



TarefALEA

N.º 6 – Número de nascimentos nos municípios de Portugal continental



Esta TarefALEA destina-se a alunos do 10.º ano de escolaridade e nela propõe-se uma sequência de atividades que interligam os conteúdos do tema “Dados”: a organização em tabelas de frequências; a elaboração de representações gráficas, como o histograma, os diagramas de extremos e quartis, as caixas de bigodes; o cálculo de estatísticas descritivas, como a média, a mediana e o desvio padrão; e a leitura e interpretação crítica dos resultados. O tema é o declínio da natalidade e os dados considerados são o total de bebés nascidos nos municípios portugueses em 1981, 2001, 2011 e 2021, por sexo.

Passo 1 – Escolha do conjunto de dados: Nados-vivos de mães residentes em Portugal, acessível em <https://www.pordata.pt/municipios/nados+vivos+de+maes+residentes+em+portugal+total+e+por+sexo-103>

Passo 2 – Os dados têm como fonte o INE (<https://www.ine.pt/xurl/ind/0008234>). Ao consultar a tabela, investigue o que significa a classificação, atribuída pelo INE, de NUTS I, NUTS II e NUTS III (Por ex. em <https://www.pordata.pt/o+que+sao+nuts>)

Passo 3 – **(A)** a unidade de observação é o Município; **(B)** a variável estatística é o número de nascimentos no Município; **(C)** a variável assume valores inteiros; **(D)** os dados são quantitativos discretos; **(E)** os dados não têm unidade de medida por se tratar de dados de contagem.

Passo 4 – Restrinja-se aos dados referentes a Portugal continental e aos anos de 1981, 2001, 2011 e 2021.

Passo 5 – Utilizando filtros convenientes, é possível obter uma tabela com os dados referentes ao número de nascimentos por município, bem como o número de nascimentos do sexo masculino, em <https://www.pordata.pt/db/municipios/ambiente+de+consulta/tabela/5832935>

Exporte a tabela para Excel, obtendo uma tabela idêntica à seguinte:

Município	1981		2001		2011		2021	
	Total	Sexo masc.	Total	Sexo masc.	Total	Sexo masc.	Total	Sexo masc.
Arcos de Valdevez	430	220	165	82	139	69	110	57
Caminha	220	117	161	73	124	72	86	53
...
Vila Real de Santo António	223	115	192	102	190	102	140	63

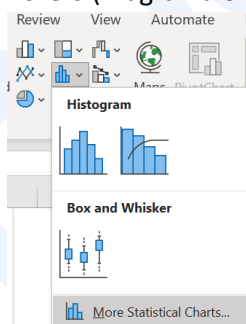
**Passo 6 – Atividades e exercícios**

6.1. Construa os diagramas de extremos e quartis paralelos, referentes ao número de nascimentos nos vários municípios, nos 4 censos considerados.

6.1.1. Interprete a representação gráfica obtida, nomeadamente no que se refere às características da distribuição de cada ano, assim como a tendência ao longo dos anos.

6.1.2. Identifica alguma situação que torna essa interpretação um pouco complexa? Seria interessante fomentar a discussão sobre a causa da complexidade na interpretação da representação gráfica obtida e sugerir um processo de transformar os dados originais em dados ponderados, quer com a dimensão da população residente, quer com o número de mulheres com idades dos 20 aos 50 anos, por exemplo.

Sugestão: Utilize a versão *Box and Whiskers* (Diagrama em caixa de bigodes) do Excel.



A representação obtida pode ser mais elaborada do que o diagrama de extremos e quartis, na medida em que, no caso de haver valores considerados *outliers*, as barras que unem o meio dos lados da caixa com o mínimo e o máximo, são substituídas por umas barras mais curtas e “pontos” que representam valores considerados *outliers*. Pode ignorar este facto e considerar as barras como se estivessem prolongadas até ao mínimo ou máximo dos pontos considerados. Para saber mais sobre a representação *Box and Whiskers*, traduzido para Diagrama em caixa de bigodes, consulte <https://www.alea.pt/images/topicos/activalea/pdf/ActivAlea20.pdf>

6.2. Considere agora apenas os dados referentes aos 2 últimos censos.

6.2.1. Calcule, para cada município, a variação (em percentagem) do número de nascimentos entre 2011 e 2021.

6.2.2. Para os dados obtidos na alínea anterior, calcule a média e a mediana e faça uma representação gráfica na forma de um histograma. Interprete os resultados obtidos, referindo-se à existência de alguma assimetria no histograma e à relação entre os valores das duas características amostrais calculadas.

6.2.3. Indique os 5 municípios onde o número de nascimentos mais cresceu e os 5 municípios onde o número de nascimentos mais diminuiu e tente apresentar, por exemplo em termos de localização territorial, razões que possam justificar esses resultados.



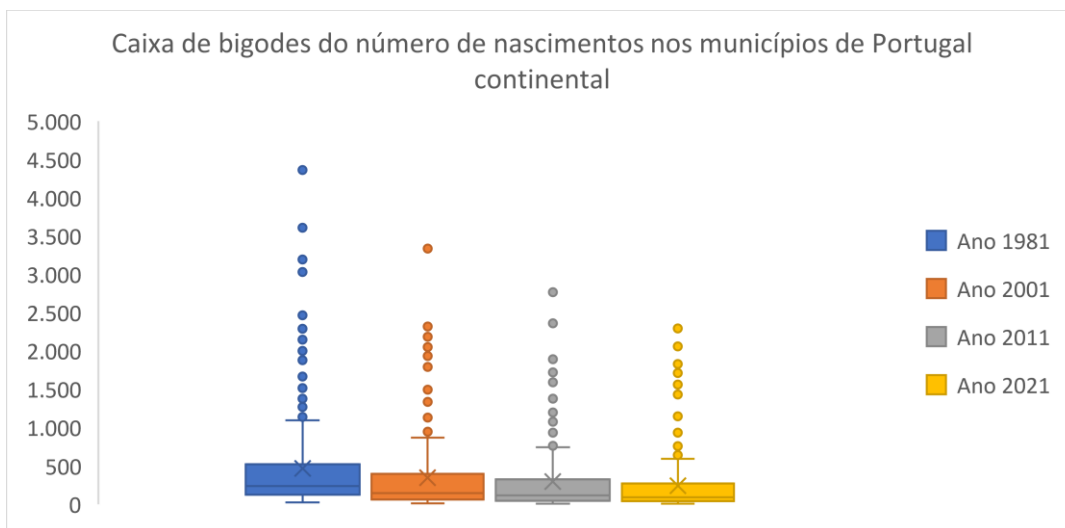
- 6.3.** Obtenha, a partir da tabela construída no passo 5, a percentagem de nados-vivos do sexo masculino para todos os municípios em 2021.
Calcule algumas características amostrais desses dados e represente-os numa caixa de bigodes e num histograma. Interprete as representações obtidas e compare-as.
- 6.4.** É usual referir que a probabilidade de nascer rapaz é igual à de nascer rapariga, ou seja 50%. Tendo em conta os resultados obtidos na questão anterior, poderemos continuar a utilizar essa referência para a população portuguesa?
- 6.5.** Elabore um pequeno relatório com as principais conclusões deste estudo. Caso necessite, elabore mais alguns cálculos ou representações gráficas que não foram pedidas nos pontos anteriores.

ALEA



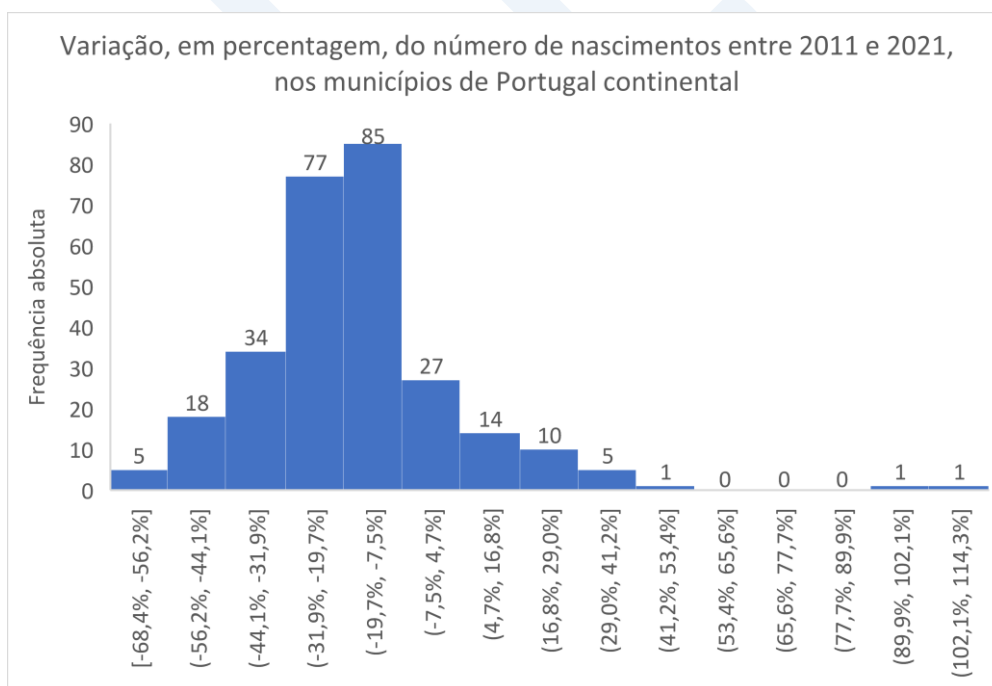
Alguns resultados utilizando a folha de cálculo Excel

6.1.



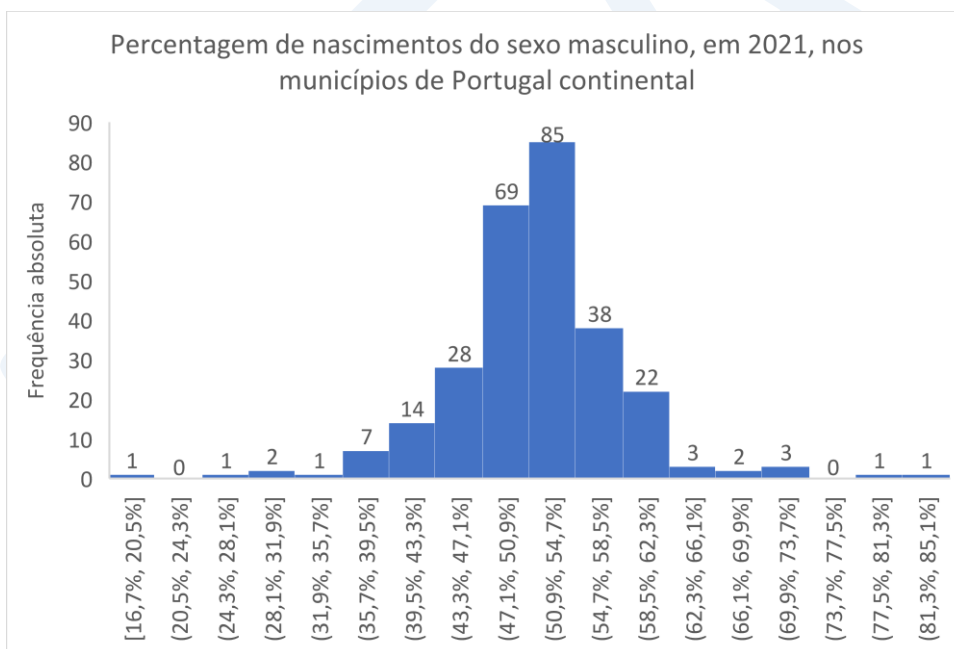
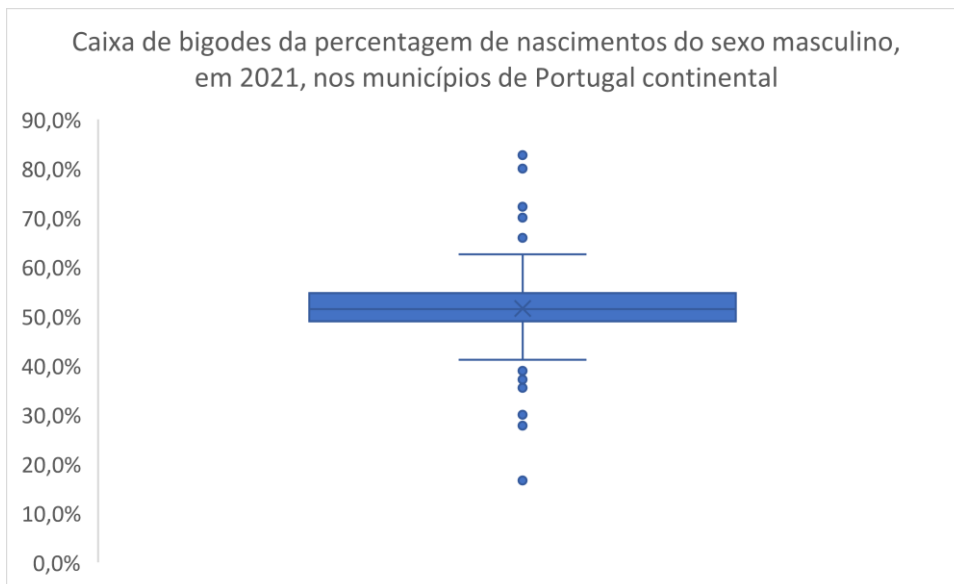
6.2.2.

Média= -17,3%
 Mediana= -19,3%



6.3.

Média= 51,6%
 Mediana= 51,4%
 Desvio padrão= 7,1%



Dica – Repare-se que o histograma anterior apresenta, para ambos os lados, algumas classes afastadas da parte central e com poucos elementos, mantendo alguma simetria. É um histograma designado de caudas longas, característico de dados que integram *outliers*. Estas particularidades também são apresentadas pelo diagrama em caixa de bigodes, onde os *outliers* são visíveis. Repare-se ainda que a simetria apresentada pelas duas representações, de certo modo, já era esperada, tendo em consideração que a média e a mediana estão próximas.

Nota 1 - Algumas considerações sobre a construção das classes, na forma de intervalos, em que se organizam os dados para construir um histograma

Quando se pretende representar um conjunto de dados, de um modo geral contínuos, através de um histograma, é necessário começar por construir as classes para elaborar a tabela de frequências. De um modo geral, consideram-se os seguintes passos.



- 1) Definição do número de classes que se deseja considerar. Neste primeiro passo, deve-se considerar um número de classes, k , que será tanto maior quanto o número de dados, mas que, de um modo geral, se aconselha não ultrapassar as 15 classes, não havendo uma regra fixa.
- 2) Definição da amplitude do intervalo de classe. Uma vez estipulado o número de classes que se deseja considerar, procede-se à obtenção da amplitude, h , do intervalo de classe. Para isso, é usual considerar a amplitude dos dados, dividir essa amplitude pelo número de classes k , e tomar para amplitude do intervalo de classe, um valor aproximado, por excesso, do quociente anterior, ou seja

$$h = \text{valor aproximado, por excesso, do quociente } \frac{\text{máximo dos dados} - \text{mínimo dos dados}}{k}$$

- 3) Uma vez obtida a amplitude de classe h , procede-se à construção das classes (todas com a mesma amplitude) utilizando uma das seguintes metodologias:

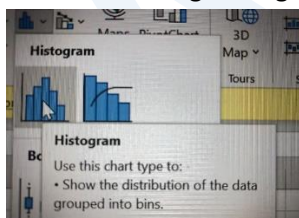
<p>Começar a partir do mínimo</p> <p>1ª classe [mínimo, mínimo+h[</p> <p>2ª classe [mínimo+h, mínimo+2h[</p> <p>...</p> <p>k-ésima classe [mínimo+(k-1)h, mínimo+kh[</p>	<p>ou Começar a partir do máximo</p> <p>k-ésima classe]máximo-h, máximo]</p> <p>(k-1)ésima classe]máximo-2h, máximo-h]</p> <p>...</p> <p>1ª classe]máximo-kh, máximo-(k-1)h]</p>
--	---

Repare-se que para a construção das classes, se começarmos a partir do mínimo (máximo), devem-se considerar os intervalos fechados (abertos) à esquerda e abertos (fechados) à direita, nomeadamente o último (primeiro).

Nota 2 – Algumas considerações sobre o histograma obtido pela folha de cálculo Excel

Com a folha de cálculo do Excel temos a possibilidade de, a partir da lista dos dados, obter imediatamente o histograma. Para isso basta seguir os seguintes passos:

- 1) Selecionar as células da coluna onde estão os dados;
- 2) No painel das funções Selecionar *Insert*;
- 3) Selecionar na figura seguinte *Histogram*



- 4) As representações gráficas das alíneas 6.2 e 6.3.3 foram obtidas seguindo esta metodologia;
- 5) Fazemos notar que a construção das classes é feita considerando como limite inferior da 1ª classe o mínimo, mas não é seguida a metodologia de considerar todas as classes fechadas à esquerda e abertas à direita;
- 6) Existe a possibilidade de o utilizador definir à partida o número de classes que pretende considerar ou, em alternativa, a amplitude de classe.