# I – Tarefas Excel

A primeira secção deste trabalho tem como objetivo explorar algumas ferramentas e funções do Excel, com o objetivo de se ter um eficiente tratamento dos dados. Este segmento do trabalho é um importante ponto de começo, não só porque dá um contacto direto com os dados e por sua vez algumas diretrizes acerca de como estes devem ser tratados tendo em conta a sua natureza, como também nos dá algum tratamento e limpeza. Para começar esta tarefa, foi criada uma nova folha vazia com as mesmas categorias da folha “dados”. Foi também adicionado códigos às restantes respostas às perguntas do inquérito.

## 1

Esta primeira tarefa pede para “Aplicar o filtro a todas as variáveis e apresentar a primeira linha fixa”. O filtro, não só ordena as linhas de qualquer variável que se queira, como também torna possível a seleção de apenas certas linhas, dependendo do pretendido. Por outro lado, a afixação da primeira linha faz com que esta fique sempre visível mesmo quando os dados são arrastados para baixo. Em conjunto, estas duas ferramentas facilitam a identificação das colunas e a filtragem dos dados.

Procedimento

1. Ir a “Dados” na aba de cima do Excel.
2. Carregar na opção “Filtrar”.

## 2

O segundo ponto pedido consiste em “Atribuir de código de não resposta a uma variável com valores omissos”, isto é, foram substituídas todas as respostas deixadas em branco por “NA”.

Procedimento

1. Utilizar a função IF com o valor lógico ISBLANK

## 3

A terceira tarefa entende-se como “Atribuir texto às categorias de uma variável que tenha códigos”. Assim, foram substituídos todos os valores codificados pela respetiva correspondência em texto, de forma a que a leitura dos dados seja mais fácil e intuitiva.

Procedimento

1. Utilizar a função IF com o valor lógico ISNUMBER, usando aqui as funções VLOOKUP e HLOOKUP quando adequado.

## 4

O quarto ponto “Corrigir possíveis erros” tem uma interpretação muito clara, pedindo então para se corrigirem os valores respondidos que são irrealistas e que não fazem sentido. Estes são considerados erros que ocorreram na inserção dos dados.

Sendo assim, para se realizar o ponto 4 do enunciado, foram alterados manualmente os valores correspondentes a estes casos. Por exemplo, na coluna da idade, foi observado que não era possível que um inquirido tivesse a idade de 700 anos. Desta observação foi concluído que, quando o inquirido colocou 700 na idade, este enganou-se e queria colocar 70. Assumindo esta possibilidade como verdadeira, foi substituído neste caso o valor 700 por 70 na coluna dos anos e a respetiva linha. Da mesma maneira, foi substituído o valor 500 por 50.

Procedimento

1. Alterar manualmente os valores.

## 5

Para se efetuar o ponto 5 do enunciado “Criar regra de validação para duas variáveis: uma quantitativa e a outra qualitativa”, foi aplicada uma regra de validação das variáveis “idades” e “urbrural”. Deste modo, foi deliberado que na idade estariam apenas valores entre 16 e 99 e que a variável rural assumiria apenas os seguintes valores de texto: “rural”, “urbano” e “suburbano”.

Procedimento

1. Pressionar a secção dados da aba superior e utilizar as ferramentas de dados.
2. Pressionar validação de dados.
3. Selecionar a lista para o urbrural e o valor numérico para as idades.
4. Escolher os valores que queríamos que este respeita-se.

## 6

O sexto ponto pedia-se para “Construir uma tabela de frequências absolutas com funções de contagem e cálculo de percentagens. Era então tencionado que se utilizassem funções de contagem numa dada variável, determinando assim a frequência (quantidade de vezes que aparecem) de certos valores numa dada coluna. Por fim, esses dados seriam utilizados para construir a tabela de frequências absolutas. A variável escolhida para ser analisada foi “urbrural” e foram considerados todos os valores da coluna respetiva (incluindo os “NA”).

Procedimento

1. Utilizar a função CONTAR.SE.
2. No primeiro parâmetro inserir a coluna referente à variável escolhida.
3. O segundo parâmetro colocar o valor que se procura
4. Para se calcular a percentagem, dividir a frequência de cada valor pelo total de linhas da coluna correspondente à variável a se analisada (no nosso caso foi 860) e meter isso em estilo percentagem (\*100).

## 7

No ponto sete, o objetivo era “Construir uma tabela de frequências dinâmica”. Como o próprio nome indica, esta baseia-se numa tabela de frequências habitual, no entanto a vantagem em relação á de frequências normal é que esta é feita automaticamente pelo Excel, bastando assim escolher a variável a ser analisada. Para a realização deste exercício, foi escolhida a variável “idades”.

Procedimento

1. Carregar em “Inserir”, na aba superior do Excel e selecionar a opção “Tabela Dinâmica”.
2. No editor da tabela dinâmica, selecionar como linhas, colunas e valores a variável em estudo (idades neste caso).

## 8

A tarefa 8 envolvia, tal como é expresso no enunciado, “Construir uma tabela de cruzamentos dinâmica”. Este último ponto da secção do “Excel”, assemelha-se muito ao anterior que visa montar uma tabela de frequências dinâmica. A diferença entre ambas é que a de cruzamentos “cruza”, como sugere o nome duas variáveis distintas enquanto que a outra tem como base apenas os valores de uma variável. Para a criação desta tabela decidiu-se as seguintes correlações para os parâmetros pedidos no editor da tabela. Para as “Linhas” foi colocada a variável “anos\_esc”, para as “Colunas” foi colocado a o “agregado” e para os “Valores” foi colocada a contagem dos agregados.

Procedimento

1. Seguir atentamente os paços 1 e 2 do procedimento relativo á construção de uma tabela de frequências dinâmica.
2. No editor anteriormente mencionado, inserir diferentes correspondências para os parâmetros: “Linhas”, “Colunas” e “Valores”.

# II – Tarefas Jamovi

Na segunda secção deste trabalho foi utilizada a ferramenta “Jamovi”, que proporcionou uma fácil leitura dos dados. Assim, várias tabelas descritivas e um gráfico com cruzamento de variáveis foram criadas no Jamovi e com recurso ao módulo R para consolidar esta análise.

Para se melhor a importância do Jamovi, digamos que esta ferramenta surgiu com o intuito de “estabelecer uma ponte entre investigador e estaticista, de uma forma livre e aberta”. Quer isto dizer, que este projeto (Jamovi) foi formado segundo a filosofia de que todo o software de ciência de dados deve ser dirigido pela comunidade, ou seja, qualquer pessoa deve conseguir publicar análises de dados.

Tabelas com medidas descritivas

No primeiro ponto da parte II é pedido para criar 3 tabelas com medidas descritivas, de forma a demonstrar a capacidade da ferramenta. As medidas descritivas vão ajudar-nos diretamente na análise dos dados, dando-nos informação sobre várias informações pertinentes come a média e a mediana de variáveis.

Passos para a importação do ficheiro Excel

1. Selecionar o menu de barras no canto superior esquerdo.
2. Escolha a opção “Abrir” no ficheiro Excel com os dados já tratados.

Tratamento dos dados

1. Introduzir uma descrição às variáveis “v12”, “v13”, “v14”, “v15”, “v16”, “v72” (para uma melhor compreensão dos dados).

Construção de tabelas de 1 variável

1. Ir para a aba “Análises” e selecionar “Exploração”.
2. Abrir uma janela de Estatística Descritiva com as variáveis da base de dados importada.
3. Selecionar a variável que se quer meter em tabela e arrastar para o separador “Variáveis”

Para a primeira tabela, foi escolhida para analisar a variável “idade” e foi esboçada a seguinte tabela:

Uma imagem com mesa

Descrição gerada automaticamente

Construção de tabelas de 2 ou mais variáveis

1. Passos 1 e 2 do procedimento anterior
2. Selecionar as variáveis que se querem tabular e mover para o separador “Variáveis”

Para a segunda tabela, optámos por juntar as duas variáveis: “anos\_esc” e “agregado”. O resultado foi a seguinte tabela:

Uma imagem com mesa

Descrição gerada automaticamente

Construção de tabelas com cruzamento de variáveis

1. Passo 1 do procedimento anterior
2. Arrastar uma das variáveis para o separador “Variáveis” e a segunda variável para o separador “Separar por”

Para a última tabela decidimos cruzar os variáveis “anos\_esc” (para meter em “Variáveis”) e “sexo” (para meter em “Separar por”), obtendo a seguinte tabela:

Uma imagem com mesa

Descrição gerada automaticamente

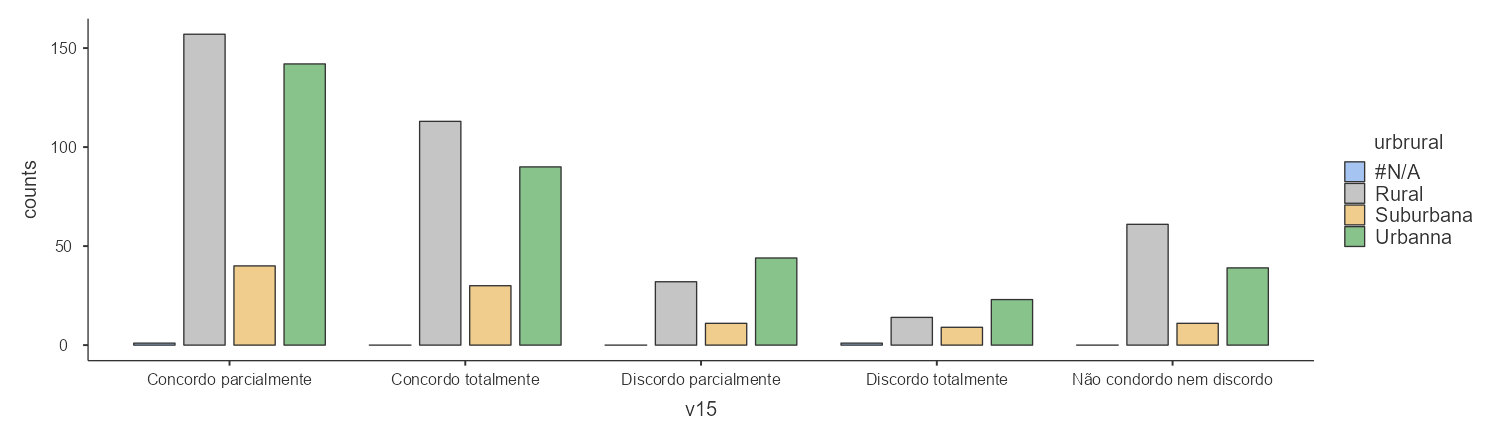
Tabelas com medidas descritivas

Para o segundo ponto, era pedido “Um gráfico com cruzamento de variáveis”. Para a resolução deste tópico decidimos utilizar as variáveis “v15” (“É correto usar animais em experiências médicas se isso contribuir para salvar vidas humanas?”) e “urbrural” (indica o meio de habitação do inquirido).

Gráfico de cruzamento de variáveis

1. Passo 1 e 2 do procedimento anterior.
2. Selecionar a opção “Gráficos” e de seguida a opção “Gráfico de barras”.

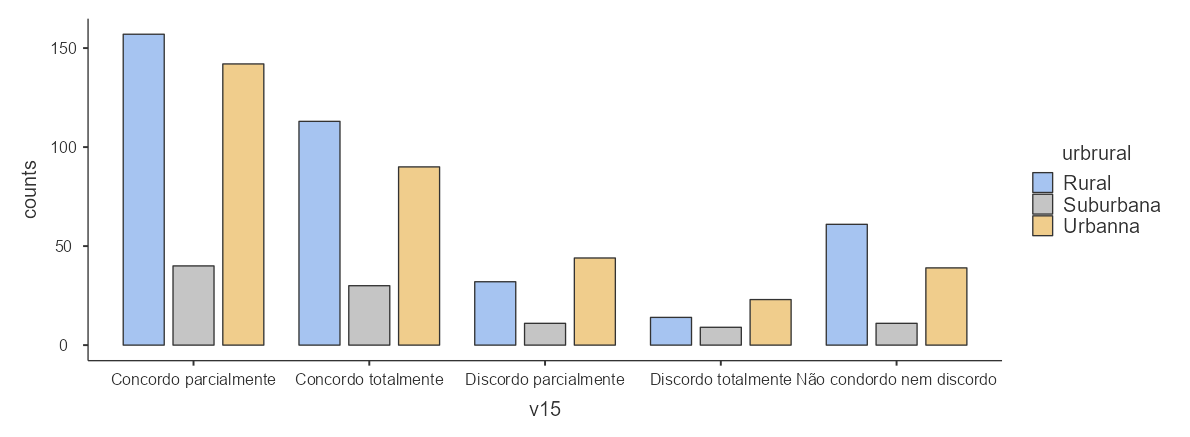
Utilizando as variáveis acima (a “v15” para meter em “Variáveis” e a “urbrural” para meter em “Separar por”), o seguinte gráfico de barras foi feito.



Eliminação dos valores omissos no gráfico

1. Ir à aba “Dados” e selecionar a coluna da variável que se quer apagar os valores omissos e “Valores omissos”.
2. Abrir a janela dos valores omissos e clicar em “Adicionar valor omisso”
3. Completar o espaço “quando $source” com “== ‘#N/A’””

Através da utilização deste procedimento na variável “urbrural”, o gráfico anterior atualiza para o seguinte:



Com uma breve análise ao gráfico, observa-se que a maioria de pessoas que concorda ser correto usar animais em experiências médicas se isso contribuir para salvar vidas são do meio rural, enquanto a maioria que discorda é proveniente do meio Urbano.

Apagar

(Para possíveis arranjos no último procedimento)

Como apareciam os valores omissos no gráfico, solucionamos com os seguintes passos:

Fomos para a aba dos “Dados”, selecionamos a coluna da variável “urbrural” e “Valores omissos”. Abrindo a janela dos valores omissos, clicamos em “Adicionar valor omisso”, completamos o espaço C“quando $source” com “== ‘#N/A’” e agora o gráfico atualiza para este: