

Como gerar uma amostra aleatória simples com o Microsoft Excel ®

Este texto complementa o conteúdo da Unidade 2 da disciplina Estatística Aplicada à Administração.

Na seção 2.1 da Unidade 2 vimos a importância da utilização de números aleatórios para a retirada de amostras. Neste texto mostraremos como usar o Microsoft Excel ® para gerar uma amostra aleatória simples. Isso pode ser feito de duas maneiras: através da ferramenta geração de números aleatórios ou da opção amostragem, presentes no suplemento “Ferramentas de Análise”. Nosso objetivo é retirar uma amostra aleatória de 40 pessoas registradas no CRA de Santa Catarina, graduados em faculdades situadas no Estado. De acordo com os textos das Unidades 1 e 2, supomos que dentre os 11676 registrados em 24/10/2007, 9000 se enquadram nesta condição.

O primeiro passo é abrir o aplicativo e instalar o suplemento “Ferramenta de Análise”, caso este ainda não esteja disponível no menu Ferramentas.

Instalando o suplemento “Ferramentas de Análise”

O Microsoft Excel ® dispõe de um suplemento com vários métodos estatísticos, chamado “Ferramentas de Análise”, que se instalado faz surgir a opção “Análise de dados” no menu Ferramentas. Dentre os vários métodos estatísticos disponíveis, há um para geração de números aleatórios (na realidade números pseudo-aleatórios). Na instalação típica do Excel este suplemento não é ativado, de modo que ele não aparece no menu Ferramentas, veja a figura a seguir.

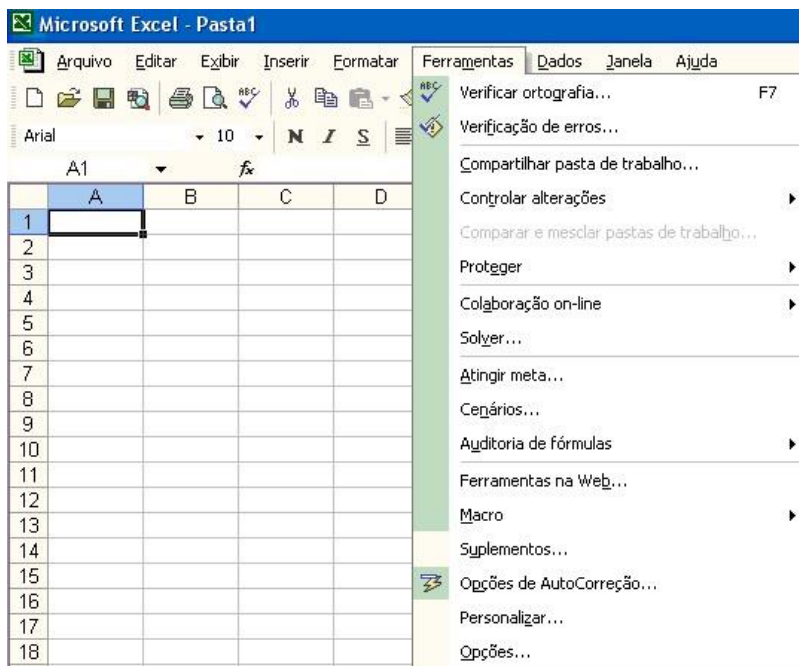


Figura 1 - Menu ferramentas do Microsoft Excel ®
Fonte: Adaptada pelo autor a partir de Microsoft ®

Observe que não aparece a opção “Análise de Dados”, indicando que o suplemento ainda não foi instalado. Para realizar a instalação devemos clicar sobre a opção “Suplementos”, que resultará na figura 2, a seguir.

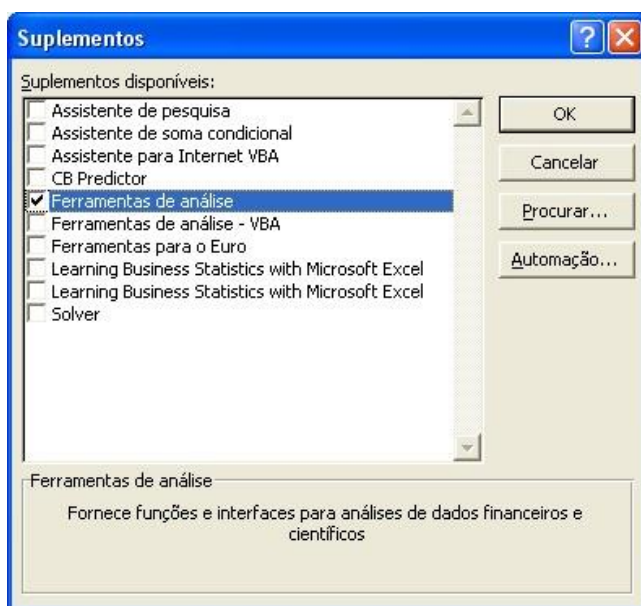


Figura 2 - Suplementos disponíveis no Microsoft Excel ®
Fonte: adaptada pelo autor de Microsoft ®

Para instalar o suplemento é necessário marcar a opção “Ferramentas de Análise” e clicar OK.

É bastante provável que o Windows ® peça o CD do Microsoft Office ® para carregar o suplemento, pois na versão típica de instalação o suplemento não é incluído.

Após a ativação do suplemento o menu “Ferramentas” apresentará a opção “Análise de Dados”:

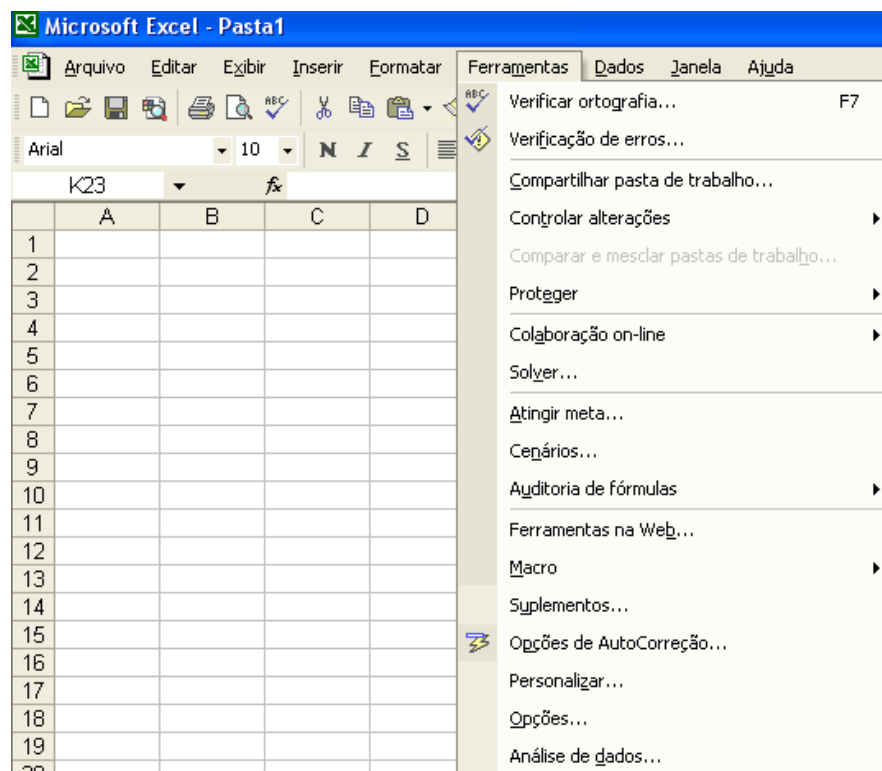


Figura 3 - Menu ferramentas com o suplemento Ferramentas de análise ativado.

Fonte: adaptada pelo autor de Microsoft ®

Agora, você pode acionar a opção Análise de dados e aparecerá a tela mostrada na figura 4.

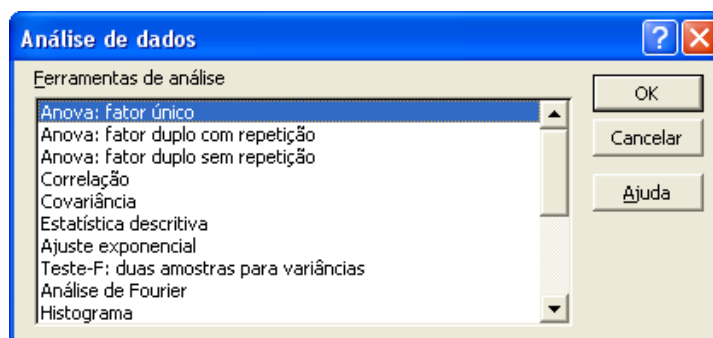


Figura 4 - Opções de Análise de Dados

Fonte: adaptada pelo autor de Microsoft ®

Observe cuidadosamente. Muitas das opções descritas serão vistas no transcorrer da disciplina de Estatística Aplicada à Administração: Estatística Descritiva na Unidade 4, Histograma na Unidade 3, entre outras. Algumas não serão estudadas na disciplina, mas com as referências bibliográficas você poderá perfeitamente utilizá-las no futuro.

Vamos mostrar a primeira maneira de obter uma amostra aleatória a partir do Microsoft Excel ®.

Obtendo uma amostra aleatória através da ferramenta “Geração de número aleatório”

Devemos procurar a opção “Geração de número aleatório”, através da barra de rolagem da figura 4, resultando na figura 5.

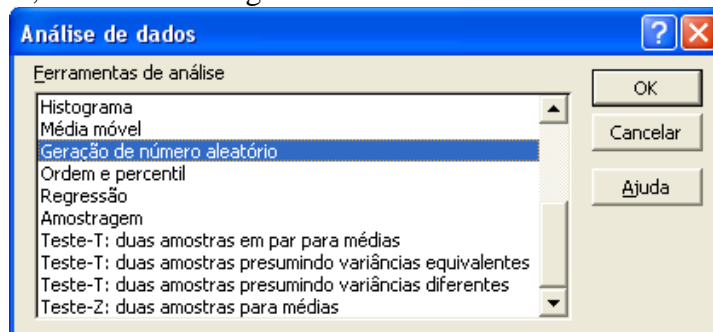


Figura 5 - Opção "Geração de número aleatório"

Fonte: adaptada pelo autor de Microsoft ®

Agora pressione OK, e vamos ter acesso às opções específicas de Geração de número aleatório. Precisamos preenchê-las com cuidado, tendo em mente os objetivos de nosso procedimento. Veja a figura 6.

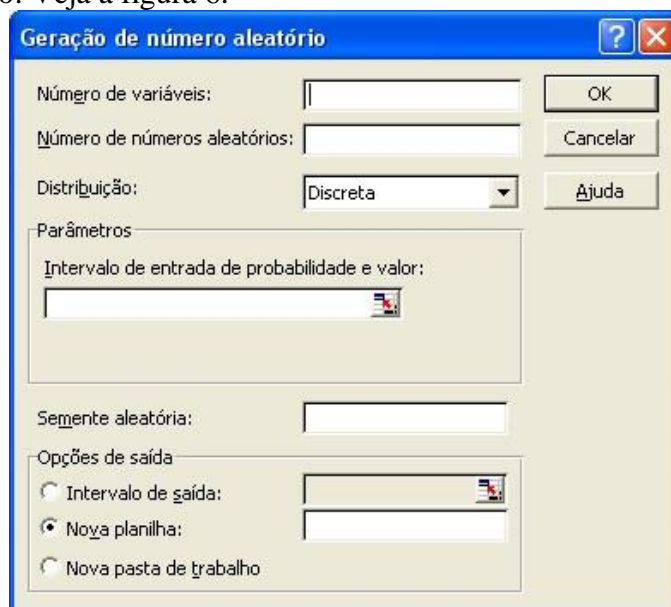


Figura 6 - Opções de geração de número aleatório

Fonte: adaptada pelo autor de Microsoft ®

Vamos analisar cada opção (se quiser, pode consultar a Ajuda do próprio Excel ®):

- Número de variáveis: define quantas colunas de uma planilha você quer “preencher” com os números aleatórios. No nosso caso vamos escolher apenas 1.
- Número de números aleatórios: quantos números você quer efetivamente gerar (se você escolheu Número de variáveis maior do que 1 o Excel ® gerará esta quantidade de números em cada uma das colunas). No nosso caso, vamos escolher 40, conforme estabelecido no início do texto.

- Distribuição: o Excel® possibilita a você escolher entre uma série de distribuições de probabilidade para gerar os números. Os conceitos de probabilidade serão vistos nas Unidades 5 e 6 da disciplina de Estatística Aplicada à Administração. O padrão do Excel é a distribuição discreta (seria uma distribuição fornecida pelo usuário em células de alguma planilha ativa, relacionando valores e probabilidades de ocorrência). Por enquanto, estamos apenas interessados em obter uma amostra aleatória simples, em que todos os elementos da população tenham a mesma probabilidade de serem sorteados: para garantir isso devemos selecionar a distribuição uniforme.
- Intervalo de entrada de probabilidade e valor: caso fosse usada uma distribuição discreta devemos especificar as células de uma planilha ativa que contém os valores da distribuição e as respectivas probabilidades.
- Semente aleatória, citando a ajuda do Microsoft Excel ®: “um valor opcional, a partir do qual números aleatórios possam ser gerados. Você pode voltar a usar este valor posteriormente para gerar os mesmos números aleatórios”. Vamos deixá-lo em branco.
- Opções de saída: você pode escolher onde quer que os resultados sejam escritos.
 - # Intervalo de saída: indicar a célula superior esquerda a partir da qual você quer que os resultados sejam escritos na planilha ativa.
 - # Nova planilha: “para adicionar uma nova planilha à pasta de trabalho atual e colar os resultados começando pela célula A1 da nova planilha”.
 - # Nova pasta de trabalho: “para criar uma nova pasta de trabalho e colar os resultados em uma nova planilha na nova pasta de trabalho”.

Como já foi dito, temos interesse em selecionar a distribuição uniforme, devemos então clicar na aba associada ao campo distribuição, tal como na figura 7.

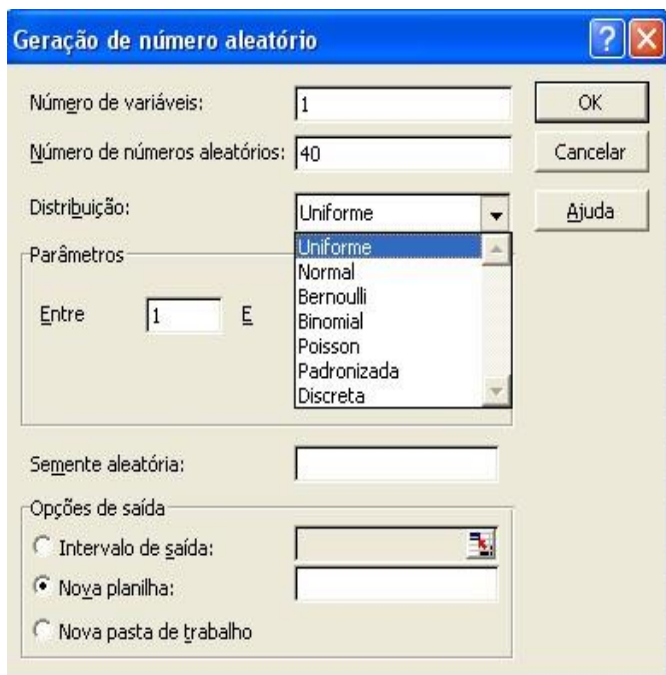


Figura 7 - Escolha da distribuição uniforme
Fonte: adaptada pelo autor de Microsoft ®

Observe que há várias opções de distribuição. Nas Unidades 5 e 6 da disciplina de Estatística Aplicada à Administração você conhecerá as suas propriedades e aplicações.

Após selecionar a distribuição uniforme é preciso informar ao Excel ® os seus parâmetros. Sem entrar em maiores detalhes, são os números iniciais e finais da distribuição, na figura 8.

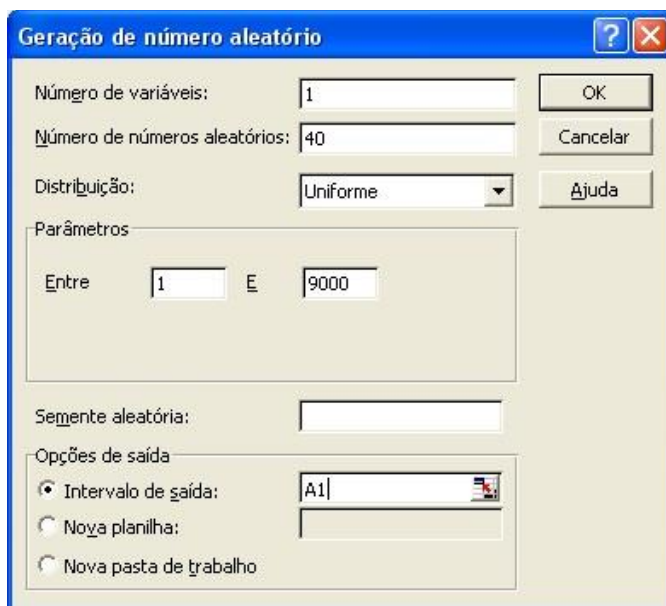


Figura 8 - Opções da distribuição uniforme
Fonte: adaptada pelo autor de Microsoft ®

Conforme dito no início do texto, queremos uma amostra aleatória de 40 elementos. Escolhemos a distribuição uniforme, e queremos que o aplicativo gere números aleatórios entre 1 e 9000 (os 9000 registrados no CRA de Santa Catarina graduados no estado. Além disso, desejamos que os resultados sejam colocados a partir da célula A1 da planilha em uso).

Uma vez selecionadas todas as opções basta clicar em OK, e os números serão gerados, como é mostrado na figura 9, onde são apresentados apenas os primeiros 10 valores.

Microsoft Excel - Pasta1

Arquivo Editar Exibir Inserir Formatar

Arial 10 N I S

A1 3438,6196478164

| | A | B | C | D |
|----|----------|---|---|---|
| 1 | 3438,62 | | | |
| 2 | 907,0244 | | | |
| 3 | 5368,762 | | | |
| 4 | 8092,053 | | | |
| 5 | 7961,601 | | | |
| 6 | 8626,22 | | | |
| 7 | 131,4521 | | | |
| 8 | 3667,391 | | | |
| 9 | 7769,356 | | | |
| 10 | 1248,122 | | | |

Figura 9 - Resultados da distribuição uniforme
Fonte: adaptada pelo autor de Microsoft ®

Você deve estar se perguntando: “Como selecionarei os elementos com tais valores ‘quebrados’?”

Há duas soluções possíveis:

- 1) Simplesmente ignorar a parte decimal: o primeiro selecionado seria o registrado 3438, o segundo será o 907, e assim por diante.
- 2) Arredondar os valores para o inteiro mais próximo através de uma função do Microsoft Excel ®.

A segunda opção demanda um pouco mais de explicação. O Excel ® dispõe de uma função estatística de arredondamento (na realidade mais de uma). Com ela podemos arredondar valores numéricos de células para tantas casas quanto quisermos. No nosso caso, queremos obter números inteiros, com zero casas decimais. Se escolhêssemos arredondar o valor que está na célula A1, e pôr o resultado na célula B1, teríamos a seguinte sintaxe de função: = ARRED(A1;0). Onde ARRED é a função de arredondamento do Excel ®, A1 é a célula que contém o valor a arredondar, e 0 é o número de casas decimais que queremos no resultado. Podemos ver a sintaxe da função na planilha através da figura 10.

Microsoft Excel - Pasta1

Arquivo Editar Exibir Inserir Formatar

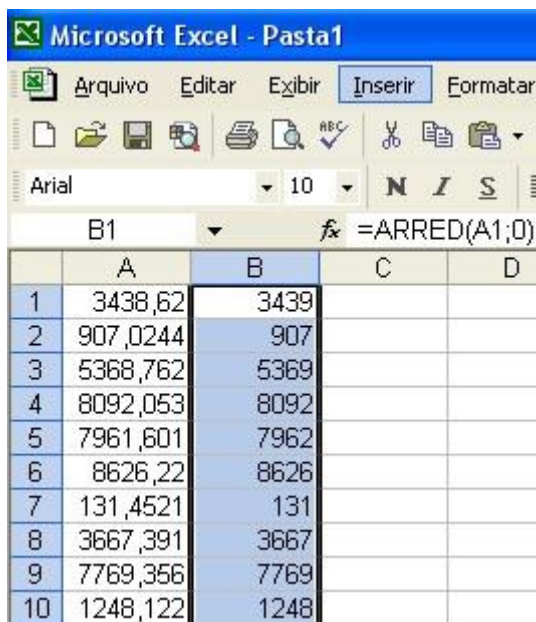
Arial 10 N I S

SOMAQUAD X ✓ fx =arred(A1;0)

| | A | B | C | D |
|---|----------|--------------|---|---|
| 1 | 3438,62 | =arred(A1;0) | | |
| 2 | 907,0244 | | | |
| 3 | 5368,762 | | | |
| 4 | 8092,053 | | | |
| 5 | 7961,601 | | | |

Figura 10 - Função de arredondamento
Fonte: adaptada pelo autor de Microsoft

Após digitarmos a fórmula (se tiver dúvidas a respeito, consulte o material da disciplina de Informática Básica), pressionamos “Enter” e o Excel ® calcula o resultado. Agora basta “arrastar” a fórmula até o final da amostra de 40 elementos, o que leva à situação mostrada na figura 11.



| | A | B | C | D |
|----|----------|------|---|---|
| 1 | 3438,62 | 3439 | | |
| 2 | 907,0244 | 907 | | |
| 3 | 5368,762 | 5369 | | |
| 4 | 8092,053 | 8092 | | |
| 5 | 7961,601 | 7962 | | |
| 6 | 8626,22 | 8626 | | |
| 7 | 131,4521 | 131 | | |
| 8 | 3667,391 | 3667 | | |
| 9 | 7769,356 | 7769 | | |
| 10 | 1248,122 | 1248 | | |

Figura 11 - Resultados do arredondamento
Fonte: adaptada pelo autor de Microsoft ®

Novamente, você vê apenas os primeiros 10 resultados.

Repare que agora o primeiro selecionado será o registrado 3439, seguido pelo 907, 5369, e assim por diante. Como o sorteio foi aleatório, limitado entre 1 e 9000 teoricamente todos os elementos da população têm chance de pertencer à amostra. Agora bastaria identificar os elementos sorteados e realizar a pesquisa com eles.

Uma forma alternativa de obter a amostra aleatória simples é através da ferramenta “Amostragem”, que veremos a seguir.

Obtendo uma amostra aleatória através da ferramenta “Amostragem”

A ferramenta de Amostragem do Excel ® seleciona apenas valores numéricos, quantitativos. Além disso, os valores precisam estar em uma planilha ativa para que possa ser selecionada a amostra. Assim, para o caso da nossa pesquisa, em que há 9000 elementos na população devemos ter uma listagem com 9000 números em uma planilha, tal como mostrado na figura 12 (apenas os primeiros 10 valores).

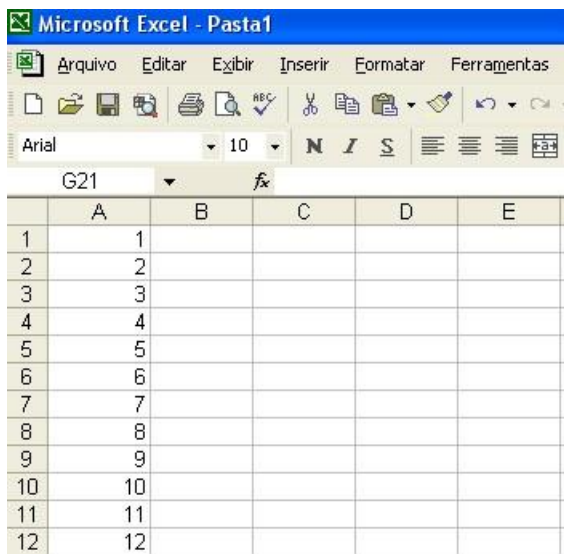


Figura 12 - Listagem de elementos para Amostragem

Fonte: adaptada pelo autor de Microsoft ®

Teoricamente você precisa escrever os 9000 números. Mas graças à característica de “arrastar” do Excel ® basta escrever os dois primeiros (1 na célula A1, e 2 na célula A2), selecionar as duas células e arrastá-los até a linha 9000. O aplicativo automaticamente porá os números de 3 a 9000 nas respectivas células.

A listagem está pronta para ser amostrada.

O próximo passo é selecionar a opção Análise de dados no menu Ferramentas (ver figura 3), e procurar pela ferramenta Amostragem, tal como mostrado na figura 13:

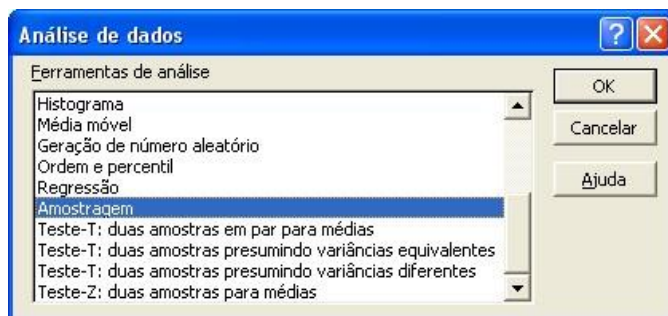


Figura 13 - Ferramenta Amostragem

Fonte: adaptada pelo autor de Microsoft ®

Ao selecionar a ferramenta Amostragem veremos a tela da figura 14:

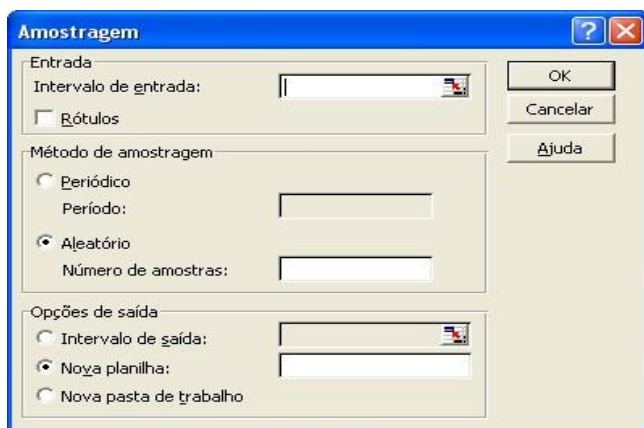


Figura 14 - Opções da ferramenta de amostragem

Fonte: adaptada pelo autor de Microsoft ®

O intervalo de entrada contém a listagem da qual queremos retirar amostra: da célula A1 à célula A9000.

O método de amostragem é o Aleatório (a nossa amostragem aleatória simples), sendo que o Número de amostras deve ser 40.

Veja a figura 15.

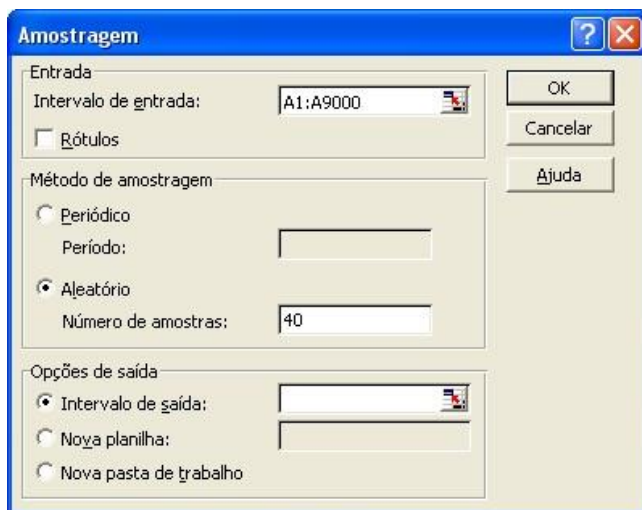


Figura 15 - Opções da ferramenta de amostragem

Fonte: adaptada pelo autor de Microsoft ®

Resta definir onde serão colocados os resultados. Podemos escolher que eles serão escritos a partir da célula B1 da planilha em uso.

Os primeiros 10 elementos da amostra selecionada podem ser vistos na figura 16.

| | A | B | C |
|----|----|------|---|
| 1 | 1 | 6966 | |
| 2 | 2 | 2483 | |
| 3 | 3 | 6071 | |
| 4 | 4 | 6737 | |
| 5 | 5 | 629 | |
| 6 | 6 | 4780 | |
| 7 | 7 | 3598 | |
| 8 | 8 | 7796 | |
| 9 | 9 | 6033 | |
| 10 | 10 | 5442 | |

Figura 16 - Amostra produzida pela ferramenta amostragem

Fonte: adaptada pelo autor de Microsoft ®

Você vê aqui os primeiros 10 elementos sorteados: o registrado número 6966, seguido pelo 2483, 6071 e assim por diante. Basta agora identificar pelos números quem são as pessoas e entrevistá-las.

Teoricamente você pode realizar este processo com populações de até 65 mil elementos, que é o limite máximo de linhas do Excel ®