

Instalação de Biblioteca

2

```
pip install TextBlob
```

Importação de Bibliotecas

4

```
from pyspark.sql import SparkSession
from pyspark.sql.functions import col, count, month, year, to_date, trim, lower, mean, to_timestamp, split, when
from pyspark.sql.types import TimestampType, IntegerType
from pyspark.sql import DataFrame
import matplotlib.pyplot as plt
from textblob import TextBlob
import pandas as pd
import numpy as np
import seaborn as sns
import itertools
```


Criação de sessão Spark

6

```
spark = SparkSession.builder.appName("Análise de Dados PDL").getOrCreate()
```

Leitura dos Dados

```
df = spark.read.option("sep", ",") \  
    .option("quote", "\"") \  
    .csv('/FileStore/tables/AvaliacaoDeReacaoPDL.csv', header=True, inferSchema=True)
```

►  df: pyspark.sql.connect.dataframe.DataFrame = [Carimbo de data/hora: string, Como você avalia a organização do PDL como um todo? [O evento foi baseado na necessidade real das(os) líderes e de suas(seus) substitutos.]: string ... 27 more fields]

Renomeação das Colunas

```


df_renomeado = df \
    .withColumnRenamed('Carimbo de data/hora', 'DataHora') \
    .withColumnRenamed('Como você avalia a organização do PDL como um todo? [O evento foi baseado na necessidade real das(os) líderes e de suas(seus) substitutos.]', 'Avaliação_Necessidade') \
    .withColumnRenamed('Como você avalia a organização do PDL como um todo? [A comunicação foi assertiva e me auxiliou na adesão ao PDL]', 'Avaliação_Comunicação') \
    .withColumnRenamed('Como você avalia a organização do PDL como um todo? [As dinâmicas foram bem elaboradas e conduzidas]', 'Avaliação_Dinâmicas') \
    .withColumnRenamed('Como você avalia a organização do PDL como um todo? [Senti que me desenvolvi enquanto liderança]', 'Avaliação_Desenvolvimento') \
    .withColumnRenamed('Como você avalia a organização do PDL como um todo? [Estou satisfeita(o) com a experiência geral do treinamento.]', 'Avaliação_Satisfação') \
    .withColumnRenamed('Algum comentário sobre a organização geral do PDL?', 'Comentário_Organização') \
    .withColumnRenamed('Experiência com a Metodologia e com as(os) Instrutoras(es) da Fornecedora [Os conteúdos abordados foram relevantes para os meus objetivos de aprendizagem.]', 'Experiência_Conteúdos_Relevantes') \
    .withColumnRenamed('Experiência com a Metodologia e com as(os) Instrutoras(es) da Fornecedora [O material do curso foi objetivo e bem organizado.]', 'Experiência_Material_Objetivo') \
    .withColumnRenamed('Experiência com a Metodologia e com as(os) Instrutoras(es) da Fornecedora [As informações apresentadas são aplicáveis no meu ambiente de trabalho.]', 'Experiência_Informações_Aplicáveis') \
    .withColumnRenamed('Experiência com a Metodologia e com as(os) Instrutoras(es) da Fornecedora [Há conteúdos que considero imediatamente úteis.]', 'Experiência_Conteúdos_Uteis') \
    .withColumnRenamed('Experiência com a Metodologia e com as(os) Instrutoras(es) da Fornecedora [Os métodos de ensino foram eficazes para manter o engajamento.]', 'Experiência_Métodos_Engajamento') \
    .withColumnRenamed('Experiência com a Metodologia e com as(os) Instrutoras(es) da Fornecedora [O(a) instrutor(a) demonstrou conhecimento profundo do assunto.]', 'Experiência_Conhecimento_Profundo') \
    .withColumnRenamed('Experiência com a Metodologia e com as(os) Instrutoras(es) da Fornecedora [O(a) instrutor(a) foi acessível e encorajou perguntas e discussões.]', 'Experiência_Acessibilidade') \
    .withColumnRenamed('Algum comentário sobre a Fornecedora ou sobre os(as) instrutores(as)?', 'Comentário_Fornecedora') \
    .withColumnRenamed('Experiência com o espaço coworking [As instalações do espaço proporcionaram um ambiente confortável e adequado para as atividades realizadas.]', 'Experiência_Espaço_Confortável') \
    .withColumnRenamed('Experiência com o espaço coworking [Os equipamentos e recursos disponíveis atenderam às necessidades do PDL]', 'Experiência_Equipamentos') \
    .withColumnRenamed('Experiência com o espaço coworking [Acessível e fácil de localizar/navegar para todas(os) as(os) participantes.]', 'Experiência_Acessibilidade_Coworking') \
    .withColumnRenamed('Experiência com o espaço coworking [O lanche foi adequado para o momento]', 'Experiência_Lanche_Adequado') \
    .withColumnRenamed('Experiência com o espaço coworking [Minha experiência geral com o espaço coworking foi positiva.]', 'Experiência_Geral_Coworking') \
    .withColumnRenamed('Algum comentário sobre o espaço da coworking?', 'Comentário_Espaço_Coworking') \
    .withColumnRenamed('Evento de Encerramento do PDL (4/12) [O Rancho foi adequado para o evento de encerramento.]', 'Evento_Rancho_Adequado') \
    .withColumnRenamed('Evento de Encerramento do PDL (4/12) [A comunicação prévia e durante o evento foi clara e eficaz.]', 'Evento_Comunicação_Eficaz') \
    .withColumnRenamed('Evento de Encerramento do PDL (4/12) [Considero a organização do evento adequada para o

```

```

propósito.],'Evento_Organização_Adequada') \
.withColumnRenamed('Evento de Encerramento do PDL (4/12) [As atividades programadas foram envolventes e
apropriadas para o evento de encerramento.],'Evento_Atividades_Envolventes') \
.withColumnRenamed('Evento de Encerramento do PDL (4/12) [O evento ajudou a reforçar o propósito do Programa de
Desenvolvimento de Liderança.],'Evento_Reforço_Propósito') \
.withColumnRenamed('Evento de Encerramento do PDL (4/12) [A alimentação fornecida no Rancho (café da manhã,
almoço e lanche) foi adequada para o momento.],'Evento_Alimentação_Adequada') \
.withColumnRenamed('Evento de Encerramento do PDL (4/12) [Estou satisfeita(o) com a experiência geral do evento
de encerramento.],'Evento_Satisfação_Geral') \
.withColumnRenamed('Gostaria de registrar algum comentário geral sobre o PDL 2024?',
'Comentário_Geral_PDL_2024')

```

►  df_renomeado: pyspark.sql.connect.dataframe.DataFrame = [DataHora: string, Avaliação_Necessidade: string ... 27 more fields]

Tratamento dos Dados

12

```

def tratar_dados(df):

    # Preenche valores nulos
    for column in df.columns:
        if dict(df.dtypes)[column] == 'int':
            media = df.select(mean(col(column))).first()[0]
            df = df.fillna({column: media})
        elif dict(df.dtypes)[column] == 'string':
            df = df.fillna({column: 'N/A'})

    # Remove espaços em branco
    for column in df.columns:
        if dict(df.dtypes)[column] == 'string':
            df = df.withColumn(column, trim(col(column)))

    # Normaliza texto
    for column in df.columns:
        if dict(df.dtypes)[column] == 'string':
            df = df.withColumn(column, lower(col(column)))

    return df

```

13

```
# Aplicando o tratamento
df_tratado = tratar_dados(df_renomeado)
```

► df_tratado: pyspark.sql.connect.dataframe.DataFrame = [DataHora: string, Avaliação_Necessidade: string ... 27 more fields]

14

```
# Trata coluna de data
df_tratado = df_tratado.withColumn("DataHora",
    when(col("DataHora").rlike(r"^\d{1}/\d{1,2}/\d{4} \d{1,2}:\d{2}:\d{2}$"),
        to_timestamp(trim(col("DataHora")), "d/M/yyyy H:mm:ss"))
    .when(col("DataHora").rlike(r"^\d{2}/\d{1,2}/\d{4} \d{1,2}:\d{2}:\d{2}$"),
        to_timestamp(trim(col("DataHora")), "dd/M/yyyy H:mm:ss"))
    .when(col("DataHora").rlike(r"^\d{1,2}/\d{2}/\d{4} \d{1,2}:\d{2}:\d{2}$"),
        to_timestamp(trim(col("DataHora")), "d/MM/yyyy H:mm:ss"))
    .otherwise(to_timestamp(trim(col("DataHora")), "dd/MM/yyyy HH:mm:ss"))
)
```

► df_tratado: pyspark.sql.connect.dataframe.DataFrame = [DataHora: timestamp, Avaliação_Necessidade: string ... 27 more fields]

15

```
df_tratado.printSchema()
```

```
-- Experiência_Informações_Aplicáveis: string (nullable = false)
-- Experiência_Conteúdos_Uteis: string (nullable = false)
-- Experiência_Métodos_Engajamento: string (nullable = false)
-- Experiência_Conhecimento_Profundo: string (nullable = false)
-- Experiência_Acessibilidade: string (nullable = false)
-- Comentário_Fornecedora: string (nullable = false)
-- Experiência_Espaço_Confortável: string (nullable = false)
```

```
|-- Evento_Alimentacao_Adequada: string (nullable = false)
|-- Evento_Satisfacao_Geral: string (nullable = false)
|-- Comentario_Geral_PDL_2024: string (nullable = false)
```

Análises Iniciais

17

```
df_tratado.show(5)
```

18

```
#Calcula a Contagem de Feedbacks por mês
feedbacks_por_mes = df_tratado.groupBy(year(col("DataHora")).alias("Ano"), month(col("DataHora")).alias("Mês")) \
    .agg(count("*").alias("Contagem_Feedbacks")) \
    .orderBy("Ano", "Mês")
```

►  feedbacks_por_mes: pyspark.sql.connect.dataframe.DataFrame = [Ano: integer, Mês: integer ... 1 more field]

19

```
# Apresenta o resultado do Cálculo de Feedbacks ao Longo dos Meses
print("Contagem de Feedbacks ao Longo do Mês:")
feedbacks_por_mes.show(truncate=False)
```

Contagem de Feedbacks ao Longo do Mês:

```
+---+---+-----+
|Ano |Mês|Contagem_Feedbacks|
+---+---+-----+
|2024|12 |36                |
+---+---+-----+
```

20

```
# Analisa Comentários sobre a Organização
comentarios = df_renomeado.select('Comentário_Organização').filter(col('Comentário_Organização') != 'Não informado').collect()
for comentario in comentarios:
    print(f'Comentário: {comentario["Comentário_Organização"]}')

```

o das pessoas e fortalecimento da liderança. Que possamos realmente colocar em prática todos os ensinamentos compartilhados. Muito obrigada a todos! Foi incrível viver essa experiência com vocês.

Comentário: Foi difícil conciliar com as demais atividades

Comentário: A organização do PDL foi sensacional. Tudo foi bem planejado e executado, o que garantiu que o evento transcorresse de forma eficiente e agradável para todos os participantes. Parabéns a toda a equipe da GGE pelo excelente evento.

Comentário: Os eventos fora das paredes do Postalís trazem novas nuances, especialmente o estreitamento e humanização das relações. Fazer os treinamentos e ter a praça de aprendizagem fecham um ciclo de evolução, concatenando a teoria com a prática trazendo exemplos do cotidiano. Eu acredito que uma das evoluções desse treinamento poderia ser um grupo de discussões das nossas vulnerabilidades como gestores. Se estou passando alguma dificuldade com o feedbacks de um trabalhador da minha área que tem X comportamentos ou Y dificuldades, ter a maturidade e confiança de compartilhar isso com outros gerentes/coordenadores e, no compartilhamento de formas de abordagem, evoluir profissionalmente como líder.

Comentário: Essa última data ficou muito complicada, período de folgas. Sugiro que não ocorra no mês de dezembro. Temos folgas de abono e o pagamento é antecipado para o dia 20/12.

Comentário: Infelizmente perdi 2 encontros presenciais por motivo de férias. Poderia ser oferecida participação em outra turma

Comentário: O programa foi muito legal. Uma observação, é conseguirmos fazer ele em uma escala mensal, para não sobrecarregar as atividades de rotina das áreas. De igual forma, montar uma rotina, para que seja um programa contínuo em todos os anos. Parabéns aos envolvidos, saímos do chumbo para o ouro muito rápido.

Comentário: A equipe estava empenhada e tornou o projeto dinâmico

Comentário: Excelente organização!

Comentário: O evento foi muito organizado. Parabéns a GGE por toda dedicação e comprometimento.

21

```
# Analisa Comentários sobre a Fornecedora e Instrutores
comentario_fornecedora = df_renomeado.select('Comentário_Fornecedora').filter(col('Comentário_Fornecedora') != 'Não informado').collect()
for comentario_fornecedora in comentario_fornecedora:
    print(f'comentario_fornecedora: {comentario_fornecedora["Comentário_Fornecedora"]}')

```

comentario_fornecedora: O conteúdo era muito extenso, sempre após o intervalo notava uma queda de participação e um cansaço geral. Especialmente a terceira aula, achei muito cansativa e pouco proveitosa.

comentario_fornecedora: Gostei do conteúdo e os instrutores foram muito acessíveis.

comentario_fornecedora: Os instrutores da Fornecedora foram sensacionais, conduziram o treinamento de uma forma simples e muito eficiente. Parabéns a toda a equipe.

comentario_fornecedora: Adriana e Wallace dialogaram muito bem com o grupo. Têm muita experiência e tiveram muita paciência e empatia com os gerentes.

comentario_fornecedora: Já conhecia um pouco da empresa e o treinamento confirmou a visão de uma boa empresa para treinamentos corporativos. Ótimo fornecedor.

comentario_fornecedora: O professor foi didático e assertivo nas suas colocações. Além disso, atuou de forma excelente ao criar um espaço confortável para os participantes se exporem.

comentario_fornecedora: Palestrantes muito bons!

comentario_fornecedora: Ótima abordagem e comunicação

comentario_fornecedora: Acredito que uma abordagem mais fechada em nossa realidade seria mais aproveitado. Teve muitas abordagens que não temos como aplicar no Postalis.

comentario_fornecedora: Na minha percepção, poderíamos ter tido somente um único instrutor. Eu senti uma diferença de didática entre eles, o que me causou um pouco de desânimo nos últimos encontros. Informo que não estou criticando a Adriana, mas pelo fato de terem diferentes

comentario_fornecedora: Excelentes instrutores.

comentario_fornecedora: Instrutores capacitados e trouxeram informações interessantes para aplicação no dia a dia.

comentario_fornecedora: O instrutor despertou a atenção de todos, mas alguns alunos ainda se sentem desconfortáveis ao

22

```
# Analisa os Comentários do Espaço Coworking
comentario_coworking = df_renomeado.select('Comentário_Espaço_Coworking').filter(col('Comentário_Espaço_Coworking')
!= 'Não informado').collect()
for comentario_coworking in comentario_coworking:
    print(f'comentario_coworking: {comentario_coworking["Comentário_Espaço_Coworking"]}')

```

comentario_coworking: Na coworking ambiente muito quente. Temos esse espaço dentro da instituição, e acho desnecessário pagar por outro. Trabalhamos com adultos, e é importante que todos compreendam que o horário de treinamento deve ser respeitado.

comentario_coworking: Acessível. As cadeiras poderiam ser mais confortáveis.

comentario_coworking: O espaço é muito bom, perto do Postalis o que facilita muito. Só tenho uma crítica quanto ao espaço, achei o banheiro com cheiro muito forte de urina.

comentario_coworking: Espaço acolhedor e muito bem escolhido e pela GGE, me fez lembrar os tempos de faculdade. Sensação maravilhosa.

comentario_coworking: Foi incrível estar naquela sala. Sentimento de sala de aula moderna e discussões maduras, sem perder a criança interior que não tem medo de errar, trazendo leveza para os assuntos abordados.

comentario_coworking: Cadeiras desconfortáveis e banheiros sujo no ultimo encontro.

comentario_coworking: Precisa melhorar o café e a água, inclusive melhorando a disponibilidade. Sempre oferecido em pouca quantidade e não parecia ser uma prioridade para o local cuidar do suporte ao cliente.

comentario_coworking: .

comentario_coworking: O banheiro do espaço é bem ruim .

comentario_coworking: Só o fato de realizarmos o treinamento fora do ambiente de trabalho já proporciona uma imersão significativa. A coworking entregou um bom espaço, instalações e equipamentos adequados.

comentario_coworking: Estaço bom.

comentario_coworking: Um bom local acessível e próximo ao Postalis.

23


```
# Analisa os Comentários Gerais sobre o PDL
comentarios_gerais = df_renomeado.select('Comentário_Geral_PDL_2024').filter(col('Comentário_Geral_PDL_2024') !=
'Não informado').collect()
for comentarios_gerais in comentarios_gerais:
    print(f'comentarios_gerais: {comentarios_gerais["Comentário_Geral_PDL_2024"]}')

```

comentarios_gerais: Achei legal a experiência, mas ao meu ver serviu mais para inteiração que para reforçar o treinamento, o que também acho interessante. Único ponto que observei foi que nem todos os funcionários tinham interesse em engajar em algumas atividades, especialme

comentarios_gerais: Memorável e muito agregador

comentarios_gerais: O evento foi maravilhoso, e a GGE está de parabéns pela organização. Gostaria apenas de compartilhar um a observação: as atividades do evento coincidiram bastante com outras demandas que ocupavam boa parte das manhãs ou tardes, como o Planejamento Estratég

comentarios_gerais: Só tenho a agradecer ao Postalis e à GGE pelo excelente treinamento. Saio do treinamento muito capacitado e cheio de insights, com uma compreensão mais profunda dos processos e ferramentas de liderança, o que me permitirá aplicar os novos conhecimentos de

comentarios_gerais: Referente ao rancho posso opinar sobre o local que já conheço, mas o evento infelizmente não pude participar por questões de saúde.

comentarios_gerais: Acho que pode ter um fechamento com uma análise das dinâmicas. Refeições estavam incríveis, o lugar era maravilhoso e senti o carinho e cuidado da GGE.

comentarios_gerais: Parabenizar a equipe da GGE.

comentarios_gerais: Achei o ambiente bastante quente e bastante mosquito (apesar de sermos abusados de levar repelente). O banheiro de onde teve o café da manhã era extremamente quente sem ventilação.

comentarios_gerais: Não consegui participar por questões de demandas internas (fechamento dos balancetes do mês de out/2024)

comentarios_gerais: Foi um ótimo programa. Um verdadeiro choque de realidade com o que estava acostumado na empresa. Se tiver continuidade, pode ser o ponto de inflexão que mudará positivamente a cultura da empresa.

Definição de Categorias

```

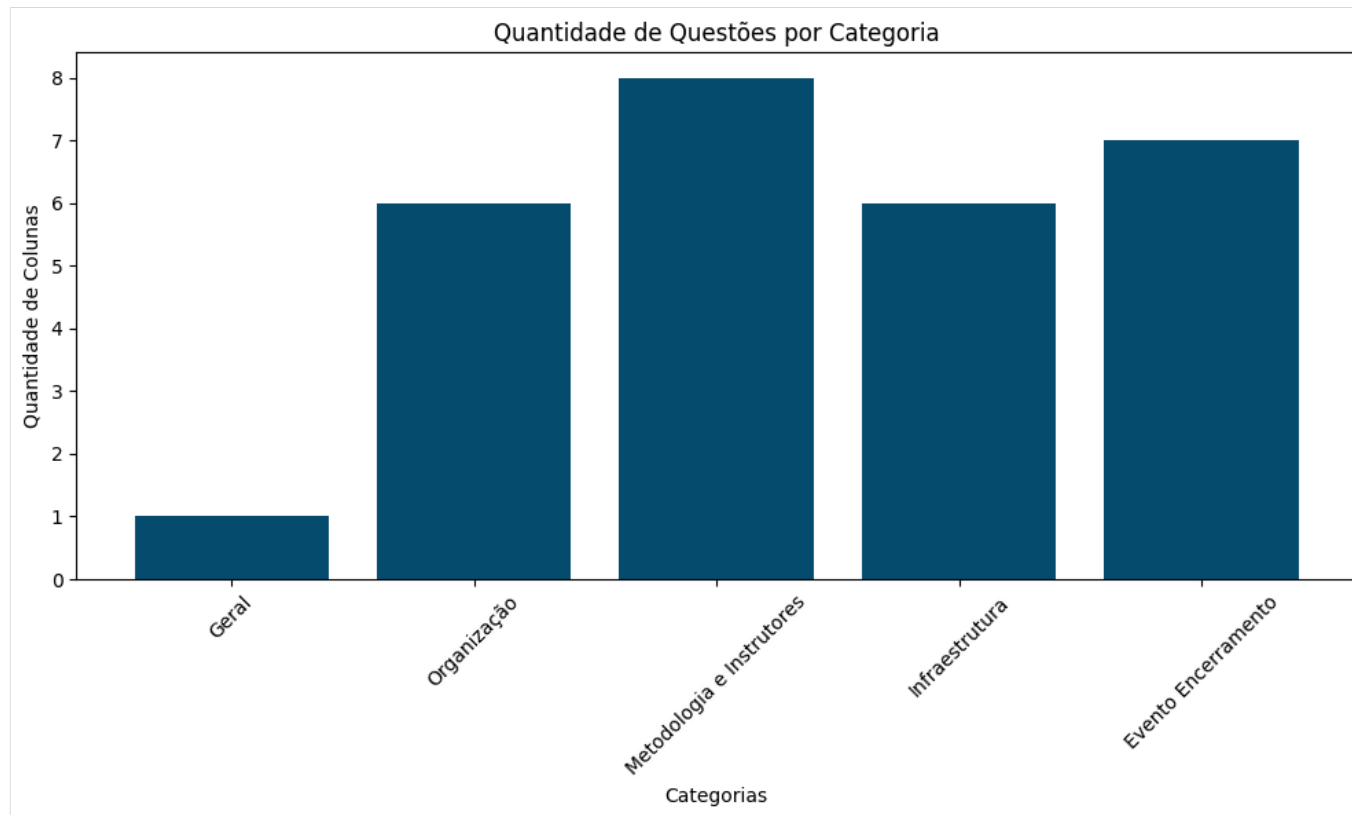
# Define as categorias com suas respectivas colunas
categorias = {
  "Geral": ["Comentário_Geral_PDL_2024"],
  "Organização": [
    "Avaliação_Necessidade",
    "Avaliação_Comunicação",
    "Avaliação_Dinâmicas",
    "Avaliação_Desenvolvimento",
    "Avaliação_Satisfação",
    "Comentário_Organização"
  ],
  "Metodologia e Instrutores": [
    "Experiência_Conteúdos_Relevantes",
    "Experiência_Material_Objetivo",
    "Experiência_Informações_Aplicáveis",
    "Experiência_Conteúdos_Uteis",
    "Experiência_Métodos_Engajamento",
    "Experiência_Conhecimento_Profundo",
    "Experiência_Acessibilidade",
    "Comentário_Fornecedora"
  ],
  "Infraestrutura": [
    "Experiência_Espaço_Confortável",
    "Experiência_Equipamentos",
    "Experiência_Acessibilidade_Coworking",
    "Experiência_Lanche_Adequado",
    "Experiência_Geral_Coworking",
    "Comentário_Espaço_Coworking"
  ],
  "Evento Encerramento": [
    "Evento_Rancho_Adequado",
    "Evento_Comunicação_Eficaz",
    "Evento_Organização_Adequada",
    "Evento_Atividades_Envolventes",
    "Evento_Reforço_Propósito",
    "Evento_Alimentação_Adequada",
    "Evento_Satisfação_Geral"
  ]
}

```

```
# Conta o número de colunas em cada categoria
quantidade_colunas = {categoria: len(colunas) for categoria, colunas in categorias.items()}
```

27


```
# Cria o gráfico
plt.figure(figsize=(10, 6))
plt.bar(quantidade_colunas.keys(), quantidade_colunas.values(), color='#084d6e')
plt.title('Quantidade de Questões por Categoria')
plt.xlabel('Categorias')
plt.ylabel('Quantidade de Colunas')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



Mapeando as respostas para a escala Likert

```
likert_columns = [
    "Avaliação_Necessidade", "Avaliação_Comunicação", "Avaliação_Dinâmicas",
    "Avaliação_Deenvolvimento", "Avaliação_Satisfação",
    "Experiência_Conteúdos_Relevantes", "Experiência_Material_Objetivo",
    "Experiência_Informações_Aplicáveis", "Experiência_Conteúdos_Uteis",
    "Experiência_Métodos_Engajamento", "Experiência_Conhecimento_Profundo",
    "Experiência_Acessibilidade", "Experiência_Espaço_Confortável",
    "Experiência_Equipamentos", "Experiência_Acessibilidade_Coworking",
    "Experiência_Lanche_Adequado", "Experiência_Geral_Coworking",
    "Evento_Rancho_Adequado", "Evento_Comunicação_Eficaz",
    "Evento_Organização_Adequada", "Evento_Atividades_Envolventes",
    "Evento_Reforço_Propósito", "Evento_Alimentação_Adequada",
    "Evento_Satisfação_Geral"
]

for column in likert_columns:
    df_tratado = df_tratado.withColumn(column,
        when(col(column) == "concordo totalmente", 4)
        .when(col(column) == "concordo", 3)
        .when(col(column) == "discordo", 2)
        .when(col(column) == "discordo totalmente", 1)
        .otherwise(None)
    )
)
```

►  df_tratado: pyspark.sql.connect.dataframe.DataFrame = [DataHora: timestamp, Avaliação_Necessidade: integer ... 27 more fields]

30

```
# Altera o tipo das colunas para numérico
for column in likert_columns:
    df_tratado = df_tratado.withColumn(column, col(column).cast("integer"))
```

►  df_tratado: pyspark.sql.connect.dataframe.DataFrame = [DataHora: timestamp, Avaliação_Necessidade: integer ... 27 more fields]

31

```
# Exibe o esquema para verificar as alterações
df_tratado.printSchema()
```

```

|-- Experiencia_Conhecimento_Profundo: integer (nullable = true)
|-- Experiencia_Acessibilidade: integer (nullable = true)
|-- Comentário_Fornecedora: string (nullable = false)
|-- Experiencia_Espaço_Confortável: integer (nullable = true)
|-- Experiencia_Equipamentos: integer (nullable = true)
|-- Experiencia_Acessibilidade_Coworking: integer (nullable = true)
|-- Experiencia_Lanche_Adequado: integer (nullable = true)
|-- Experiencia_Geral_Coworking: integer (nullable = true)
|-- Comentário_Espaço_Coworking: string (nullable = false)
|-- Evento_Rancho_Adequado: integer (nullable = true)
|-- Evento_Comunicação_Eficaz: integer (nullable = true)
|-- Evento_Organização_Adequada: integer (nullable = true)
|-- Evento_Atividades_Envolventes: integer (nullable = true)
|-- Evento_Reforço_Propósito: integer (nullable = true)
|-- Evento_Alimentação_Adequada: integer (nullable = true)
|-- Evento_Satisfação_Geral: integer (nullable = true)
|-- Comentário_Geral_PDL_2024: string (nullable = false)

```

32

```

# Exibe as primeiras linhas do DataFrame tratado
df_tratado.show(5)

```

```

4|
|2024-12-05 15:43:05|          3|          4|          n/a|          3|          4|
4|          n/a|          4|          3|          4|          n/a|          4|
4|          4|          4|          4|          4|          n/a|          3|
3|          3|          3|          3|          3|          3|          3|
n/a|          4|          3|          3|          3|          3|          3|
3|          4|          3|          n/a|          3|          4|
|2024-12-05 15:45:53|          4|          4|          3|          4|          4|
4|          n/a|          4|          3|          3|          4|          4|
4|          4|          4|          4|          4|          n/a|          3|
2|          2|          3|          3|          3|          3|          3|
na coworking ambi...|          3|          3|          3|          3|          n/a|
3|          3|          3|          3|          3|          3|          3|
|2024-12-05 15:47:17|          3|          3|          4|          4|          4|          3|
4| apenas um elogio ...|          3|          4|          4|          4|          n/a|
4|          3|          4|          4|          4|          3|          3|
3|          3|          3|          3|          3|          3|          3|
n/a|          3|          3|          4|          3|          3|          3|
3|          4|          4|          n/a|          3|          4|          4|
|2024-12-05 15:47:19|          4|          4|          3|          4|          4|          4|
4|          n/a|          4|          4|          4|          4|          4|

```

Calcular a Média das Avaliações de cada categoria

34

```
media_por_categoria = {}
for categoria, colunas in categorias.items():
    # Calcula a média das colunas da categoria
    media = df_tratado.select([mean(col(column)).alias(column) for column in colunas if column in
df_tratado.columns]).first()

    # Verifica se a média não é None antes de calcular a média geral
    media_categoria = 0
    count_valid = 0
    for column in colunas:
        if column in df_tratado.columns and media[column] is not None:
            media_categoria += media[column]
            count_valid += 1

    # Evita divisão por zero
    if count_valid > 0:
        media_categoria /= count_valid
    else:
        media_categoria = None

    media_por_categoria[categoria] = media_categoria
```

35

```
# Verifica se há valores None antes de gerar o gráfico
for categoria, media in media_por_categoria.items():
    if media is None:
        media_por_categoria[categoria] = 0 # Substitui None por 0
```

36

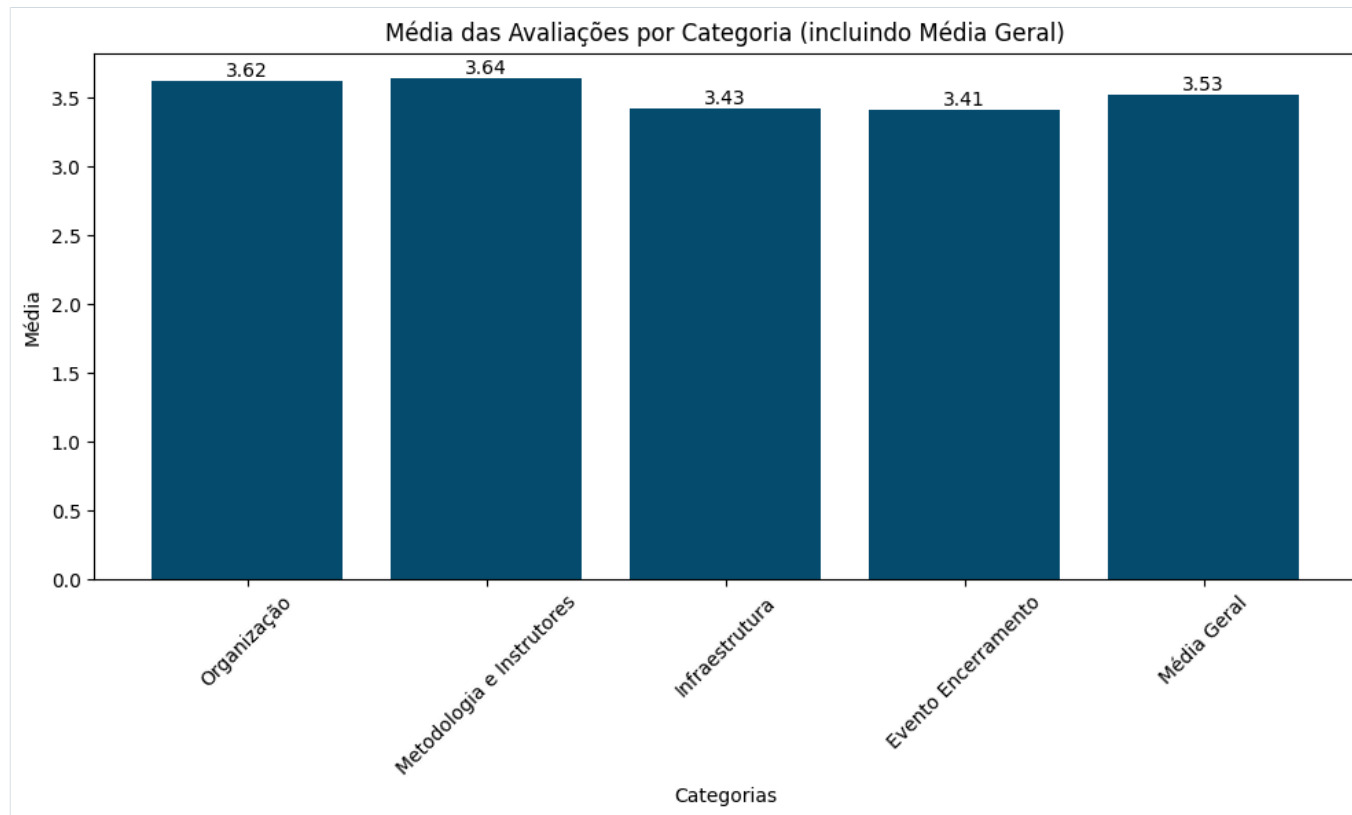
```
# Calcula a média geral
media_geral = sum(media_por_categoria_filtrado.values()) / len(media_por_categoria_filtrado)

# Adiciona a média geral ao dicionário
media_por_categoria_filtrado['Média Geral'] = media_geral

# Gera o gráfico com a média geral
plt.figure(figsize=(10, 6))
bars = plt.bar(media_por_categoria_filtrado.keys(), media_por_categoria_filtrado.values(), color='#084d6e')

# Adiciona os valores em cada barra
for bar in bars:
    yval = bar.get_height()
    plt.text(bar.get_x() + bar.get_width()/2, yval, round(yval, 2), ha='center', va='bottom')

plt.title('Média das Avaliações por Categoria (incluindo Média Geral)')
plt.xlabel('Categorias')
plt.ylabel('Média')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```

Análise de Sentimentos

38

```
def get_sentiment(text):  
    analysis = TextBlob(text)  
    return analysis.sentiment.polarity # Retorna um valor entre -1 (negativo) e 1 (positivo)
```

39

```
# Aplica a análise de sentimentos aos comentários gerais
comentarios_gerais = df_tratado.select("Comentário_Geral_PDL_2024").filter(col("Comentário_Geral_PDL_2024") !=
"N/A").collect()
comentarios_texto = [comentarios_gerais["Comentário_Geral_PDL_2024"] for comentarios_gerais in comentarios_gerais]
```

40

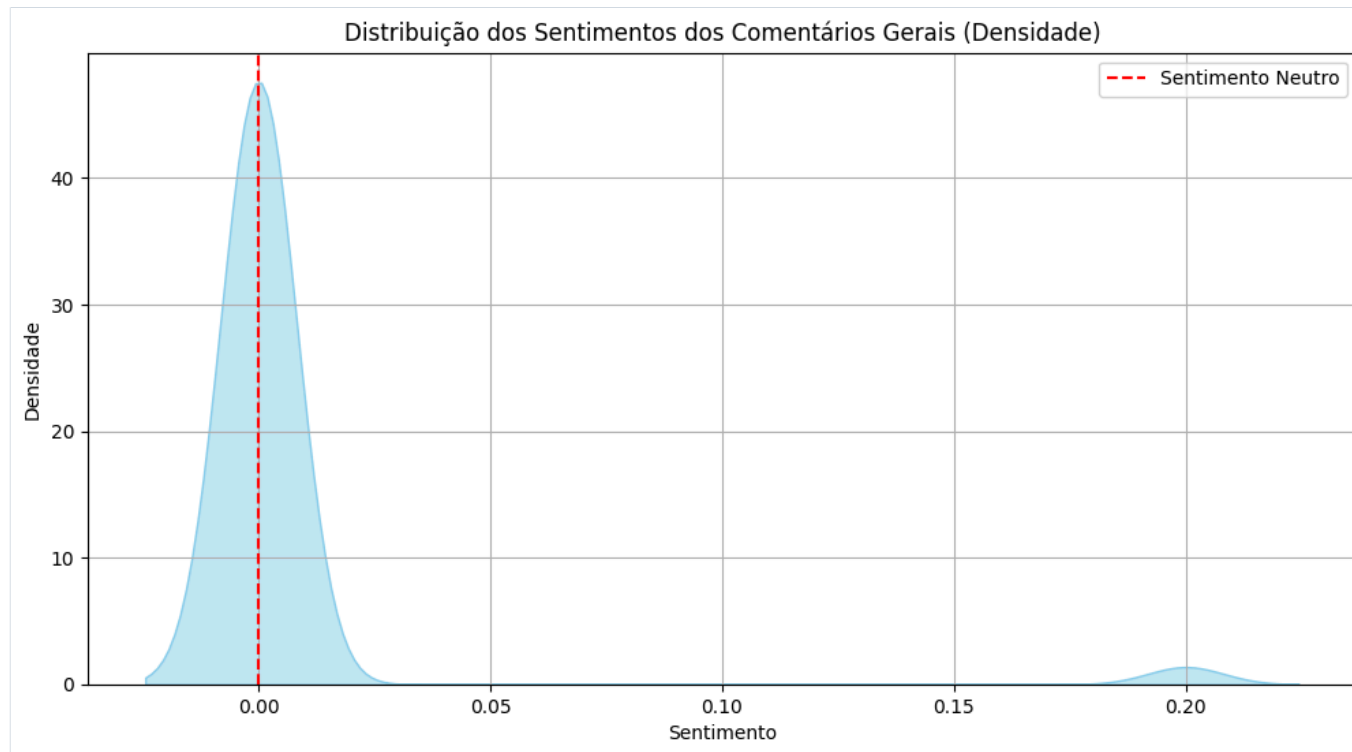
```
# Calcula o sentimento para cada comentário
sentimentos = [get_sentiment(text) for text in comentarios_texto]
```

41

```
# Cria DataFrame para os sentimentos
sentimentos_df = pd.DataFrame(sentimentos, columns=['Sentimento'])
```

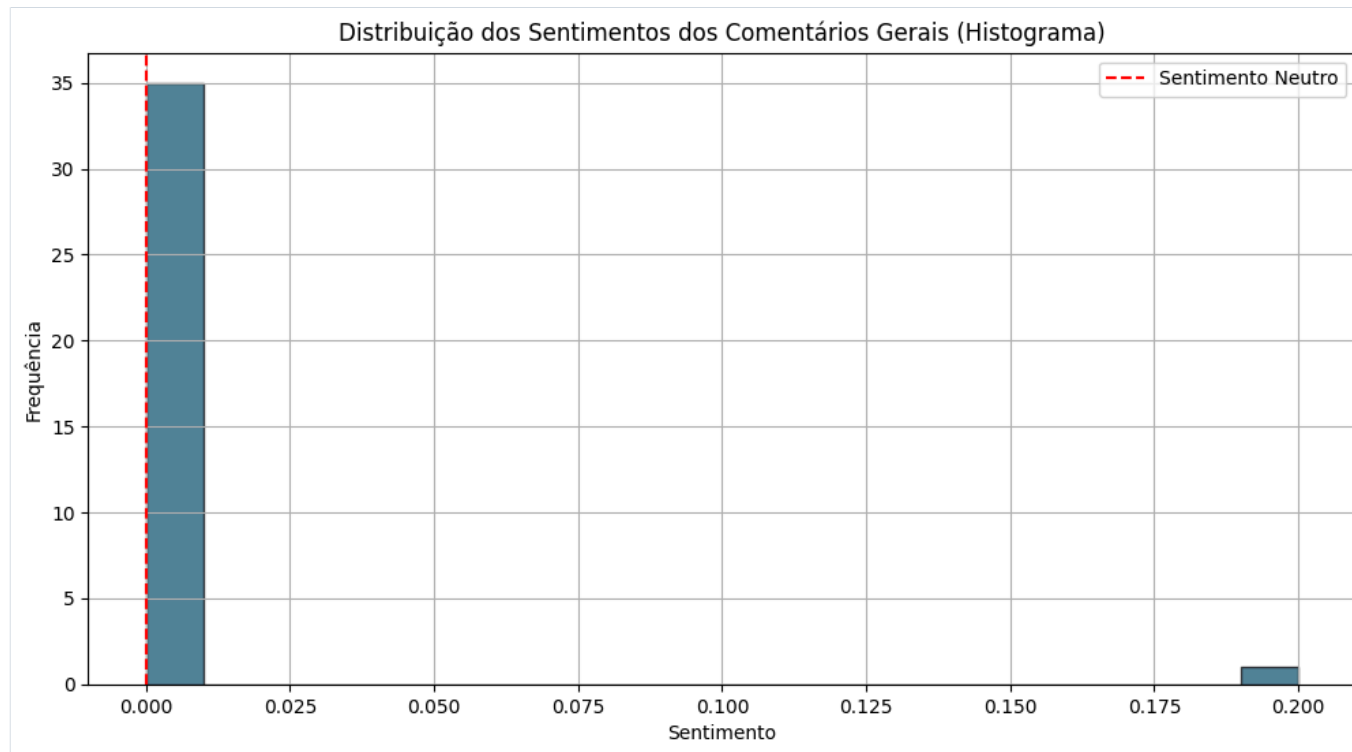
42

```
# Plota Gráfico de Densidade
plt.figure(figsize=(12, 6))
sns.kdeplot(sentimentos_df['Sentimento'], fill=True, color='skyblue', alpha=0.5, bw_adjust=0.5)
plt.axvline(x=0, color='red', linestyle='--', label='Sentimento Neutro')
plt.title('Distribuição dos Sentimentos dos Comentários Gerais (Densidade)')
plt.xlabel('Sentimento')
plt.ylabel('Densidade')
plt.legend()
plt.grid()
plt.show()
```



43

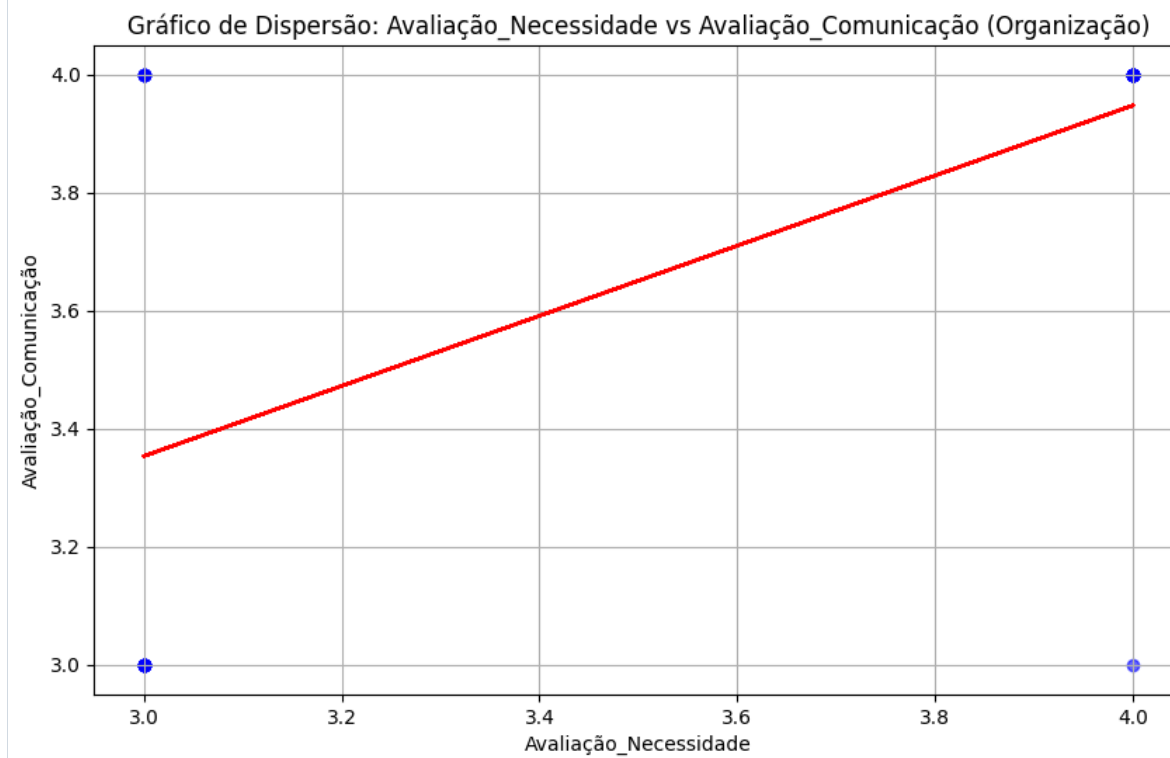
```
# Plota Gráfico de Histograma
plt.figure(figsize=(12, 6))
plt.hist(sentimentos_df['Sentimento'], bins=20, color='#084d6e', edgecolor='black', alpha=0.7)
plt.axvline(x=0, color='red', linestyle='--', label='Sentimento Neutro')
plt.title('Distribuição dos Sentimentos dos Comentários Gerais (Histograma)')
plt.xlabel('Sentimento')
plt.ylabel('Frequência')
plt.legend()
plt.grid()
plt.show()
```

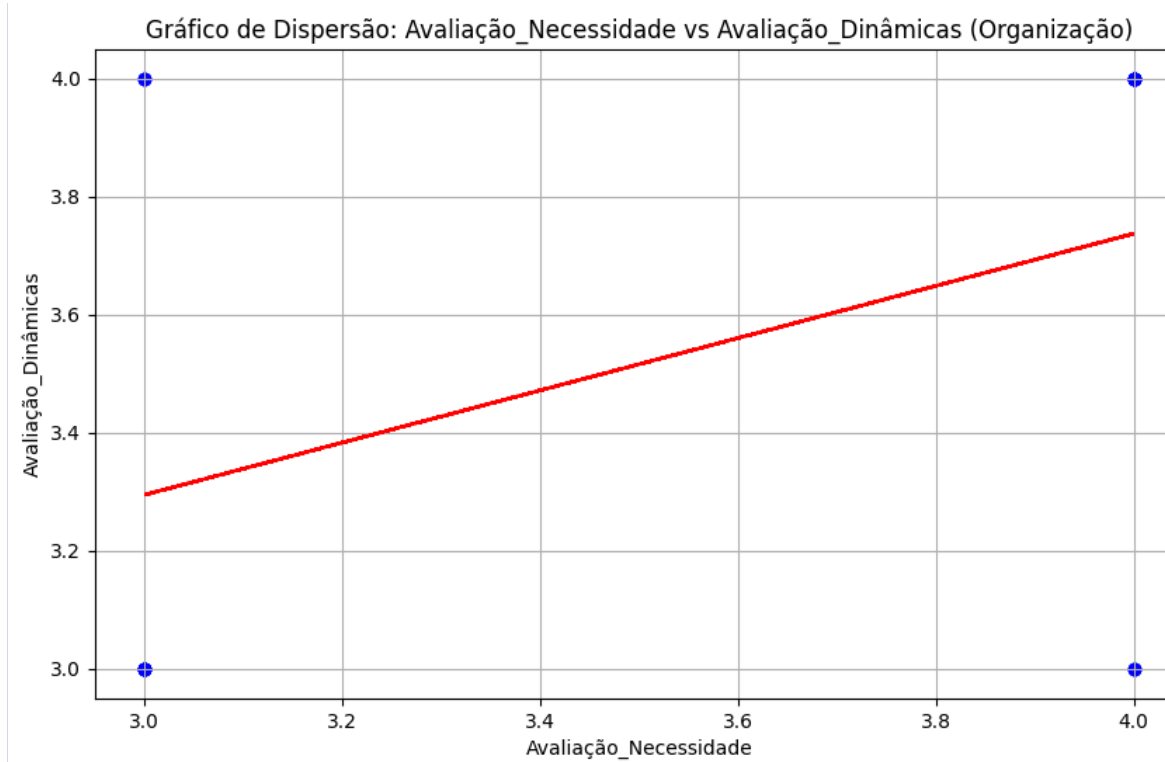


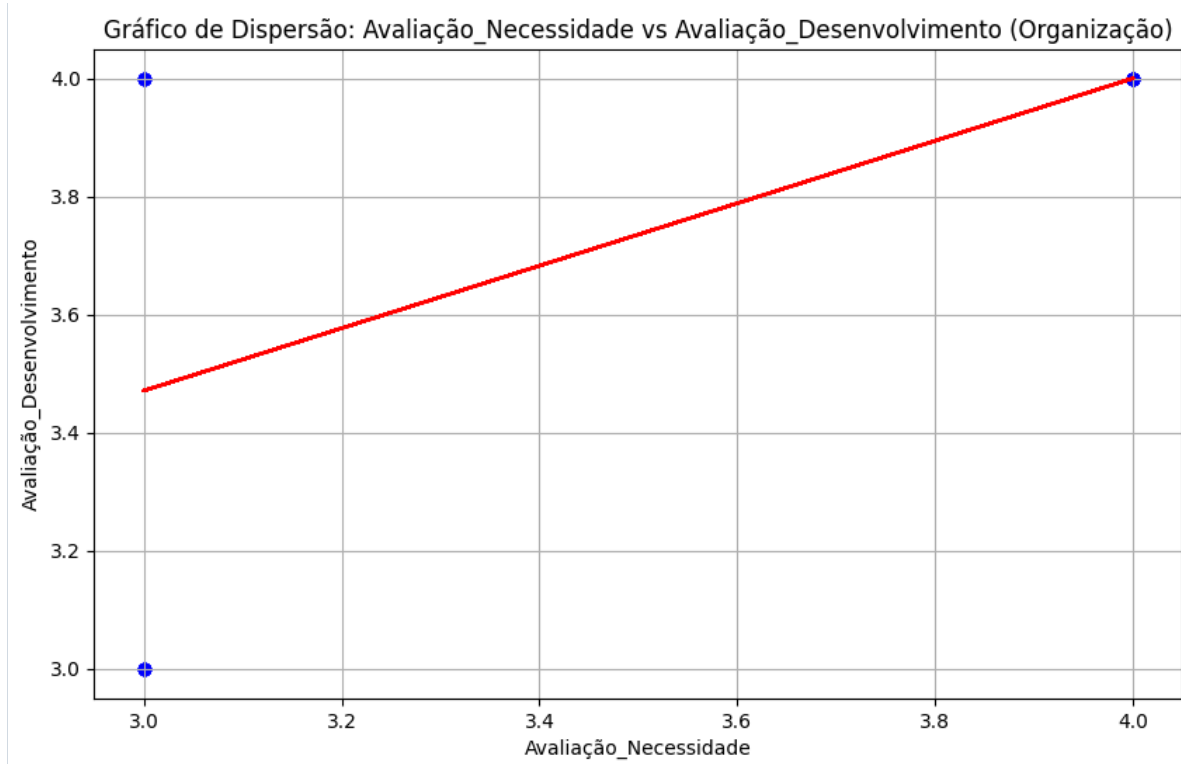
Gráficos de Dispersão

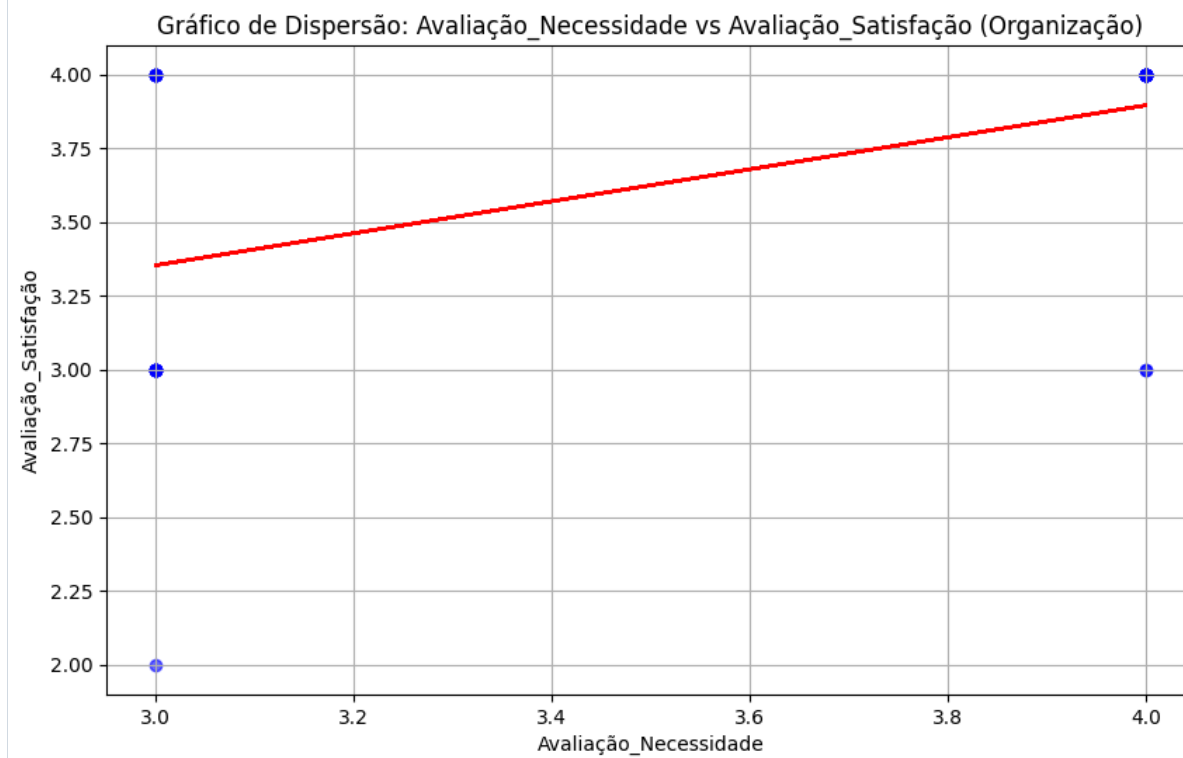
45

► df_scatter: pyspark.sql.connect.dataframe.DataFrame = [Evento_Alimentação_Adequada: double, Evento_Satisfação_Geral: double]



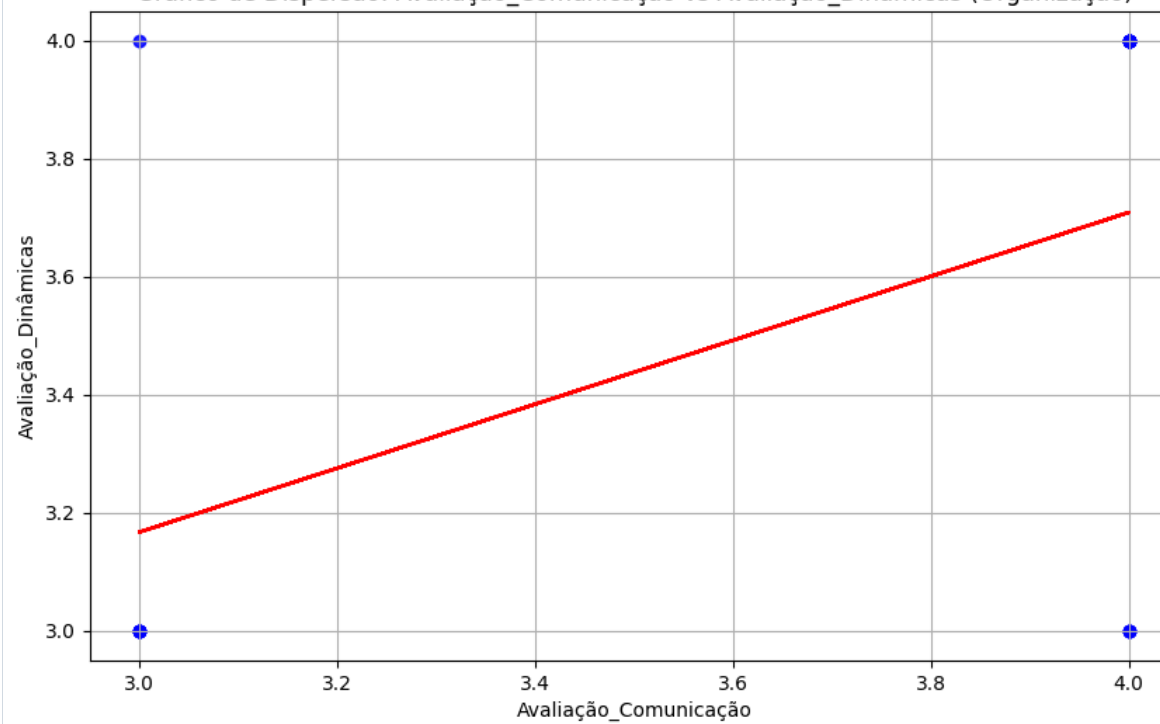


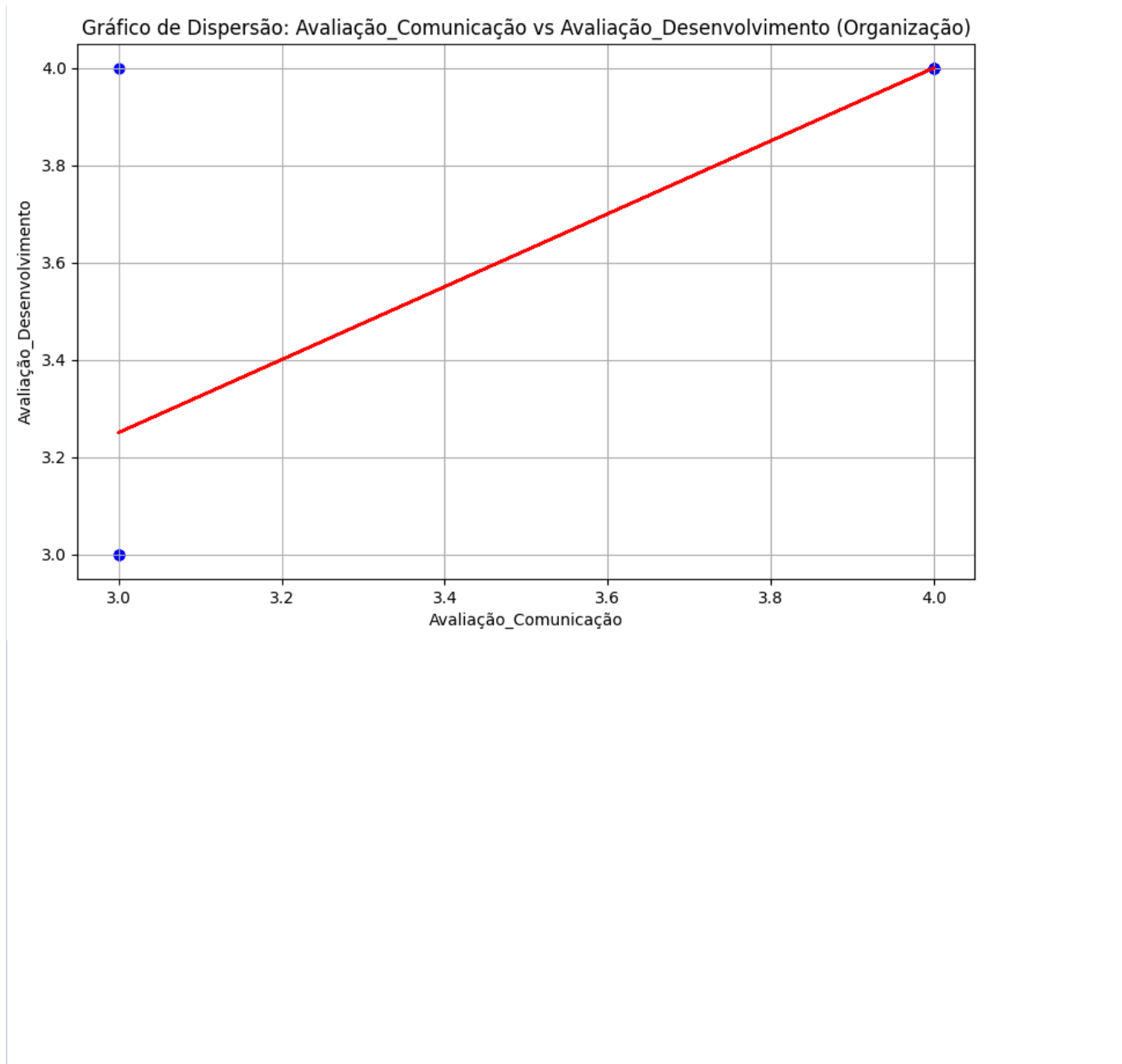


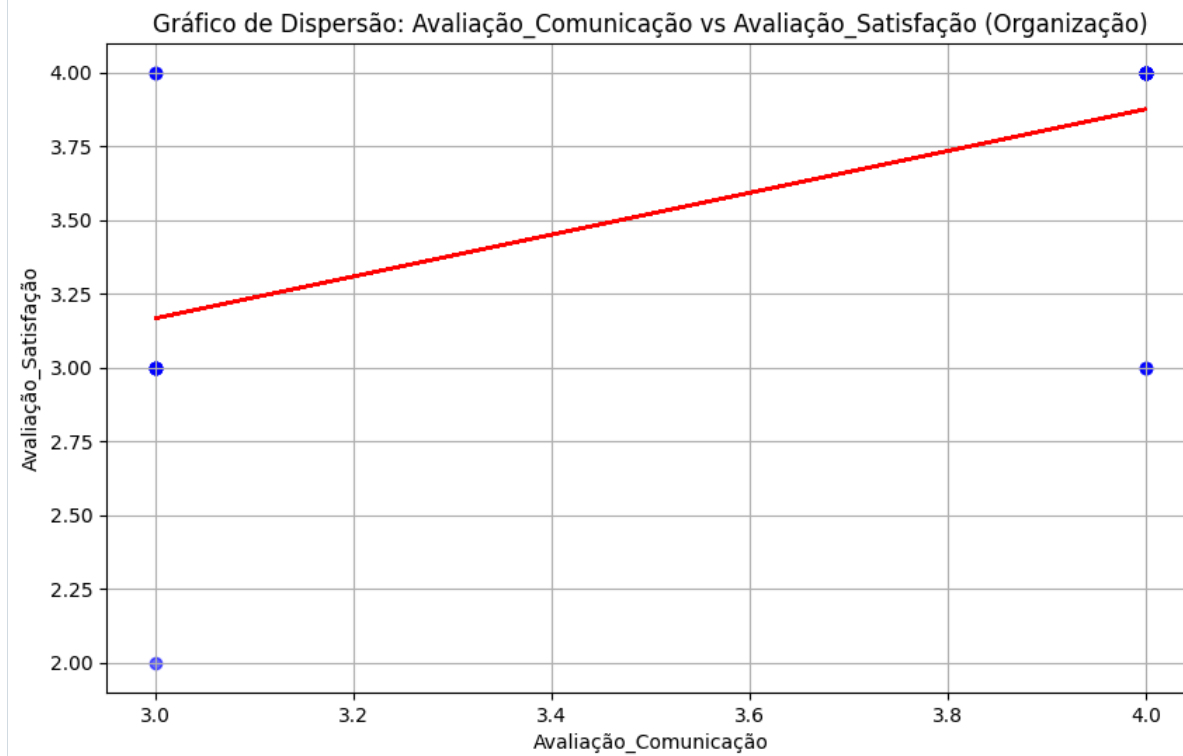


Não há dados suficientes para plotar Avaliação_Necessidade vs Comentário_Organização (Organização)

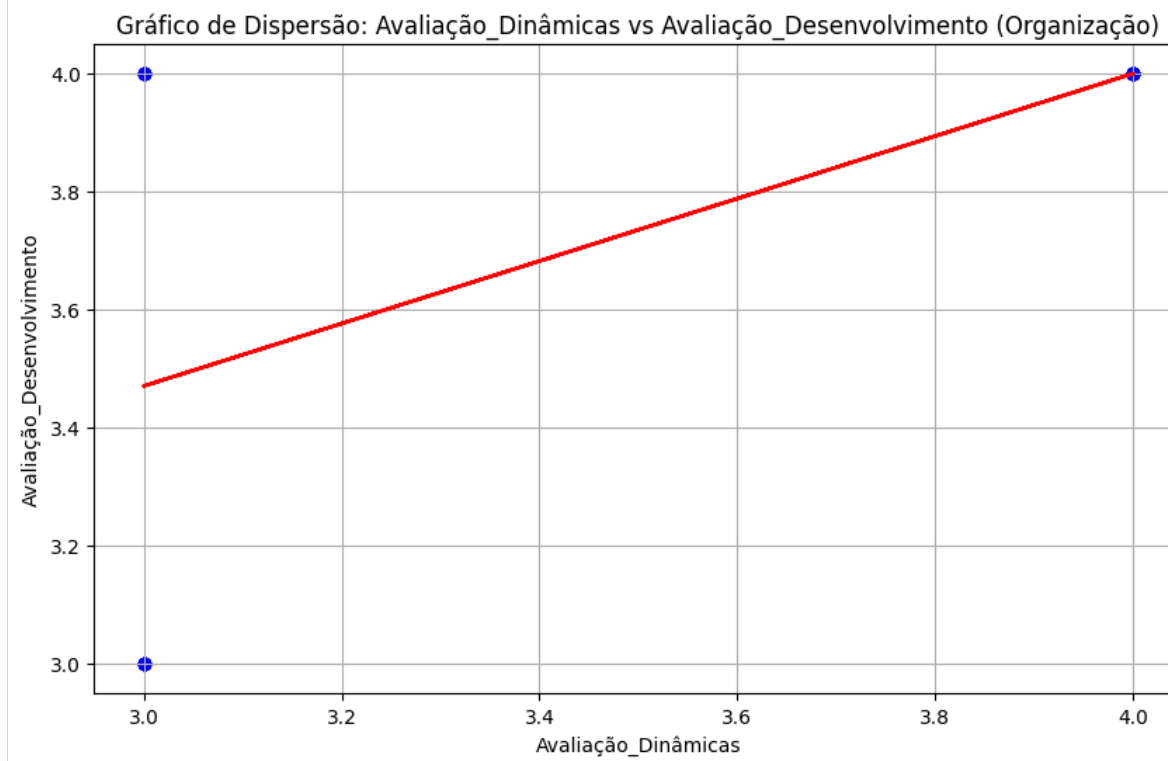
Gráfico de Dispersão: Avaliação_Comunicação vs Avaliação_Dinâmicas (Organização)

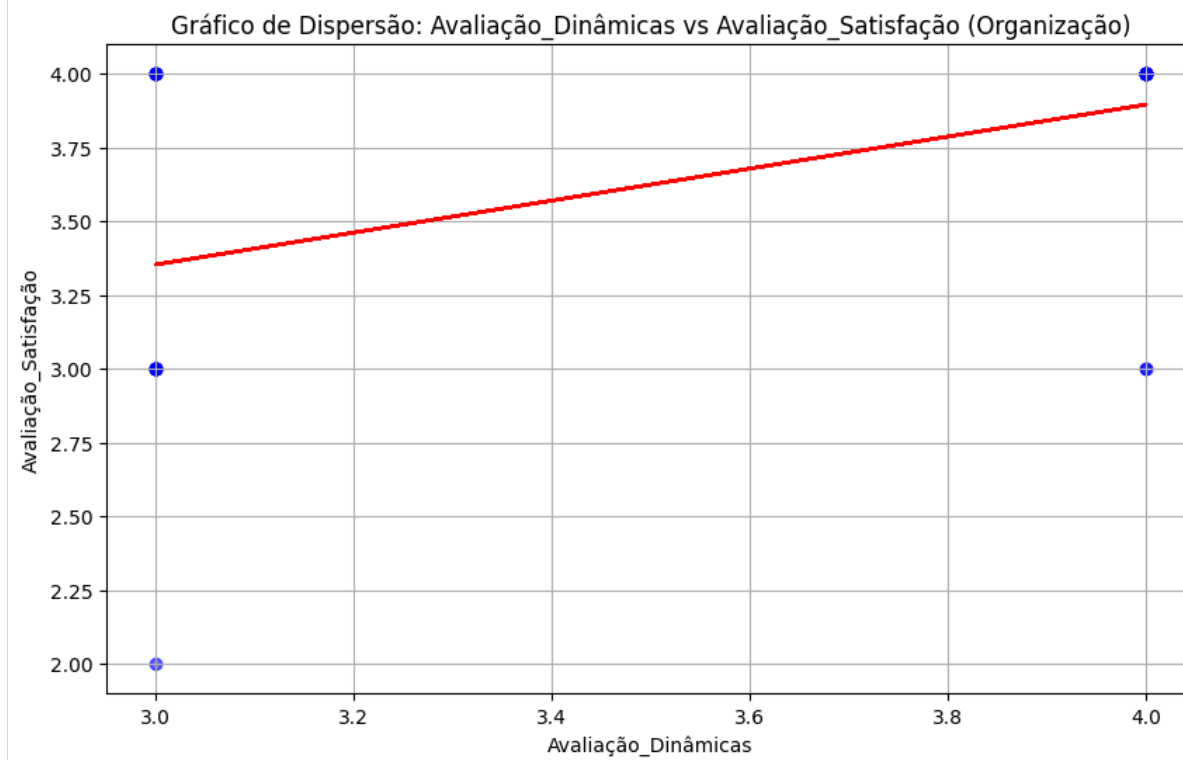




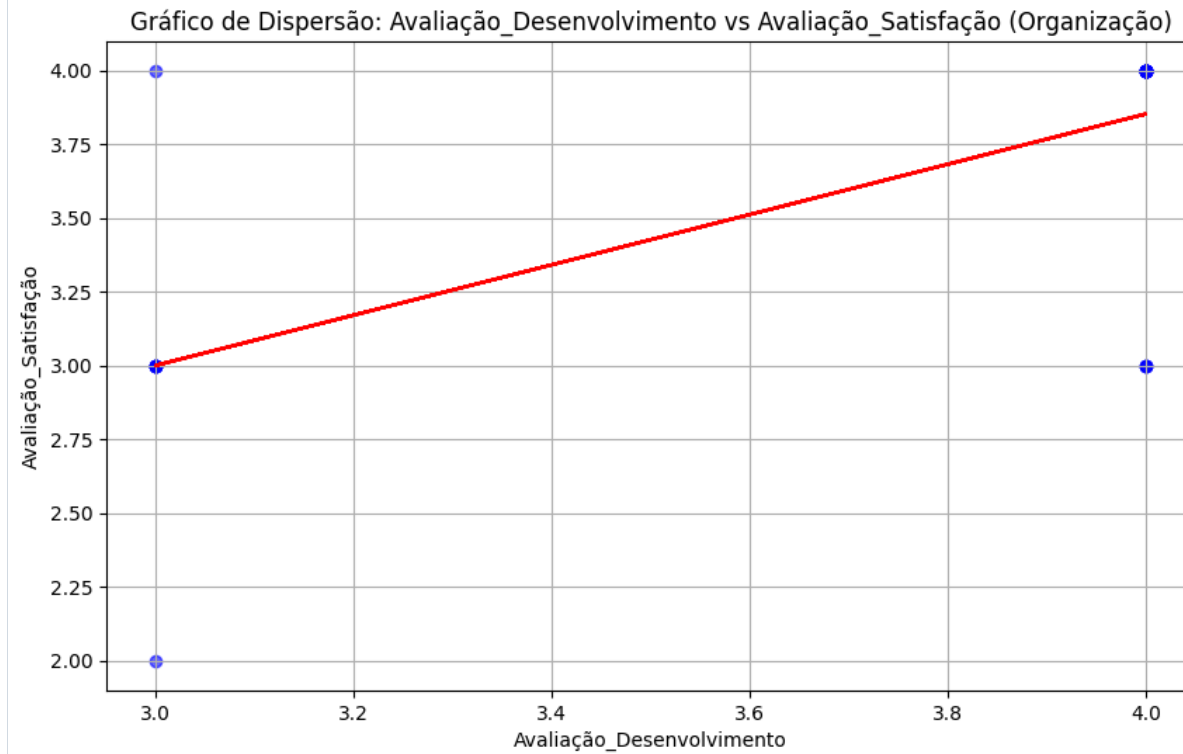


Não há dados suficientes para plotar Avaliação_Comunicação vs Comentário_Organização (Organização)





Não há dados suficientes para plotar Avaliação_Dinâmicas vs Comentário_Organização (Organização)



Não há dados suficientes para plotar Avaliação_Desenvolvimento vs Comentário_Organização (Organização)

Não há dados suficientes para plotar Avaliação_Satisfação vs Comentário_Organização (Organização)

Gráfico de Dispersão: Experiência_Conteúdos_Relevantes vs Experiência_Material_Objetivo (Metodologia e Instrutores)

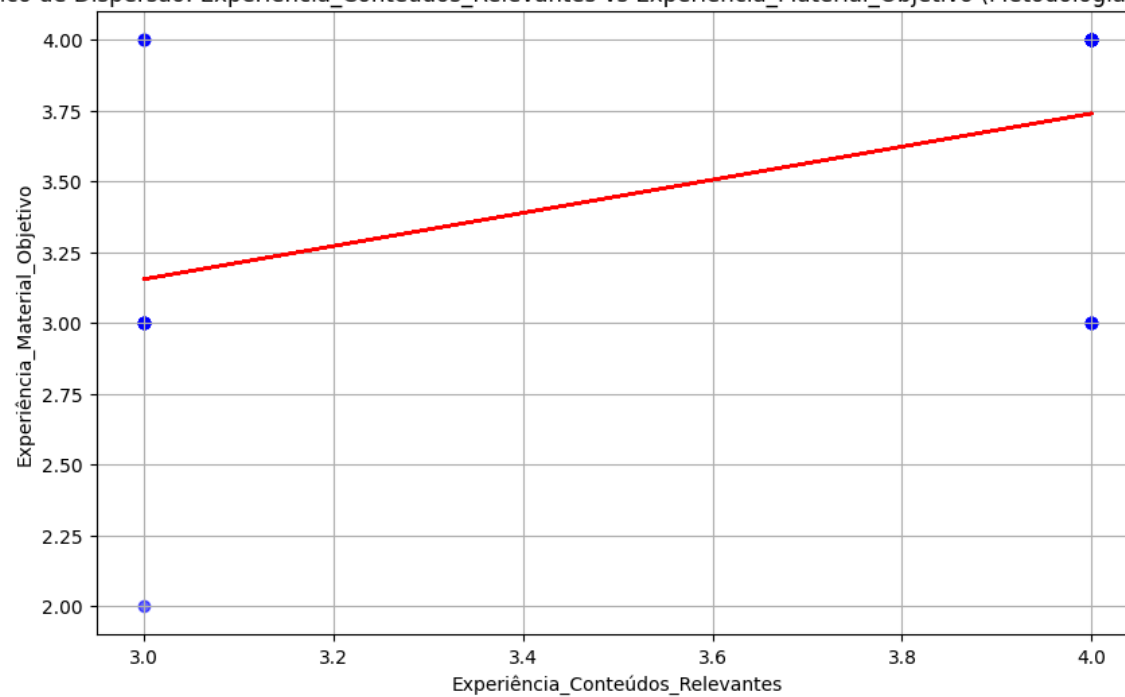


Gráfico de Dispersão: Experiência_Conteúdos_Relevantes vs Experiência_Informações_Aplicáveis (Metodologia e Instrutores)

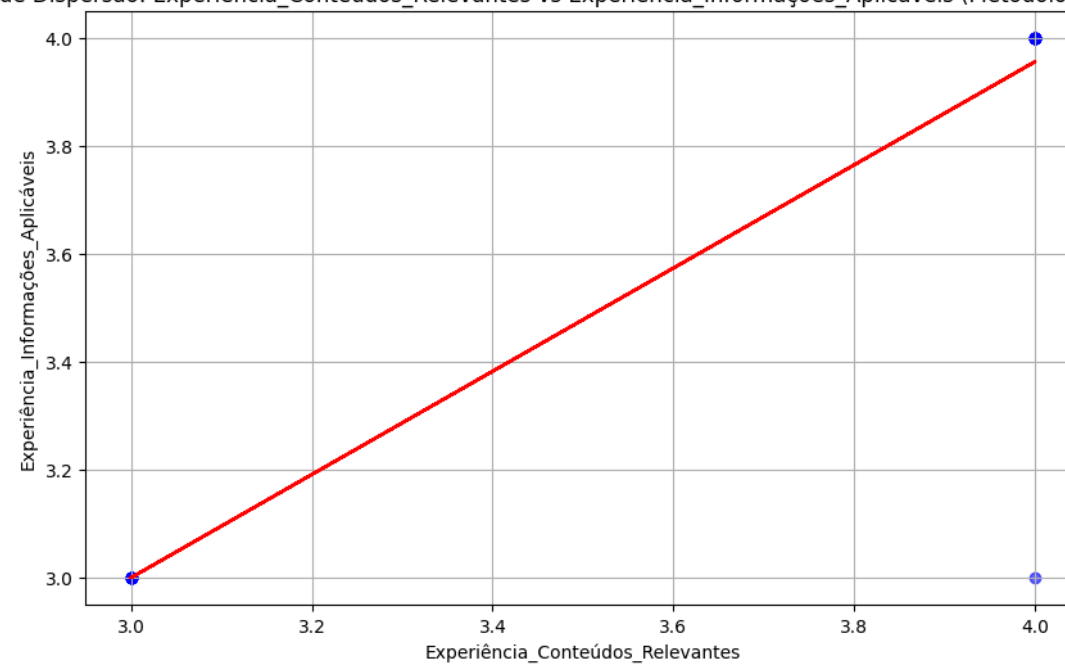


Gráfico de Dispersão: Experiência_Conteúdos_Relevantes vs Experiência_Conteúdos_Uteis (Metodologia e Instrutores)

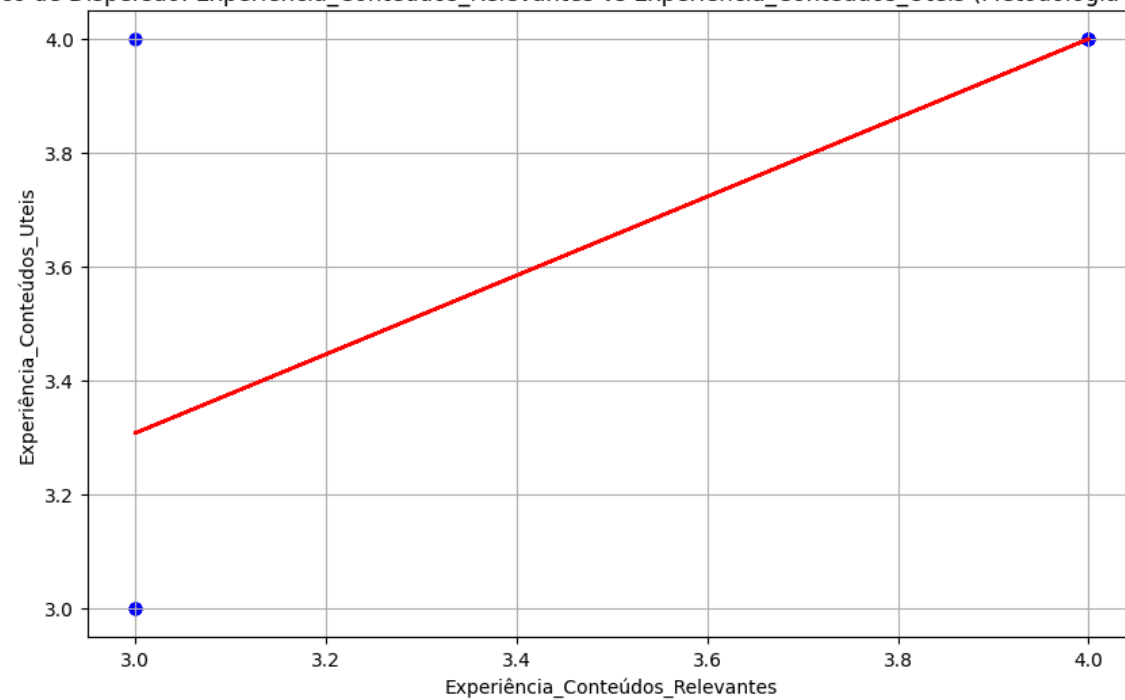


Gráfico de Dispersão: Experiência_Conteúdos_Relevantes vs Experiência_Métodos_Engajamento (Metodologia e Instrutores)

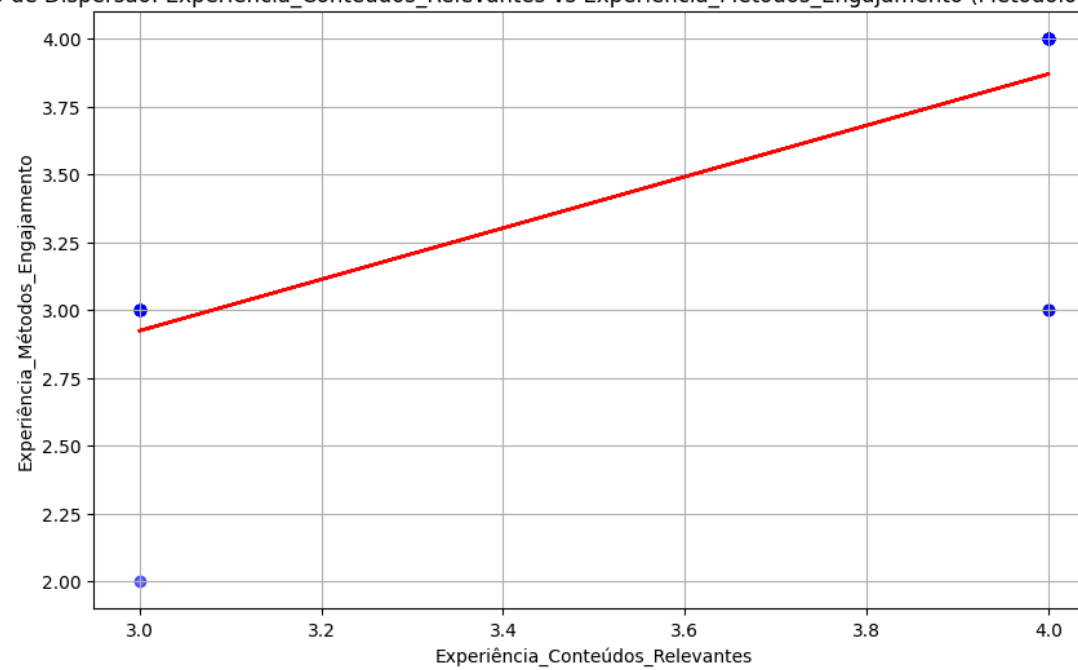


Gráfico de Dispersão: Experiência_Conteúdos_Relevantes vs Experiência_Conhecimento_Profundo (Metodologia e Instrutores)

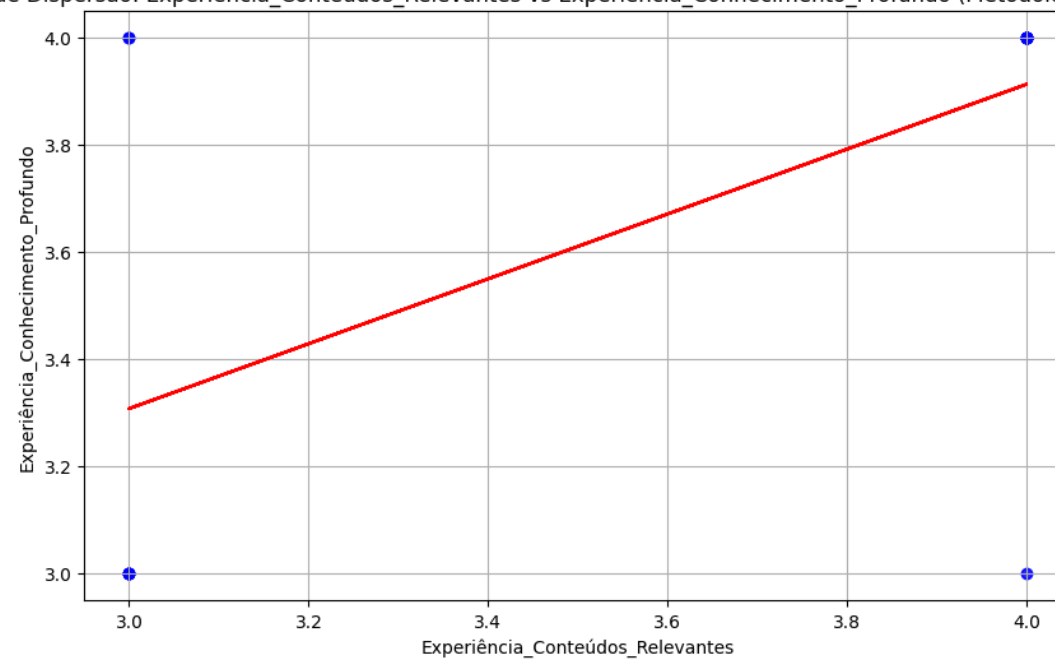
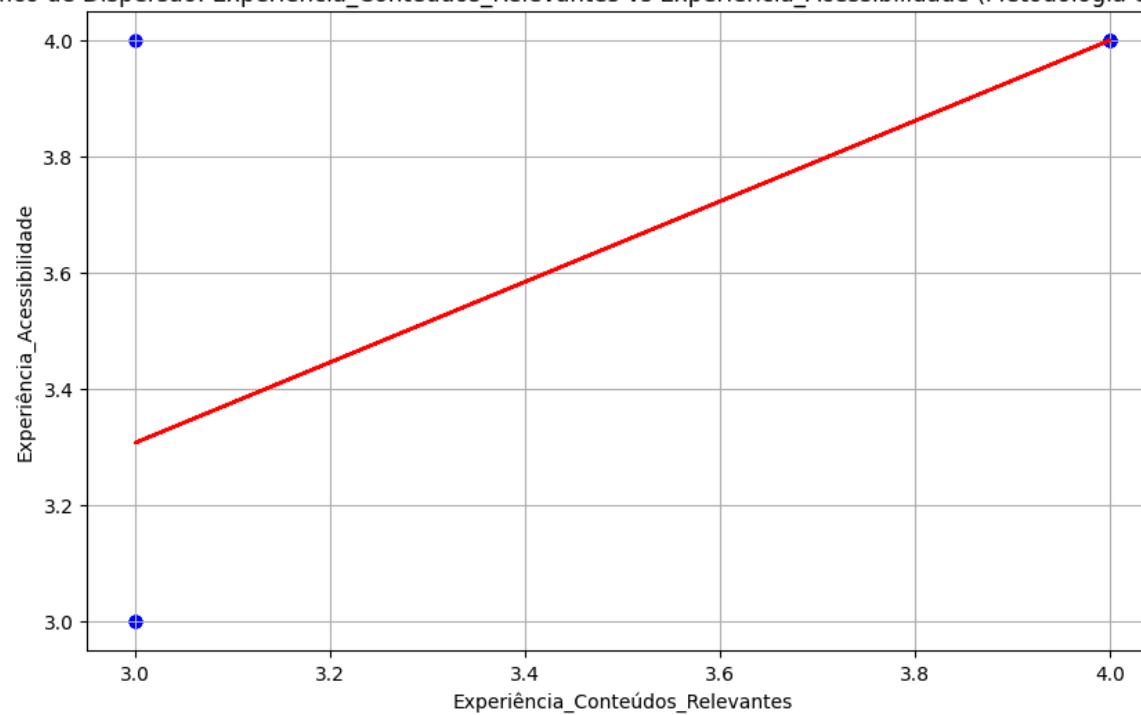


Gráfico de Dispersão: Experiência_Conteúdos_Relevantes vs Experiência_Acessibilidade (Metodologia e Instrutores)



Não há dados suficientes para plotar Experiência_Conteúdos_Relevantes vs Comentário_Fornecedora (Metodologia e Instrutores)

Gráfico de Dispersão: Experiência_Material_Objetoivo vs Experiência_Informações_Aplicáveis (Metodologia e Instrutores)

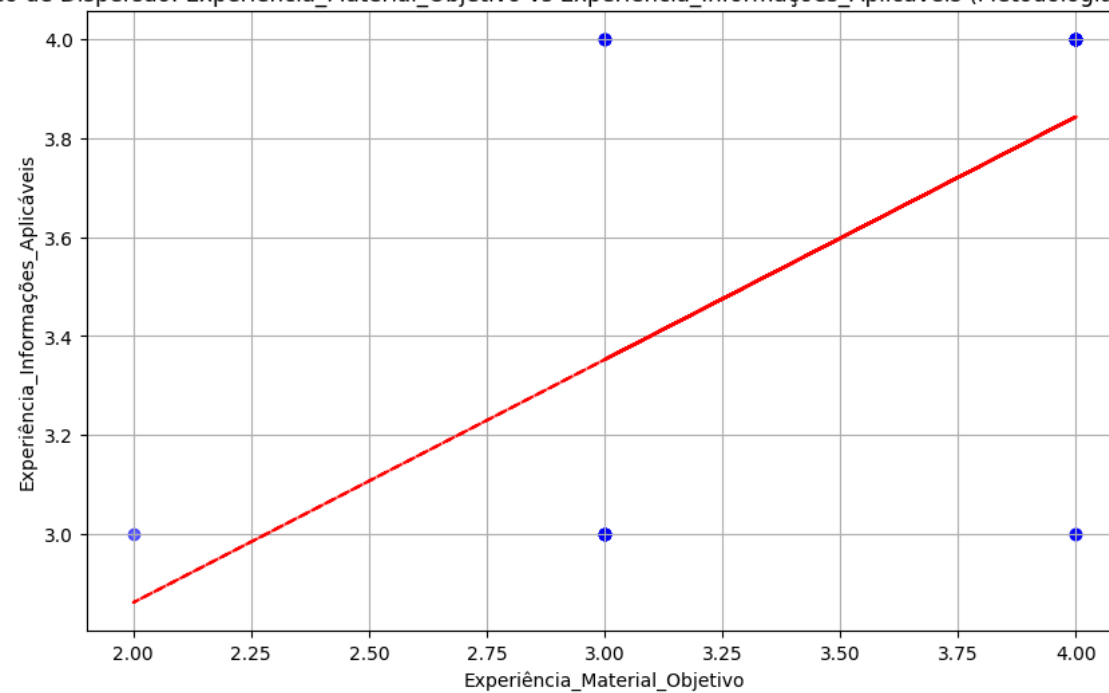


Gráfico de Dispersão: Experiência_Material_Objetoivo vs Experiência_Conteúdos_Uteis (Metodologia e Instrutores)

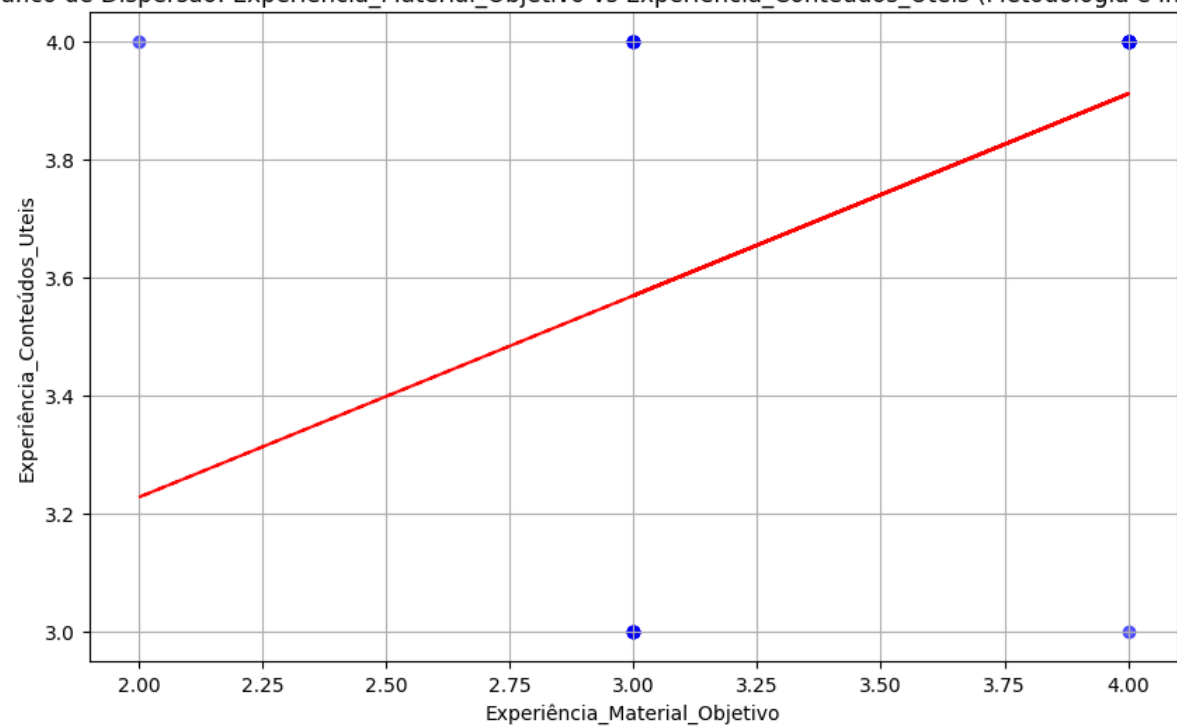


Gráfico de Dispersão: Experiência_Material_Objetivo vs Experiência_Métodos_Engajamento (Metodologia e Instrutores)

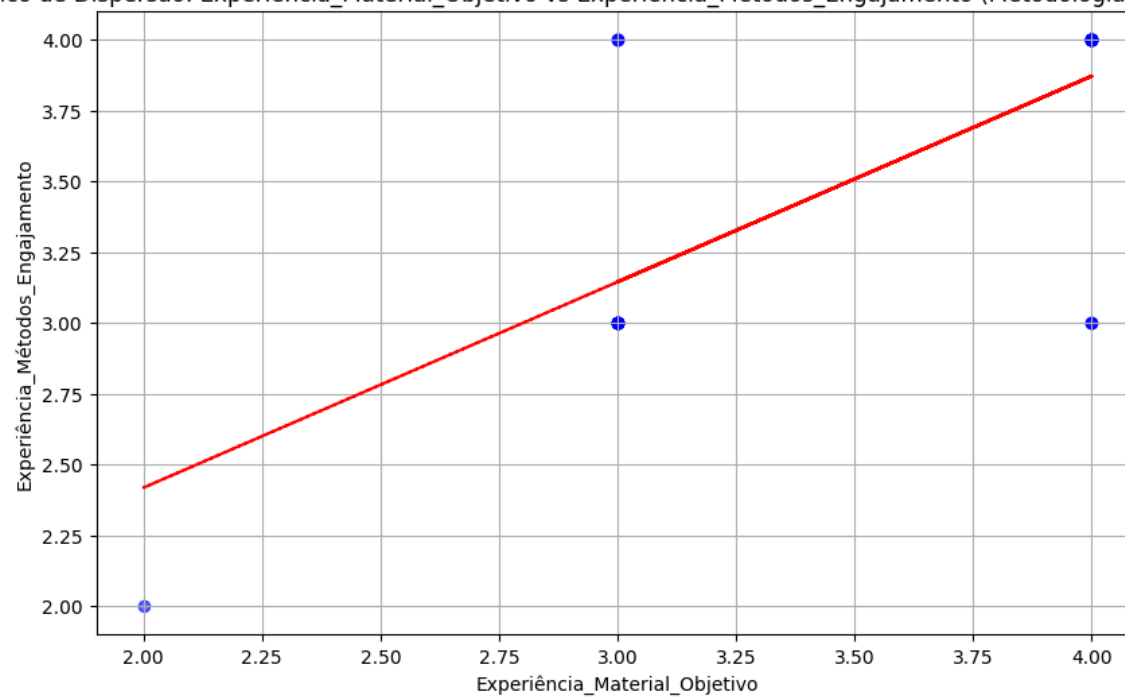


Gráfico de Dispersão: Experiência_Material_Objetivo vs Experiência_Conhecimento_Profundo (Metodologia e Instrutores)

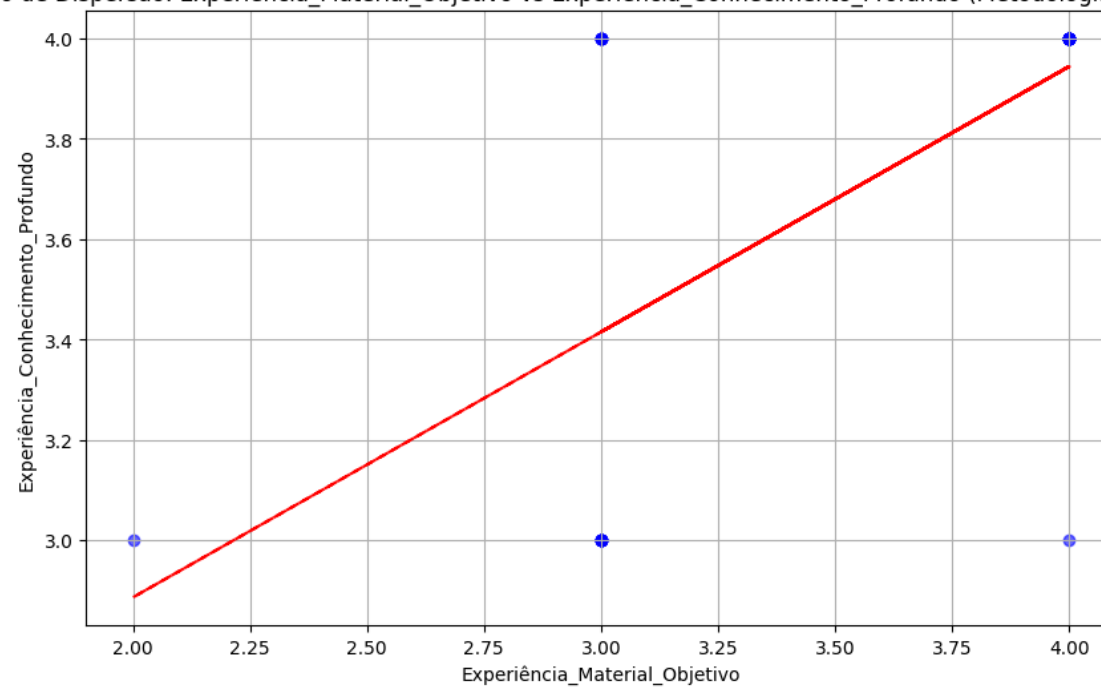
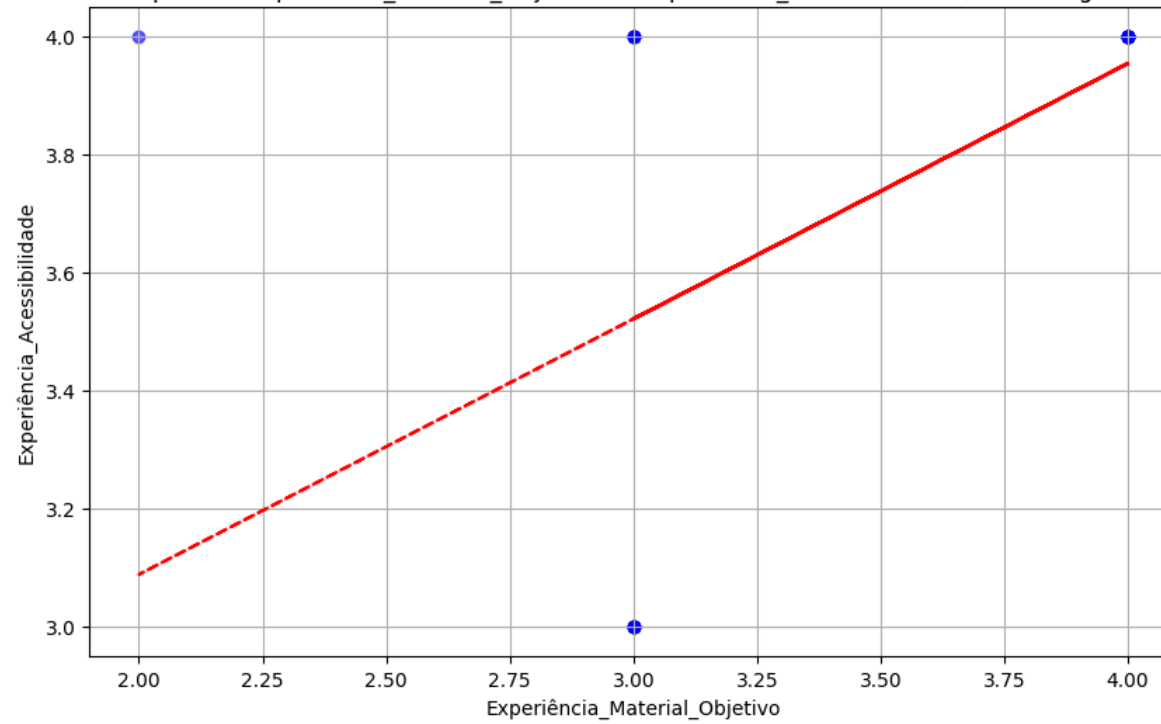


Gráfico de Dispersão: Experiência_Material_Objetoivo vs Experiência_Acessibilidade (Metodologia e Instrutores)



Não há dados suficientes para plotar Experiência_Material_Objetoivo vs Comentário_Fornecedora (Metodologia e Instrutores)

Gráfico de Dispersão: Experiência_Informações_Aplicáveis vs Experiência_Conteúdos_Uteis (Metodologia e Instrutores)

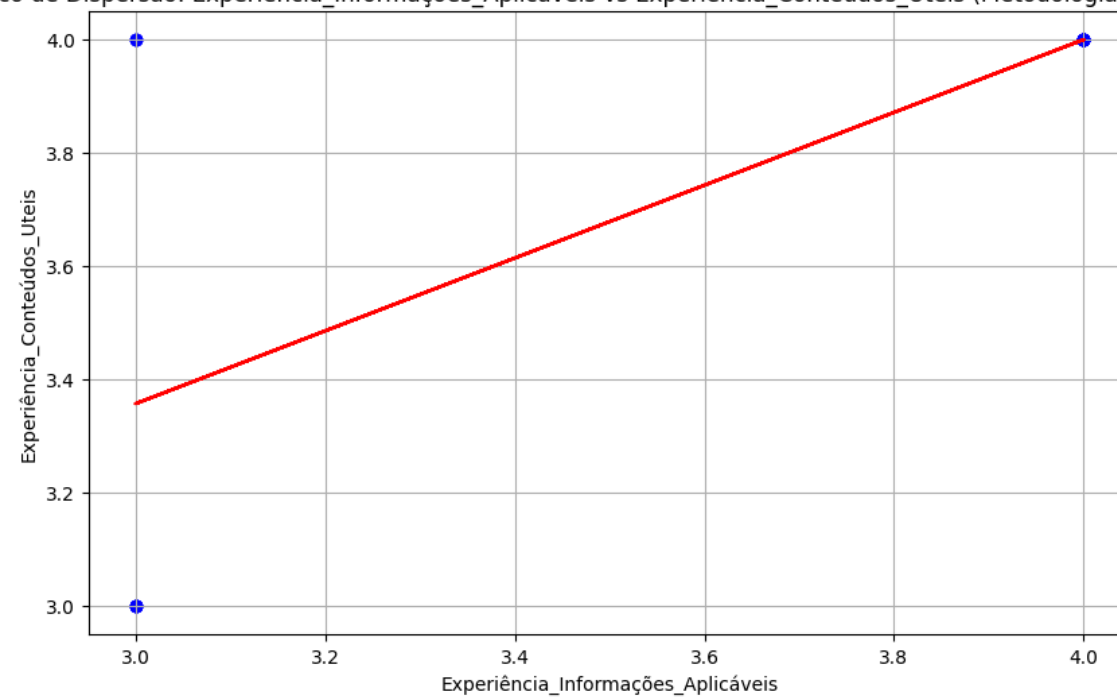


Gráfico de Dispersão: Experiência_Informações_Aplicáveis vs Experiência_Métodos_Engajamento (Metodologia e Instrutores)

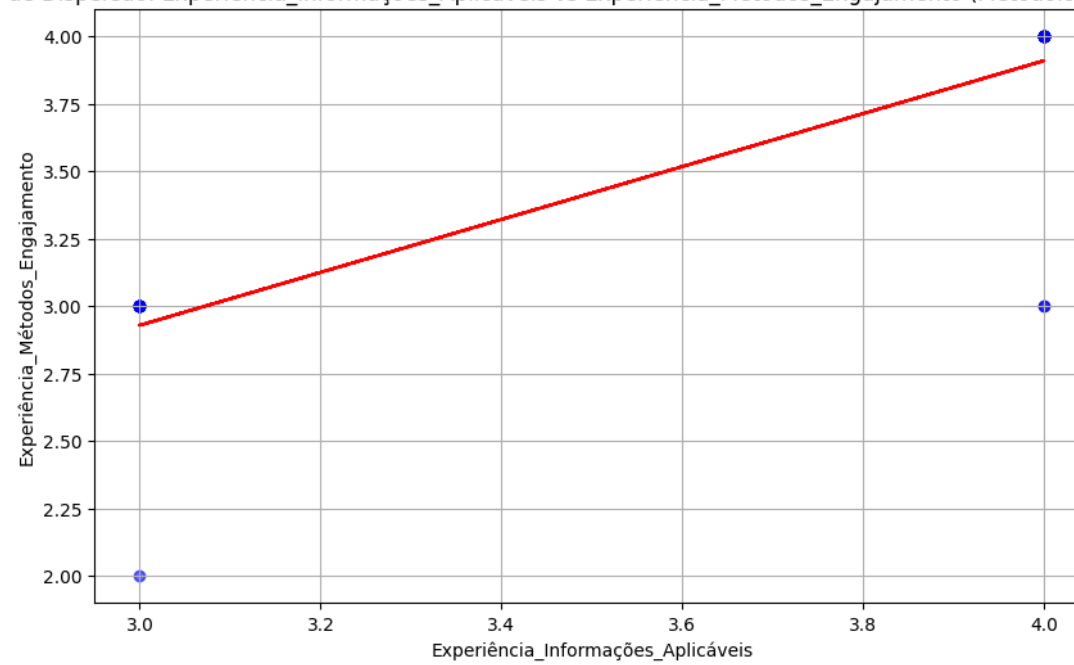


Gráfico de Dispersão: Experiência_Informações_Aplicáveis vs Experiência_Conhecimento_Profundo (Metodologia e Instrutores)

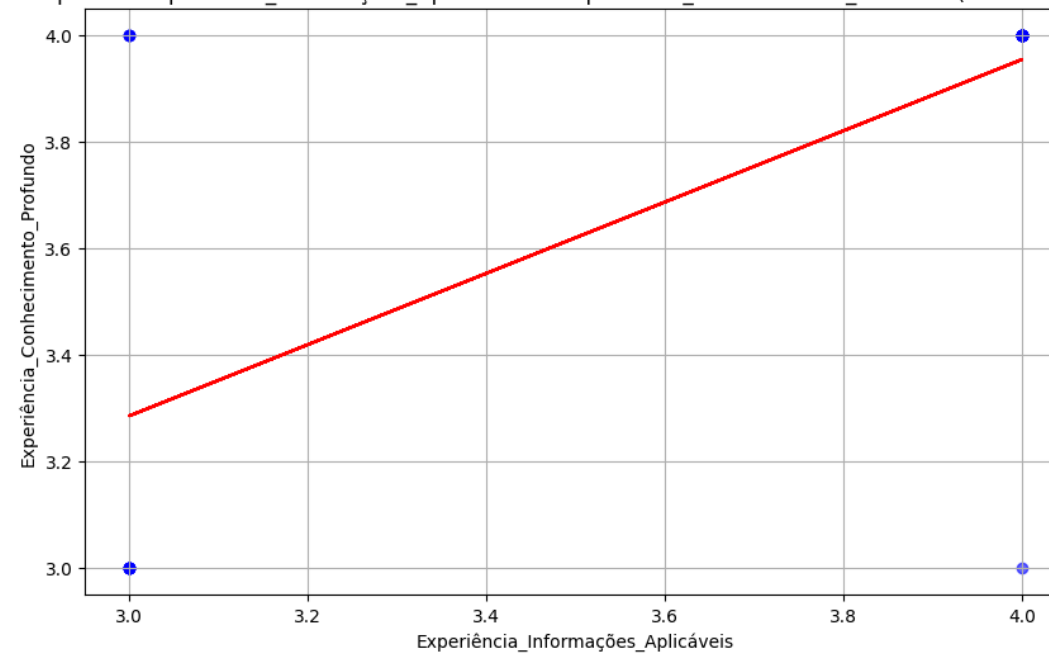
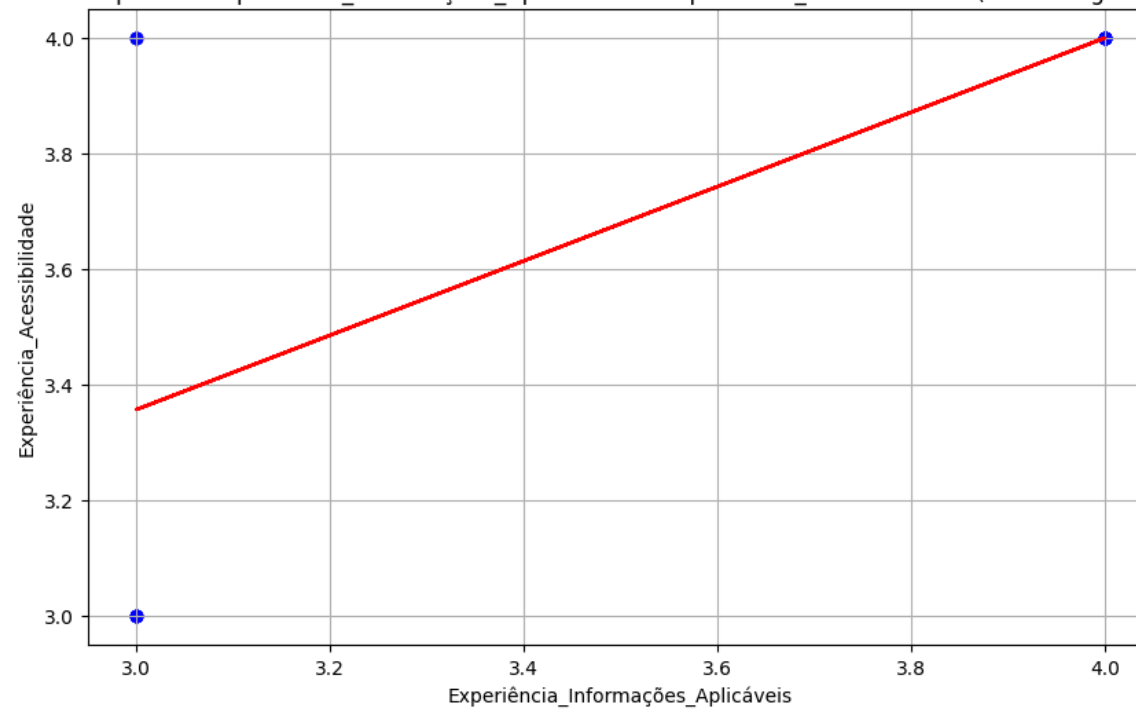


Gráfico de Dispersão: Experiência_Informações_Aplicáveis vs Experiência_Acessibilidade (Metodologia e Instrutores)



Não há dados suficientes para plotar Experiência_Informações_Aplicáveis vs Comentário_Fornecedora (Metodologia e Instrutores)

Gráfico de Dispersão: Experiência_Conteúdos_Uteis vs Experiência_Métodos_Engajamento (Metodologia e Instrutores)

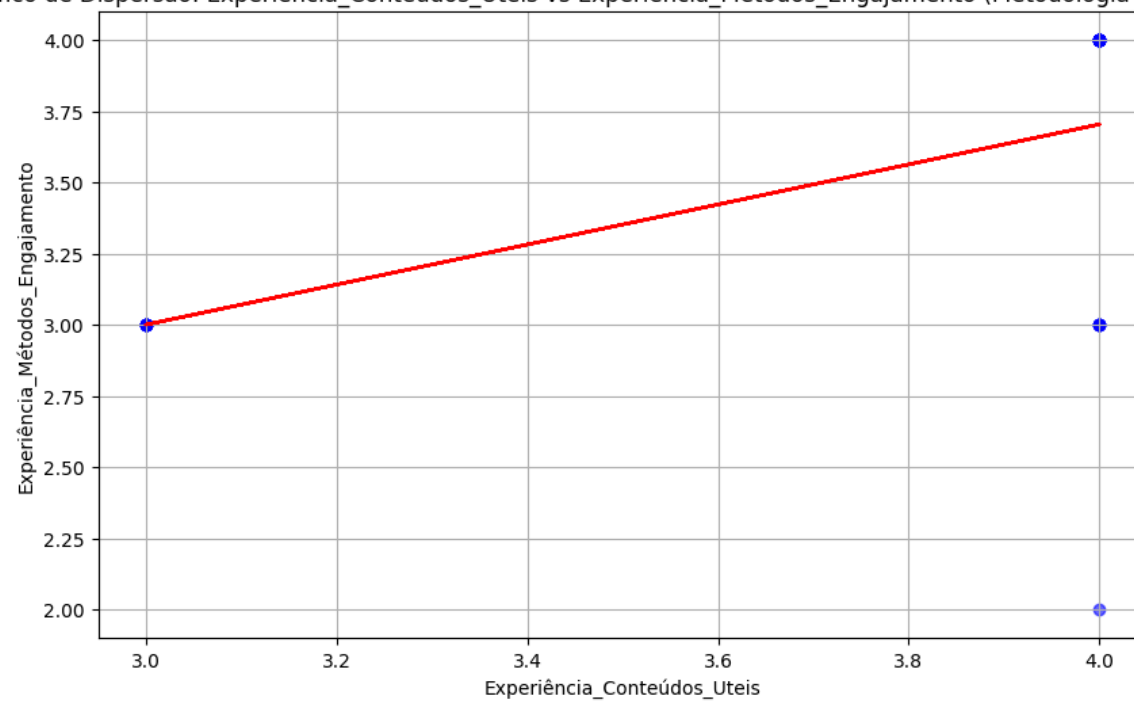


Gráfico de Dispersão: Experiência_Conteúdos_Uteis vs Experiência_Conhecimento_Profundo (Metodologia e Instrutores)

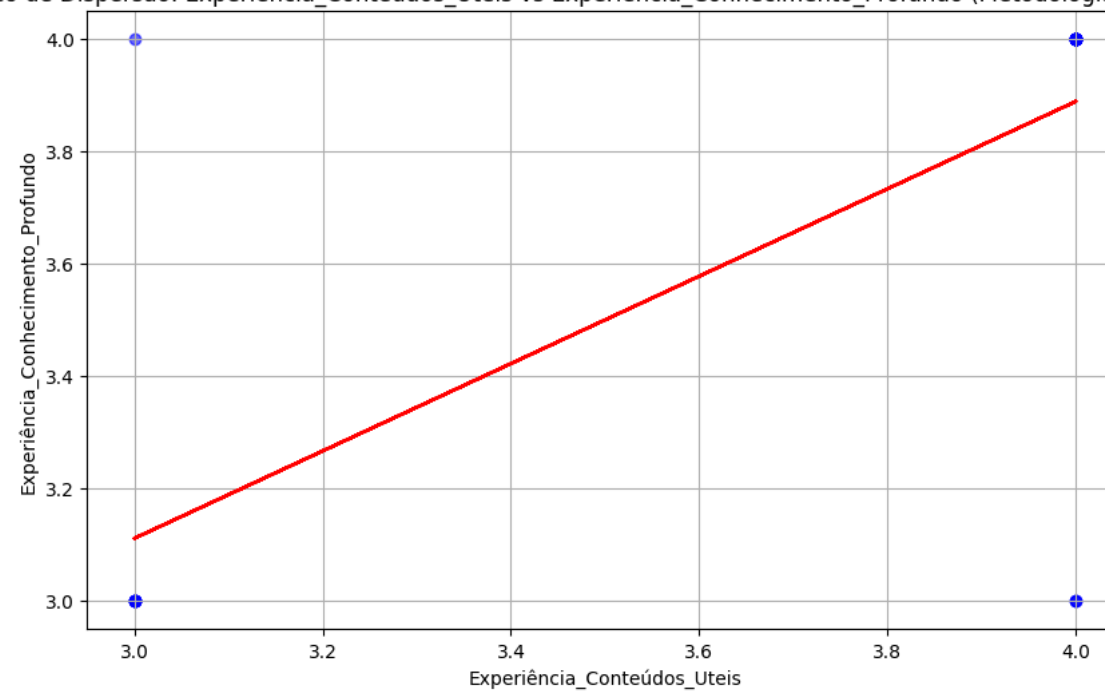
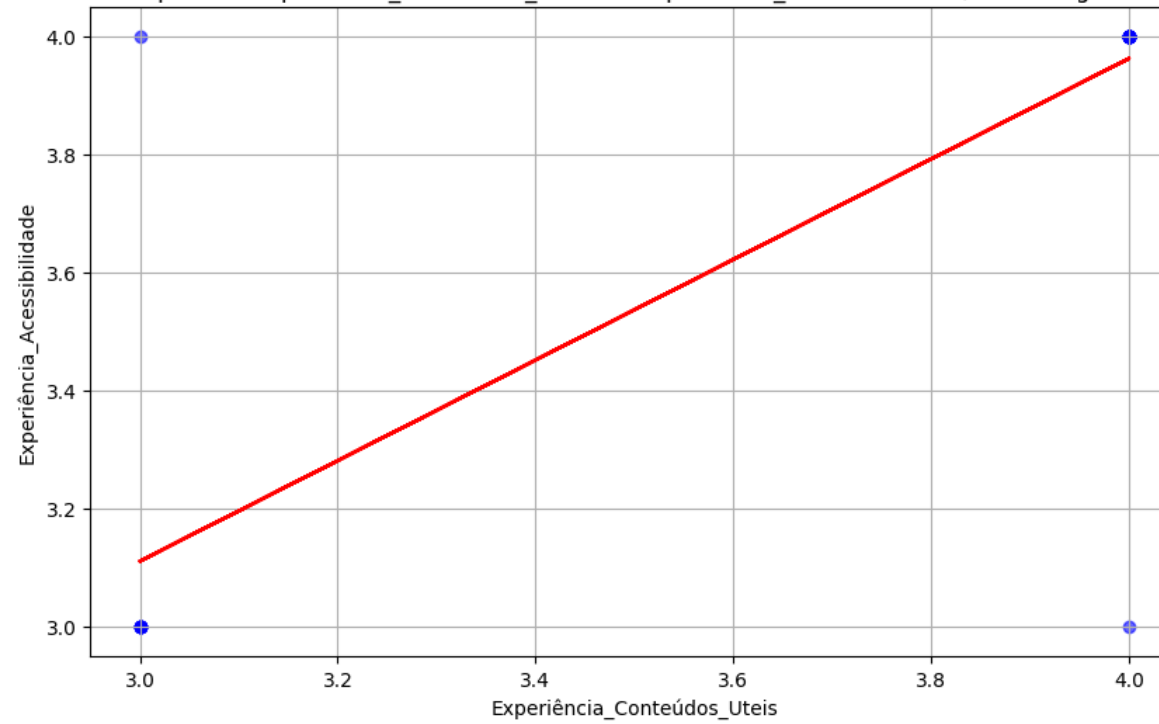


Gráfico de Dispersão: Experiência_Conteúdos_Uteis vs Experiência_Acessibilidade (Metodologia e Instrutores)



Não há dados suficientes para plotar Experiência_Conteúdos_Uteis vs Comentário_Fornecedora (Metodologia e Instrutores)

Gráfico de Dispersão: Experiência_Métodos_Engajamento vs Experiência_Conhecimento_Profundo (Metodologia e Instrutores)

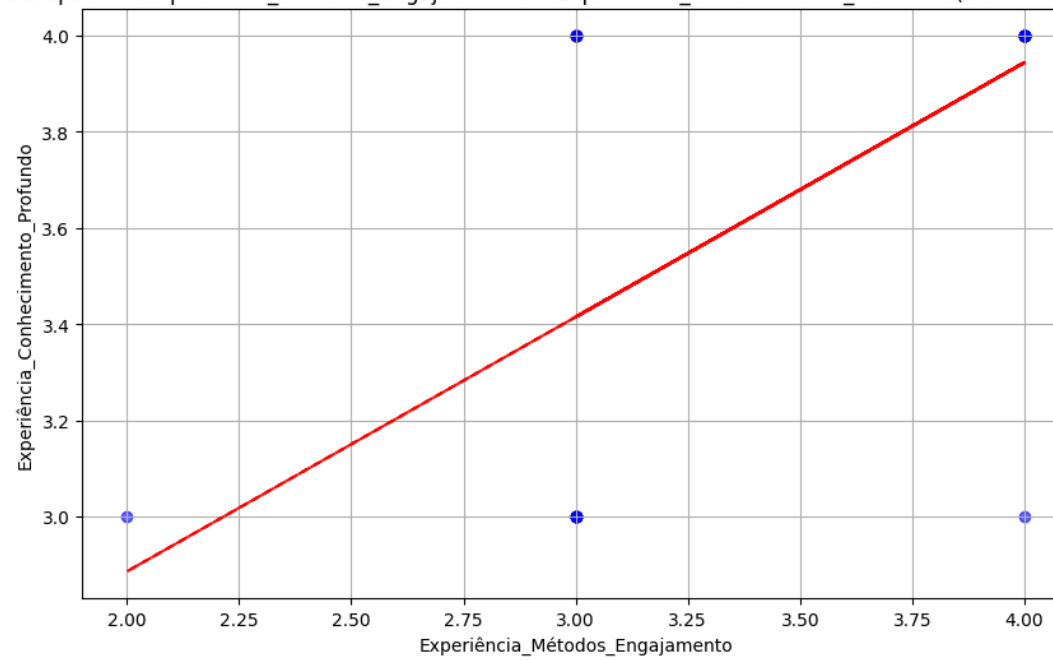
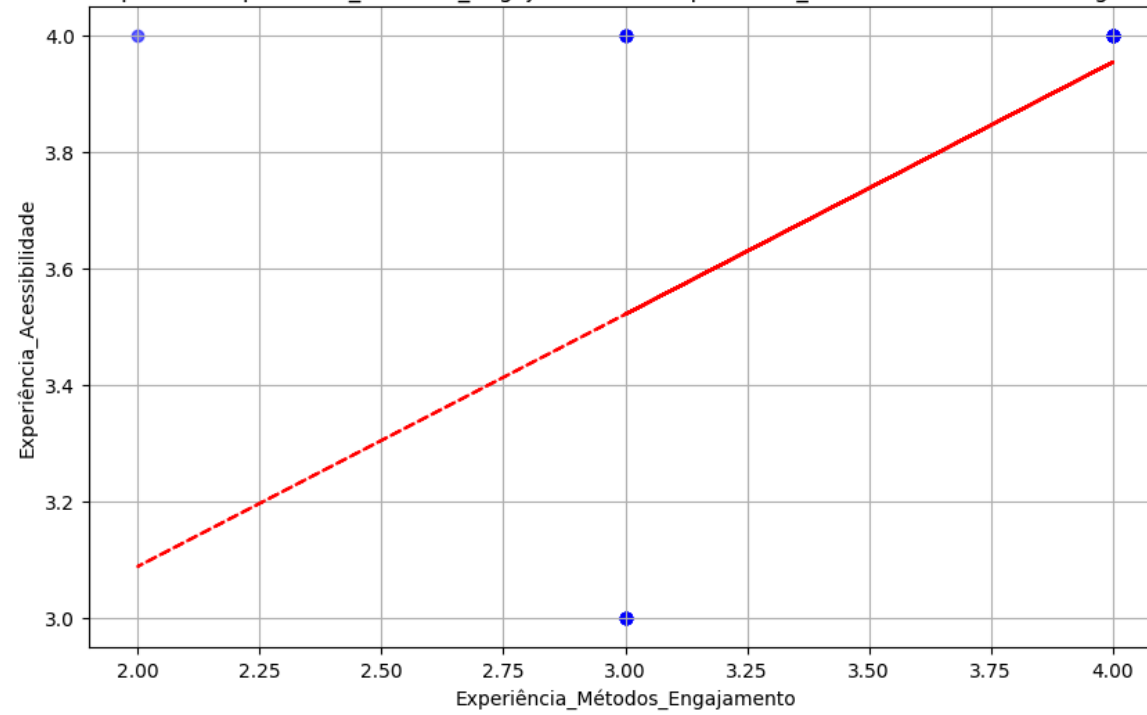
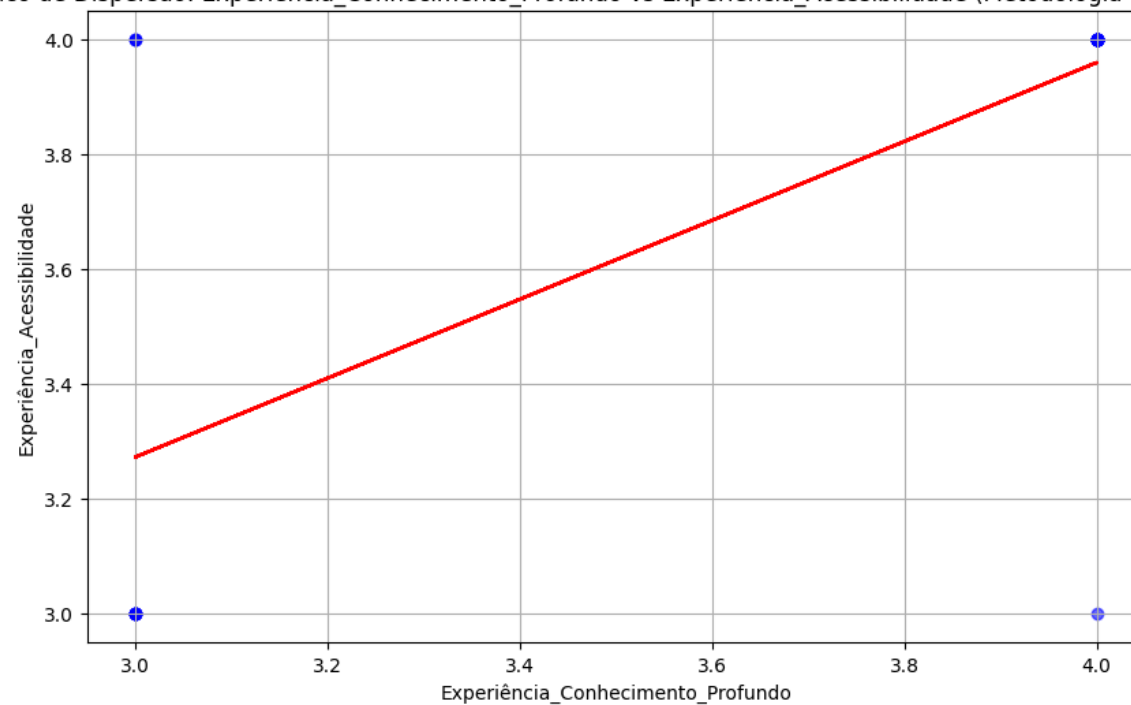


Gráfico de Dispersão: Experiência_Métodos_Engajamento vs Experiência_Acessibilidade (Metodologia e Instrutores)



Não há dados suficientes para plotar Experiência_Métodos_Engajamento vs Comentário_Fornecedora (Metodologia e Instrutores)

Gráfico de Dispersão: Experiência_Conhecimento_Profundo vs Experiência_Acessibilidade (Metodologia e Instrutores)



Não há dados suficientes para plotar Experiência_Conhecimento_Profundo vs Comentário_Fornecedora (Metodologia e Instrutores)

Não há dados suficientes para plotar Experiência_Acessibilidade vs Comentário_Fornecedora (Metodologia e Instrutores)

Gráfico de Dispersão: Experiência_Espaço_Confortável vs Experiência_Equipamentos (Infraestrutura)

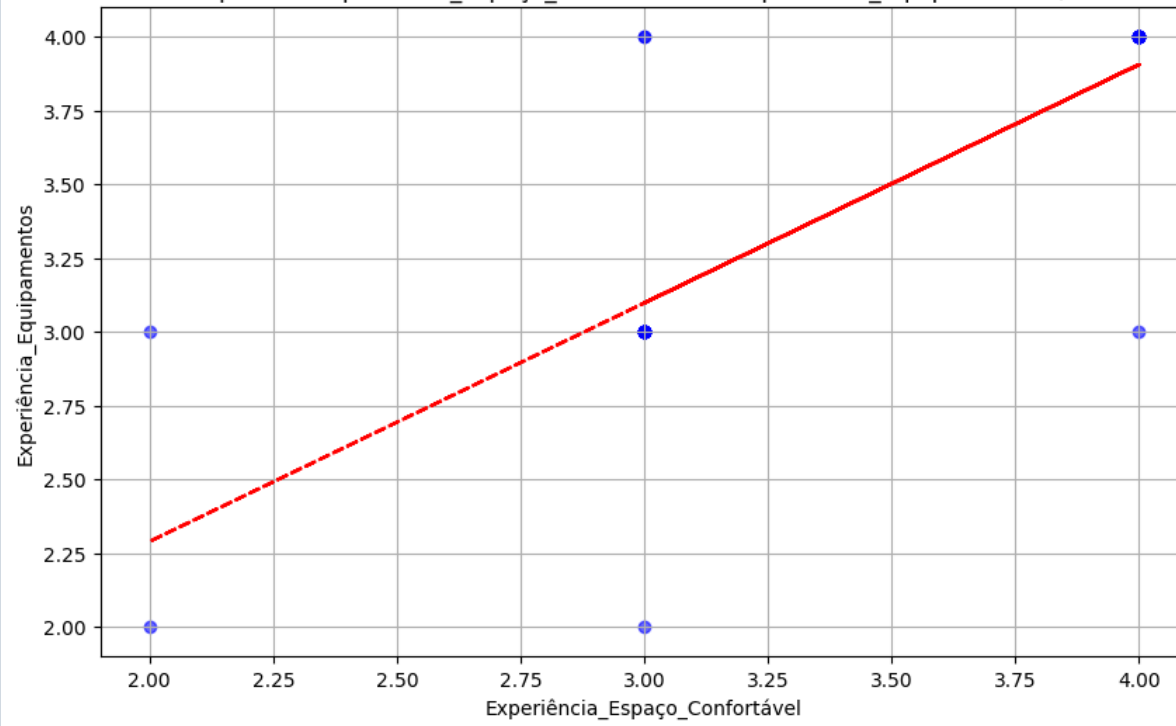


Gráfico de Dispersão: Experiência_Espaço_Confortável vs Experiência_Acessibilidade_Coworking (Infraestrutura)

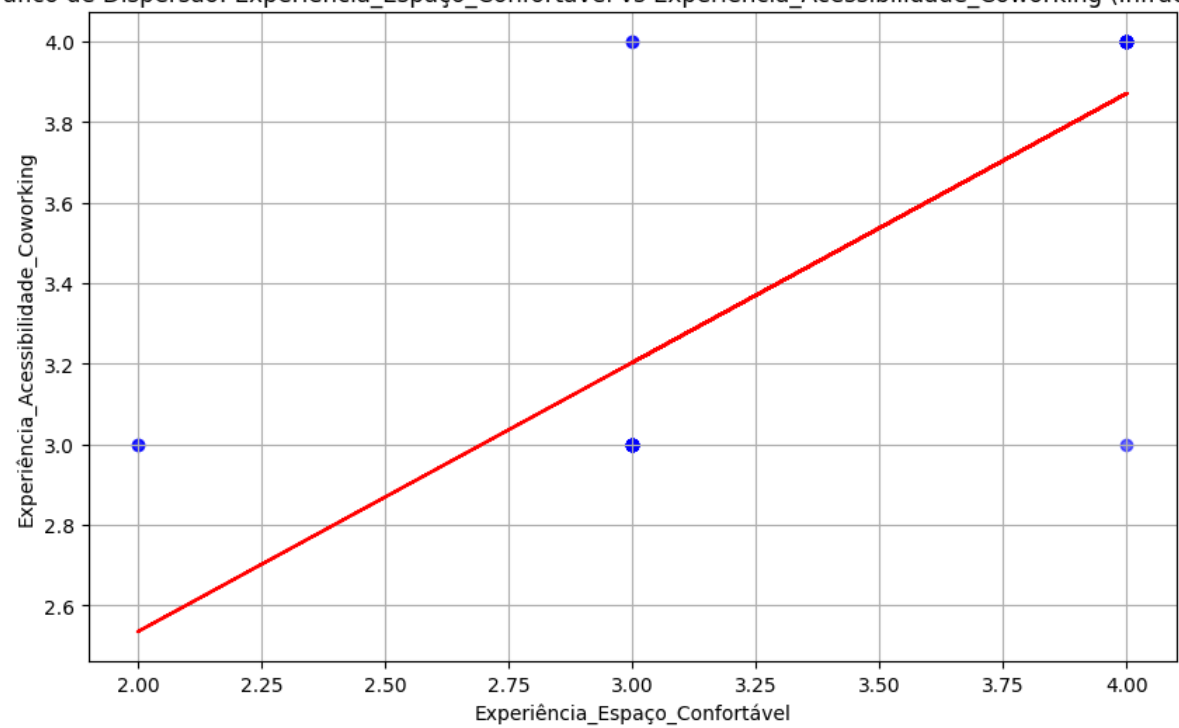


Gráfico de Dispersão: Experiência_Espaço_Confortável vs Experiência_Lanche_Adequado (Infraestrutura)

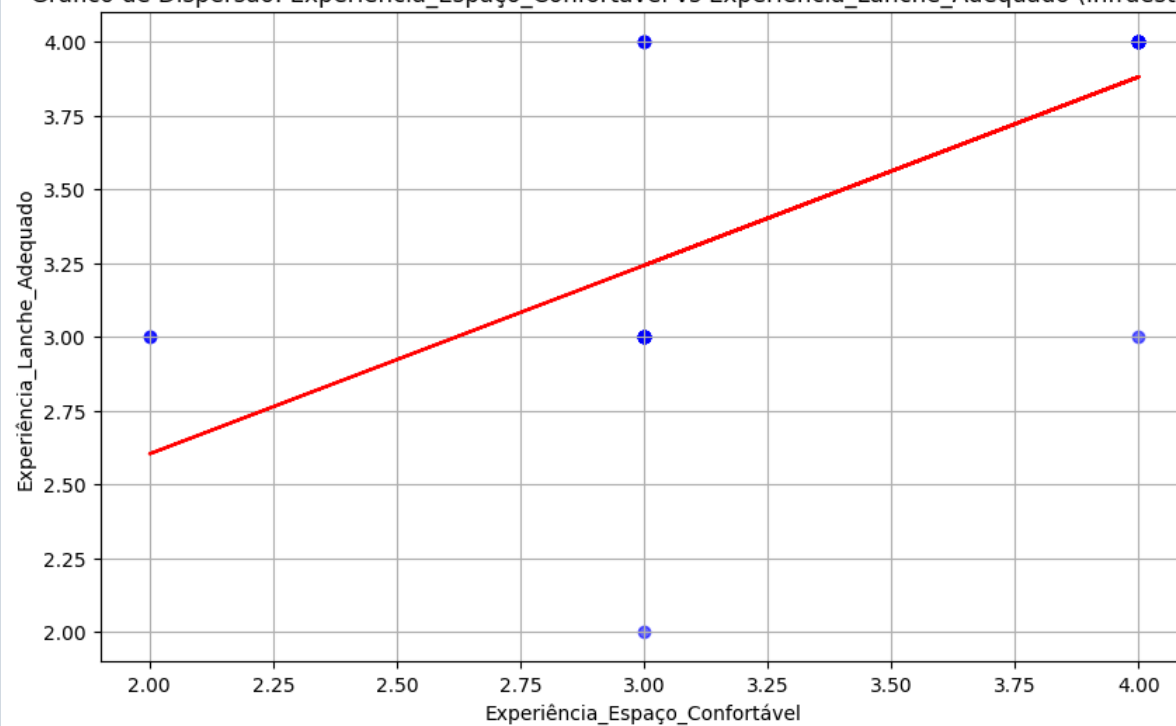
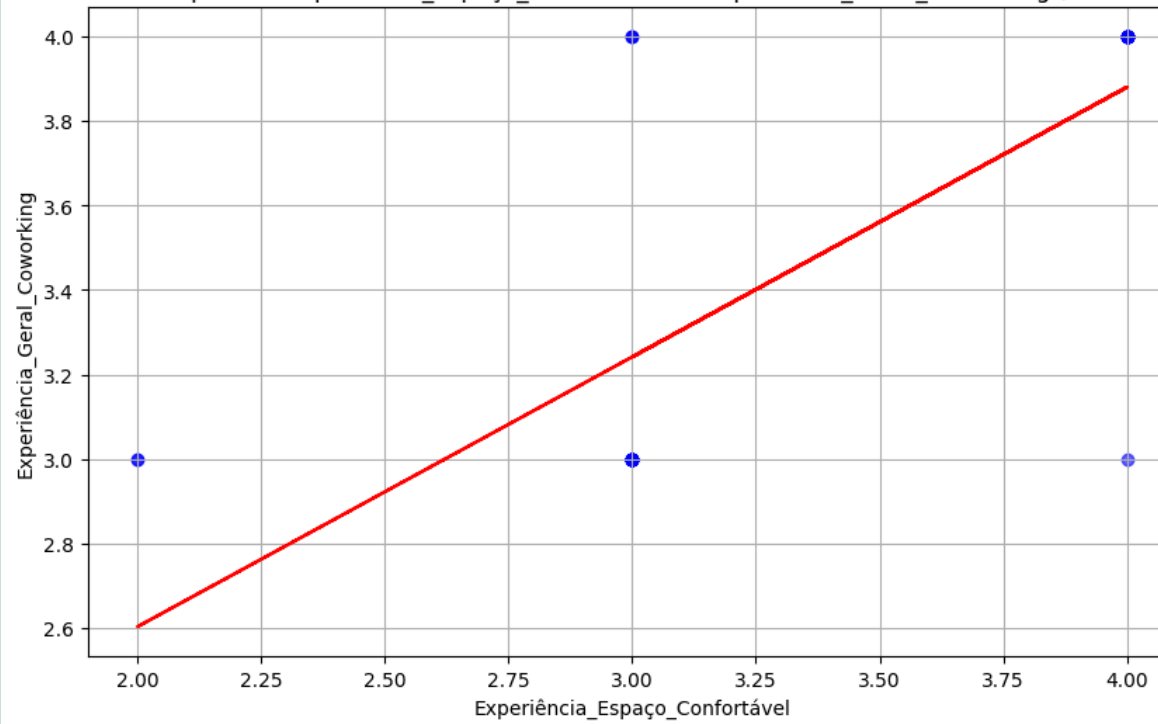


Gráfico de Dispersão: Experiência_Espaço_Confortável vs Experiência_Geral_Coworking (Infraestrutura)



Não há dados suficientes para plotar Experiência_Espaço_Confortável vs Comentário_Espaço_Coworking (Infraestrutura)

Gráfico de Dispersão: Experiência_Equipamentos vs Experiência_Acessibilidade_Coworking (Infraestrutura)

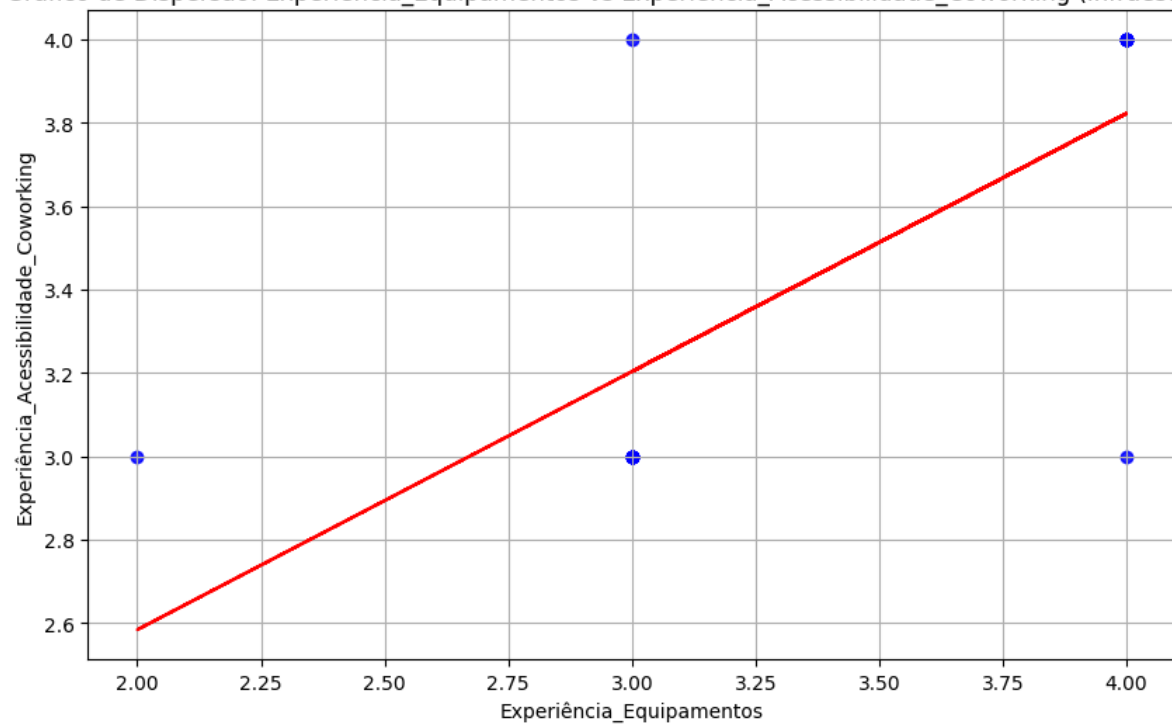
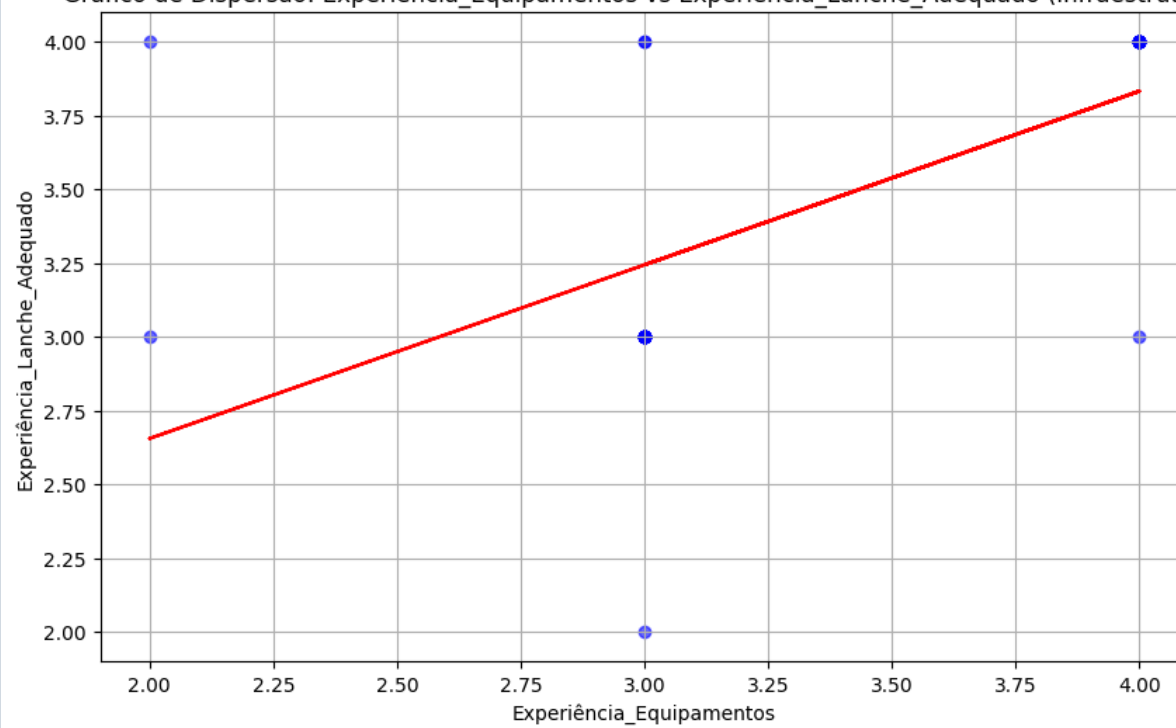
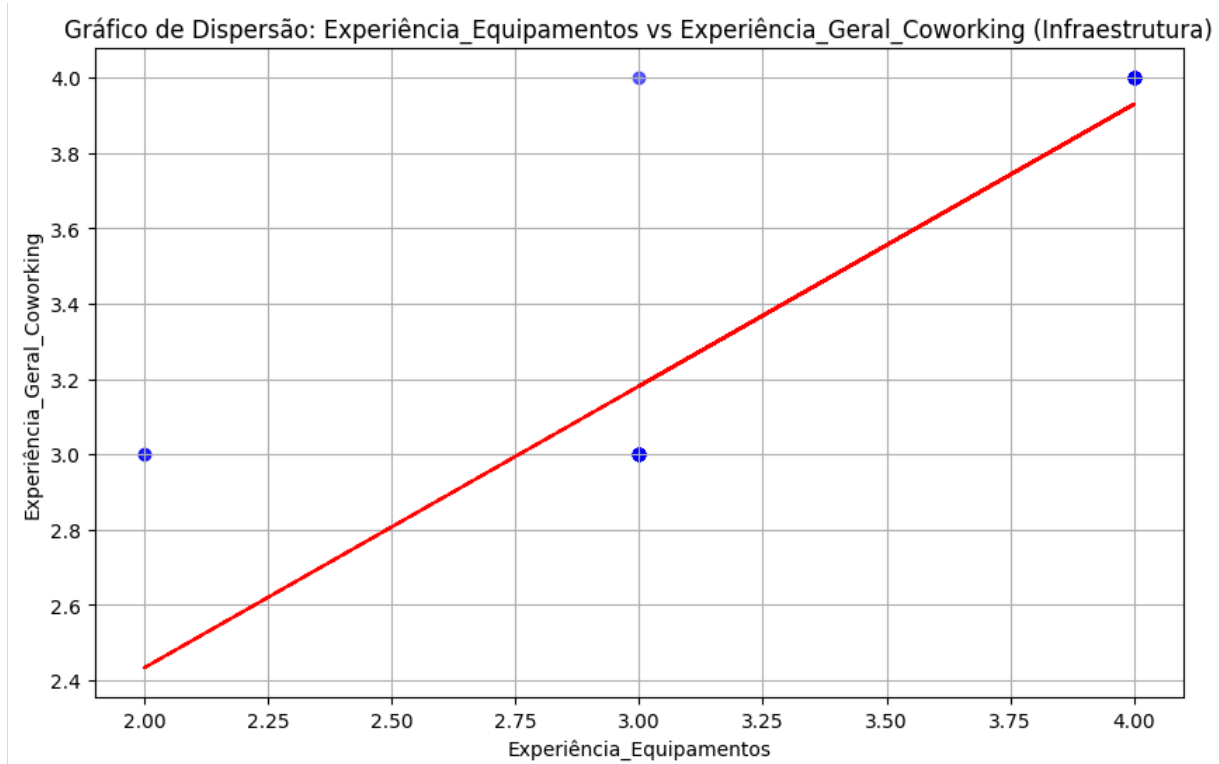


Gráfico de Dispersão: Experiência_Equipamentos vs Experiência_Lanche_Adequado (Infraestrutura)





Não há dados suficientes para plotar Experiência_Equipamentos vs Comentário_Espaço_Coworking (Infraestrutura)

Gráfico de Dispersão: Experiência_Acessibilidade_Coworking vs Experiência_Lanche_Adequado (Infraestrutura)

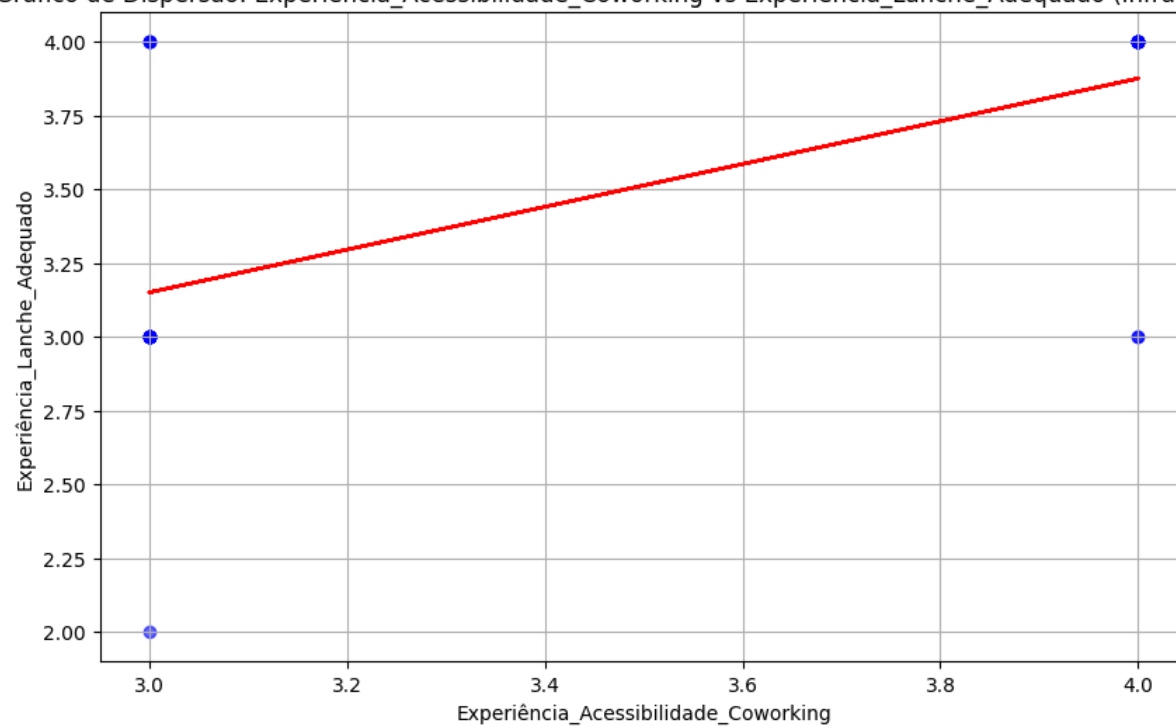
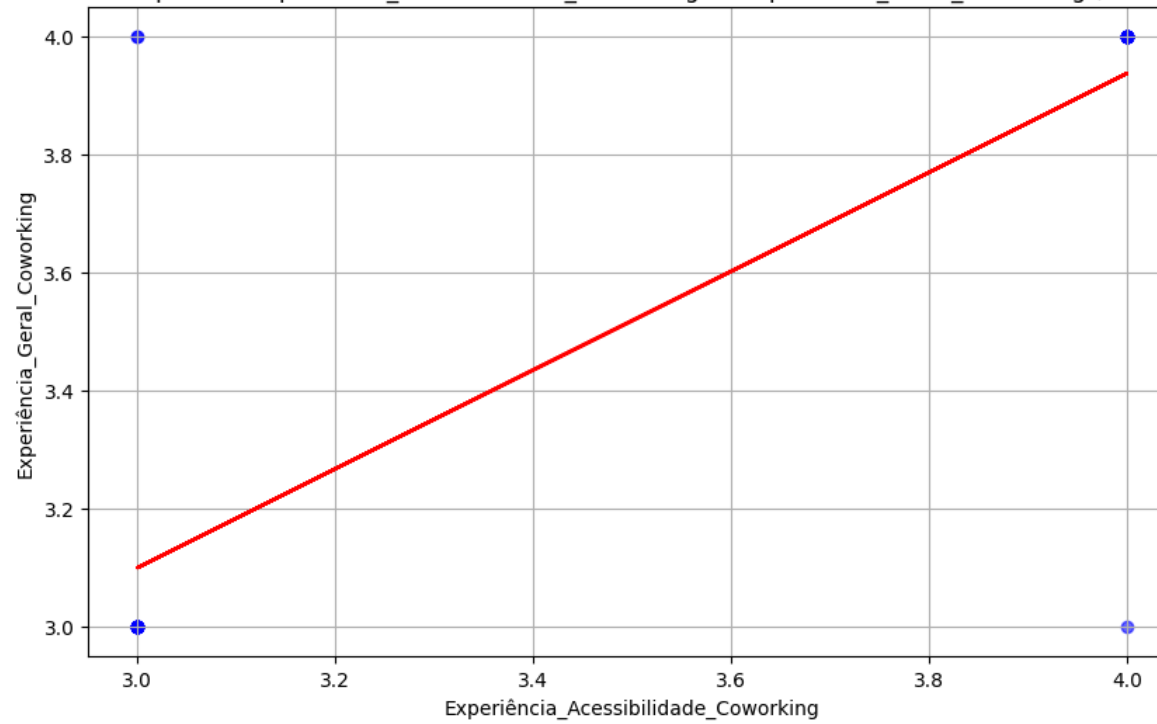
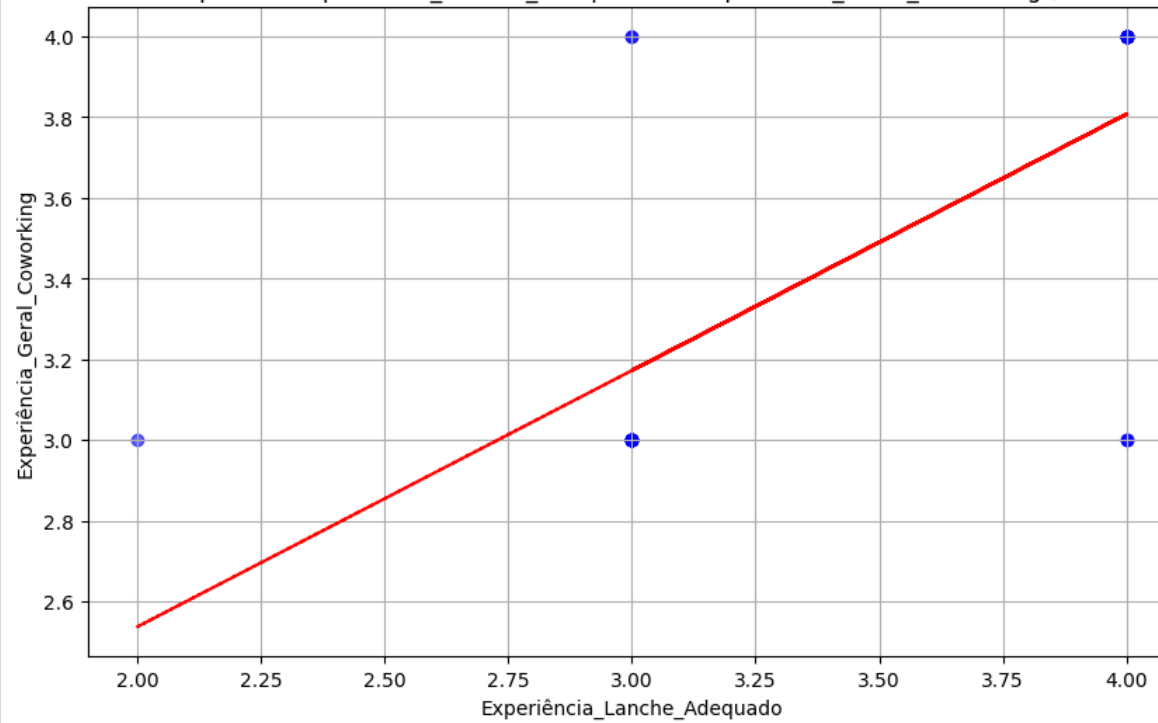


Gráfico de Dispersão: Experiência_Acessibilidade_Coworking vs Experiência_Geral_Coworking (Infraestrutura)



Não há dados suficientes para plotar Experiência_Acessibilidade_Coworking vs Comentário_Espaço_Coworking (Infraestrutura)

Gráfico de Dispersão: Experiência_Lanche_Adequado vs Experiência_Geral_Coworking (Infraestrutura)



Não há dados suficientes para plotar Experiência_Lanche_Adequado vs Comentário_Espaço_Coworking (Infraestrutura)

Não há dados suficientes para plotar Experiência_Geral_Coworking vs Comentário_Espaço_Coworking (Infraestrutura)

Gráfico de Dispersão: Evento_Rancho_Adequado vs Evento_Comunicação_Eficaz (Evento Encerramento)

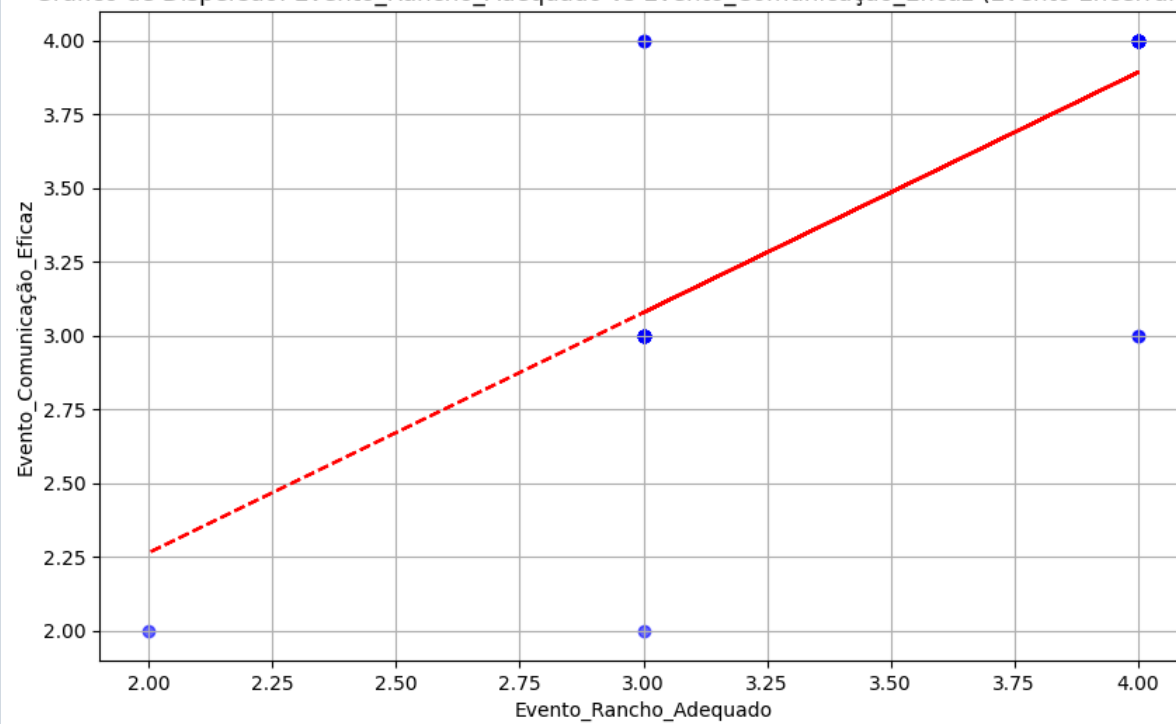


Gráfico de Dispersão: Evento_Rancho_Adequado vs Evento_Organização_Adequada (Evento Encerramento)

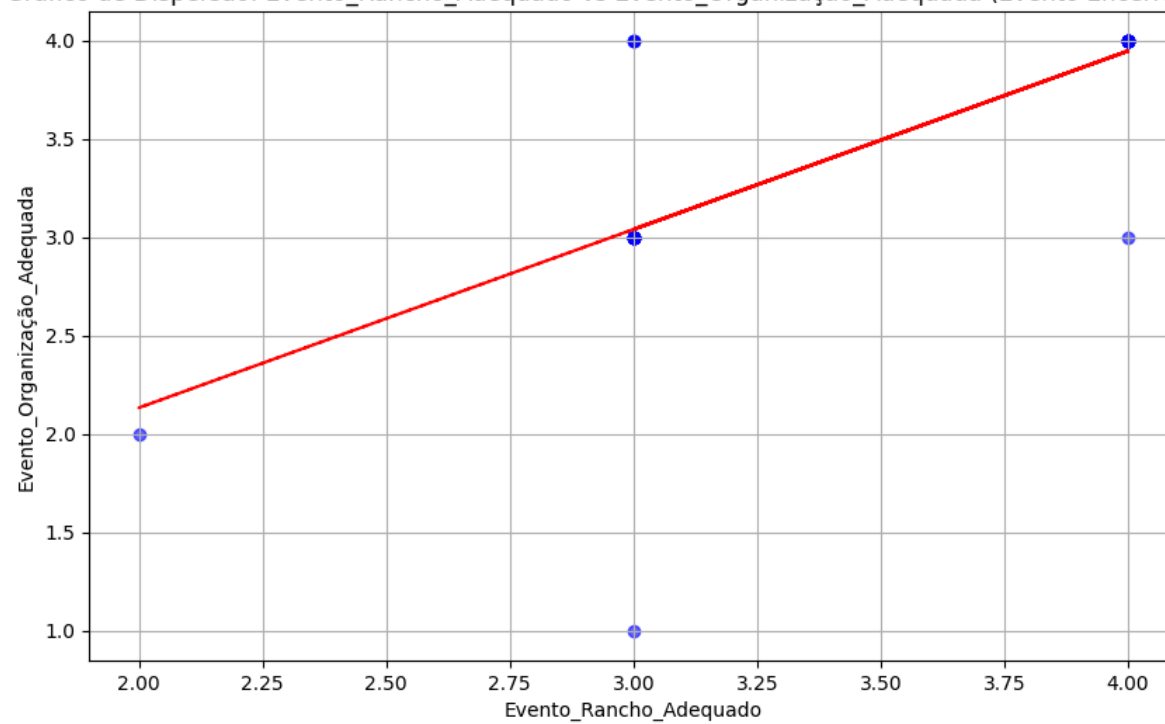


Gráfico de Dispersão: Evento_Rancho_Adequado vs Evento_Atividades_Envolventes (Evento Encerramento)

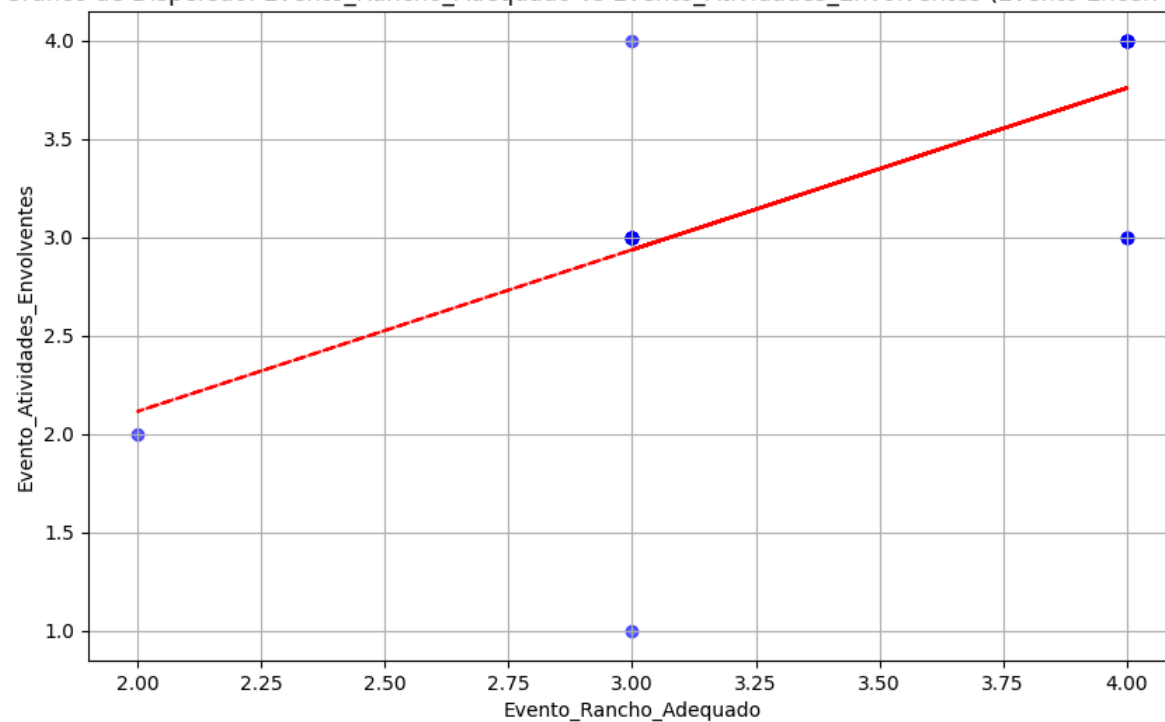


Gráfico de Dispersão: Evento_Rancho_Adequado vs Evento_Reforço_Propósito (Evento Encerramento)

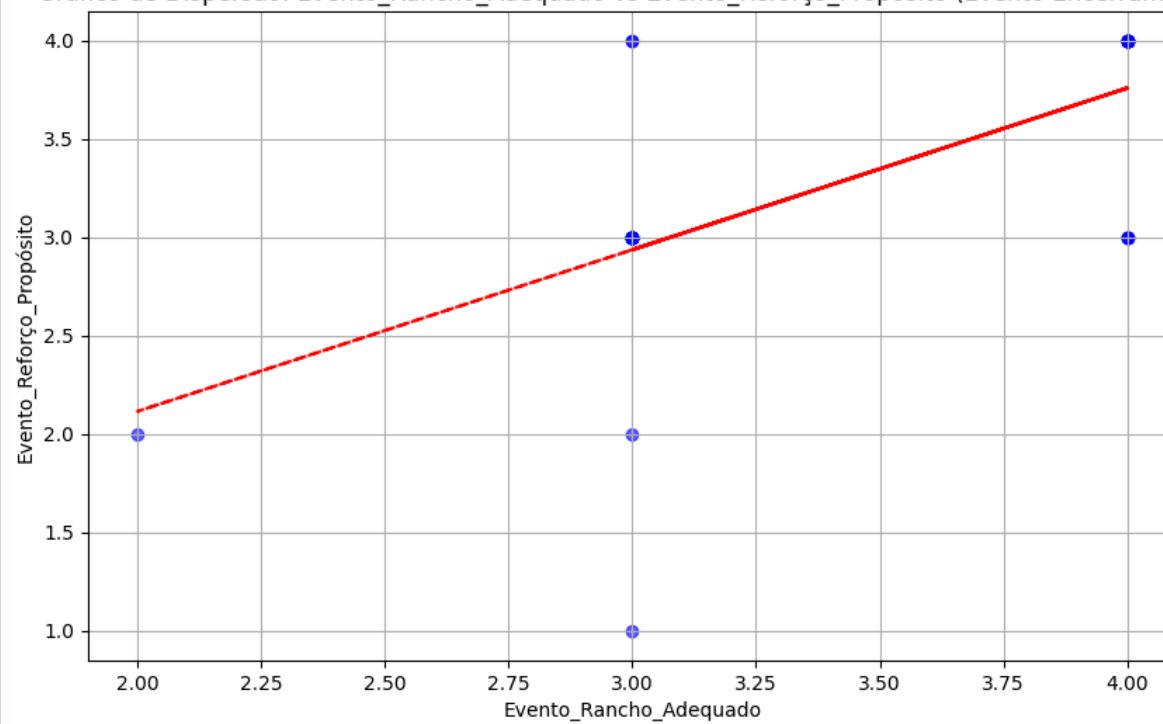


Gráfico de Dispersão: Evento_Rancho_Adequado vs Evento_Alimentação_Adequada (Evento Encerramento)

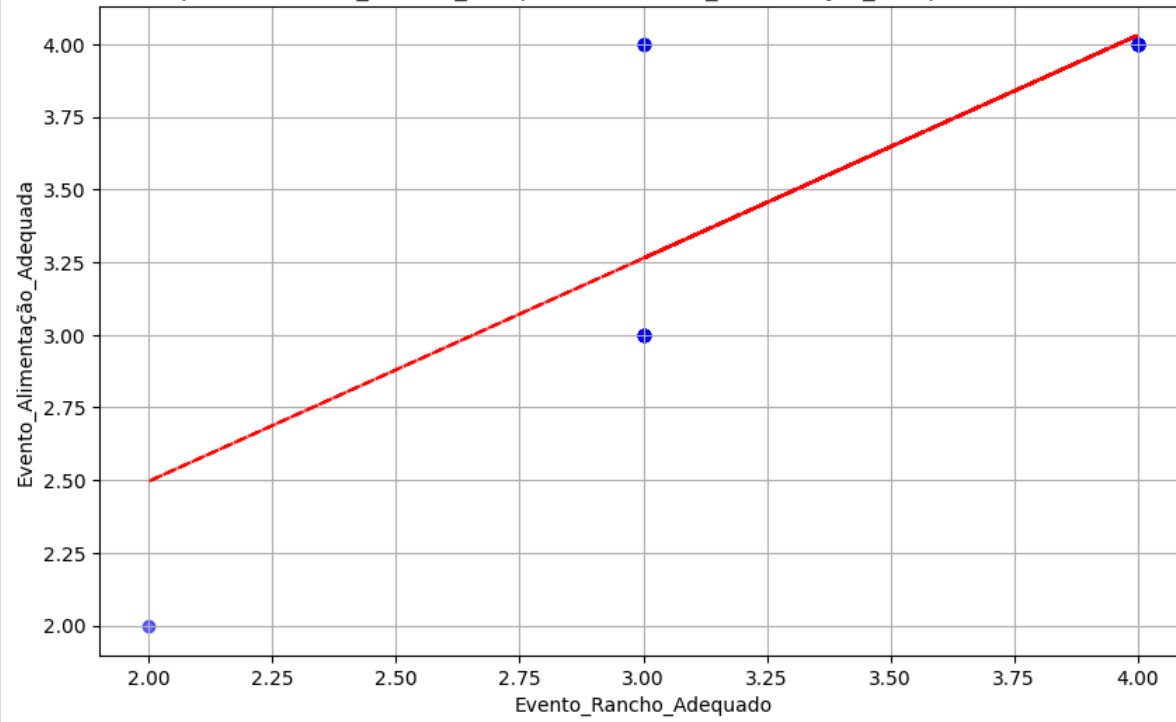


Gráfico de Dispersão: Evento_Rancho_Adequado vs Evento_Satisfação_Geral (Evento Encerramento)

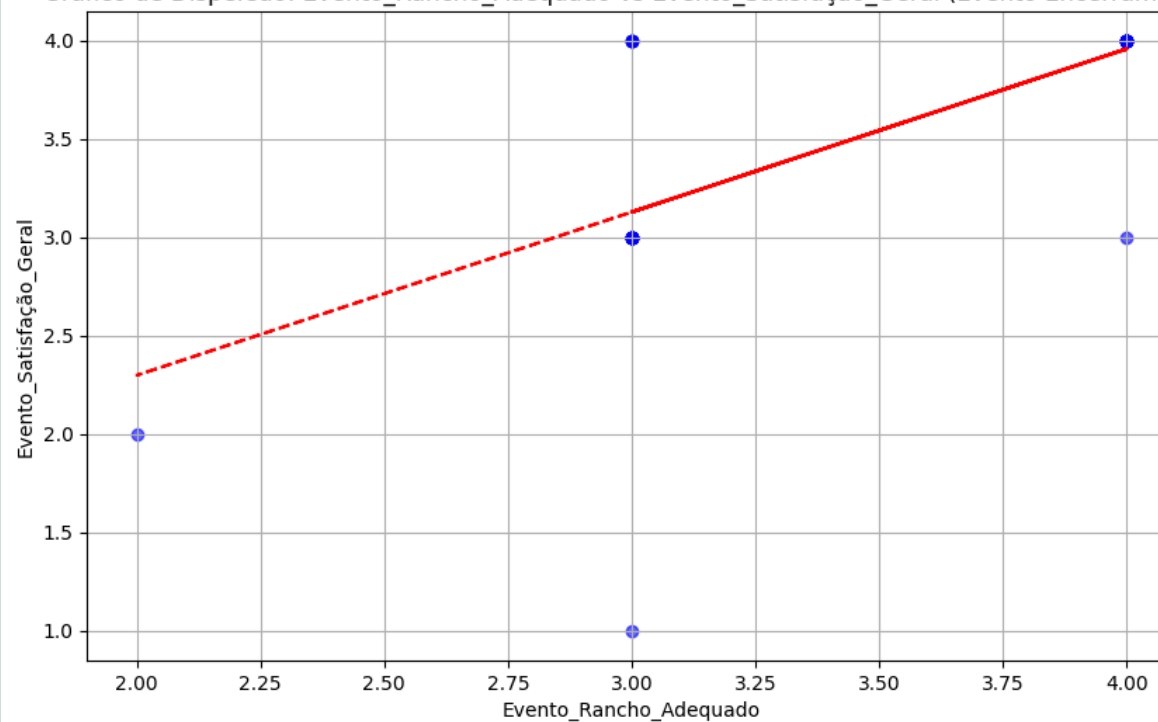


Gráfico de Dispersão: Evento_Comunicação_Eficaz vs Evento_Organização_Adequada (Evento Encerramento)

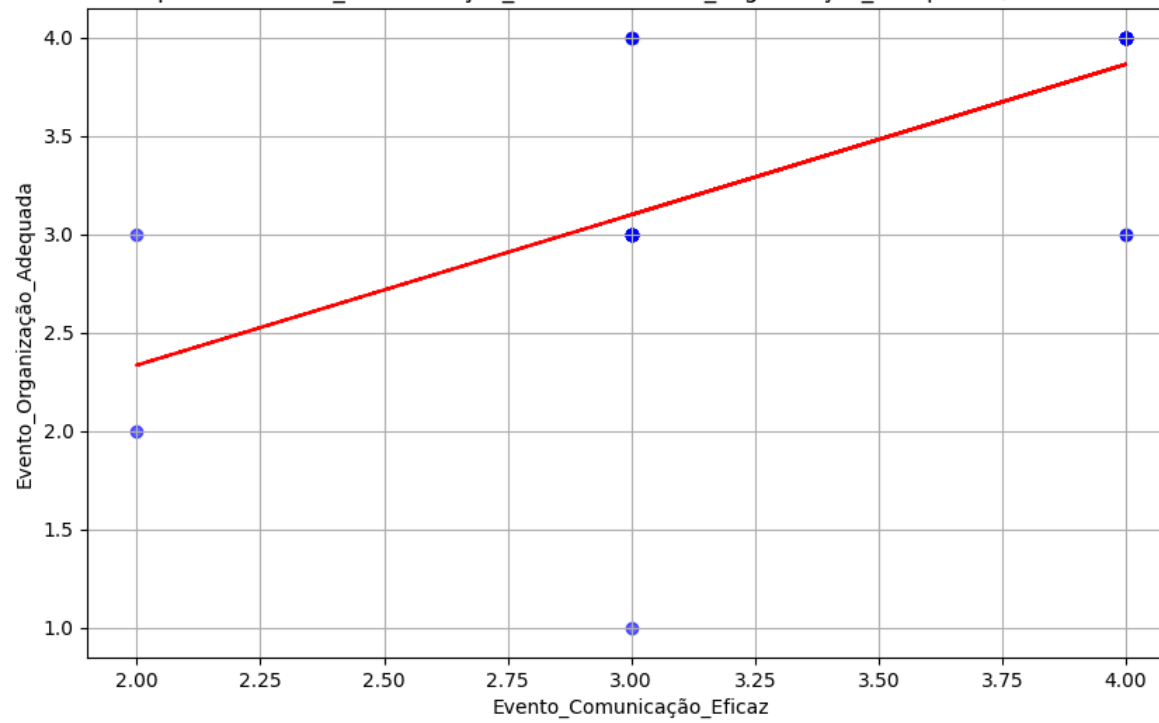


Gráfico de Dispersão: Evento_Comunicação_Eficaz vs Evento_Atividades_Envolventes (Evento Encerramento)

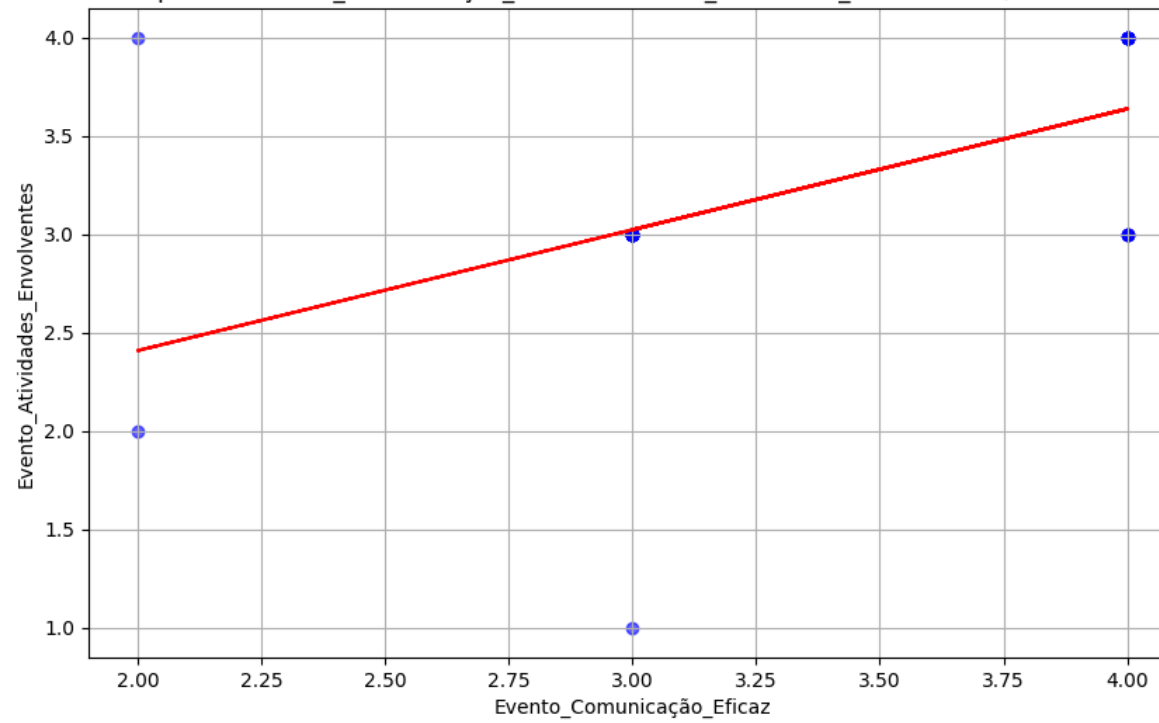


Gráfico de Dispersão: Evento_Comunicação_Eficaz vs Evento_Reforço_Propósito (Evento Encerramento)

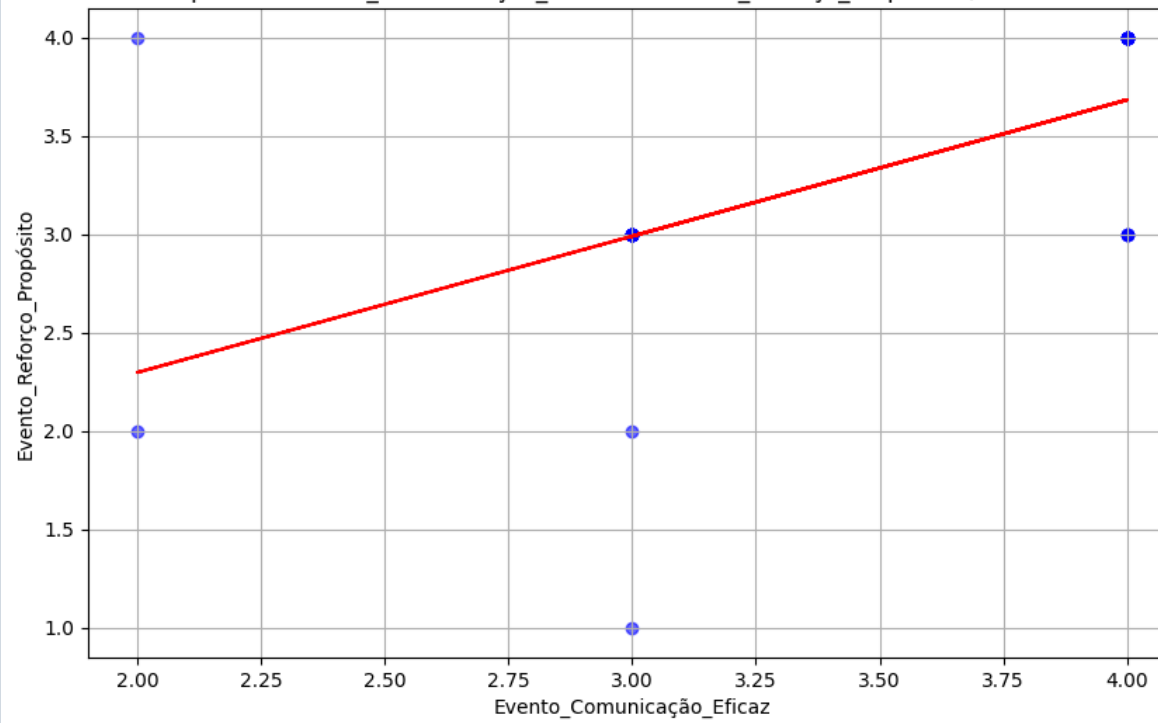


Gráfico de Dispersão: Evento_Comunicação_Eficaz vs Evento_Alimentação_Adequada (Evento Encerramento)

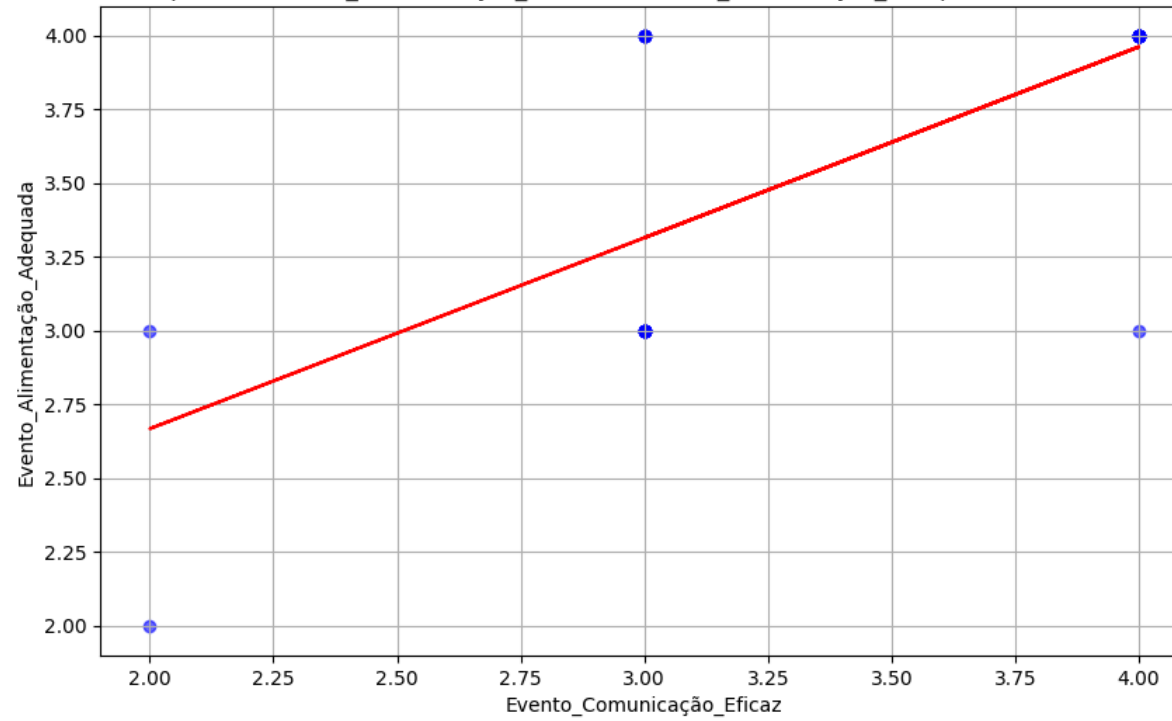


Gráfico de Dispersão: Evento_Comunicação_Eficaz vs Evento_Satisfação_Geral (Evento Encerramento)

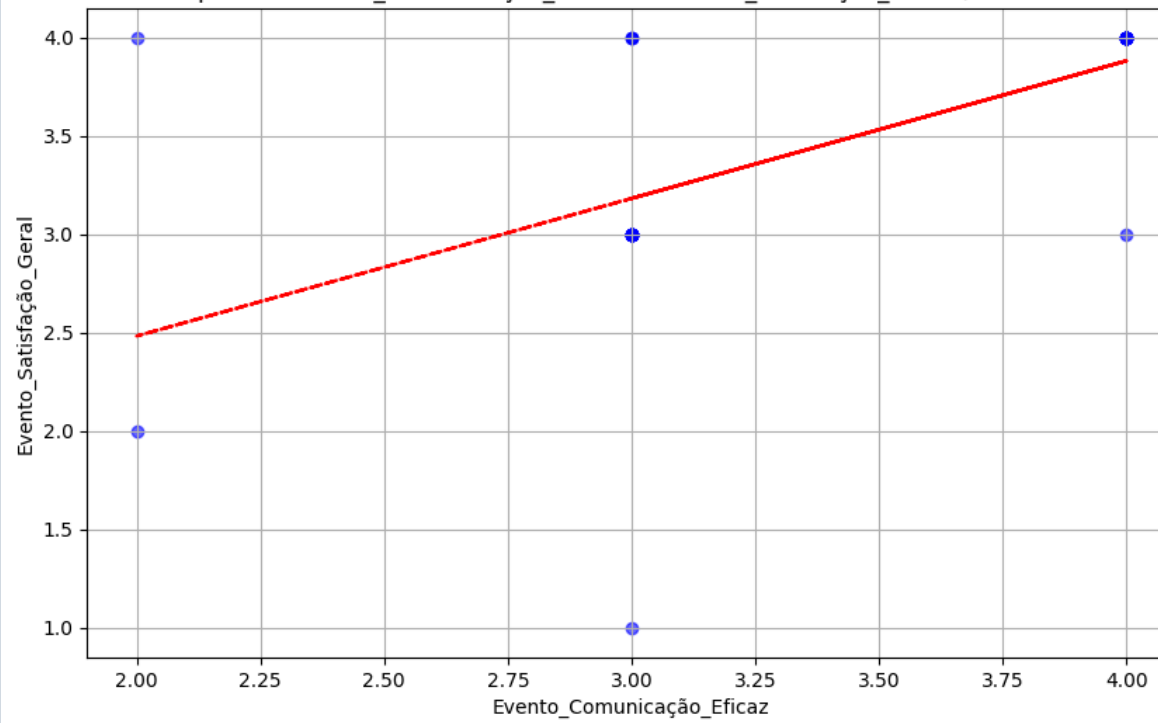


Gráfico de Dispersão: Evento_Organização_Adequada vs Evento_Atividades_Envolventes (Evento Encerramento)

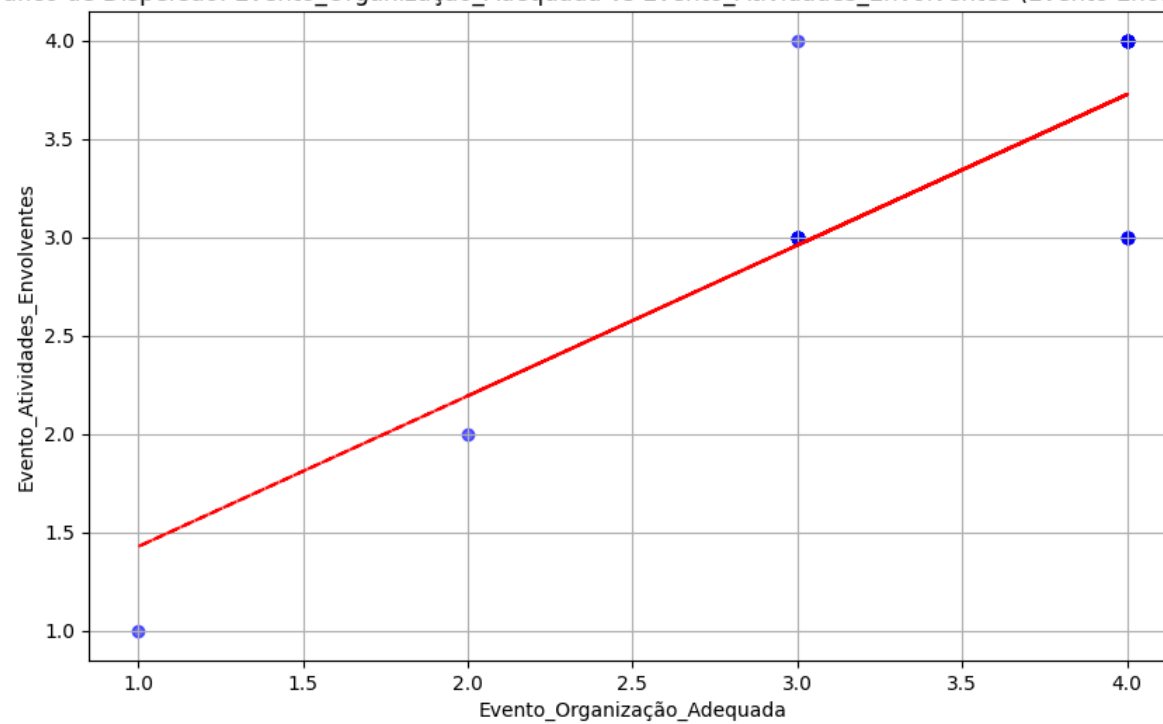


Gráfico de Dispersão: Evento_Organização_Adequada vs Evento_Reforço_Propósito (Evento Encerramento)

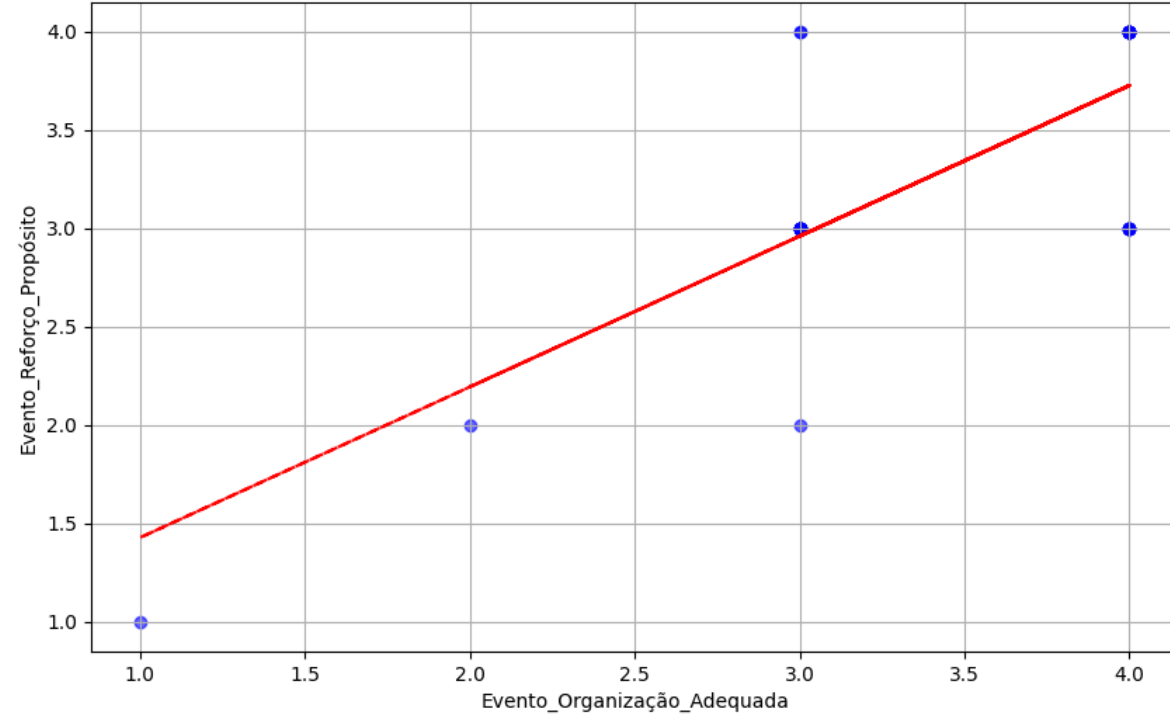


Gráfico de Dispersão: Evento_Organização_Adequada vs Evento_Alimentação_Adequada (Evento Encerramento)

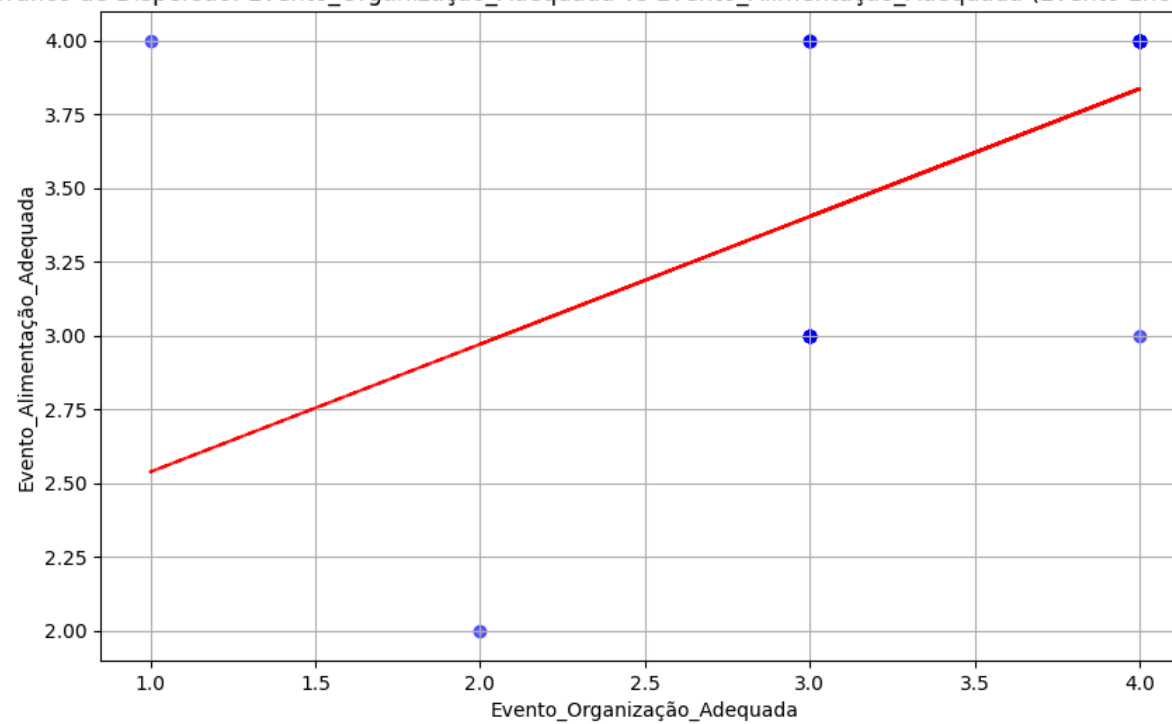


Gráfico de Dispersão: Evento_Organização_Adequada vs Evento_Satisfação_Geral (Evento Encerramento)

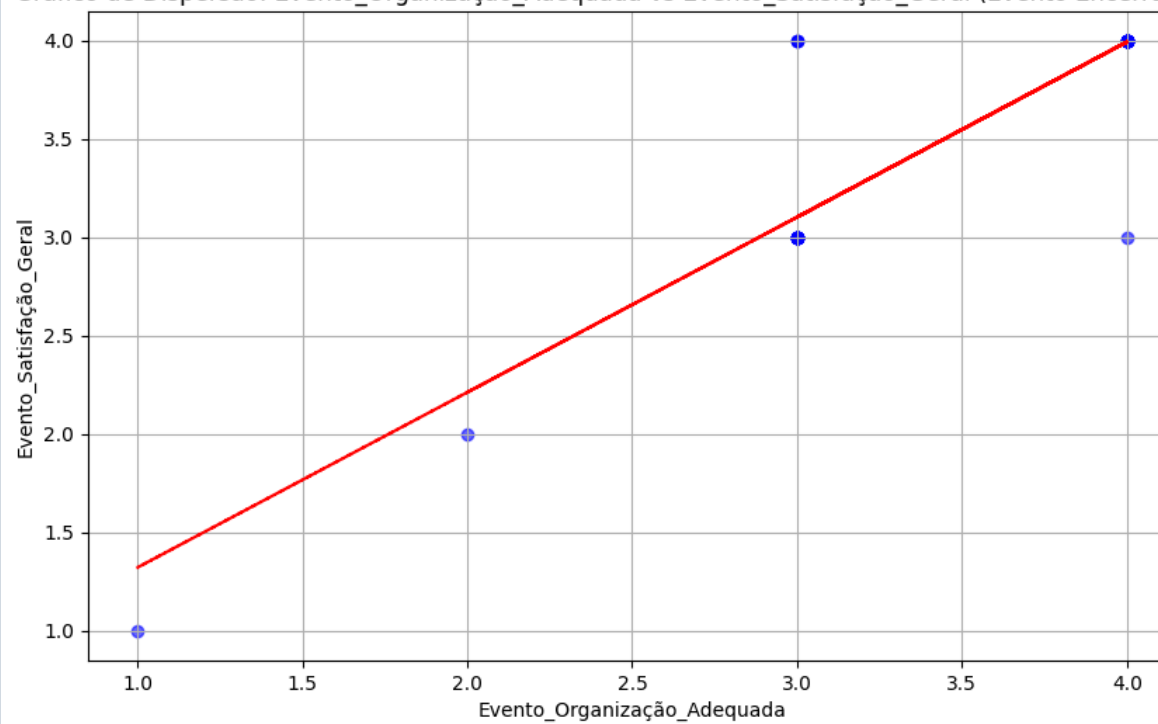


Gráfico de Dispersão: Evento_Atividades_Envolventes vs Evento_Reforço_Propósito (Evento Encerramento)

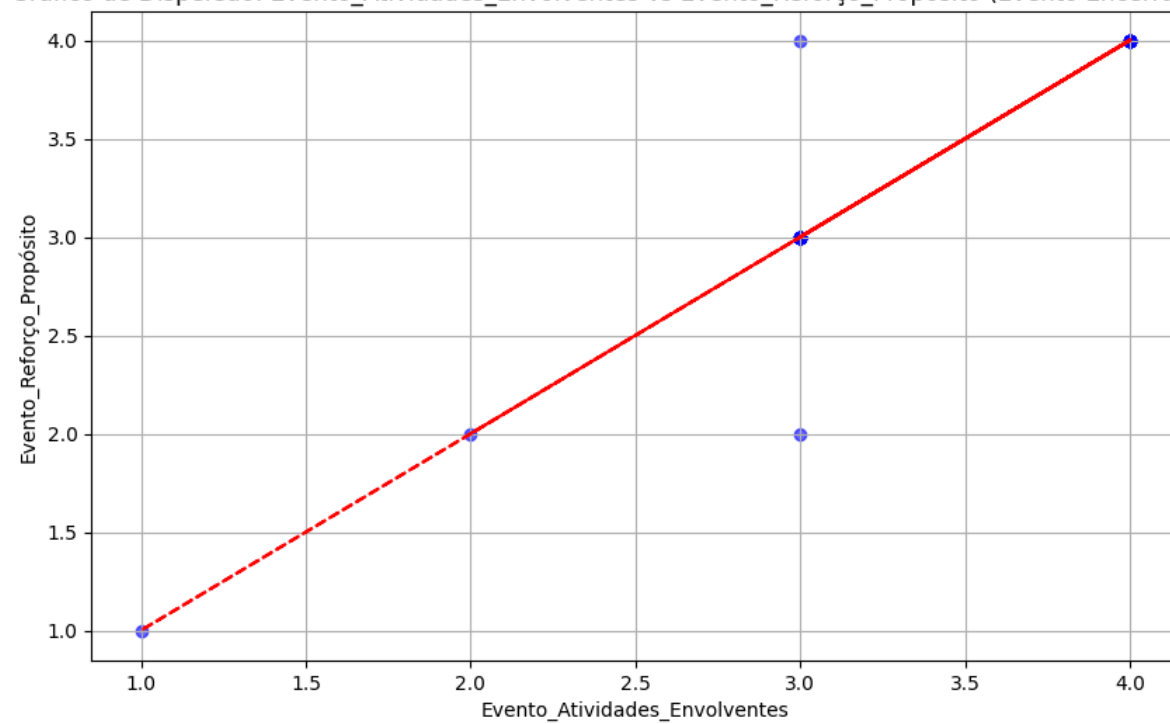


Gráfico de Dispersão: Evento_Atividades_Envolventes vs Evento_Alimentação_Adequada (Evento Encerramento)

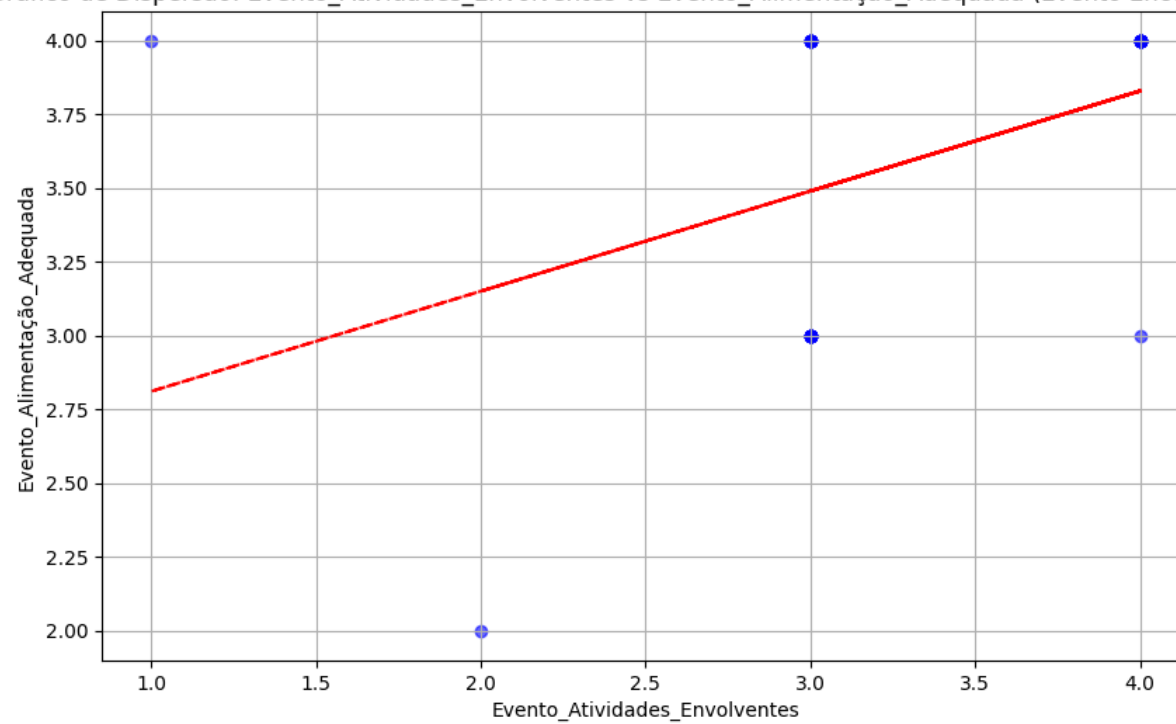


Gráfico de Dispersão: Evento_Atividades_Envolventes vs Evento_Satisfação_Geral (Evento Encerramento)

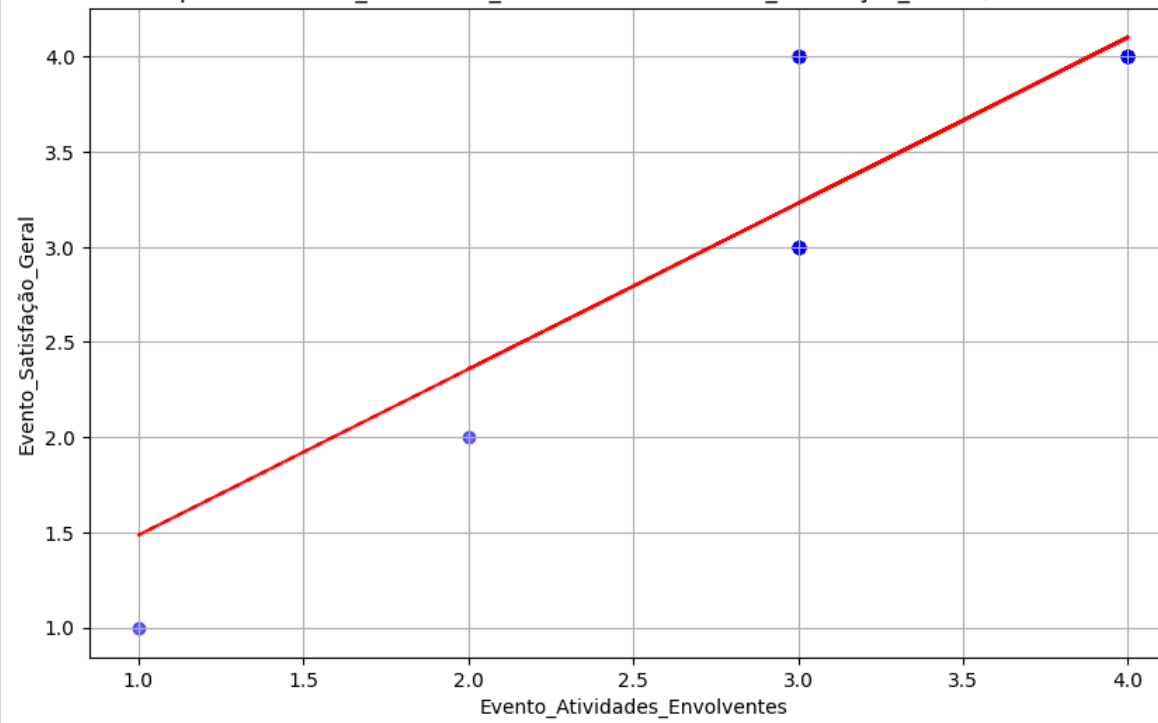


Gráfico de Dispersão: Evento_Reforço_Propósito vs Evento_Alimentação_Adequada (Evento Encerramento)

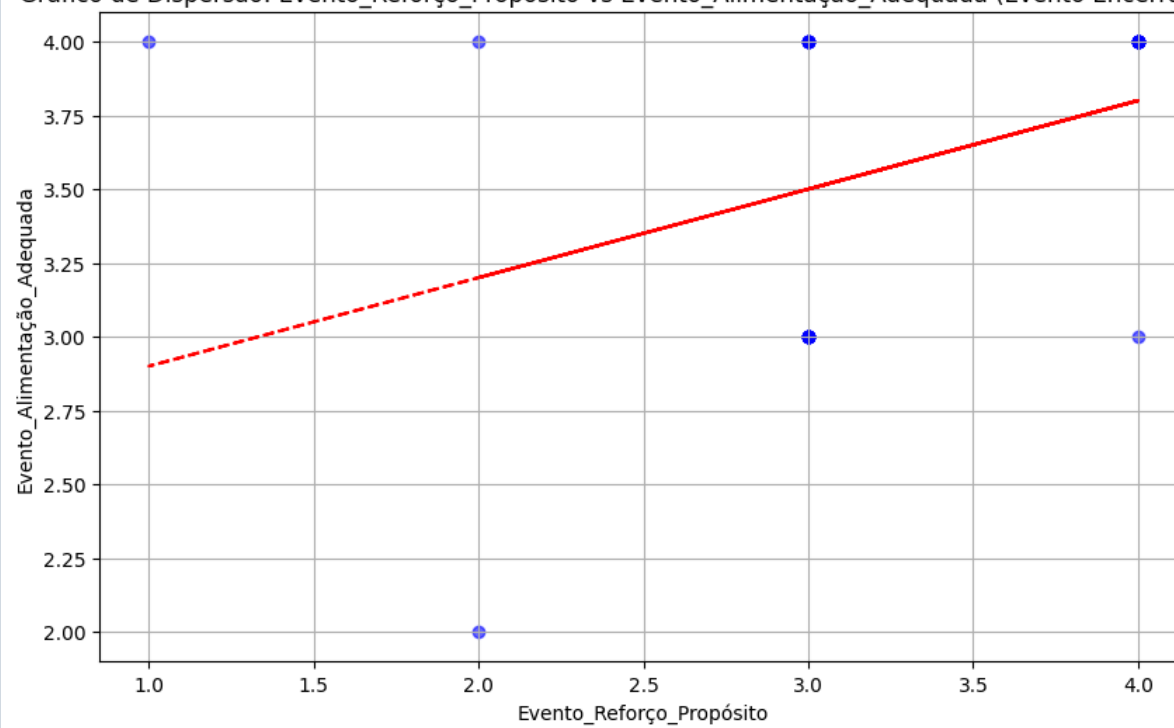


Gráfico de Dispersão: Evento_Reforço_Propósito vs Evento_Satisfação_Geral (Evento Encerramento)

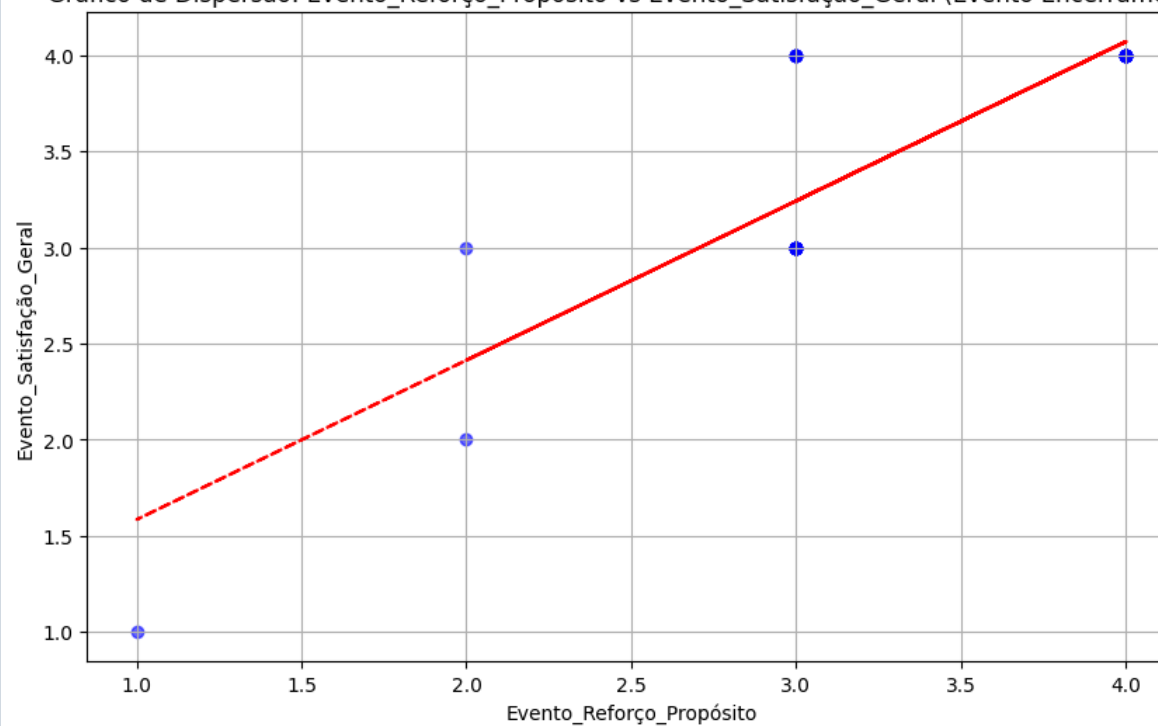


Gráfico de Dispersão: Evento_Alimentação_Adequada vs Evento_Satisfação_Geral (Evento Encerramento)

