**ESTUDO DE CASO: Base de Dados MM2022**

MARCOS VINICIUS SOKABE RIBEIRO

CIENTISTA DE DADOS

**METADADOS**

*Descrição da base de dados*

A base de dados original é um arquivo de Excel (.xlsx) de, aproximadamente, 9.5 megabytes, nomeada “BASE DE DADOS – ESTUDO DE CASO 2022”. Contém sete planilhas, registradas e nomeadas individualmente em abas: a) “BASE DE INSCRITOS”; b) “ABERTURA”; c) “Palestras”; d) “Oficinas”; e) “Respostas Avaliação do Evento”; f) “INDICAÇÃO DOS PARCEIROS”; g) “CEP BRASIL”.

*Planilhas*

1. BASE INSCRITOS: tabela de dados brutos com dados gerais sobre o público de inscritos no evento MM2022 (Mentalidades Matemáticas). Contém 759 linhas preenchidas, com 758 respostas em dezesseis colunas que representam variáveis de múltiplas respostas. Elas, em sequência: “INSCRIÇÃO” (tipo numérico sequencial, tratado como número de identificação único para cada inscrição no evento); “EMAIL” (texto com endereço de e-mail de cada inscrito); “Check abertura” (tipo numérico que confirma a participação do inscrito no evento “Abertura”); “Check Oficina 2” (tipo textual que confirma a participação do inscrito no evento “Oficina 2”, orientado por tema: “Ciência de dados”, “Economia”, “Equidade” e “Neurociência”, através do formato: *0 = não participou; nome da oficina = participou*; “Data de nascimento” (tipo numérico de datação, formato de algarismos comum, DD/MM/AAAA, com a data de nascimento dos inscritos); “Gênero” (tipo textual que classifica o gênero dos inscritos, através do formato: “MASCULINO”, “FEMININO”, “HOMEM” e “MULHER” (ver item 2: saneamento de dados); “Raça” (tipo textual que classifica a etnia dos inscritos, através do formato: “AMARELA”, “BRANCA”, “INDÍGENA”, “PARDA” e “PRETA”; “Rede” (tipo textual que classifica a rede educacional do inscrito, através do formato: “PÚBLICA MUNICIPAL”, “PÚBLICA ESTADUAL” e “OUTRA”; “Secretaria” (tipo textual e geográfico que classifica a Secretaria educacional do inscrito com base na localização, através do formato: nome da cidade( “null” ou “/” ou “-“ ou “(“) e estado (ver item 2: saneamento de dados); “CEP” (tipo numérico que classifica o código postal de origem das inscrições); “Escolaridade” (tipo textual que classifica a escolaridade dos inscritos, através do formato: “ESPECIALIZAÇÃO OU APERFEIÇOAMENTO COMPLETO”, “MESTRADO”, “ENSINO SUPERIOR COMPLETO”, “ENSINO SUPERIOR INCOMPLETO”, “ENSINO MÉDIO COMPLETO” e “DOUTORADO”; “OCUPAÇÃO” (tipo textual que classifica a ocupação profissional dos inscritos (ver item 2: saneamento de dados); “Deficiência” (tipo textual que informa se o inscrito possui ou não algum tipo de deficiência, através do formato: *NÃO = não possui; (“VISUAL”, “AUDITIVA” ou “FÍSICA”) = possui;* “Como soube” (tipo textual livre que informa o canal/origem de recebimento da notícia de realização do evento; “Cadastro” (tipo numérico que informa a data e hora do cadastro da inscrição, formato: “DD/MM/AAAA HH:MM), (ver item 2: saneamento de dados).

1. ABERTURA: tabela de dados brutos com dados sobre o evento ABERTURA. Contém 759 linhas preenchidas, com 758 respostas em quatro colunas que representam variáveis de múltiplas respostas. Elas, em sequência: “EMAIL” (texto com endereço de e-mail de cada inscrito); “ABERTURA” (tipo textual que identifica a presença dos inscritos no evento ABERTURA, através do fomato: *X = presente, null = ausente*); “INSCRIÇÃO” (tipo numérico sequencial, tratado como número de identificação único para cada inscrição no evento); “Cadastro” (tipo numérico que informa a data e hora do cadastro da inscrição, formato: “DD/MM/AAAA HH:MM), (ver item 2: saneamento de dados).
2. Palestras: tabela de dados brutos com dados sobre o evento Palestras (1, 2, 3, 4 e 5). Contém 759 linhas preenchidas, com 758 respostas em sete colunas que representam variáveis de resposta binária. Elas, em sequência: “INSCRIÇÃO” (tipo numérico sequencial, tratado como número de identificação único para cada inscrição no evento); “EMAIL” (texto com endereço de e-mail de cada inscrito); “Palestra 1” (tipo textual que identifica a presença dos inscritos no evento Palestra 1, através do fomato: *X = presente, null = ausente*); “Palestra 2” (tipo textual que identifica a presença dos inscritos no evento Palestra 2, através do fomato: *X = presente, null = ausente*); “Palestra 3” (tipo textual que identifica a presença dos inscritos no evento Palestra 3, através do fomato: *X = presente, null = ausente*); “Palestra 4” (tipo textual que identifica a presença dos inscritos no evento Palestra 4, através do fomato: *X = presente, null = ausente*); “Palestra 5” (tipo textual que identifica a presença dos inscritos no evento Palestra 5, através do fomato: *X = presente, null = ausente*).
3. Oficinas: tabela de dados brutos com dados sobre o evento Oficina (1, 2 e 3). Contém 759 linhas preenchidas, com 758 respostas em sete colunas que representam variáveis de múltiplas respostas. Elas, em sequência: “INSCRIÇÃO” (tipo numérico sequencial, tratado como número de identificação único para cada inscrição no evento); “EMAIL” (texto com endereço de e-mail de cada inscrito); “Cadastro” (tipo numérico que informa a data e hora do cadastro da inscrição, formato: “DD/MM/AAAA HH:MM), (ver item 2: saneamento de dados); “Oficina 1”, “Oficina 2” e “Oficina 3” (tipo textual que confirma a participação do inscrito no evento “Oficina 1”, “Oficina 2” e “Oficina 3”, orientado por tema: “Ciência de dados”, “Economia”, “Equidade” e “Neurociência”, através do formato: *null = não participou; nome da oficina = participou*.
4. Resposta Avaliação do Evento: tabela de dados brutos com dados de sentimento sobre o evento. Contém 692 linhas preenchidas, com 691 respostas em dezesseis colunas que representam variáveis de múltiplas respostas. Elas, sem sequência: “Em uma única palavra, resuma este Evento” (tipo textual que solicita uma palavra para resumir a impressão geral do evento); “Qual a probabilidade de você indicar este evento para colegas de trabalho?” (tipo numérico escalar que avalia a disposição dos participantes em recomendar o evento para colegas de trabalho, através do formato: 0 a 10); “Como você avalia a ampliação dos seus conhecimentos sobre as temáticas a partir deste Evento?” (tipo numérico escalar que avalia a extensão em que o evento contribuiu para ampliar os conhecimentos do participante, através do formato: 1 a 5); “Como você avalia seu interesse em utilizar os conteúdos, instrumentos e saberes deste Evento em sua prática?” (tipo numérico escalar que avalia o interesse do participante em aplicar o que foi aprendido no evento, em sua prática profissional, através do formato: 1 a 5); “Sobre o Evento em geral, selecione as opções que se aplicam: [Chato]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Confuso]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Útil]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Criativo]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Engajador]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Interessante]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Neutra]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Legal]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Aceitável]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Divertido]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Assustador]” (tipo textual ou marcador que avalia a impressão geral dos inscritos sobre o evento entre “Chato”, “Confuso”, “Útil”, “Criativo”, “Engajador”, “Interessante”, “Neutra”, “Legal”, “Aceitável”, “Divertido” e “Assustador”); “Use o espaço abaixo caso queira compartilhar sugestões, opiniões ou comentários” (tipo textual livre, espaço aberto para feedback adicional e comentários dos participantes”.
5. INDICAÇÃO DOS PARCEIROS: tabela de dados brutos com dados de contato (EMAIL) dos participantes indicados por parceiros. Contém 319 linhas preenchidas, com 318 respostas em uma coluna que representa uma variável de múltiplas respostas. Ela: “EMAIL” (texto com endereço de e-mail de cada inscrito advindo da indicação dos parceiros).
6. CEPS BRASIL: tabela de dados brutos com dados de contato (EMAIL) dos participantes indicados por parceiros. Contém 319 linhas preenchidas, com 318 respostas em uma coluna que representa uma variável de múltiplas respostas. Ela: “EMAIL” (texto com endereço de e-mail de cada inscrito advindo da indicação dos parceiros).
7. CEPS BRASIL: tabela de dados brutos com dados de código postal dos participantes. Contém 732.764 linhas preenchidas, com 732.763 respostas em uma coluna que representa uma variável de múltiplas respostas. Ela: “CEP” (tipo numérico com o código postal de cada inscrição); “CIDADE/ESTADO” (tipo textual que especifica a cidade e o estado de cada inscrição, em correspondência aos dados de “CEP”).

**Saneamento de dados**

Ferramentas e fórmulas de tratamento e normalização de dados utilizados: #Microsoft\_Power\_Query, #Linguagem\_M.

Métodos:

1. Normalização de dados de datação: extração dos oito primeiros dígitos dos valores das colunas “Cadastro”; substituição dos dígitos “/” por “”; extração dos primeiros dois dígitos da string sem “/” *(= Text.Start([SuaColuna], 2*); extração dos dígitos centrais da string sem “/” *(= Text.Middle([SuaColuna], 2, 2*); extração dos últimos dois dígitos da string sem “/” ( *= Text.End([SuaColuna], 2*); concatenação dos valores, em ordem correta e separadas por “/” ( *= Text.From([DD]) & "/" & Text.From([MM]) & "/" & [AAAA])*; criação de coluna personalizada “Idade” (= [Data de Nascimento] – [Cadastro\_x] / 365).
2. Normalização de classes em “Gênero” e “Ocupação”: substituição dos dígitos “HOMEM” por “MASCULINO”; substituição dos dígitos “MULHER” por “FEMININO”; substituição de “OUTRO:” por “”.
3. Normalização de classes em “Como Soube”: concatenação de diversas subclasses em classes mais abrangentes:

=if

List.Contains(

{

"GRUPO DE WHATSAPP (CÉLULA).",

"GRUPOS DE WHATSAPP",

"WHATSAPP",

"WHATSAPP",

"GRUPO WHATSAPP",

"WHATSAPP",

"WHATSAPP",

"WHATSAPP DA CÉLULA SP",

"NO GRUPO DO WHATSAPP"

},

[AA]

)

then "WHATSAPP"

else [AA]

**Estudo de tabelas**

**BASE INSCRITOS**

*Colunas da tabela “BASE INSCRITOS”*

A tabela “BASE INSCRITOS” é a parte do dataset que contém dados gerais das inscrições no evento MM2022. Seu propósito é gerenciar e organizar as inscrições dos participantes, fornecendo um registro detalhado de suas informações pessoais, educacionais, profissionais e de contato. No entanto, sua estrutura original pode ser otimizada de modo a conter apenas dados úteis. Para tal, as colunas de “Check de abertura” (Coluna C), “Check Oficina 2 (Coluna D) e Coluna E (não possui dados) foram excluídas.

*Valores da tabela “BASE INSCRITOS”*

Os valores das colunas “Data de Nascimento” e “Cadastro” foram convertidos em data, “DD/MM/AAAA”. Os valores da coluna “Cadastro”, em específico, possuem notação diferente da adotada usualmente. Optou-se pela exclusão do horário e rearranjo de formato de data para: “DD/MM/AAAA”, através de uma nova coluna chamada “Cadastro\_x”. Apenas um erro foi identificado na transformação: o valor da linha 41, “14/09/2021 21:47:34”, foi interpretado como erro de registro, já que 99% da amostra na coluna “Cadastro” se referem ao mês de março de 2021, enquanto os dias variam, majoritariamente, entre “18” a “21”. Portanto, esse dado foi tratado como erro e excluído das medições. A partir da nova coluna de cadastro e a coluna de data de nascimento, foi possível calcular a idade, em anos, dos participantes. (“Cadastro\_x” – “Data de Nascimento” = “Idade”). Na nova coluna “Idade”, notou-se o erro de cálculo nas linhas 4, 438 e 507 e que indicavam idades sobre-humanas ou zero. Esses três valores foram excluídos da medição.

Na coluna “Gênero”, há quatro classes diferentes de definição de gênero do inscrito. “MASCULINO”, “FEMININO”, “HOMEM” E “MULHER”. Entendeu-se que “HOMEM” e “MULHER” podem ser classificados em “MASCULINO” e “FEMININO” também, de modo a normalizar os dados.

Na coluna “Ocupação”, várias profissões a preenchem, no entanto, algumas classes representam o mesmo significado. Para que os dados fossem normalizados e, ainda assim, fossem mantidas as diferenças e detalhamento de classes, o especificador “OUTRO:” foi excluído das linhas. Entende-se que esse especificador diminui classes populares da amostra, como “Professor” e “Professora”, diminuindo sua abrangência.

A coluna “Como Soube” trata-se do conjunto de dados sobre canais de divulgação do evento. Devido à enorme variedade de classes, sendo muitas apenas diferenças entre dígitos, optou-se por criar 12 categorias mais abrangentes para medição dos dados. Elas: AMIGOS E COLEGAS DE TRABALHO, E-MAIL, FACEBOOK, INDICAÇÃO, INSTAGRAM, INSTITUIÇÕES EDUCACIONAIS, PREFEITURAS E SECRETARIAS, OUTRAS REDES SOCIAIS, JÁ CONHECIA, JO BOALER, OUTRAS INSTITUIÇÕES, OUTROS, PARCEIRO.

**ABERTURA**

*Colunas da tabela “ABERTURA”*

A tabela “ABERTURA” é a parte do dataset que contém dados de participantes do evento de Abertura do MM2022. Seu propósito é conferir quais, dentre os inscritos, compareceram ao avento de Abertura.

*Valores da tabela “ABERTURA”*

Os valores da coluna “Cadastro” foram convertidos em data, “DD/MM/AAAA”. Os valores da coluna “Cadastro”, em específico, possuem notação diferente da adotada usualmente. Optou-se pela exclusão do horário e rearranjo de formato de data para: “DD/MM/AAAA”, através de uma nova coluna chamada “Cadastro\_x”.

**Oficinas**

*Colunas da tabela “Oficinas”*

A tabela “Oficinas” é a parte do dataset que contém dados de participantes dos eventos “Oficina 1”, “Oficina 2” e “Oficina 3” do MM2022. Seu propósito é conferir quais, dentre os inscritos, compareceram aos eventos de Oficina oferecidos.

*Valores da tabela “Oficinas”*

Os valores da coluna “Cadastro” foram convertidos em data, “DD/MM/AAAA”. Os valores da coluna “Cadastro”, em específico, possuem notação diferente da adotada usualmente. Optou-se pela exclusão do horário e rearranjo de formato de data para: “DD/MM/AAAA”, através de uma nova coluna chamada “Cadastro\_x”.

**Respostas Avaliação do Evento**

*Colunas da tabela “Respostas Avaliação do Evento”*

A tabela “Oficinas” é a parte do dataset que contém dados de avaliação dos participantes do MM2022. Seu propósito é conferir a impressão geral dos participantes através de análise qualitativa.

*Valores da tabela “Respostas Avaliação do Evento”*

Os valores das colunas “Sobre o Evento em geral, selecione as opções que se aplicam: [Confuso]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Útil]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Criativo]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Engajador]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Interessante]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Neutra]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Legal]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Aceitável]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Divertido]”, “Sobre o Evento em geral, selecione as opções que se aplicam: [Assustador]” tiveram seu marcador substituído de “.” por “X”, a fim de obter mais clareza.

**Estatística descritiva e análise exploratória**

**Público (tabela BASE INSCRITOS)**

*Público*

Para definição de perfil do público de inscritos do evento MM2022, a tabela “BASE INSCRITOS” foi utilizada como conjunto de dados mais relevante. Havendo um marcador que singulariza cada inscrito no evento, podemos compilar as respostas dentre as diversas categorias de variáveis para produzir insights interessantes sobre o público. Ao todo, foram 754 inscrições de pessoas com variação de idade, gênero, raça, secretaria, CEP, escolaridade, ocupação profissional e caracterização de pessoa com deficiência ou não, após saneamento de dados.

*Idade*

Segundo a base de dados, 53 números de idades distintos foram identificados, com 5 valores exclusivos contendo apenas um registro. A idade dos participantes variou do máximo de 75 anos até o mínimo de 19 anos, tendo uma média de 42,45 anos, aproximadamente. Através da análise de distribuição de frequência, em um intervalo de 20 anos, nota-se que a faixa de idade mais frequente entre os participantes encontra-se de 33 a 53 anos, representando cerca de 66% da amostra. Isolando apenas uma classe de idade, o tipo com maior representação foi 34 anos de idade, com 30 registros, cerca de 3,97%, seguido de 38 anos, 42 anos e 44 anos, cada qual contendo 29 registros, 3,84% da amostra. As idades com menor frequência foram 71 anos e 67 anos com apenas um registro cada e, 19 anos e 64 anos, com apenas dois registros, representando menos de 1% do total geral.

Gráfico, Histograma

Descrição gerada automaticamente

Figura 1: Histograma que cruza os dados de número de inscrições por idade em intervalos de classe

Gênero

O gênero dos participantes possuí dois valores textuais para identificação, MASCULINO e FEMININO. Do total de 754 inscrições, cerca de 80,10% foi composto pelo gênero FEMININO, classe majoritária, enquanto 19,90% foi composto de inscritos do gênero MASCULINO.

Raça

A raça dos participantes possuí cinco valores textuais para identificação, BRANCA, PARDA, PRETA, AMARELA E INDÍGENA. Do total de 754 inscrições, a classe majoritária dos inscritos foi BRANCA, representando 57,42% da amostra, aproximadamente, com 433 registros, seguida de PARDA, 213 valores, 28,24% da amostra, PRETA, 91 valores, 12,06%, AMARELA, 14 valores, 1,85% e, INDÍGENA, 3 valores, 0,3%.

Figura 2. Gráfico de Funil, aqui, utilizado para mensurar a distribuição de raça por inscritos

Rede

Quanto à coluna “Rede” que identifica a rede de ensino, ela possuí três valores textuais para identificação, PÚBLICA MUNICIPAL, PÚBLICA ESTADUAL e OUTRA. Do total de 754 inscrições, a classe majoritária dos inscritos foi a rede PÚBLICA MUNICIPAL, com 326 valores registrados, cerca de 43,23% dos inscritos. Seguidamente, a classe OUTRA com 299 registros, com 39,65% e, PÚBLICA ESTADUAL com 129 registros, aproximadamente 17,10% da amostra.

Gráfico, Gráfico de pizza

Descrição gerada automaticamente

Figura 3. Gráfico de Setores que mostra a distribuição das redes de ensino

Secretaria

A coluna Secretaria identifica a secretaria de ensino a qual o inscrito pertence. Trata-se de uma identificação a partir da cidade/estado de origem do inscrito. Segundo a base dados, a Secretaria com maior número de inscritos foi da cidade de São Paulo, com 101 valores registrados, representando cerca de 13,39% de inscritos, seguida por Ponta Grossa, cidade do Paraná, com 35 registros, 4,64% da amostra. Candeias aparece em terceiro, com 31 registros, 4,11% da amostra. Completam o top5 secretaria mais frequentes: Rio de Janeiro, 2,78% e, Gurupi, 2,38%, aproximadamente. Em relação aos estados com maior quantidade de secretarias registradas, SP aparece em primeiro com cerca de 30% da amostra.

Escolaridade

A coluna Escolaridade identifica o nível de escolaridade que cada inscrito possui. Trata-se de uma identificação de seis classes de variáveis, DOUTORADO, ENSINO MÉDIO COMPLETO, ENSINO SUPERIOR COMPLETO, ENSINO SUPERIOR INCOMPLETO, ESPECIALIZAÇÃO OU APERFEIÇOAMENTO COMPLETO e MESTRADO. A classe de escolaridade mais frequente entre os inscritos foi ESPECIALIZAÇÃO OU APERFEIÇOAMENTO COMPLETO, representando cerca de 47,61% da amostra, 359 valores registrados, seguida de: ENSINO SUPERIOR COMPLETO, 21,35%; MESTRADO, 19,23%; DOUTORADO, 6,23%; ENSINO SUPERIOR INCOMPLETO, 4,90% E ENSINO MÉDIO COMPLETO, 0,6%.

Gráfico, Gráfico de barras

Descrição gerada automaticamente

Figura 4. Gráfico de Barras é utilizado para mostrar a escolaridade dos inscritos

Ocupação

A coluna Ocupação identifica a ocupação profissional que cada inscrito possui. Trata-se de uma identificação de várias classes de variáveis. Devido à alta variedade de classes descritas na base dados e a dificuldade inerente de expandir a sua classificação de modo mais generalista, optou-se por utilizar um N superior = 10 do conjunto da coluna para identificar as cinco classes com maior frequência no dataset. Professora aparece como a profissão de maior frequência, representando cerca de 19% da amostra. Nota-se que, dentro do conjunto, há variáveis que, em tese, representariam de modo mais amplo a mesma classe, como “Professora de Matemática” = “Professora”, “Professor de Matemática = Professor”. No entanto, entende-se, também, que “Professora de Matemática” e “Professor de Matemática” compõem subgrupos relevantes que não devem ser ignorados na medição. A amplitude do saber profissional se estende para inúmeras derivações. A subclasse de Professoras de Matemática representa cerca de 1,19% da amostra, enquanto Professor de Matemática representa cerca de 0,53%. Em segundo, a classe profissional com maior frequência entre os inscritos foi Coordenador/a Pedagógico, com 18% da amostra, seguido de Professor, 14%, Dirigente/Superintendente/Supervisor de Ensino, 10% e, Diretor/a/Vice-Diretor/a, com 3%.

Dentre as subclasses, há, ainda categorias notáveis, como Estudante, 0,92% da amostra, Professora de Educação Básica, 0,13% e, Formador, com 0,66%.

Gráfico, Gráfico de barras

Descrição gerada automaticamente

Figura 5. Gráfico de Barras é utilizado para mostrar as profissões mais frequentes dos inscritos, com N Superior = 10 (dez maiores classes). Dados percentuais calculados em relação apenas às dez maiores classes

Deficiência

A coluna Deficiência indica se o inscrito pode ser caracterizado como Pessoa com Deficiência. Trata-se de uma identificação de quatro classes de variáveis. Segundo a base de dados, a imensa maioria dos inscritos não possui deficiência, com 98% das respostas. A deficiência visual aparece como mais frequente, representando 0,66% da amostra, seguido por deficiência auditiva, 0,53% e, deficiência física, com 0,13% da amostra.

Como Soube

A coluna Como Soube indica o canal de recebimento da informação do evento MM2022. É uma variável importante para mensurar o sucesso dos processos publicitários. Seus valores, originalmente, divergiam em dígitos e subclassificações. Após saneamento de dados e transformação de subcategorias em categorias mais abrangentes, obtivemos as informações: E-MAIL demonstrou-se como canal de divulgação mais eficiente, representando 22% das escolhas dos inscritos; seguido por INSTITUIÇÕES EDUCACIONAIS, PREFEITURAS E SECRETARIAS, com 21% dos registros. AMIGOS E COLEGAS DE TRABALHO representaram 13% da amostra, como terceiro melhor canal de conhecimento do evento. Em relação às redes sociais, INSTAGRAM apareceu como 6% da amostra e, FACEBOOK com 5%, aproximadamente. Com menos de 1%, WHATSAPP se mostrou como a rede social especificada menos eficiente na divulgação para os inscritos. Quando desconsiderados a especificidade dessas três redes sociais, OUTRAS REDES SOCIAIS aparece como terceira maior fonte de divulgação, com 7% da amostra.

Gráfico, Gráfico de mapa de árvore

Descrição gerada automaticamente

Figura 6. Treemap, utilizado aqui para demonstrar quais canais de divulgação atraíram mais inscritos. Dados calculados em relação apenas aos canais preenchidos

**Agrupamentos**

*Grupos com 3 variáveis: gênero, raça e idade*

Isoladamente, as medições por colunas ajudam a entender o perfil do público presente no evento MM2022. No entanto, é possível determinar os grupos mais frequentes com base em combinações de variáveis. Para a primeira análise, foi utilizada uma combinação de três classes: gênero, raça e idade.

A combinação dominante entre os participantes entre 20 e 30 anos, foi o grupo de mulheres (FEMININO) da raça BRANCA, representando 40% da amostra recortada pela faixa de idade. Seguidamente, aparecem mulheres pardas, representando 17,14%, aproximadamente. Mulheres pretas se mostram como o terceiro grupo jovem mais popular, com cerca de 15,23% da amostra fatiada pela faixa de idade. Homens (MASCULINO) e brancos aparecem em terceiro, com 12 combinações entre três variáveis registradas, com 11,42%. Nessa faixa etária, os grupos com menores representação foram: mulheres amarelas, homens pretos, homens amarelos e homens indígenas, com nenhum registro de combinação encontrado.

Quando mudamos a faixa de idade para participantes entre 31 e 41 anos, o grupo de mulheres brancas se mantém como majoritário, representando cerca de 40,83% da amostra fatiada. Mulheres pardas continua como segundo maior grupo, com 23,75%. O terceiro maior grupo, no entanto, muda: homens brancos aparecem com 30 registros, representando 12,5%, enquanto mulheres pretas, logo em sequência, com pouco menos de 11% da amostra. Os grupos com menos registros encontrados foram mulheres indígenas e homens amarelos, 0,42%, aproximadamente, seguido por homens indígenas, classe que não teve registro.

Considerada a faixa etária de maior frequência segundo análise anterior de padrão de distribuição de idade por inscrições, adaptada aqui como 42 anos a 62 anos, mulheres brancas continuam sendo maioria substancial do público de participantes do evento, com 207 registros encontrados, 53,35%, aproximadamente. Mulheres pardas, também, se mantém como segunda maior classe combinatória, representando 85 registros e 21,90% da amostra fatiada. Homens brancos e mulheres pretas aparecem, respectivamente, com 7,73% e 7,21%. É notável que os grupos de homens e mulheres indígenas continua como classe menos frequente. Nesta faixa de idade, por exemplo, ambas as combinações estão zeradas.

A última faixa de idade considerada nessa medição compreendia entre participantes com idade igual ou superior a 63 anos. Mulheres brancas continuam como classe majoritária, representando cerca de 52,63% da amostra recortada. Homens pardos assumem a segunda posição, com 21% de representatividade, aproximadamente. Homens brancos aparecem em terceiro, com cerca de 15,78%. Homens e mulheres pretos, amarelos e indígenas não apresentaram registros.

*Grupos com 2 variáveis: idade e canal*

Para esta análise, foi utilizada uma combinação de duas classes: idade e canal. A combinação dos dois marcadores nos possibilita a chance de entender quais canais são mais efetivos de acordo com a idade do participante.

Entre 20 a 35 anos, foram encontrados 200 registros de perfis. O canal de divulgação mais frequente foi E-MAIL com 20% da amostra, seguido de INSTITUIÇÕES EDUCACIONAIS, PREFEITURAS E SECRETARIAS, representando 18%. INSTAGRAM e FACEBOOK, juntos, representam 18,5%, enquanto AMIGOS E COLEGAS DE TRABALHO aparecem com 11,5% da amostra. Quando concatenamos todas as outras classes de canais, JÁ CONHECIA, OUTRAS INSTITUIÇÕES, OUTRAS REDES SOCIAIS, PARCEIRO e WHATSAPP, elas representam 12% da frequência do conjunto.

Entre 36 e 51 anos, foram encontrados 397 registros de perfis. O canal de divulgação mais frequente foi INSTITUIÇÕES EDUCACIONAIS, PREFEITURAS E SECRETARIAS, com 21,15% da amostra, aproximadamente. O segundo canal com maior frequência foi E-MAIL, representando 19,39%. INSTAGRAM, FACEBOOK e WHATSAPP, juntos, representam 12,09% da amostra. Como é notável, participantes entre 36 e 51 anos, ao contrário dos jovens, preferem receber ou recebem naturalmente informações de instituições educacionais, prefeituras e secretarias em relação às redes sociais. A diminuição significativa dos canais, mesmo concatenados entre instagram, facebook e whatsapp, nos sugere que as redes sociais possuiu menor relevância no recebimento da informação do evento. Na faixa mais jovem, facebook e instagram representavam quase 20% da amostra, enquanto na faixa mais velha, ainda que incluindo a classe whatsapp, representou apenas 12%.

Os participantes com idade superior a 51 anos apresentaram 84 registros. O canal de divulgação mais eficiente foi E-MAIL, com 29,76% da amostra. Em segundo, aparece INSTITUIÇÕES EDUCACIONAIS, PREFEITURAS E SECRETARIAS, representando 21,42%. FACEBOOK e INSTAGRAM, juntos, representaram apenas 11,90%, aproximadamente. AMIGOS E COLEGAS DE TRABALHO aparece com 13,09% da representatividade do conjunto de dados. A tendência de a relevância de redes sociais diminuir conforme aumento de idade parece continuar, apesar da diminuição menos significativa entre jovens e adultos entre 36 a 51 anos de idade. Nesse sentido, INSTITUIÇÕES EDUCACIONAIS, PREFEITURAS E SECRETARIAS aparece, novamente, com grande relevância e segundo canal de maior frequência da análise de combinação.

Nota-se que o canal E-MAIL se mantém como o meio mais eficiente de divulgação, sendo a classe majoritária na faixa mais jovem (20-35), a segunda na segunda faixa de idade (36 a 51) e, primeira, novamente, na faixa que comporta os participantes com idade superior a 51 anos.

*Grupos com 5 variáveis: gênero, raça, rede, escolaridade e ocupação*

Para esta análise, foi utilizada uma combinação de cinco classes: gênero, raça, rede, escolaridade e ocupação. Combinações com maior número de variáveis tendem a espraiar a definição de novas classes devido ao alto número de possibilidades combinatórias. Essa observação é importante para entendermos que perfis combinatórios encontrados aqui, apesar de possuírem menor representatividade percentual, são significativos.

O perfil combinatório entre as cinco classes mais frequentes foi mulheres (FEMININO), da raça BRANCA, advinda da rede PÚBLICA MUNICIPAL, de escolaridade ESPECIALIZAÇÃO OU APERFEIÇOAMENTO COMPLETO, de ocupação COORDENADOR/A PEDAGÓGICO, representando 3,31% da amostra, com 25 registros encontrados. Seguido por mulheres brancas, da rede pública municipal, com especialização ou aperfeiçoamento completo, e professoras, com 2,78%, aproximadamente.

Dentre o top10 perfis mais frequentes encontradas nessa análise combinatória, coordenador/a pedagógico aparece como ocupação de maior frequência, representando cerca de 47,55% do conjunto de dados. Professora aparece em sequência, com 44% da amostra. Dirigente, Superintendente, Supervisor de Ensino aparece com 8,39%, aproximadamente. Ainda nesse recorte, a rede pública municipal se mostra como classe majoritária, representando 93 registros e 65% da amostra, aproximadamente. OUTRA aparece com 50 registros, aproximadamente 35%. Rede pública estadual, nesse fatiamento, não possui registros. Em relação à escolaridade, 100% dos primeiros dez perfis, isto é, mais frequentes, são catalogados por Especialização ou Aperfeiçoamento Completo.

Outros perfis de interesse, ainda que menos frequentes, podem trazer insights interessantes sobre a análise combinatória. A escolaridade de nível de pós-graduação, concatenada pela agregação de MESTRADO e DOUTORADO, representa 25,46% da amostra, enquanto a agregação ENSINO MÉDIO COMPLETO, SUPERIOR COMPLETO e SUPERIOR INCOMPLETO representa 26,92%, aproximadamente. Ambos os marcadores individualmente inferiores à quantidade percentual de representação de ESPECIALIZAÇÃO OU APERFEIÇOAMENTO COMPLETO, que compõe, sozinha, 47,61% da amostra.

Trezentos e três (303) combinações exclusivas foram encontradas na análise de combinação, ou seja, trezentos e três perfis possuem combinações que não se repetem, portanto, com apenas uma frequência. Para citar algumas: mulheres pretas com ensino superior completo, advindas de outra rede e Educadora; mulheres indígenas com ensino superior incompleto, advindas da rede pública estadual e universitária; homens amarelos com doutorado, advindos de rede pública estadual e professor.

**Engajamento (tabela ABERTURA, tabela Palestras e Tabela Oficinas)**

ABERTURA

Para medição do engajamento de inscritos no evento Abertura, a tabela “ABERTURA” foi utilizada como conjunto de dados mais relevante. Ao todo, foram 758 inscrições de participantes no evento Abertura do MM2022. O determinador de comparecimento, na tabela, é identificado como “X”.

Em uma análise simples, nota-se que, dentre as inscrições, 19,78% do público compareceu ao evento Abertura, com 150 confirmações pelo determinador. Portanto, 80,21% dos inscritos, aproximadamente, não compareceram ao evento de Abertura.

Palestras

Para medição do engajamento de inscritos nos eventos de Palestras, a tabela “Palestras” foi utilizada como conjunto de dados mais relevante. Ao todo, foram 758 inscrições de participantes no evento Palestras. O determinador de comparecimento, na tabela, é identificado como “X”.

A Palestra 1 aparece como evento de maior sucesso, obtendo uma taxa de comparecimento de 51,45%, seguida pela Palestra 5 com 381 participantes e uma taxa de 50,26%. As palestras menos populares foram as Paletras 4 e 2, com 45,38% e 43,53% de comparecimento, respectivamente. A média de comparecimento dos inscritos às palestras foi de 47,27%, aproximadamente. Em números absolutos, cerca de 358,4 inscritos compareceram, em média, aos eventos de palestras no MM2022.

Cento e oitenta e oito (188) inscritos não compareceram a nenhuma palestra, representando 24,88% da amostra. Cento e setenta um (171), cerca de 22,55%, compareceram a todas às palestras. Cerca de 5,14% dos inscritos compareceram apenas à Palestra 5, enquanto, 5,01% compareceram apenas à Palestra 1. Cerca de 40,89% dos inscritos compareceram entre uma e três palestras, enquanto 34,43% comparecem entre quatro e cinco palestras.

Oficinas

Para medição do engajamento de inscritos nos eventos de Oficinas, a tabela “Oficinas” foi utilizada como conjunto de dados mais relevante. Ao todo, foram 758 inscrições de participantes no evento Oficinas. O determinador de comparecimento, na tabela, é identificado como nome da Oficina comparecida, subclassificado através de temáticas.

A Oficina 1 aparece como evento de maior engajamento, com 290 comparecimentos, cerca de 38,25%. Em segundo, a Oficina 3 possuiu 34,69% de comparecimento, enquanto a Oficina 2 se mostrou como o evento com menor engajamento, cerca de 25,98%. A média da taxa de comparecimento foi de 32,97% e, em números absolutos, de 250 participantes nas oficinas.

Em relação às temáticas das oficinas, Neurociência aparece como o tema de oficina com maior engajamento, representando frequência de 39,70%. Em segundo, o tema Equidade aparece com 20,18% da amostra, aproximadamente. Ciência de Dados e Economia, respectivamente, apresentaram frequência de 19,78% e 19,26%. Dentro de Oficina 1, Neurociência se mostrou como o tema de oficina de maior procura, representando cerca de 23% da escolha dos participantes. Economia e Equidade aparecem com valores iguais, cerca de 5%, enquanto Ciência de Dados se mostrou apenas com 3%. Na Oficina 2, Equidade aparece como tema mais popular, representando cerca de 7% dos comparecimentos, enquanto Economia e Neurociência aparecem em terceiro e quarto lugar, respectivamente, com valores próximos a 6%. Ciência de Dados, novamente, aparece em último, com 5,54%. Na Oficina 3, Ciência de Dados aparece como categoria majoritária, representando 10% dos participantes. Neurociência, em segundo, possuiu 9%, enquanto Equidade e Economia se mostraram próximos a 7%.

Trezentos e cinquenta e seis (356) inscritos não compareceram a nenhuma oficina, representando 46,96% da amostra. Cento e dezoito (118), cerca de 15,56%, compareceram a todas às oficinas. Cerca de 6,59% dos inscritos compareceram apenas em uma oficina. Dentre os participantes que compareceram em apenas uma oficina, o tema mais frequente foi Neurociência com 61,8%, seguido por Equidade, 17,91%, Ciência de Dados, 11,56% e, Economia, com 6,35% da amostra fatiada.

Entre os participantes que frequentaram todas as oficinas, a sequência de escolhas de temas de oficina mais popular foi Oficina 1 = Neurociência, Oficina 2 = Economia, Oficina 3 = Equidade. Seguidamente, a tríade Oficina 1 = Neurociência, Oficina 2 = Equidade, Oficina 3 = Economia. Nesse mesmo recorte, Neurociência aparece como tema mais popular essa faixa, com 111 registros em oficinas, seguido Equidade, 79, Economia, 70 e Ciência de Dados, 62. O total dessa faixa é de 118 comparecimentos.

**Análise de sentimento (tabela Respostas Avaliação do Evento)**

Nuvem de palavras (coluna “Em uma única palavra, resuma este Evento.”)

A nuvem de palavras permite que entendamos as descrições de sentimento mais frequentes na avaliação dos participantes do evento MM2022. Através de uma concatenação de palavras e termos similares que puderam ser abrangidos por termos de mais amplitude, criamos uma visualização das palavras mais frequentes, expressas pelos tamanhos com que aparecem na nuvem.

Texto preto sobre fundo branco

Descrição gerada automaticamente

Figura 7. Nuvem de palavras expressa as palavras mais frequentemente usadas para descrever o evento

A palavra “Desafiador” aparece como termo mais usado para descrever o sentimento geral sobre o evento MM2022, representando 8,68% da amostra. Em seguida, “Ótimo” aparece com 5,49%. Completando o top três palavras mais utilizadas, aparece “Inspiração” com 4,48%.

**NPS (NET PROMOTER SCORE) para medição de Probabilidade de Indicação do Evento para Colegas de Trabalho**

O Net Promoter Score (NPS) é uma métrica para avaliar a satisfação e a lealdade dos participantes em relação a um evento específico. Esta métrica é obtida através de uma pergunta direta: "Qual a probabilidade de você indicar este evento para colegas de trabalho?", em que as respostas variam de 0 a 10.

Com base nos dados coletados, o NPS para a pergunta sobre a probabilidade de indicação do evento para colegas de trabalho foi calculado da seguinte forma:

Promotores (Notas 9-10): 643 participantes, representando 91,4% do total;

Neutros (Notas 7-8): 31 participantes, representando 4,4% do total;

Detratores (Notas 0-6): 18 participantes, representando 2,6% do total.

Promotores (Notas 9-10): Com a vasta maioria dos participantes (91,4%) dando notas altas, é evidente que o evento recebeu um alto nível de satisfação e é altamente recomendado. Estes participantes são os defensores mais entusiasmados do evento e são essenciais para impulsionar a promoção e a participação em futuras edições.

Neutros (Notas 7-8): A proporção de participantes neutros é relativamente baixa, representando apenas 4,4% do total. Isso indica que a satisfação geral foi muito positiva, com poucos participantes tendo opiniões ambivalentes ou neutras sobre a recomendação do evento.

Detratores (Notas 0-6): A presença de detratores é mínima, representando apenas 2,6% do total de participantes. Isso é um excelente indicativo, pois sugere que houve uma recepção positiva e que as críticas negativas ou insatisfações foram praticamente inexistentes, não comprometendo a reputação do evento.

O NPS de mais de 90 para a pergunta sobre a probabilidade de indicação do evento para colegas de trabalho é extremamente positivo e demonstra a eficácia, qualidade e impacto positivo do evento nas temáticas abordadas.

Para as próximas edições, é fundamental manter o foco na qualidade das temáticas e conteúdos apresentados, bem como continuar a promover práticas e abordagens que foram bem recebidas pelos participantes. A comunicação transparente e o engajamento contínuo com os participantes são cruciais para sustentar e possivelmente aumentar os altos níveis de satisfação e lealdade demonstrados.

**Estatística descritiva das colunas “Como você avalia a ampliação dos seus conhecimentos sobre as temáticas a partir deste Evento?” e “Como você avalia seu interesse em utilizar os conteúdos, instrumentos e saberes deste Evento em sua prática?”**

*“Como você avalia a ampliação dos seus conhecimentos sobre as temáticas a partir deste Evento?”*

A média de 4,51 está próxima do valor máximo (5), o que sugere que, em média, os participantes sentiram que o evento contribuiu significativamente para a ampliação dos seus conhecimentos sobre as temáticas abordadas. Com uma mediana de 5, também próxima do valor máximo, reforça-se a ideia de que a maioria dos participantes avaliou positivamente a ampliação de conhecimentos proporcionada pelo evento. Com a moda de 5 sendo o valor que mais se repete, há indicação de que o valor máximo (5) é o mais comum entre as respostas, reforçando a avaliação positiva geral. O desvio padrão indica a dispersão dos dados em relação à média. Um desvio padrão relativamente baixo sugere que as respostas estão concentradas em torno da média, indicando uma consistência nas avaliações.

*“Como você avalia seu interesse em utilizar os conteúdos, instrumentos e saberes deste Evento em sua prática?”*

A média de 4,66 está próxima do valor máximo (5), o que sugere que, em média, os participantes têm um alto interesse em aplicar os conteúdos, instrumentos e saberes adquiridos no evento em suas práticas profissionais ou pessoais. A mediana de 5, assim como na primeira coluna, também é próxima do valor máximo, indicando um interesse geral elevado em utilizar os conhecimentos adquiridos. A moda de 5 novamente é o valor máximo (5), indicando que o valor máximo é o mais comum entre as respostas. O desvio padrão de 0.55 é relativamente baixo, sugerindo que as respostas estão consistentes e próximas da média, o que significa que há um consenso sobre o alto interesse em aplicar os conhecimentos adquiridos.

Os resultados sugerem que o evento foi bem avaliado pelos participantes, tanto em relação à ampliação dos conhecimentos sobre as temáticas abordadas quanto ao interesse em aplicar esses conhecimentos na prática. A alta média, mediana e moda próximas ao valor máximo, juntamente com desvios padrão relativamente baixos, indicam uma avaliação positiva e consistente por parte dos participantes. Isso sugere que o evento foi eficaz em fornecer conhecimentos relevantes e despertar interesse em aplicá-los, o que é um indicativo positivo de sua qualidade e relevância.

**Avaliação do evento (Chato; Confuso; Útil; Criativo; Engajador; Interessante; Neutra; Legal; Aceitável; Divertido; Assustador)**

Dados: Chato:= 7 votos; Confuso = 34 votos; Útil = 405 votos; Criativo = 537 votos; Engajador = 446 votos; Interessante = 564 votos; Neutra = 9 votos; Legal = 286 votos; Aceitável = 95 votos; Divertido = 333 votos; Assustador = 28 votos.

As categorias "Interessante", "Criativo" e "Engajador" têm as maiores contagens, com 564, 537 e 446 participantes respectivamente. Isso indica que a maioria dos participantes teve uma experiência positiva e envolvente com o evento. "Útil" e "Divertido" também receberam avaliações significativas, com 405 e 333 participantes respectivamente. Isso sugere que o evento não apenas foi percebido como valioso em termos de conteúdo, mas também como uma experiência agradável. Categorias como "Neutra", "Chato", "Confuso" e "Assustador" têm as menores contagens, com 9, 7, 34 e 28 participantes respectivamente. Isso indica que uma minoria dos participantes teve uma experiência neutra ou negativa em relação ao evento. As categorias "Legal" e "Aceitável" têm contagens intermediárias, com 286 e 95 participantes respectivamente. Isso sugere uma percepção geralmente positiva, mas menos intensa em comparação com as categorias mais altas.

É importante notar que um participante pode votar em várias opções de avaliação. Por exemplo, alguém que achou o evento "Interessante" também pode ter achado "Criativo" e "Engajador". Isso significa que as contagens não são mutuamente exclusivas e podem se sobrepor.

As avaliações altamente positivas (Interessante, Criativo, Engajador) sugerem que o evento foi bem-sucedido em envolver e interessar a maioria dos participantes. Embora as avaliações neutras e negativas sejam minoritárias, elas ainda existem. É importante analisar feedbacks específicos desses participantes para identificar áreas de melhoria.

A presença de uma ampla gama de avaliações indica que o evento teve aspectos que agradaram diferentes tipos de participantes. Isso é positivo, mas também destaca a importância de personalizar e diversificar futuros eventos para atender a diferentes expectativas e interesses.

Em resumo, a base de dados sugere que o evento foi geralmente bem recebido, com a maioria dos participantes tendo uma experiência positiva, criativa e engajadora. No entanto, é essencial continuar coletando feedback e ajustando futuros eventos para maximizar o valor e a satisfação do público-alvo.

**Perguntas fundamentais**

*Qual perfil do público que participou do evento? Ele corresponde ao público-alvo desejado?*

O perfil do público que participou do evento MM2022 pode ser caracterizado da seguinte forma: Em média, os participantes tinham aproximadamente 42,45 anos, com a faixa etária mais predominante situada entre 35 e 51 anos, representando cerca de 54,64% da amostra. No que diz respeito ao gênero, a maioria era do gênero feminino, correspondendo a 80,10% das inscrições, enquanto o gênero masculino representava 19,90%.

Em relação à raça, os participantes se distribuíram majoritariamente como brancos, totalizando 57,42% da amostra. Os inscritos autodeclarados como pardos constituíram 28,24%, seguidos pelos pretos com 12,06%. As categorias amarela e indígena apresentaram representações menores, com 1,85% e 0,3%, respectivamente.

No contexto educacional, a maioria dos participantes tinha concluído uma especialização ou aperfeiçoamento, representando 47,61% da amostra. Aqueles com ensino superior completo constituíram 21,35%, enquanto os mestres e doutores representaram 19,23% e 6,23%, respectivamente. A ocupação profissional mais frequente entre os participantes foi a de professora, com 19% da amostra, seguida por coordenador/a pedagógico com 18%, professor com 14%, e dirigente/superintendente/supervisor de ensino com 10%.

Quanto à rede de ensino, a maioria dos inscritos pertencia à rede pública municipal, com 43,23%, seguida por 39,65% na categoria "outra" e 17,10% na rede pública estadual. Em relação à origem dos participantes, a cidade de São Paulo foi a mais representada, com 13,39% das inscrições, seguida por Ponta Grossa (4,64%) e Candeias (4,11%).

A análise combinatória de variáveis como gênero, raça e idade mostrou que o grupo mais frequente entre os participantes de 20 a 30 anos era composto por mulheres brancas, representando 40% da amostra nessa faixa etária. Para os participantes entre 31 e 41 anos, as mulheres brancas também predominaram, com 40,83% da amostra. Na faixa etária de 42 a 62 anos, as mulheres brancas continuaram sendo a maioria, com 53,35% da amostra.

Em relação ao canal de divulgação do evento, o e-mail foi o mais eficiente, sendo responsável por atrair 22% das inscrições. As instituições educacionais, prefeituras e secretarias também desempenharam um papel significativo na divulgação, representando 21% das inscrições, seguidas por amigos e colegas de trabalho com 13%.

Em resumo, o perfil do público que participou do evento MM2022 é predominantemente feminino, com idade média em torno de 42 anos, pertencente majoritariamente à raça branca e com formação em especialização ou aperfeiçoamento. A ocupação mais comum entre os participantes é a de professora, e a maioria dos inscritos provém da rede pública municipal. O evento conseguiu atingir seu público-alvo desejado, considerando as características observadas nos participantes.

*Qual a avaliação geral do Evento?*

Com base nos dados fornecidos, a avaliação geral do Evento MM2022 é extremamente positiva. A análise de sentimento revelou que as palavras mais frequentemente usadas para descrever o evento foram "Desafiador", "Ótimo" e "Inspiração", indicando uma experiência predominantemente positiva e enriquecedora para os participantes.

O Net Promoter Score (NPS) também reforça essa visão positiva, com um resultado impressionante de mais de 90 para a probabilidade de indicação do evento para colegas de trabalho. Este valor é indicativo de um alto nível de satisfação e lealdade dos participantes, com a vasta maioria sendo promotores entusiasmados do evento.

As estatísticas descritivas sobre a ampliação dos conhecimentos e o interesse em aplicar os conteúdos do evento também corroboram essa avaliação positiva. As médias, medianas e modas próximas ao valor máximo, juntamente com desvios padrão relativamente baixos, indicam que os participantes sentiram que o evento foi eficaz em fornecer conhecimentos relevantes e despertar interesse em aplicá-los em suas práticas profissionais ou pessoais.

A avaliação do evento em termos de categorias específicas também é majoritariamente positiva. As categorias "Interessante", "Criativo" e "Engajador" lideram as contagens, seguidas por "Útil" e "Divertido". Embora haja algumas avaliações neutras ou menos positivas, elas são minoritárias e não comprometem a impressão geral altamente positiva do evento.

Em resumo, os dados coletados sugerem que o Evento MM2022 foi bem-sucedido em proporcionar uma experiência enriquecedora, inspiradora e engajadora para a maioria dos participantes. A alta satisfação, evidenciada pela análise de sentimento, NPS, estatísticas descritivas e avaliações específicas, demonstra a eficácia, qualidade e impacto positivo do evento. Para futuras edições, é crucial manter o foco na qualidade das temáticas e conteúdos apresentados, bem como continuar a promover práticas e abordagens que foram bem recebidas pelos participantes. A escuta ativa e o engajamento contínuo com os participantes também são fundamentais para sustentar e possivelmente aumentar os altos níveis de satisfação e lealdade demonstrados.

*As indicações dos parceiros foram efetivas?*

Dos 319 indicados pelos parceiros, 123 se tornaram inscrições, o que representa uma conversão de aproximadamente 38,56%. Em outras palavras, cerca de 38,56% das pessoas indicadas pelos parceiros realmente se inscreveram no serviço ou programa em questão. Ao analisar esses números, podemos tirar algumas conclusões importantes. Primeiramente, a taxa de conversão de 38,56% indica que as indicações dos parceiros têm um impacto significativo nas inscrições. Isso sugere que as indicações feitas pelos parceiros são, em grande parte, efetivas em atrair novos inscritos para o serviço ou programa.

No entanto, é crucial considerar o contexto completo. Dos 754 inscritos totais, apenas 123 vieram das indicações dos parceiros. Isso significa que as indicações dos parceiros representam aproximadamente 16,3% do total de inscrições. Embora a taxa de conversão seja alta entre as indicações dos parceiros, o volume de inscrições provenientes dessas indicações ainda é relativamente baixo em comparação com o total de inscrições.

Isso nos leva a outra análise importante: existe um potencial significativo para aumentar o número de inscrições provenientes das indicações dos parceiros. Se conseguirmos otimizar e aumentar a eficácia das indicações dos parceiros, há uma oportunidade clara de impulsionar ainda mais as inscrições totais. Além disso, é essencial considerar estratégias para incentivar e motivar os parceiros a fazerem mais indicações. Isso pode envolver a implementação de programas de incentivo, treinamentos ou campanhas de marketing direcionadas aos parceiros para aumentar o engajamento e a eficácia das suas indicações.

Em resumo, enquanto as indicações dos parceiros mostraram ser efetivas com uma alta taxa de conversão de 38,56%, ainda há espaço para otimização e crescimento. Focar em estratégias para aumentar o volume e a eficácia das indicações dos parceiros pode ser uma abordagem promissora para impulsionar as inscrições totais e maximizar o impacto das parcerias no crescimento do serviço ou programa.

**Conclusão**

A análise exploratória da tabela "BASE INSCRITOS" do evento MM2022 revelou insights valiosos sobre o perfil dos participantes. A idade média dos inscritos foi de aproximadamente 42,45 anos, com a faixa etária mais comum situando-se entre 35 e 51 anos. O gênero feminino predominou significativamente, com 80,10% das inscrições, enquanto a raça branca foi a mais representada, totalizando 57,42% dos inscritos. Quanto à rede de ensino, a maioria pertencia à rede pública municipal (43,23%), seguida pela categoria "outra" (39,65%) e rede pública estadual (17,10%).

No que diz respeito à origem geográfica dos participantes, a cidade de São Paulo liderou com 13,39% das inscrições, seguida por Ponta Grossa (4,64%) e Candeias (4,11%). Em relação à escolaridade, a maioria possuía especialização ou aperfeiçoamento completo (47,61%), seguida por ensino superior completo (21,35%) e mestrado (19,23%).

A ocupação profissional mais comum entre os inscritos foi a de professora, representando 19% da amostra. Nota-se uma variedade de subcategorias dentro dessa profissão, como "Professora de Matemática" e "Professor de Matemática", que juntas compõem cerca de 1,72% da amostra. Outras ocupações notáveis incluíram coordenador/a pedagógico (18%), professor (14%), e dirigente/superintendente/supervisor de ensino (10%).

Quanto à identificação de pessoas com deficiência, a grande maioria (98%) dos inscritos não possuía deficiência. Entre aqueles que tinham, a deficiência visual foi a mais comum, representando 0,66% da amostra.

Por fim, em relação ao canal de divulgação do evento, o e-mail foi o mais eficaz, responsável por 22% das inscrições. Instituições educacionais, prefeituras e secretarias foram o segundo meio mais eficiente (21%), seguido por amigos e colegas de trabalho (13%). As redes sociais tiveram uma contribuição menor na divulgação, com Instagram representando 6% das inscrições, Facebook 5%, e WhatsApp menos de 1%. Considerando todas as redes sociais, incluindo outras não especificadas, elas representaram 7% das inscrições.

Em relação aos agrupamentos com três variáveis - gênero, raça e idade - observou-se que entre os participantes com idades entre 20 e 30 anos, as mulheres brancas predominaram, representando 40% da amostra. Em seguida, mulheres pardas e mulheres pretas foram o segundo e terceiro grupos mais numerosos, com 17,14% e 15,23% respectivamente. Homens brancos apareceram como o terceiro grupo mais frequente nessa faixa etária, com 11,42% das combinações registradas. Por outro lado, mulheres e homens indígenas, amarelos e pretos, bem como homens amarelos e pretos, foram os grupos menos representados nesta faixa etária.

Para participantes com idades entre 31 e 41 anos, o perfil manteve-se relativamente semelhante, com mulheres brancas representando 40,83% da amostra, seguidas por mulheres pardas e homens brancos com 23,75% e 12,5% respectivamente. Novamente, mulheres e homens indígenas e amarelos tiveram representações quase nulas.

Já na faixa etária de 42 a 62 anos, mulheres brancas continuaram sendo a maioria expressiva, com 53,35% da amostra. Mulheres pardas e homens brancos foram os outros grupos mais frequentes, com 21,90% e 7,73% respectivamente. É notável a ausência de participantes indígenas, amarelos e pretos nesta faixa etária.

Por fim, entre os participantes com 63 anos ou mais, mulheres brancas mantiveram a maioria com 52,63%. Homens pardos e brancos ocuparam as posições seguintes, com 21% e 15,78% respectivamente. Novamente, não houve registros de participantes indígenas, amarelos ou pretos.

No que diz respeito aos agrupamentos com duas variáveis - idade e canal de divulgação -, o e-mail foi o canal mais eficaz para alcançar participantes entre 20 e 35 anos, com 20% da amostra. Instituições educacionais, prefeituras e secretarias mostraram-se mais eficazes para a faixa etária de 36 a 51 anos, com 21,15% da amostra. Já para os participantes com mais de 51 anos, o e-mail se destacou novamente, com 29,76% da amostra.

Quanto aos agrupamentos com cinco variáveis - gênero, raça, rede, escolaridade e ocupação - identificou-se que coordenadores pedagógicos eram os mais frequentes, representando 47,55% dos perfis mais comuns. Mulheres brancas da rede pública municipal, com especialização completa e ocupando cargos de coordenador/a pedagógico, foram as mais representativas, totalizando 3,31% da amostra.

Em relação ao engajamento dos participantes, os dados foram coletados em três categorias principais: Abertura, Palestras e Oficinas. Na Abertura, apenas 19,78% dos inscritos compareceram ao evento. Já nas Palestras, a Palestra 1 obteve a maior taxa de comparecimento, com 51,45%, seguida pela Palestra 5 com 50,26%. No entanto, nas Oficinas, a Oficina 1 liderou com 38,25% de participação. Em média, os participantes compareceram a cerca de 47,27% das palestras e 32,97% das oficinas.

Em relação à análise de sentimento, a palavra "Desafiador" foi a mais utilizada para descrever o evento MM2022, representando 8,68% das respostas, seguida por "Ótimo" com 5,49% e "Inspiração" com 4,48%. Finalmente, o Net Promoter Score (NPS) revelou um alto nível de satisfação entre os participantes. Dos participantes, 91,4% foram classificados como promotores do evento, enquanto 4,4% foram neutros e apenas 2,6% foram detratores.

A avaliação do evento revelou insights significativos sobre a percepção dos participantes. As categorias "Interessante", "Criativo" e "Engajador" emergiram como as mais populares, com 564, 537 e 446 votos, respectivamente. Isso indica uma experiência predominantemente positiva e envolvente para a maioria dos participantes. Além disso, "Útil" e "Divertido" também receberam avaliações positivas, com 405 e 333 votos, sugerindo que o evento foi percebido como valioso em termos de conteúdo e agradável em sua execução.

Por outro lado, as categorias "Neutra", "Chato", "Confuso" e "Assustador" tiveram contagens mais baixas, destacando uma minoria que teve uma experiência neutra ou negativa. As categorias "Legal" e "Aceitável" apresentaram avaliações intermediárias, com 286 e 95 votos, indicando uma percepção geralmente positiva, porém menos intensa que as categorias principais.

Apesar das avaliações majoritariamente positivas, é essencial prestar atenção aos feedbacks específicos dos participantes com avaliações neutras ou negativas. Essas informações podem fornecer insights valiosos para identificar áreas de melhoria e fazer ajustes em eventos futuros, buscando sempre maximizar o valor e a satisfação do público-alvo.

Dos 319 indicados pelos parceiros, 123 se inscreveram, resultando em uma taxa de conversão de 38,56%. Embora essa taxa seja alta, as inscrições provenientes das indicações representam apenas 16,3% do total de 754 inscrições. Isso indica um potencial para otimizar as indicações dos parceiros e aumentar as inscrições totais. Estratégias de incentivo e treinamento para os parceiros podem ser eficazes para impulsionar o volume e a eficácia das indicações, maximizando o impacto das parcerias no crescimento do serviço ou programa.