**Exercício 1**

(a) Unidade estatística: cada jogador da NBA

População: todos os jogadores da NBA

(b) Posição: qualitativa nominal

Peso: quantitativa contínua

Altura: quantitativa contínua

Idade\_grupo: qualitativa ordinal

(c) Número de pontos obtidos obtidos na época anterior

(d) Dimensão da amostra n = 505

(e)

Data – Active – Save NÃO ESQUECER DE GRAVAR A BASE EM CADA ITERAÇÃO

Para carregar a base de dados, Data – Load

(f) Clicar em Edit Data

(g)

Observações omissas – valores em falta na variável

Statistics – summaries – count missing

sapply(DadosNBA2013, function(x)(sum(is.na(x)))) # NA counts

Posição Peso Altura Idade\_grupo V5

0 3 1 0 506

(h) Para criar novas variáveis temos a opção “compute” para “contas” e o “recode” para novas atribuições

Data – manage – compute

(i) Regra de Sturges: o menor inteiro k tal que 2k >= nvalido, onde nvalido = n – casos omissos = 506 – 1 = 505

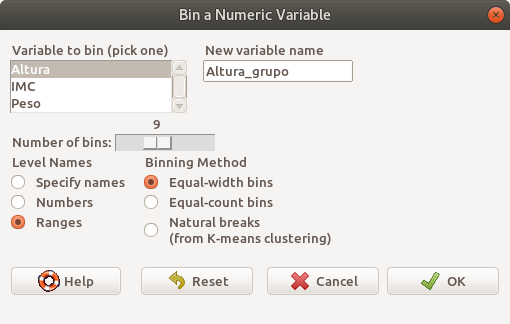
2k >= 505

k = 9 (2^9 = 512 que é >= 505)

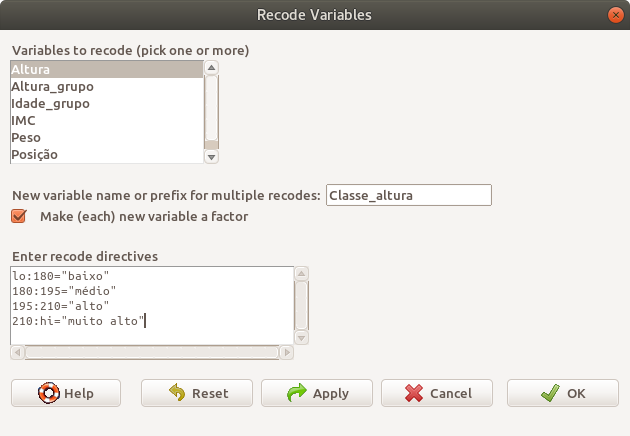
ou

2k = 505 (=) k = log2(505) (=) k = 8.98014

No R: Data – Manage variables in active data set – Bin numeric variable



(j) Data – manage – recode



Os pontos de mudança de classe são alocados à primeira linha onde surgem. Como não está como pretendido, a solução é trocar as linhas.

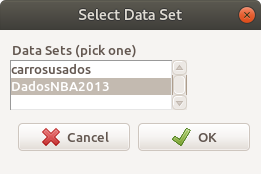
(k) Gravar a base

Data – Active data set – Save active data set

**Exercício 2**

(a) Data – load – carros usados

Podemos intercalar entre bases de dados abertas



População – todas as viaturas disponíveis para venda num stand em 2014

Dimensão da amostra – n = 61

(b) Marca - Qualitativa nominal

Ano – Quantitativa discreta

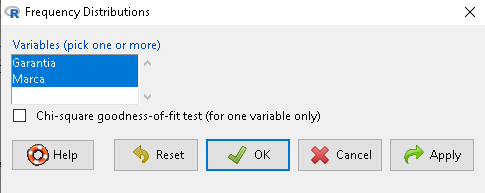
Km – Quantitativa contínua

Cavalos – Quantitativa contínua

Garantia – Qualitativa ordinal

Preço – Quantitativa discreta

(c) Statistics – summaries – frequency distribuitions



counts:

Marca

Citroen Fiat Ford Opel Peugeot Renault Smart

7 14 6 9 7 9 9

percentages:

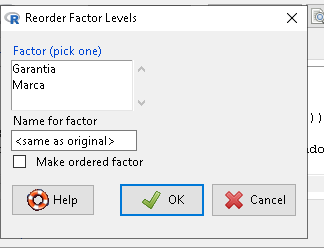
Marca

Citroen Fiat Ford Opel Peugeot Renault Smart

11.48 22.95 9.84 14.75 11.48 14.75 14.75

Para ordenar as categorias

Data – manage variables in data set – reoder factor levels



Volta-se a fazer o quadro de frequências e é mostrado com as categorias ordenadas.

counts:

Marca

Citroen Fiat Ford Opel Peugeot Renault Smart

7 14 6 9 7 9 9

percentages:

Marca

Citroen Fiat Ford Opel Peugeot Renault Smart

11.48 22.95 9.84 14.75 11.48 14.75 14.75

counts:

Garantia

completa nenhuma parcial

11 21 29

percentages:

Garantia

completa nenhuma parcial

18.03 34.43 47.54

(d) A percentagem pedida é de 47.54 + 18.03 = 65.57 %.

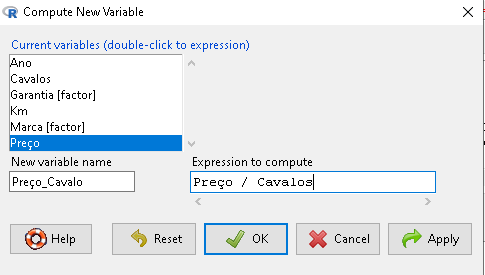
(e) O número de veículos Peugeot é 7.

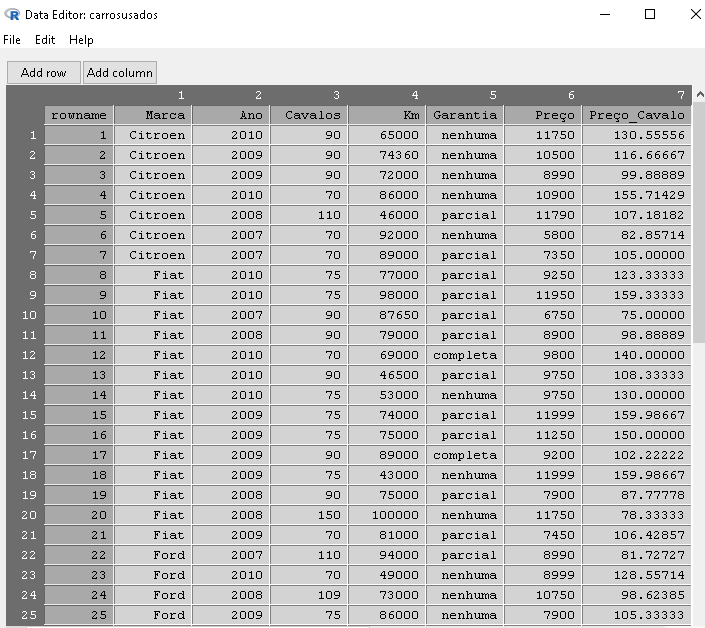
(f) Moda → Modalidade (categoria) mais frequente.

Moda da garantia: garantia parcial

Moda da marca: fiat

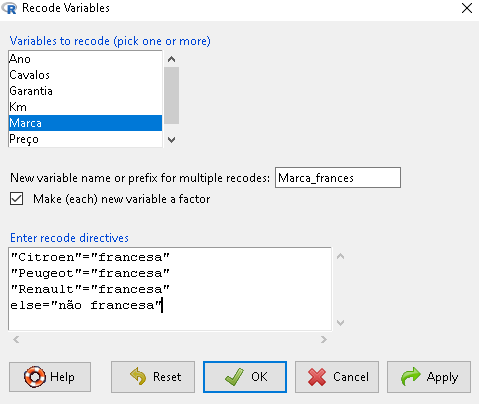
(g) Data – manage – compute



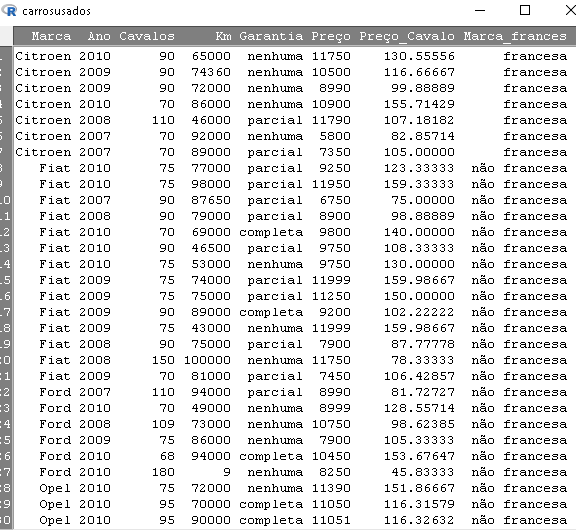


GRAVAR A BASE!!!!

(h) Data – manage - recode

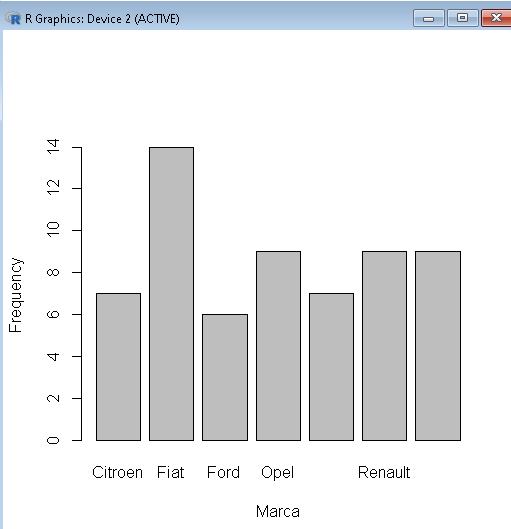


O comando eles só funciona se não houverem missings, havendo este comando transformando as omissas em “não francesas”.



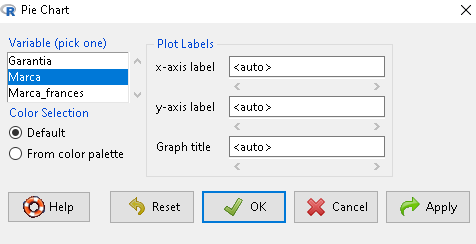
(i) Marca

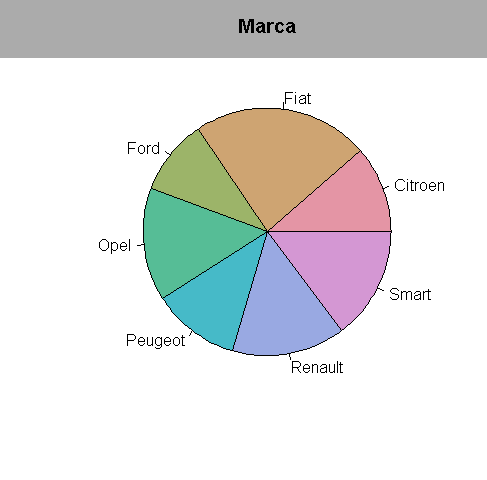
Diagrama de barras



Podemos verificar que a marca Fiat é a mais frequente, e a Ford a menos representada. Parece que a Opel, a Renault e a Smart aparentam ter o mesmo número de veículos.

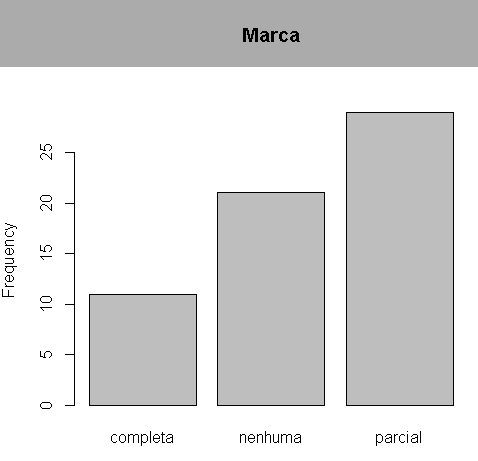
Diagrama de setores





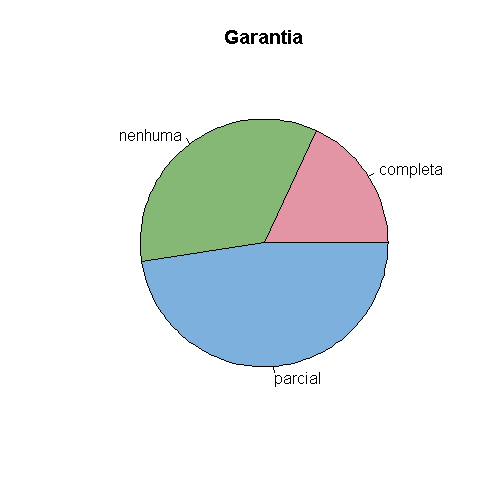
Garantia

Gráfico de barras



Podemos verificar que a garantia parcial é a mais frequente.

Gráfico de setores

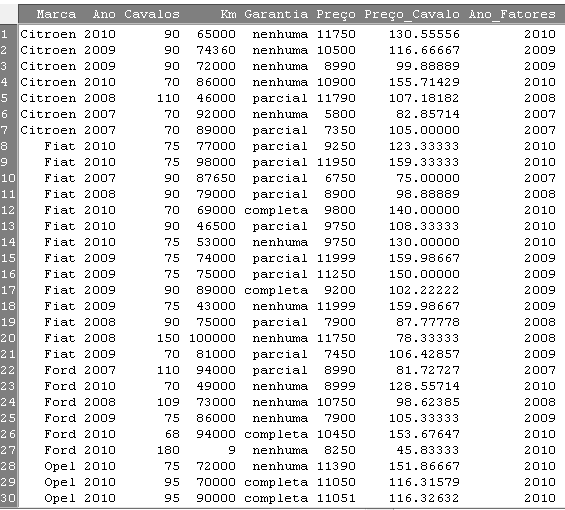


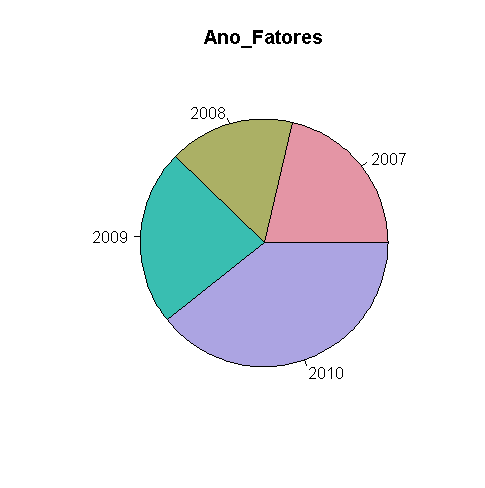
Podemos verificar que a garantia parcial é a mais frequente.

Questão extra

Converter variável ano em fatores

No R: Data – manage





(j)