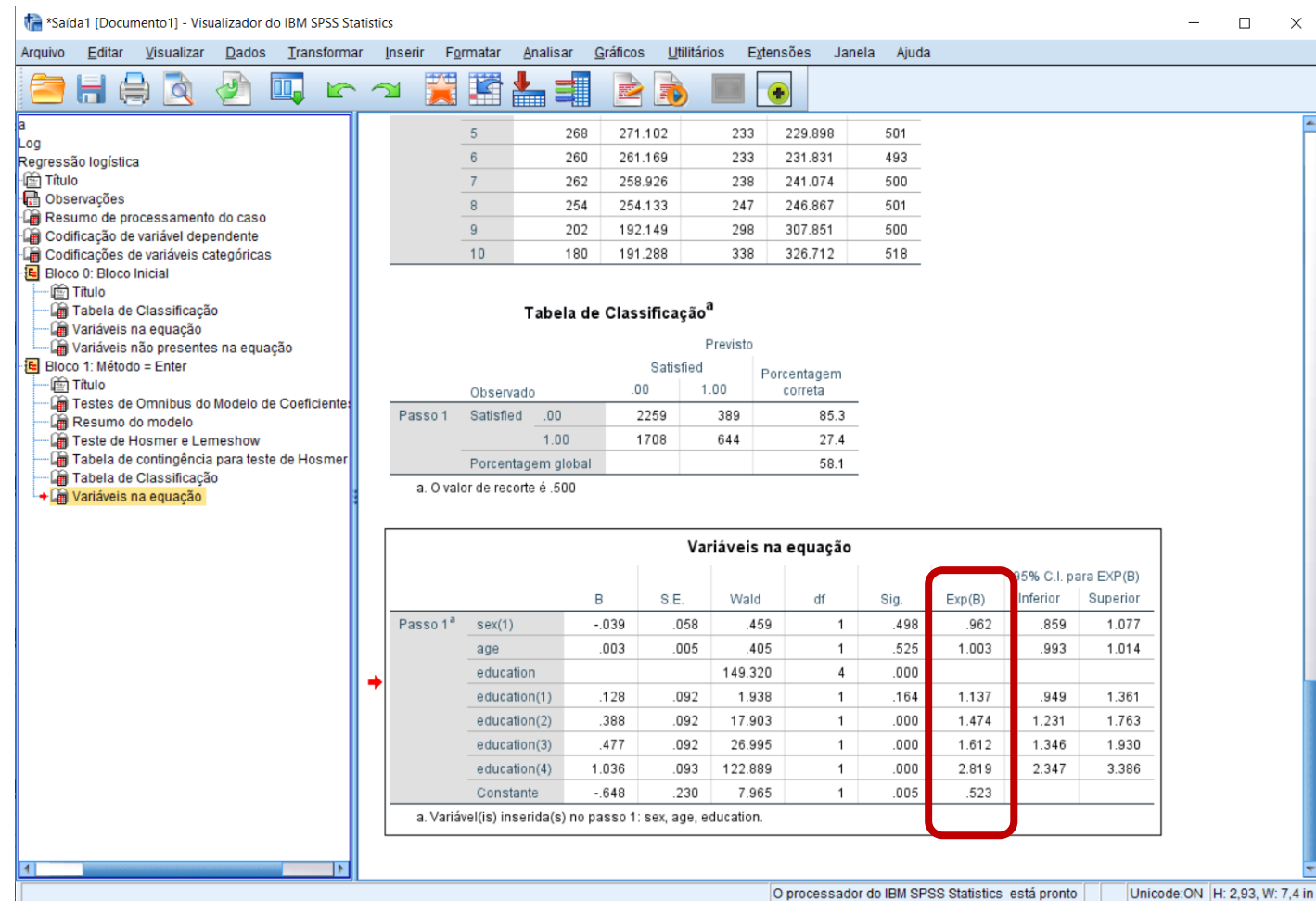


Tabela de Classificação^a

		Previsto		Porcentagem correta
		Satisfied	1.00	
Passo 1	Satisfied	.00	2259	85.3
	1.00	1708	644	27.4
Porcentagem global				58.1

a. O valor de recorte é .500

Variáveis na equação

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. para EXP(B)	
								Inferior	Superior
Passo 1 ^a	sex(1)	-.039	.058	.459	1	.498	.962	.859	1.077
	age	.003	.005	.405	1	.525	1.003	.993	1.014
	education			149.320	4	.000			
	education(1)	.128	.092	1.938	1	.164	1.137	.949	1.361
	education(2)	.388	.092	17.903	1	.000	1.474	1.231	1.763
	education(3)	.477	.092	26.995	1	.000	1.612	1.346	1.930
	education(4)	1.036	.093	122.889	1	.000	2.819	2.347	3.386
	Constante	-.648	.230	7.965	1	.005	.523		

a. Variável(is) inserida(s) no passo 1: sex, age, education.

Implementação do Modelo de Regressão Logística

6. Validar pressupostos do modelo de regressão logística

- **Diagnósticos do modelo de regressão logística**
 - **Linearidade da relação entre os variáveis independentes e a variável dependente**
 - **Existência de Observações Influentes**
 - **Ausência de multicolinearidade**



The End