

COMPONENTE CURRICULAR:	PROJETO APLICADO I - 51667 - 2024.1		
NOMES COMPLETOS DOS ALUNOS:	Andre Gustavo Monteiro Dos Santos F – RA 10424359		
	Felipe Bernardi Da Silva	– RA 10328856	
	Milan Mirco Moraes Mazur	– RA 10363757	
	Raul Santos Lages	– RA 10424621	

PROJETO APLICADO I

GRUPO: PROJETO APLICADO – MACKENZIE

Sumário

Glossário.....	3
Objetivo do estudo: Apresentação da empresa.....	3
Objetivo do estudo: Problema de Pesquisa.....	4
Cronograma de atividades e responsabilidades.....	7
Apresentação dos metadados e Análise Exploratória de Dados.....	7
Exemplo de saída de código em Python 1 – Fonte dos Valores.....	14
Exemplo de saída de código em Python 2 – Script Inicial.....	15
Exemplo de saída de código em Python 3 – Resultado vs Menor Valor.....	16
Exemplo de saída de código em Python 4 – Resultado vs Menor Valor Saneado.....	18
Exemplo de saída de código em Python 5 – Resultado vs Média.....	20
Exemplo de saída de código em Python 6 – Resultado vs Média Saneada.....	22
Exemplo de saída de código em Python 7 – Resultado vs Mediana.....	24
Exemplo de saída de código em Python 8 – Resultado vs Mediana Saneada.....	26
Exemplo de saída de código em Python 9 – Soma das Discrepâncias.....	28
Exemplo de saída de código em Python 10 – Soma das Discrepâncias.....	29
Exemplo de saída de código em Python 11 – Fonte de Valor vs Resultado.....	32
Data Storytelling.....	34
Grupo.....	37
Github do projeto.....	37
Referências bibliográficas e sites consultados:.....	38

Glossário

Licitação fracassada: licitação sem oferecimento de propostas pelas empresas

Estudo Técnico Preliminar: ETP – documento obrigatório realizado como preparação de licitação em que se analisa o mercado e as possibilidades de atendimento da demanda, considerando os custos envolvidos

NOC: Network Operation Center ou Centro de Operação de Rede

PDTIC: Plano Diretor de Tecnologia da Informação e Comunicação

PEI: Planejamento Estratégico Institucional

TIC: Tecnologia da Informação e Comunicações

TRT2/SP: Tribunal Regional do Trabalho da 2ª Região

Objetivo do estudo: Apresentação da empresa

A empresa escolhida é o Tribunal Regional do Trabalho da 2ª Região – SP, com o estudo se restringindo aos contratos licitados da área de TIC do órgão.

A missão do TRT2/SP como instituição, segundo o Planejamento Estratégico 2021-2026 é “realizar justiça”, a visão é “Poder Judiciário efetivo e ágil na garantia dos direitos e que contribua para a pacificação social e o desenvolvimento do país” e os valores são: “Acessibilidade, agilidade, credibilidade, eficiência, ética, imparcialidade, inovação, integridade, segurança jurídica, sustentabilidade, transparência e responsabilização”.

O PDTIC (Plano Diretor de Tecnologia da Informação e Comunicação) está alinhado ao que foi definido no PEI do TRT2 e detalha o desdobramento dos objetivos institucionais que demandam ações de TIC.

Além do PEI e do PDTIC, existem outros planos no âmbito da TIC há outros planos destinados a contribuir com o alcance dos objetivos do TRT2, como o Plano de Contratações de TIC, que traz uma série de contratos a serem firmados, quase na

totalidade, por meio de licitações públicas que demandam a estimativa de custos como fase obrigatória da sua preparação.

O TRT2/SP é o maior Tribunal Regional do Trabalho do país, apesar de abranger apenas a área da Grande São Paulo e da Baixada Santista, possuindo cerca de 20% das ações trabalhistas do Brasil. Em virtude da intensa informatização das atividades, possui muitos contratos ligados à área de TIC (Tecnologia da Informação e Comunicações) – cerca de 80 (oitenta) contratos vigentes – e realiza um grande número de contratações anuais, muitas vezes nacionais (em que diversos órgãos públicos diferentes compram em conjunto). Como mencionado anteriormente, e por se tratar de órgão público, realiza compras de produtos e serviços majoritariamente por meio de licitações.

O número total de pessoas que desempenham suas funções no TRT2/SP é de aproximadamente 7.000 pessoas entre juízes e servidores, sendo que a área de TIC do Tribunal é composta por cerca de 160 servidores efetivos, além de cerca de 80 trabalhadores terceirizados atuando em postos de atendimento no service desk e em outras funções como monitoramento 24x7 (NOC).

Os dados relativos aos processos de compras são públicos e facilmente obtidos por meio de sistemas públicos (p.ex. compras governamentais <https://www.gov.br/compras/pt-br/aceso-a-informacao/consulta-detalhada> e ainda na aba “transparência” do site oficial do Tribunal – <https://ww2.trt2.jus.br/transparencia/licitacoes-compras-e-contratos/licitacoes>).

As iniciativas na área de Data Science estão concentradas em processos judiciais, notadamente no PJe (Processo Judicial Eletrônico), e presentes em portais como do CNJ (ver links nas Referências bibliográficas e Sites consultados).

Objetivo do estudo: Problema de Pesquisa

Considerando as regras do processo licitatório, há necessidade de realizar estimativas de custo dos produtos ou serviços a serem adquiridos ou contratados, em etapa preparatória da compra. Assim, por exemplo, há contratos de serviços como Service Desk, manutenção de Switches, de Servidores, de Desktops, links de comunicação dedicados, etc, assim como há aquisições de equipamentos como soluções de hiperconvergência, notebooks, entre outros.

Há questões que atraem grande atenção e preocupação nos processos de aquisições e de contratações e uma delas é a questão da estimativa de custo. Isso porque o valor estimado gera grandes consequências, como a reserva do valor no orçamento previsto, que pode levar a repasseamentos orçamentários desnecessários caso se estime um valor mais alto do que deveria, além de servir como parâmetro de aceitação da proposta no momento da licitação, de forma que se estimado abaixo do que deveria acaba ocasionando a falta de proponentes no certame (licitação deserta). Outro problema é a contratação acima do valor de mercado, mesmo em caso de licitação, que pode ocorrer se a estimativa não estiver bem calculada e consentânea com os valores de mercado.

Não há, considerando as decisões e recomendações dos órgãos de controle, uma forma única de cálculo da estimativa; entretanto, é consenso que ela deva ser heterogênea, e valer-se de resultados de outros certames como forma de se aproximar o máximo possível do valor real, que se espera obter no resultado da licitação. Aceita-se, assim, que os órgãos tenham alguma liberdade na escolha da definição dessa estimativa, razão pela qual surgem diversas tentativas relacionadas à utilização de medianas saneadas ou outras métricas valendo-se de métodos estatísticos.

Propõe-se, desta maneira, o estudo (análise exploratória) dos dados de licitações no ano de 2023 (período ajustado no Encontro Síncrono conforme ponderações do Professor Felipe Albino dos Santos) pois já se encontram completos os dados relativos a esse ano em comparação com 2024, permitindo assim que a massa de dados seja maior.

Essas compras de TIC do TRT2/SP serão analisadas no sentido de obter informações sobre as relações entre as possíveis formas de cálculo da estimativa comparando-as com os resultados das licitações para permitir a busca por uma solução adequada para o equacionamento dos problemas apresentados (elaboração de uma estimativa mais precisa, feita por meio de um programa em Python que reflita a forma de cálculo mais adequada conforme resultado das análises), tudo por meio das ferramentas disponibilizadas ao cientista de dados. Nesse particular daremos especial atenção a modelo de apuração de mediana saneada utilizada recentemente no Governo Federal, que utiliza a seguinte solução (que será aplicada aos valores obtidos na fase preliminar da licitação e comparada, com outras formas de obtenção da estimativa e também com os dados dos resultados das licitações):

Passo 1) $CV = \text{Coeficiente de Variação de uma amostra} / \text{Coeficiente de Pearson} / \text{Mede a homogeneidade da amostra}$

$CV = \text{Desvio Padrão (A)} / \text{Média (A)}$

$CV \geq 25\% \rightarrow$ Amostra não homogênea. // Descartes de preços discrepantes.

$CV < 25\% \rightarrow$ Amostra homogênea.

Passo 2) $Score Z = (P1 - Média(A)) / DP$.

O maior dos valores em magnitude é descartado.

Passo 3) Repete passo 1.

Parada: ao se obter os três últimos valores ou quando o CV for menor que 25%.

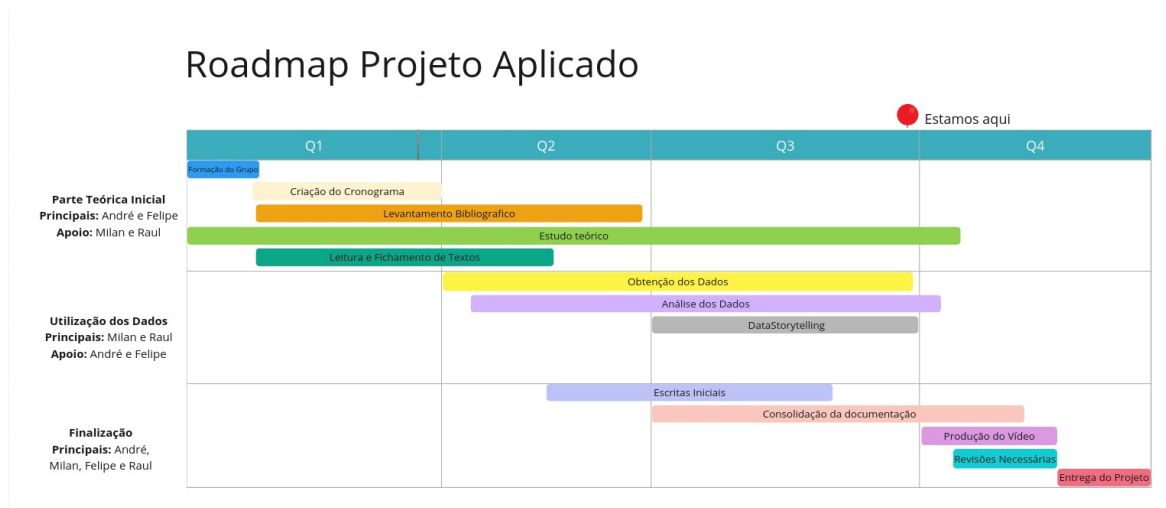
Assim pode-se afirmar que os objetivos gerais são:

- Alcançar maior precisão nas estimativas realizadas em processos licitatórios em órgãos públicos;
- Automatizar, por meio de software, a tarefa de tratar os valores obtidos no sentido de alcançar uma estimativa mais precisa;
- Comparar as diversas possibilidades de estimativa com os resultados das licitações públicas para permitir que se possa traçar estratégias;

E os objetivos específicos são:

- Conseguir uma base de dados representativa de um ano completo de licitações públicas, incluindo valores utilizados para as estimativas e também os resultados de licitações e realizar análise exploratória dos dados;
- Produzir um software em python que permita o descarte de valores outliers, valendo-se de recursos estatísticos como desvio padrão, mediana, média, Score Z, entre outros;
- Ajustar o software de maneira a considerar eventuais fatores de correção aplicáveis aos valores obtidos em fase de estimativa (como p.ex. propostas comerciais e outros) no sentido de aproximar a estimativa dos valores obtidos nos resultados dos certames (considerando as peculiaridades das compras públicas, p.ex. índice de licitações desertas ou fracassadas que podem indicar eventual falta de ajuste na estimativa);

Cronograma de atividades e responsabilidades



Apresentação dos metadados e Análise Exploratória de Dados

Conforme mencionado os dados relativos aos processos de compras são públicos e facilmente obtidos por meio de sistemas públicos (p.ex. compras governamentais e ainda na aba “transparência” do site oficial do Tribunal – www.trt2.jus.br).

O dataset está em formato xlsx, não havendo dados sensíveis e/ou de uso restrito pois são todos processos de licitação públicos.

Propõe-se basicamente três tabelas para os dados:

Tabela 1

Objeto	Valor	Fonte
Subscricoes_Linux	25333	Licitacao_terceiros
Subscricoes_Linux	27588	Licitacao_anterior
Subscricoes_Linux	33097	Proposta_Comercial
Subscricoes_Linux	35097	Proposta_Comercial
Hub_USB	23950	Proposta_Comercial
Hub_USB	22750	Proposta_Comercial
Hub_USB	26250	Proposta_Comercial
Hub_USB	40984	Proposta_Comercial
Hub_USB	32500	Licitacao_terceiros
Internet_Movel	309753	Licitacao_anterior
Link_Internet_Cubatao	1605	Licitacao_anterior
Link_Internet_Cubatao	2912	Licitacao_anterior
Link_Internet_Cubatao	3000	Proposta_Comercial
Service_Desk	22535077	Proposta_Comercial
Service_Desk	28165024	Proposta_Comercial
Service_Desk	21754548	Proposta_Comercial
Service_Desk	29667092	Proposta_Comercial
Impressora_Cracha	122760	Proposta_Comercial
Impressora_Cracha	123362	Proposta_Comercial
Impressora_Cracha	108640	Proposta_Comercial
Impressora_Cracha	96760	Licitacao_terceiros
Impressora_Cracha	100760	Licitacao_terceiros
Certificados_Digitais	2979	Proposta_Comercial
Certificados_Digitais	5128	Proposta_Comercial
Certificados_Digitais	1420	Proposta_Comercial
Certificados_Digitais	1885	Proposta_Comercial
Certificados_Digitais	1884	Licitacao_terceiros
Firewall	16010254	Proposta_Comercial

Firewall	27343039	Proposta_Comercial
Firewall	42113807	Proposta_Comercial
Suporte_Switches_SAN	460794	Proposta_Comercial
Suporte_Switches_SAN	233449	Proposta_Comercial
Suporte_Switches_SAN	147827	Proposta_Comercial
Suporte_VTL	395165	Proposta_Comercial
Suporte_VTL	741600	Proposta_Comercial
Suporte_VTL	415610	Licitacao_terceiros
Spectrum_Protect	75434	Proposta_Comercial
Spectrum_Protect	133000	Proposta_Comercial
Spectrum_Protect	80791	Licitacao_anterior
Spectrum_Protect	69928	Licitacao_terceiros
Licenças_Office	1800915	Licitacao_terceiros
Licenças_Office	1468500	Proposta_Comercial
Criptografia_fitras_IBM	32826	Proposta_Comercial
Criptografia_fitras_IBM	32148	Proposta_Comercial
Criptografia_fitras_IBM	55920	Proposta_Comercial
Criptografia_fitras_IBM	76360	Proposta_Comercial
Suporte_Switches_Borda	513188	Proposta_Comercial
Suporte_Switches_Borda	528583	Proposta_Comercial
Suporte_Switches_Core_Topo	498479	Proposta_Comercial
Suporte_Switches_Core_Topo	565907	Proposta_Comercial
Licencas_Microsoft_SQL	1594906	Proposta_Comercial
Licencas_Microsoft_SQL	1715199	Proposta_Comercial
Licencas_Microsoft_SQL	1024801	Licitacao_terceiros

Tabela 2

Objeto	Result_Licitacao	Min_Valor_pre_resultado	Min_valor_saneado	Media_pre	Media_saneada_pre	Mediana_pre	Mediana_saneada_pre
Subscricoes_Linux	28780	25333	25333	30278	30278	30342	30342
Hub_USB	8250	22750	22750	29286	29286	26250	26250
Internet_Movel	245310	309753	309753	309753	309753	309753	309753
Link_Internet_Cubatao	3000	1605	1605	2505	2505	2912	2912
Service_Desk	13889987	21754548	21754548	25530435	25530435	25350050	25350050
Impressora_Cracha	81880	96760	96760	110456	110456	108640	108640
Certificados_Digitais	986	1420	1420	2659	1729	1885	1884
Firewall	24371018	16010254	16010254	28489033	28489033	27343039	27343039
Suporte_Switches_SAN	203999	147827	147827	280690	280690	233449	233449
Suporte_VTL	379000	395165	395165	517458	517458	415610	415610
Spectrum_Protect	70087	69928	69928	89788	75384	78112	75434
Licenças_Office	1507085	1468500	1468500	1634707	1634707	1634707	1634707
Criptografia_fitras_IBM	19990	32148	32148	49313	40298	44373	32826
Suporte_Switches_Borda	364813	513188	513188	520890	520890	520890	520890
Suporte_Switches_Core_Topo	476781	498479	498479	532193	532193	532193	532193
Licencas_Microsoft_SQL	1043231	1024801	1024801	1444968	1444968	1594906	1594906

Tabela 3

Tipo_de_calculo	Soma_discrepancias
Menor Valor	-46,41
Menor Valor Saneado	-46,41
Média	-438,13
Média Saneada	-394,2
Mediana	-389,32
Mediana Saneada	-370,26

A Tabela 1 traz 53 linhas e 3 colunas. As colunas são “Objeto”, “Valor” e “Fonte” e já possuem todos os dados completos. Cada linha da coluna “Objeto” corresponde a um valor de certo objeto de licitação, sendo que cada valor da coluna “Valor”, possui valores correspondentes a Reais, utilizado/utilizável na estimativa de custos e proveniente de uma certa fonte, conforme coluna “Fonte”.

Os dados constantes da Tabela1 já permitem algumas análises descritivas, como por exemplo a proporção de cada fonte de valor (ver gráfico de fontes de valores mais adiante) utilizada e a quantidade de valores obtidos para cada item (“objeto de licitação”).

O cruzamento dos dados obtidos a partir da Tabela 1, com os dados da coluna “Resultado” da Tabela 2 (com 16 linhas e 8 colunas), correspondentes aos resultados das licitações (em valores correspondentes a Reais), permitirão que se observem padrões que podem levar a soluções para a melhora das estimativas.

Assim, por meio de um programa em Python (Código no GitHub), formulou-se médias e médias saneadas (com a exclusão de outliers conforme método já apresentado), medianas e medianas saneadas (com a exclusão de outliers conforme método já apresentado), além de menor preço e menor preço saneado (com a exclusão de outliers conforme método já apresentado), para permitir novas análises de dados e verificar qual método tem menor discrepância em relação aos resultados das licitações, e ao final propor a adoção de um deles, ainda que se utilize algum fator de correção (a ser apurado). Tudo isso, com a entrega de um software formulado em Python que fornecerá um texto pronto para eventual uso nos Estudos Técnicos Preliminares e que contará com um gráfico com as propostas utilizadas.

O código formulado em Python, é uma conversão de planilha usada no Poder Executivo Federal, e será usado inicialmente para obter simulações de valores e testar hipóteses e, após o resultado da análise dos dados poderá eventualmente sofrer “recalibração” e para permitir adequação às conclusões deste Projeto (nos termos acima descritos).

Assim, neste momento do projeto, e após todas as iterações realizadas pelo programa em cada objeto realizadas conforme constante dados constantes da Tabela 2, parte-se para uma nova análise no sentido de se comparar todas as formas de obtenção de estimativa para verificar qual mais se aproximou dos resultados da licitação e, nesse sentido permitir eventuais ajustes no programa ora proposto. A seguir temos dois exemplos de utilização de Python, o primeiro para preparação dos dados a serem

analisados e o segundo com nova análise sobre os dados preparados, sendo que novas análises serão feitas com a finalidade de obtenção das diferenças de precisão na formulação da estimativa entre todas as formas propostas constantes da Tabela 2. Isso é feito na Tabela 3, que possui a soma de todas as discrepâncias obtidas conforme o tipo de cálculo da estimativa (Menor Valor, Menor Valor Saneado, Média, Média Saneada, Mediana e Mediana Saneada).

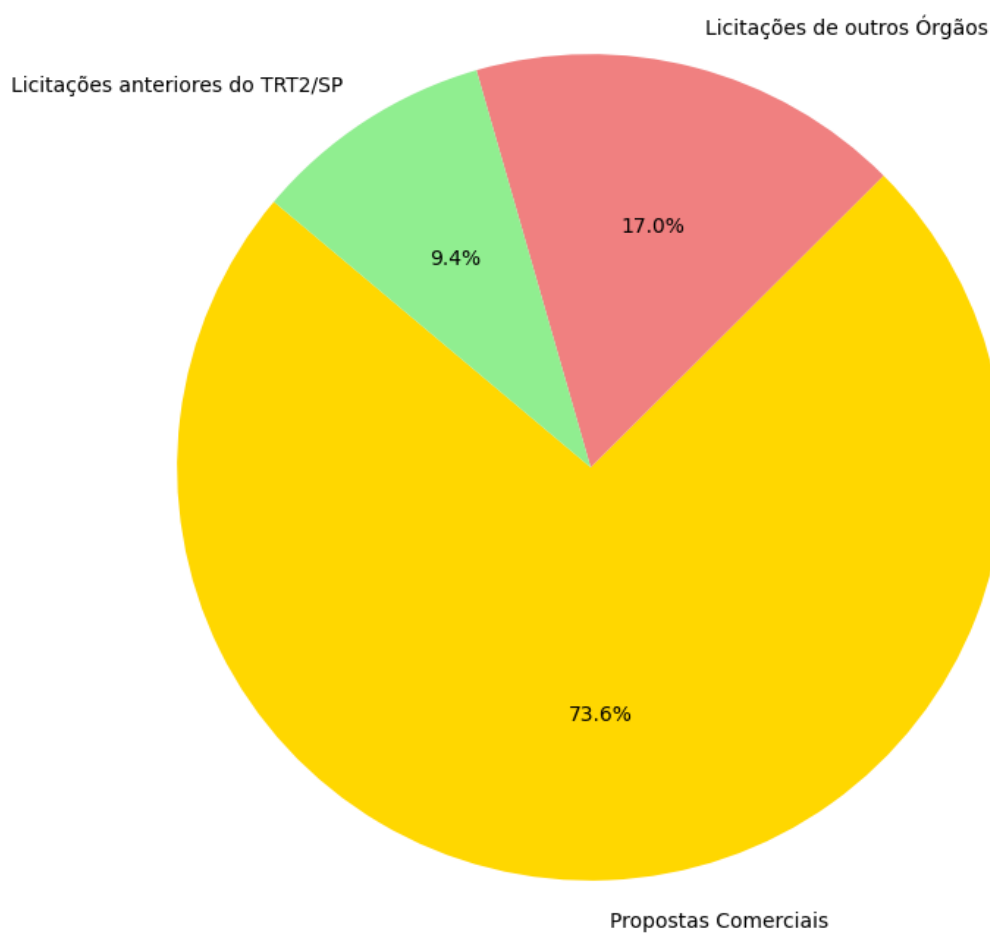
Após, essas análises pode-se ainda imaginar outras análises, como por exemplo a filtragem de cada uma das fontes dos valores (propostas comerciais, licitações de outros órgãos ou licitações anteriores do TRT2/SP) e a comparação de cada grupo de fontes com os resultados das licitações para apurar se algum grupo de fonte de valores possui discrepância menor ou maior que outro(s) grupo(s) de fontes de valores.

Exemplo de saída de código em Python 1 – Fonte dos Valores

Nome do código: Distribuicao_por_fonte.py

Dados utilizados: Coluna “Fonte” (Tabela 1)

Distribuição das fontes de valores obtidos antes da licitação



Exemplo de saída de código em Python 2 – Script Inicial

Nome do código: Analisa Estimativa por Objeto - v1

Dados utilizados: Licitação de Spectrum_Protect (Tabela 1)

TEXTO:

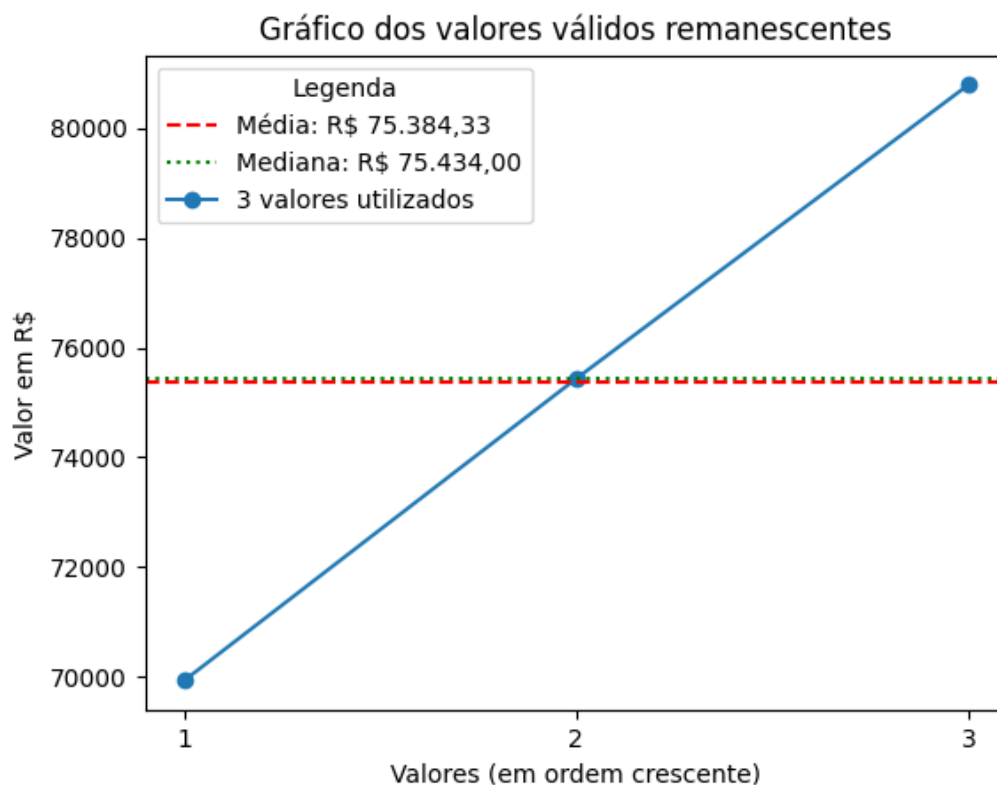
Os valores fornecidos, em número de 4, são, em ordem crescente: R\$ 69.928,00, R\$ 75.434,00, R\$ 80.791,00, R\$ 133.000,00. O desvio padrão resulta em R\$ 25.242,23. A média dos valores é de R\$ 89.788,25, a mediana corresponde a R\$ 78.112,50 e o coeficiente de variação é de 0.28.

O resultado da divisão das diferenças entre os valores e a média dos valores, pelo desvio padrão, resulta, para cada valor, na ordem crescente, em: -0.7867867951648486, -0.5686602310894893, -0.35643647450545357, 1.7118835007597915 de forma que os módulos obtidos são: 0.7867867951648486, 0.5686602310894893, 0.35643647450545357, 1.7118835007597915.

Tendo em vista que houve um coeficiente de variação geral maior que 0.25, o valor mais discrepante, que é de R\$ 133.000,00, será descartado.

Os valores remanescentes, em número de 3, são, em ordem crescente: R\$ 69.928,00, R\$ 75.434,00, R\$ 80.791,00. O desvio padrão resulta em R\$ 4.434,94. A média dos valores é de R\$ 75.384,33, a mediana corresponde a R\$ 75.434,00 e o coeficiente de variação é de 0.06. O índice encontra-se dentro dos parâmetros desejados e não é recomendado desconsiderar qualquer valor ora elencado.

GRÁFICO:



Exemplo de saída de código em Python 3 – Resultado vs Menor Valor

Nome do código: Dif_e_discrep_resultado_menorvalor

Dados utilizados: Resultados_de_Licitacoes_e_Conjecturas_v2
(Tabela 2)

TEXTO:

Diferença entre o Resultado e o Menor valor pré resultado da licitação e Discrepância:

Objeto: Subscricoes_Linux, Diferença: R\$ 3.447,00, Percentual de Discrepância: 13.61%

Objeto: Hub_USB, Diferença: R\$ -14.500,00, Percentual de Discrepância: -63.74%

Objeto: Internet_Movel, Diferença: R\$ -64.443,00, Percentual de Discrepância: -20.80%

Objeto: Link_Internet_Cubatao, Diferença: R\$ 1.395,00, Percentual de Discrepância: 86.92%

Objeto: Service_Desk, Diferença: R\$ -7.864.561,00, Percentual de Discrepância: -36.15%

Objeto: Impressora_Cracha, Diferença: R\$ -14.880,00, Percentual de Discrepância: -15.38%

Objeto: Certificados_Digitais, Diferença: R\$ -434,00, Percentual de Discrepância: -30.56%

Objeto: Firewall, Diferença: R\$ 8.360.764,00, Percentual de Discrepância: 52.22%

Objeto: Suporte_Switches_SAN, Diferença: R\$ 56.172,00, Percentual de Discrepância: 38.00%

Objeto: Suporte_VTL, Diferença: R\$ -16.165,00, Percentual de Discrepância: -4.09%

Objeto: Spectrum_Protect, Diferença: R\$ 159,00, Percentual de Discrepância: 0.23%

Objeto: Licenças_Office, Diferença: R\$ 38.585,00, Percentual de Discrepância: 2.63%

Objeto: Criptografia_fitras_IBM, Diferença: R\$ -12.158,00, Percentual de Discrepância: -37.82%

Objeto: Suporte_Switches_Borda, Diferença: R\$ -148.375,00, Percentual de Discrepância: -28.91%

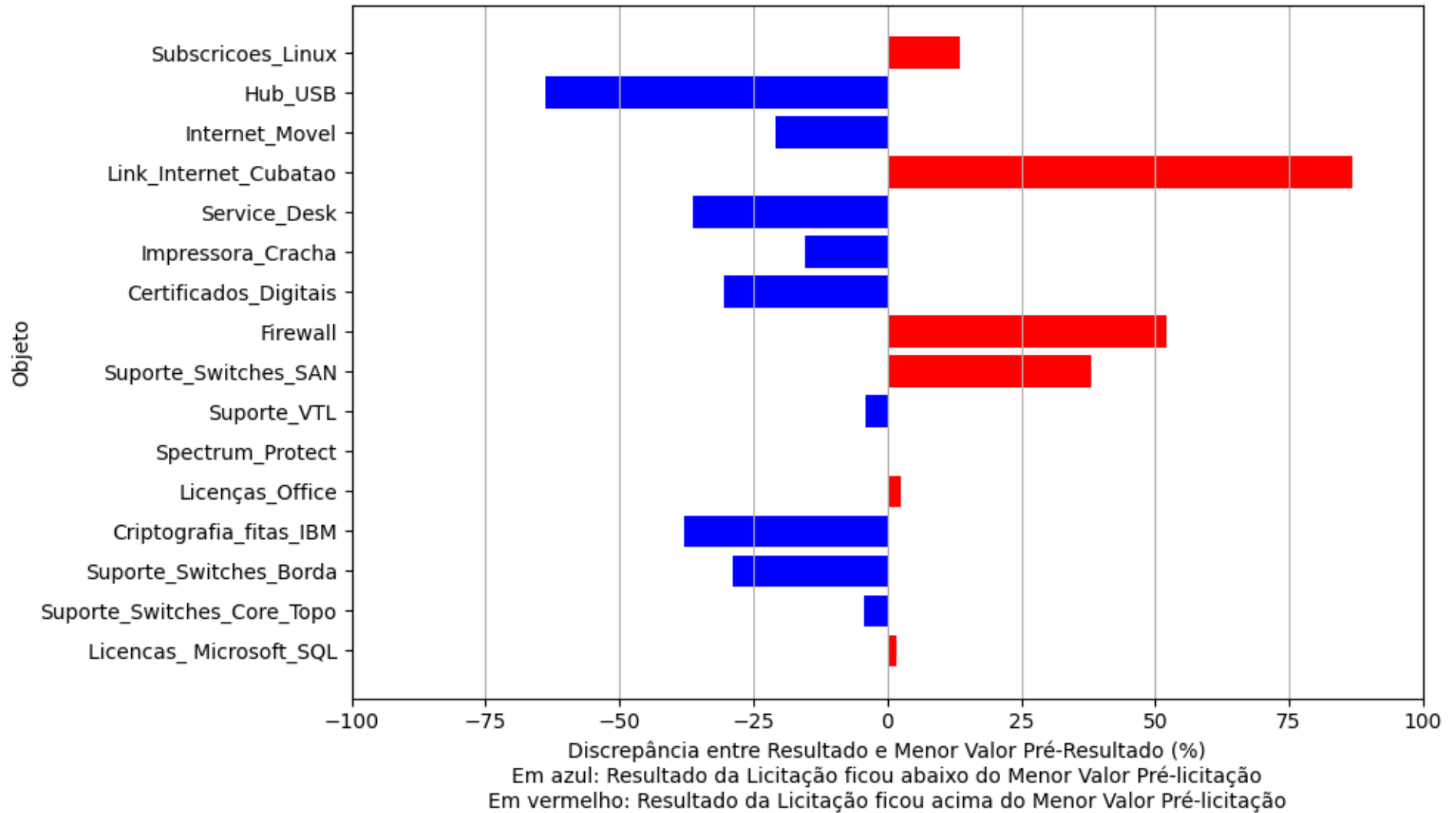
Objeto: Suporte_Switches_Core_Topo, Diferença: R\$ -21.698,00, Percentual de Discrepância: -4.35%

Objeto: Licencas_Microsoft_SQL, Diferença: R\$ 18.430,00, Percentual de Discrepância: 1.80%

Soma das Discrepâncias entre Resultados e os Menores Valores Pré-Resultados: -46.41%

GRÁFICO

Discrepância entre Resultado da Licitação e Menor Valor Pré-Resultado por Objeto



Exemplo de saída de código em Python 4 – Resultado vs Menor Valor Saneado

Nome do código: Dif_e_discrep_resultado_menorvalorsaneado

Dados utilizados: Resultados_de_Licitacoes_e_Conjecturas_v2
(Tabela 2)

TEXTO:

Diferença entre o Resultado e o Menor valor Saneado pré resultado da licitação e Discrepância:

Objeto: Subscricoes_Linux, Diferença: R\$ 3.447,00, Percentual de Discrepância: 13.61%

Objeto: Hub_USB, Diferença: R\$ -14.500,00, Percentual de Discrepância: -63.74%

Objeto: Internet_Movel, Diferença: R\$ -64.443,00, Percentual de Discrepância: -20.80%

Objeto: Link_Internet_Cubatao, Diferença: R\$ 1.395,00, Percentual de Discrepância: 86.92%

Objeto: Service_Desk, Diferença: R\$ -7.864.561,00, Percentual de Discrepância: -36.15%

Objeto: Impressora_Cracha, Diferença: R\$ -14.880,00, Percentual de Discrepância: -15.38%

Objeto: Certificados_Digitais, Diferença: R\$ -434,00, Percentual de Discrepância: -30.56%

Objeto: Firewall, Diferença: R\$ 8.360.764,00, Percentual de Discrepância: 52.22%

Objeto: Suporte_Switches_SAN, Diferença: R\$ 56.172,00, Percentual de Discrepância: 38.00%

Objeto: Suporte_VTL, Diferença: R\$ -16.165,00, Percentual de Discrepância: -4.09%

Objeto: Spectrum_Protect, Diferença: R\$ 159,00, Percentual de Discrepância: 0.23%

Objeto: Licenças_Office, Diferença: R\$ 38.585,00, Percentual de Discrepância: 2.63%

Objeto: Criptografia_fitras_IBM, Diferença: R\$ -12.158,00, Percentual de Discrepância: -37.82%

Objeto: Suporte_Switches_Borda, Diferença: R\$ -148.375,00, Percentual de Discrepância: -28.91%

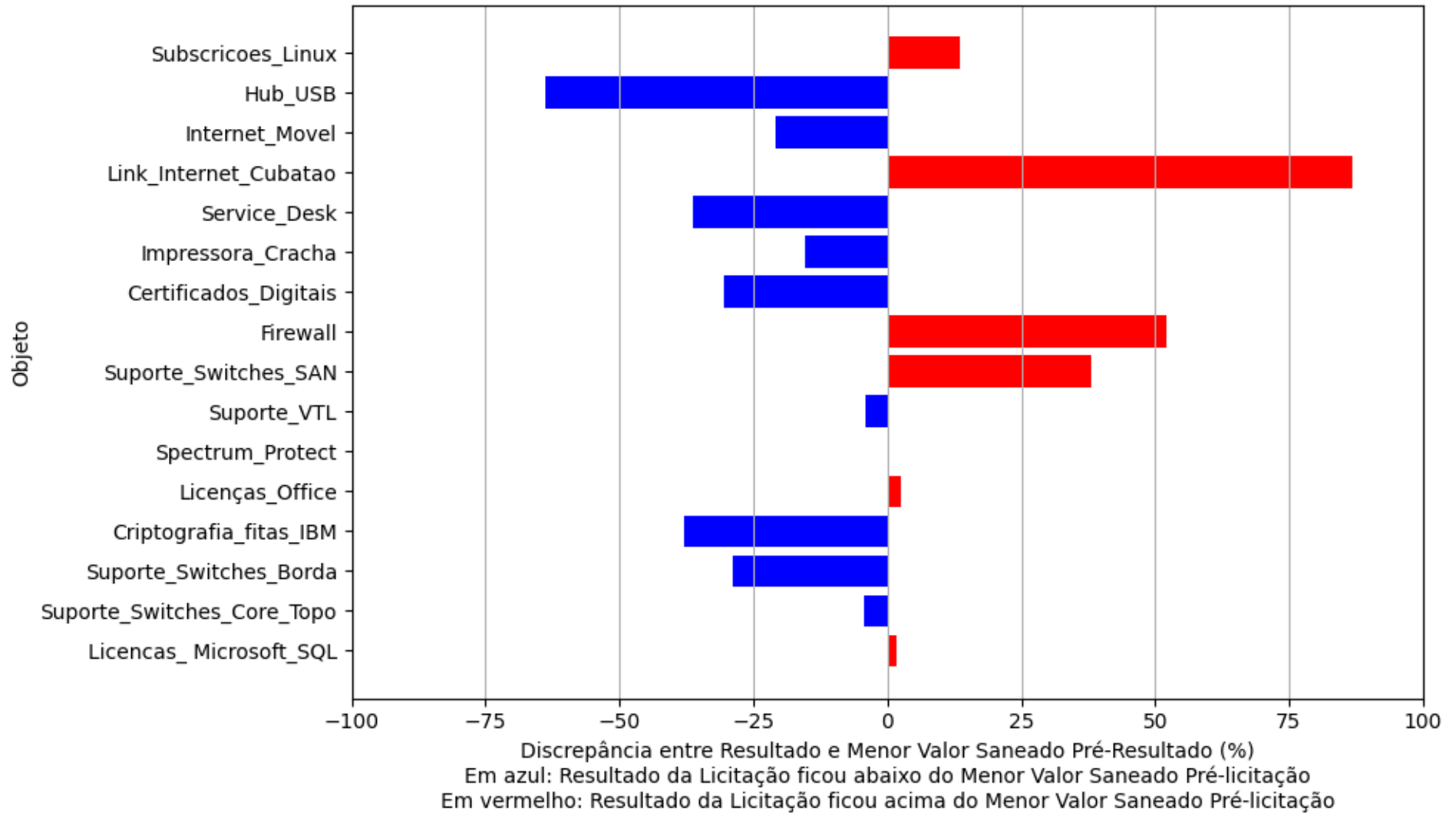
Objeto: Suporte_Switches_Core_Topo, Diferença: R\$ -21.698,00, Percentual de Discrepância: -4.35%

Objeto: Licencas_Microsoft_SQL, Diferença: R\$ 18.430,00, Percentual de Discrepância: 1.80%

Soma das Discrepâncias entre Resultados e Menores Valores Saneados Pré-Resultados: -46.41%

GRÁFICO

Discrepância entre Resultado da Licitação e Menor Valor Saneado Pré-Resultado por Objeto



Exemplo de saída de código em Python 5 – Resultado vs Média

Nome do código: Dif_e_discrep_resultado_media

Dados utilizados: Resultados_de_Licitacoes_e_Conjecturas_v2
(Tabela 2)

TEXTO:

Diferença entre o Resultado e a Média pré resultado da licitação e Discrepância:

Objeto: Subscricoes_Linux, Diferença: R\$ -1.498,00, Percentual de Discrepância: -4.95%

Objeto: Hub_USB, Diferença: R\$ -21.036,00, Percentual de Discrepância: -71.83%

Objeto: Internet_Movel, Diferença: R\$ -64.443,00, Percentual de Discrepância: -20.80%

Objeto: Link_Internet_Cubatao, Diferença: R\$ 495,00, Percentual de Discrepância: 19.76%

Objeto: Service_Desk, Diferença: R\$ -11.640.448,00, Percentual de Discrepância: -45.59%

Objeto: Impressora_Cracha, Diferença: R\$ -28.576,00, Percentual de Discrepância: -25.87%

Objeto: Certificados_Digitais, Diferença: R\$ -1.673,00, Percentual de Discrepância: -62.92%

Objeto: Firewall, Diferença: R\$ -4.118.015,00, Percentual de Discrepância: -14.45%

Objeto: Suporte_Switches_SAN, Diferença: R\$ -76.691,00, Percentual de Discrepância: -27.32%

Objeto: Suporte_VTL, Diferença: R\$ -138.458,00, Percentual de Discrepância: -26.76%

Objeto: Spectrum_Protect, Diferença: R\$ -19.701,00, Percentual de Discrepância: -21.94%

Objeto: Licenças_Office, Diferença: R\$ -127.622,00, Percentual de Discrepância: -7.81%

Objeto: Criptografia_fitras_IBM, Diferença: R\$ -29.323,00, Percentual de Discrepância: -59.46%

Objeto: Suporte_Switches_Borda, Diferença: R\$ -156.077,00, Percentual de Discrepância: -29.96%

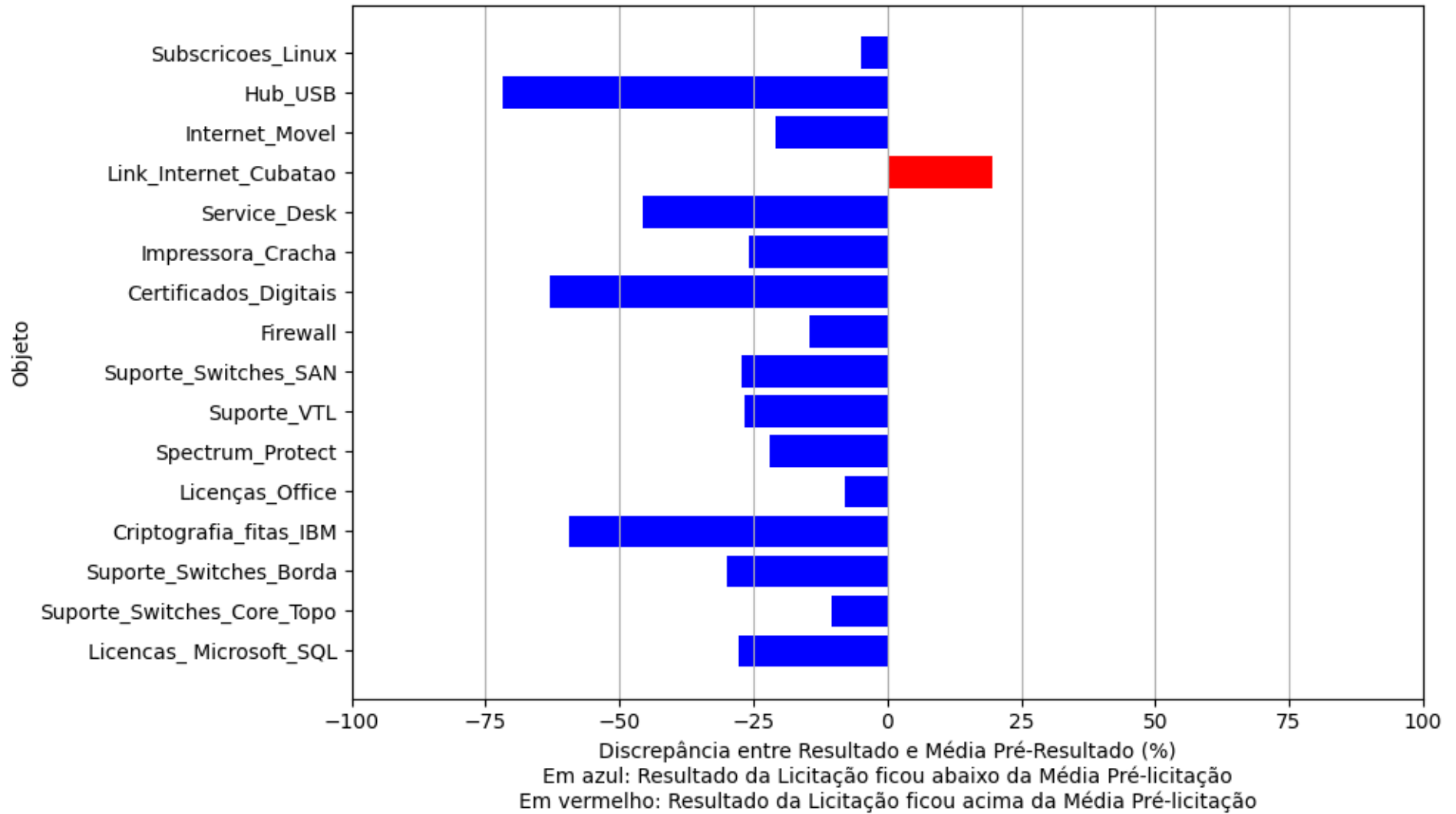
Objeto: Suporte_Switches_Core_Topo, Diferença: R\$ -55.412,00, Percentual de Discrepância: -10.41%

Objeto: Licencas_Microsoft_SQL, Diferença: R\$ -401.737,00, Percentual de Discrepância: -27.80%

Soma das Discrepâncias entre Resultados e as Médias Pré-Resultados: -438.13%

GRÁFICO

Discrepância entre Resultado da Licitação e Média Pré-Resultado por Objeto



Exemplo de saída de código em Python 6 – Resultado vs Média Saneada

Nome do código: Dif_e_discrep_resultado_mediasaneada

Dados utilizados: Resultados_de_Licitacoes_e_Conjecturas_v2
(Tabela 2)

TEXTO:

Diferença entre o Resultado e a Média Saneada pré resultado da licitação e Discrepância:

Objeto: Subscricoes_Linux, Diferença: R\$ -1.498,00, Percentual de Discrepância: -4.95%

Objeto: Hub_USB, Diferença: R\$ -21.036,00, Percentual de Discrepância: -71.83%

Objeto: Internet_Movel, Diferença: R\$ -64.443,00, Percentual de Discrepância: -20.80%

Objeto: Link_Internet_Cubatao, Diferença: R\$ 495,00, Percentual de Discrepância: 19.76%

Objeto: Service_Desk, Diferença: R\$ -11.640.448,00, Percentual de Discrepância: -45.59%

Objeto: Impressora_Cracha, Diferença: R\$ -28.576,00, Percentual de Discrepância: -25.87%

Objeto: Certificados_Digitais, Diferença: R\$ -743,00, Percentual de Discrepância: -42.97%

Objeto: Firewall, Diferença: R\$ -4.118.015,00, Percentual de Discrepância: -14.45%

Objeto: Suporte_Switches_SAN, Diferença: R\$ -76.691,00, Percentual de Discrepância: -27.32%

Objeto: Suporte_VTL, Diferença: R\$ -138.458,00, Percentual de Discrepância: -26.76%

Objeto: Spectrum_Protect, Diferença: R\$ -5.297,00, Percentual de Discrepância: -7.03%

Objeto: Licenças_Office, Diferença: R\$ -127.622,00, Percentual de Discrepância: -7.81%

Objeto: Criptografia_fitras_IBM, Diferença: R\$ -20.308,00, Percentual de Discrepância: -50.39%

Objeto: Suporte_Switches_Borda, Diferença: R\$ -156.077,00, Percentual de Discrepância: -29.96%

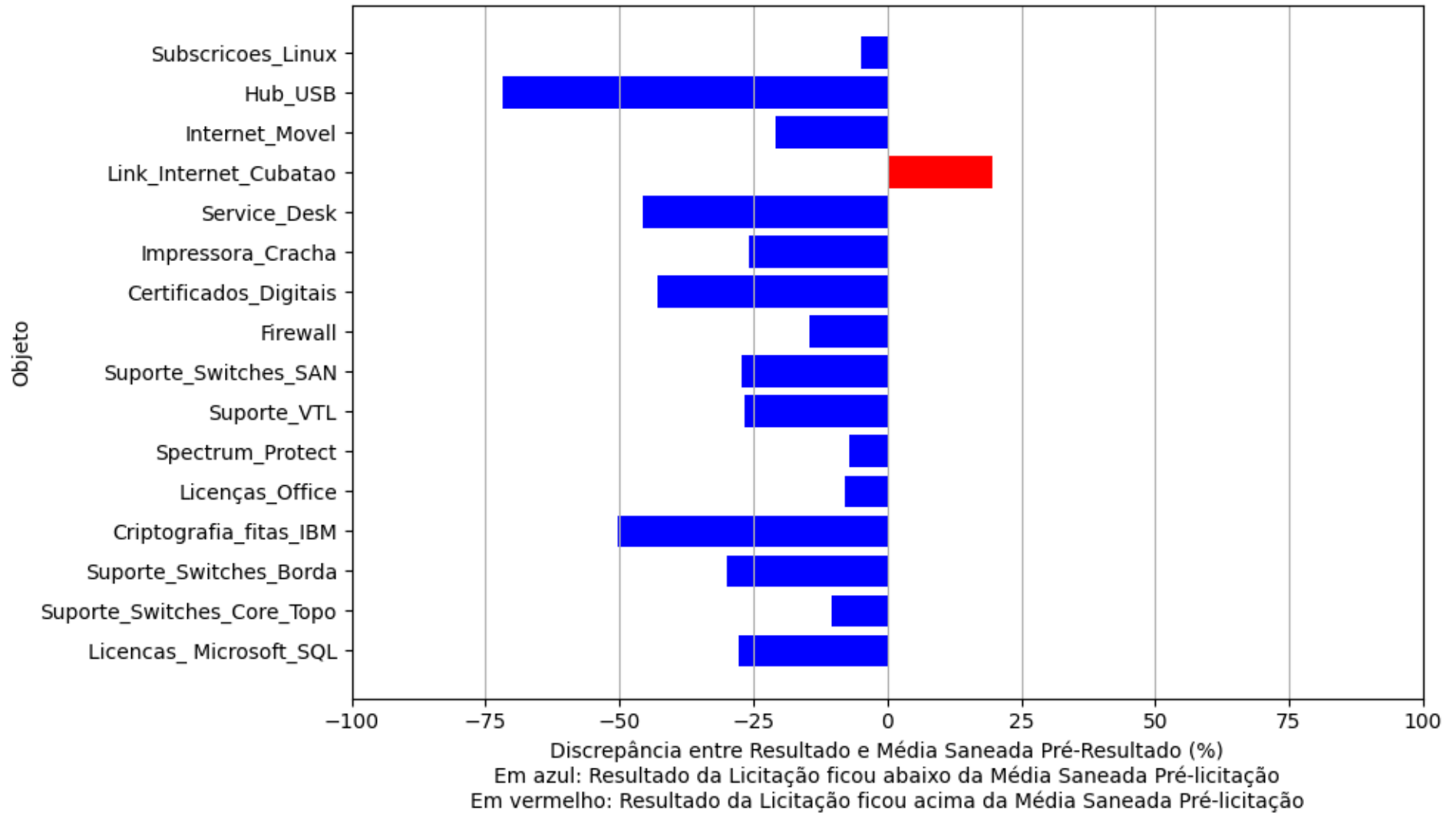
Objeto: Suporte_Switches_Core_Topo, Diferença: R\$ -55.412,00, Percentual de Discrepância: -10.41%

Objeto: Licencas_Microsoft_SQL, Diferença: R\$ -401.737,00, Percentual de Discrepância: -27.80%

Soma das Discrepâncias entre Resultados e as Médias Saneadas Pré-Resultados: -394.20%

GRÁFICO

Discrepância entre Resultado da Licitação e Média Saneada Pré-Resultado por Objeto



Exemplo de saída de código em Python 7 – Resultado vs Mediana

Nome do código: Dif_e_discrep_resultado_mediana

Dados utilizados: Resultados_de_Licitacoes_e_Conjecturas_v2
(Tabela 2)

TEXTO:

Diferença entre o Resultado e a Mediana pré resultado da licitação e Discrepância:

Objeto: Subscricoes_Linux, Diferença: R\$ -1.562,00, Percentual de Discrepância: -5.15%

Objeto: Hub_USB, Diferença: R\$ -18.000,00, Percentual de Discrepância: -68.57%

Objeto: Internet_Movel, Diferença: R\$ -64.443,00, Percentual de Discrepância: -20.80%

Objeto: Link_Internet_Cubatao, Diferença: R\$ 88,00, Percentual de Discrepância: 3.02%

Objeto: Service_Desk, Diferença: R\$ -11.460.063,00, Percentual de Discrepância: -45.21%

Objeto: Impressora_Cracha, Diferença: R\$ -26.760,00, Percentual de Discrepância: -24.63%

Objeto: Certificados_Digitais, Diferença: R\$ -899,00, Percentual de Discrepância: -47.69%

Objeto: Firewall, Diferença: R\$ -2.972.021,00, Percentual de Discrepância: -10.87%

Objeto: Suporte_Switches_SAN, Diferença: R\$ -29.450,00, Percentual de Discrepância: -12.62%

Objeto: Suporte_VTL, Diferença: R\$ -36.610,00, Percentual de Discrepância: -8.81%

Objeto: Spectrum_Protect, Diferença: R\$ -8.025,00, Percentual de Discrepância: -10.27%

Objeto: Licencas_Office, Diferença: R\$ -127.622,00, Percentual de Discrepância: -7.81%

Objeto: Criptografia_fitras_IBM, Diferença: R\$ -24.383,00, Percentual de Discrepância: -54.95%

Objeto: Suporte_Switches_Borda, Diferença: R\$ -156.077,00, Percentual de Discrepância: -29.96%

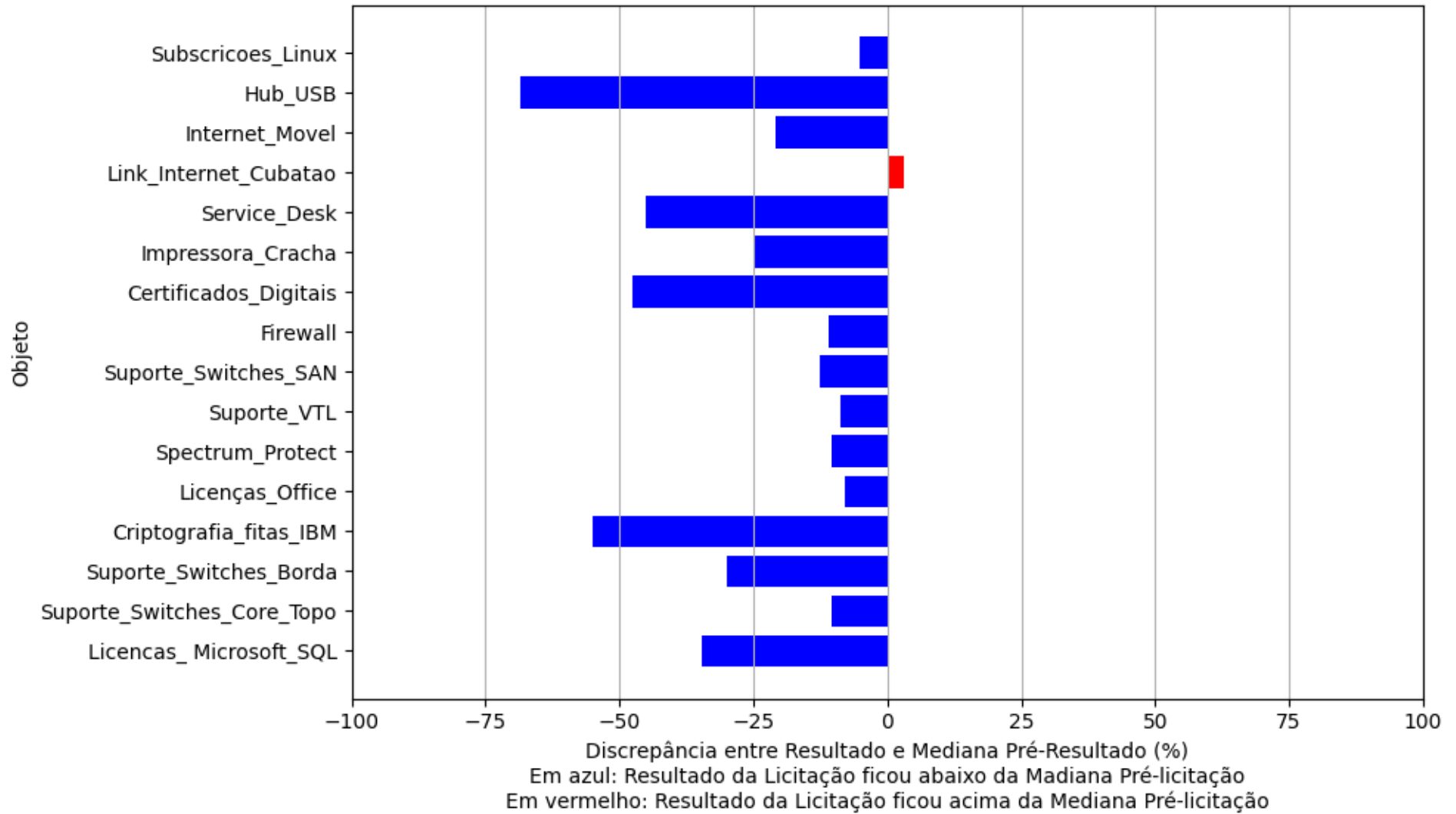
Objeto: Suporte_Switches_Core_Topo, Diferença: R\$ -55.412,00, Percentual de Discrepância: -10.41%

Objeto: Licencas_Microsoft_SQL, Diferença: R\$ -551.675,00, Percentual de Discrepância: -34.59%

Soma das Discrepâncias entre Resultados e as Medianas Pré-Resultados: -389.32%

GRÁFICO

Discrepância entre Resultado da Licitação e Mediana Pré-Resultado por Objeto



Exemplo de saída de código em Python 8 – Resultado vs Mediana Saneada

Nome do código: Dif_e_discrep_resultado_mediana_saneada

Dados utilizados: Resultados_de_Licitacoes_e_Conjecturas_v2
(Tabela 2)

TEXTO:

Diferença entre o Resultado e a Mediana Saneada pré resultado da licitação e Discrepância:

Objeto: Subscricoes_Linux, Diferença: R\$ -1.562,00, Percentual de Discrepância: -5.15%

Objeto: Hub_USB, Diferença: R\$ -18.000,00, Percentual de Discrepância: -68.57%

Objeto: Internet_Movel, Diferença: R\$ -64.443,00, Percentual de Discrepância: -20.80%

Objeto: Link_Internet_Cubatao, Diferença: R\$ 88,00, Percentual de Discrepância: 3.02%

Objeto: Service_Desk, Diferença: R\$ -11.460.063,00, Percentual de Discrepância: -45.21%

Objeto: Impressora_Cracha, Diferença: R\$ -26.760,00, Percentual de Discrepância: -24.63%

Objeto: Certificados_Digitais, Diferença: R\$ -898,00, Percentual de Discrepância: -47.66%

Objeto: Firewall, Diferença: R\$ -2.972.021,00, Percentual de Discrepância: -10.87%

Objeto: Suporte_Switches_SAN, Diferença: R\$ -29.450,00, Percentual de Discrepância: -12.62%

Objeto: Suporte_VTL, Diferença: R\$ -36.610,00, Percentual de Discrepância: -8.81%

Objeto: Spectrum_Protect, Diferença: R\$ -5.347,00, Percentual de Discrepância: -7.09%

Objeto: Licenças_Office, Diferença: R\$ -127.622,00, Percentual de Discrepância: -7.81%

Objeto: Criptografia_fitras_IBM, Diferença: R\$ -12.836,00, Percentual de Discrepância: -39.10%

Objeto: Suporte_Switches_Borda, Diferença: R\$ -156.077,00, Percentual de Discrepância: -29.96%

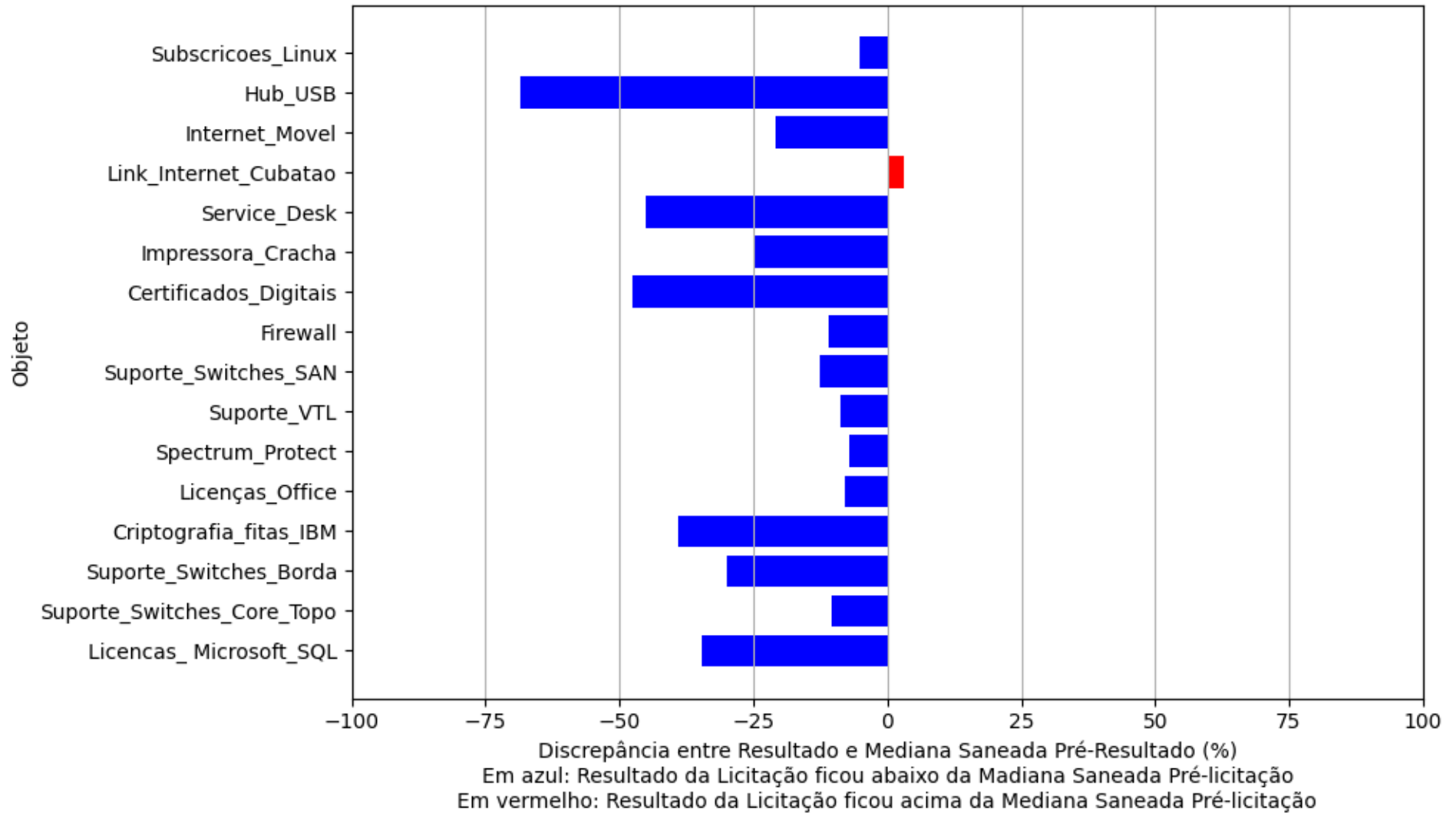
Objeto: Suporte_Switches_Core_Topo, Diferença: R\$ -55.412,00, Percentual de Discrepância: -10.41%

Objeto: Licencas_Microsoft_SQL, Diferença: R\$ -551.675,00, Percentual de Discrepância: -34.59%

Soma das Discrepâncias entre Resultados e as Medianas Saneadas Pré-Resultados: -370.26%

GRÁFICO

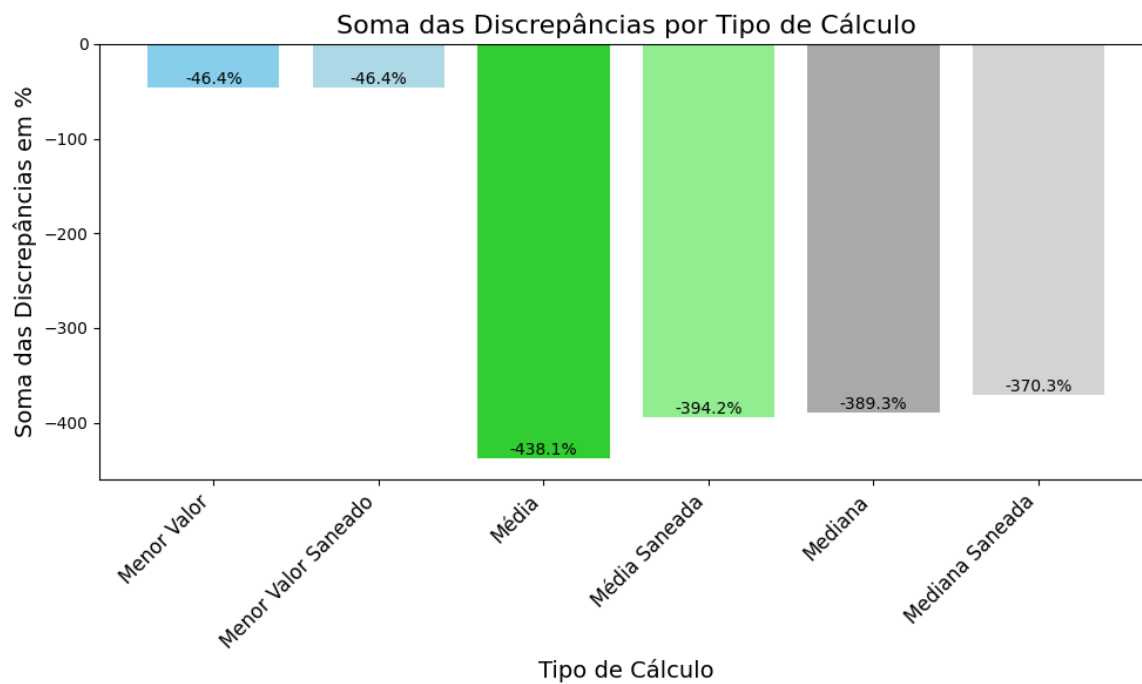
Discrepância entre Resultado da Licitação e Mediana Saneada Pré-Resultado por Objeto



Exemplo de saída de código em Python 9 – Soma das Discrepâncias

Nome do código: soma_das_discrepancias

Dados utilizados: Colunas “Tipo_de_Calculo” e “Soma_discrepancias”
(Tabela 3)



Exemplo de saída de código em Python 10 – Soma das Discrepâncias

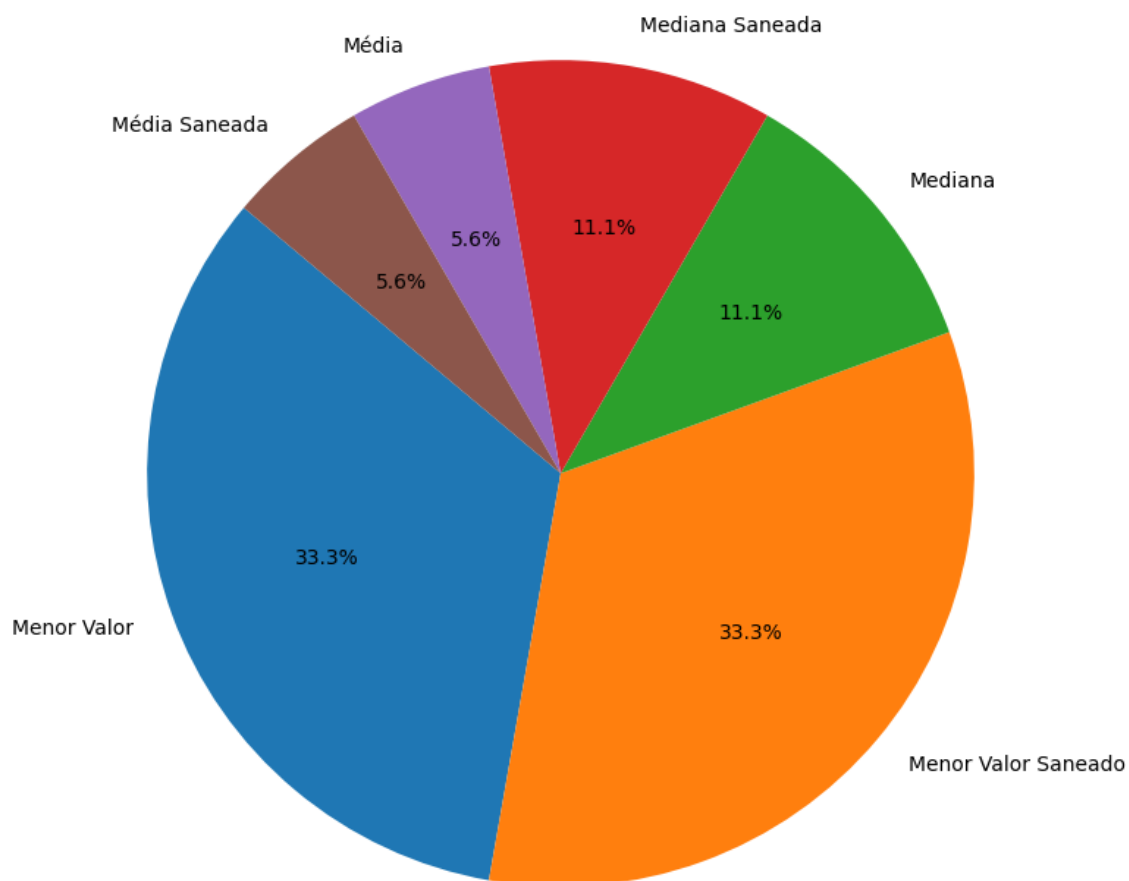
Nome do código: Apurar Menor discrepancia por objeto e Tipo v2

Dados utilizados: Colunas “Objeto” e “Tipo de Calculo” (Tabela 2)

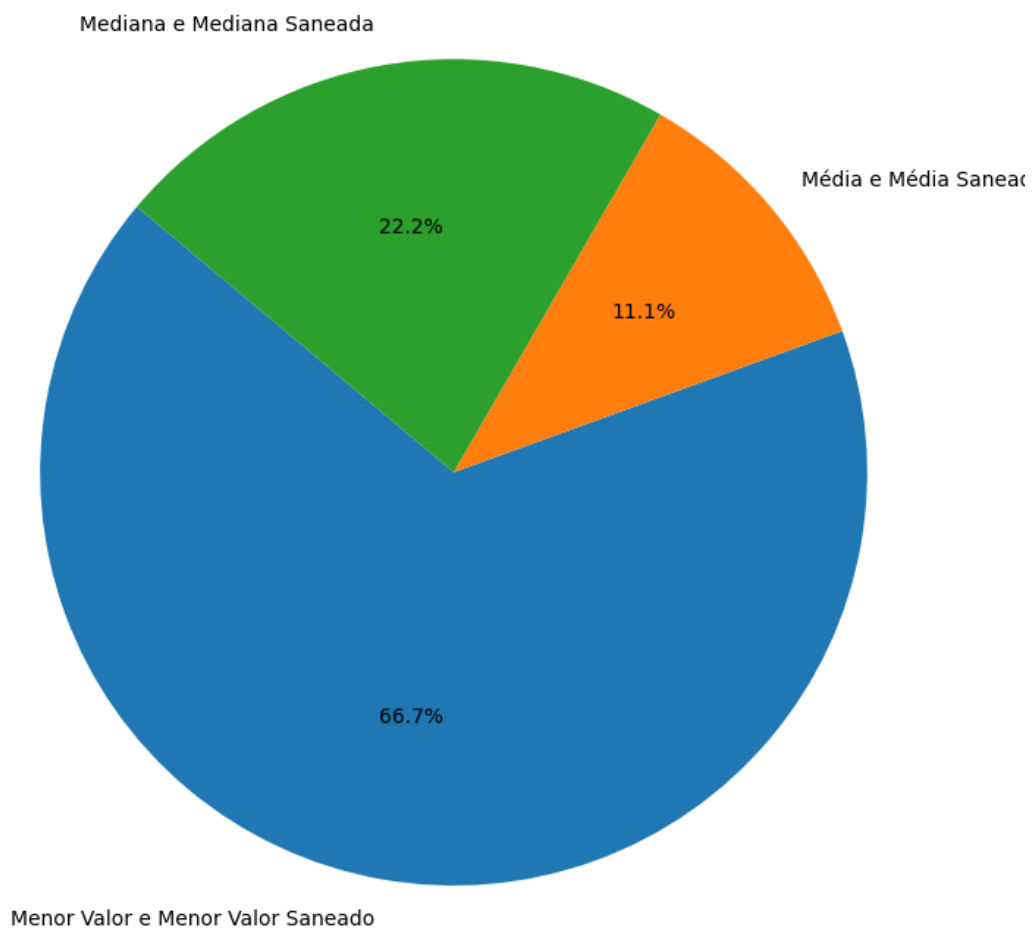
Texto:				
	Objeto	Menor Discrepância	Tipo de Cálculo	Positivo/Negativo
0	Subscricoes_Linux	-4.947487	[Média, Média Saneada]	negativo
1	Hub_USB	-63.736264	[Menor Valor, Menor Valor Saneado]	negativo
2	Internet_Movel	-20.804641	[Menor Valor, Menor Valor Saneado, Média, Méd...]	negativo
3	Link_Internet_Cubatao	3.021978	[Mediana, Mediana Saneada]	positivo
4	Service_Desk	-36.151342	[Menor Valor, Menor Valor Saneado]	negativo
5	Impressora_Cracha	-15.378255	[Menor Valor, Menor Valor Saneado]	negativo
6	Certificados_Digitais	-30.563380	[Menor Valor, Menor Valor Saneado]	negativo
7	Firewall	-10.869388	[Mediana, Mediana Saneada]	negativo
8	Suporte_Switches_SAN	-12.615175	[Mediana, Mediana Saneada]	negativo
9	Suporte_VTL	-4.090696	[Menor Valor, Menor Valor Saneado]	negativo
10	Spectrum_Protect	0.227377	[Menor Valor, Menor Valor Saneado]	positivo
11	Licenças_Office	2.627511	[Menor Valor, Menor Valor Saneado]	positivo
12	Criptografia_fitras_IBM	-37.818838	[Menor Valor, Menor Valor Saneado]	negativo
13	Suporte_Switches_Borda	-28.912406	[Menor Valor, Menor Valor Saneado]	negativo
14	Suporte_Switches_Core_Topo	-4.352841	[Menor Valor, Menor Valor Saneado]	negativo
15	Licencas_Microsoft_SQL	1.798398	[Menor Valor, Menor Valor Saneado]	positivo

GRÁFICOS:

Distribuição dos Tipos de Cálculo com menor discrepância



Totais somados por Tipo



Exemplo de saída de código em Python 11 – Fonte de Valor vs Resultado

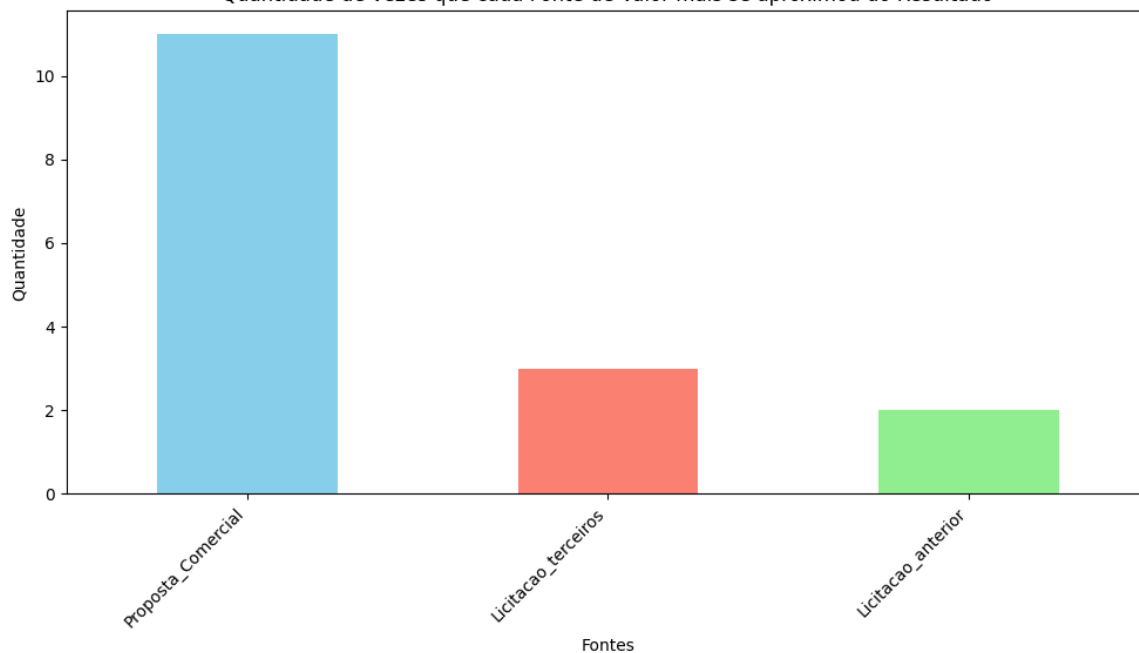
Nome do código: Comparar fonte e resultados v2

Dados utilizados: Colunas “Objeto”, “Valores” e “Fonte” (Tabela 1) e “Resultado” (Tabela 2)

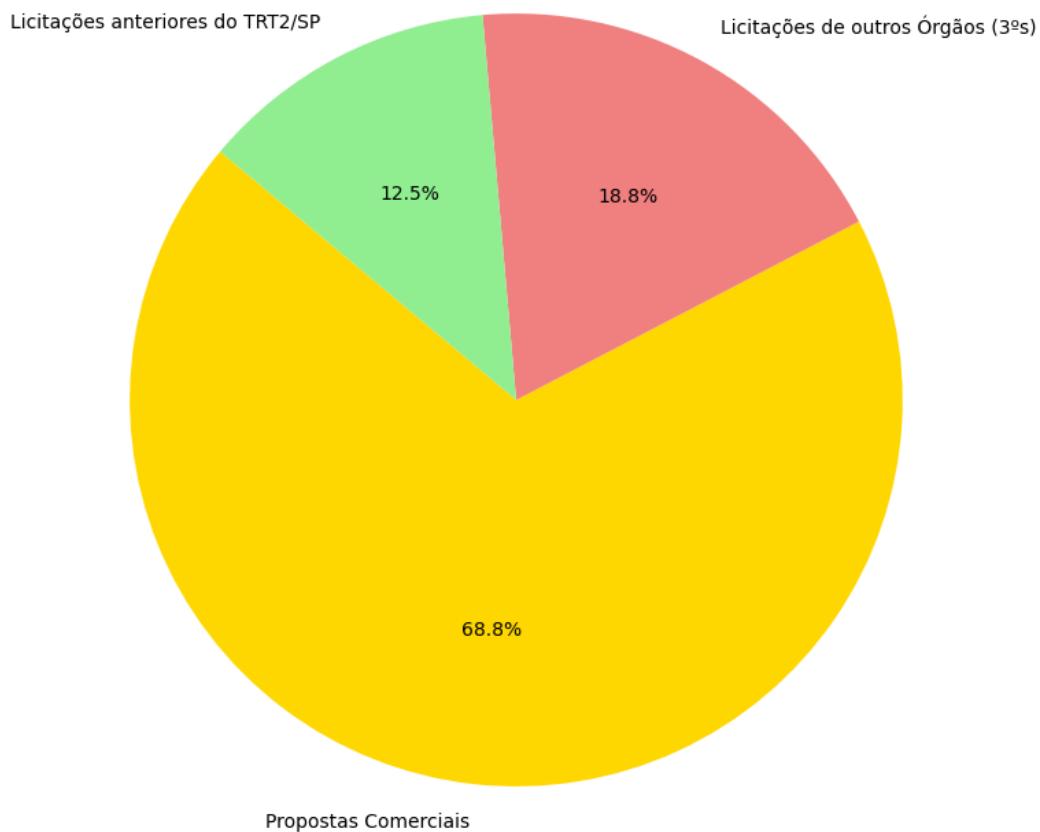
	Objeto	Discrepancia	Fonte	Resultado	Positivo/Negativo
0	Subscricoes_Linux	4.320719	Licitacao_anterior	28780	positivo
1	Hub_USB	63.736264	Proposta_Comercial	8250	negativo
2	Internet_Movel	20.804641	Licitacao_anterior	245310	negativo
3	Link_Internet_Cubatao	0.000000	Proposta_Comercial	3000	positivo
4	Service_Desk	36.151342	Proposta_Comercial	13889987	negativo
5	mpressora_Cracha	15.378255	Licitacao_terceiros	81880	negativo
6	Certificados_Digitais	30.563380	Proposta_Comercial	986	negativo
7	Firewall	10.869388	Proposta_Comercial	24371018	negativo
8	Suporte_Switches_SAN	12.615175	Proposta_Comercial	203999	negativo
9	Suporte_VTL	4.090696	Proposta_Comercial	379000	negativo
10	Spectrum_Protect	0.227377	Licitacao_terceiros	70087	positivo
11	Licenças_Office	2.627511	Proposta_Comercial	1507085	positivo
12	Criptografia_fitras_IBM	37.818838	Proposta_Comercial	19990	negativo
13	Suporte_Switches_Borda	28.912406	Proposta_Comercial	364813	negativo
14	Suporte_Switches_Core_Topo	4.352841	Proposta_Comercial	476781	negativo
15	Licencas_Microsoft_SQL	1.798398	Licitacao_terceiros	1043231	positivo

GRÁFICOS

Quantidade de vezes que cada Fonte de Valor mais se aproximou do Resultado



Proporção em que cada Fonte de Valor mais se aproximou do Resultado



Data Storytelling

Conforme visto na videoaula do Professor Fabio Lopes: “o Data Storytelling é uma abordagem estruturada para comunicar insights de dados, combinando três elementos: dados, narrativa e visualização.”

Nesse sentido, em resumo apresentamos, em estrutura de tópicos, o seguinte:

- Dados: foram obtidos dados de estimativas e de resultados de licitações relativas ao TRT2/SP e ao ano de 2023;
- Os dados foram preparados para análise tendo sido aglutinados em tabelas, tendo passado por diversas simulações de possíveis estimativas, valendo-se de ferramenta desenvolvida para formulação de médias, medianas e remoção de outliers em que se obteve valores saneados (com base em método utilizado em licitações pelo Poder Executivo Federal);
- O problema é delimitado e corresponde à verificação de diversas possibilidades de formulação da estimativa de custos de um certame, considerando as diversas implicações deste ato na seara administrativa das instituições sujeitas à licitação;
- Nesse sentido a problemática da formação do preço estimado revelou-se multifacetada pois envolve grande quantidade de valores heterogêneos, provenientes de várias fontes, sendo que uma estimativa imprecisa gera consequências indesejáveis, pois não permite corretos planejamento de gastos e manejo da verba orçamentária, ao mesmo tempo que pode ensejar (quanto muito alta) contratação por valor mais alto que o preço de mercado ou (quando muito baixa) uma licitação deserta ou fracassada;
- Some-se a isso a falta de uma métrica definida ou definitiva pré-definida que seja válida para todos os casos;
- Mas a análise dos dados envolvidos na formulação das estimativas pode indicar caminhos para que se tenha mais segurança na fixação de parâmetros para sua formulação e com mais precisão;
- Assim, durante a análise dos dados algumas hipóteses foram se confirmando como, por exemplo, a circunstância dos valores obtidos na fase preparatória estarem com valores, em geral, acima do valor final da licitação, no caso das licitações do TRT2/SP, uma vez que, os comparativos que envolviam a média ou mediana e os resultados das licitações mostraram, em geral, que os valores finais obtidos nos certames eram mais baixos (mais estudos são necessários para verificar se isso é característica somente deste órgão público ou se isso se repete em outros órgãos, principalmente em virtude de localização geográfica, quantidade de fornecedores locais, economia de escala, dentre outras variáveis);

- A visualização das análises envolve tanto gráficos exploratórios quanto aqueles destinados à apresentação final dos resultados. Procurou-se, em cada análise, obter uma visualização clara do que se estava encontrando, especialmente em virtude da utilização de scripts produzidos na linguagem Python;
- Apesar das diversas dúvidas iniciais, após as análises realizadas, pode-se dizer que há fortes indícios que apontam para uma estimativa mais equilibrada e mais próxima do resultado das licitações, quando se utiliza o Menor Valor obtido. Em algumas circunstâncias, no entanto, o valor estimado ficou abaixo do resultado da licitação, o que pode ser um problema, mas apesar disso, nota-se que os gráficos com distribuição mais centralizada e, portanto, com menos diferenças em relação aos resultados das licitações, correspondem ao Menor Valor e ao Menor Valor Saneado. Isso não significa que se pode desprezar as demais formas de cálculo, pois ainda representam um terço do total mais próximo do resultado, sendo que em algumas situações a média ou a mediana dos valores seriam soluções mais próximas do resultado das licitações, conforme gráfico do código em Python nº 10;
- Nesse cenário, o órgão precisa analisar em cada contratação, a despeito das conclusões gerais aqui obtidas, qual seria a forma mais prudente considerando outras informações sobre a contratação (como duração do contrato, multas previstas, entre outras). Pode-se, também, adotar fórmulas intermediárias como a mescla do menor valor (ou do menor valor saneado) com a mediana (ou com a mediana saneada);
- Verifica-se, também, que as fontes de valor relacionadas às licitações anteriores do órgão e de terceiros guardam valores mais próximos ao resultado das licitações quando comparadas com as propostas comerciais. Isso pode ser constatado comparando-se os gráficos de pizza dos códigos em Python de nº 1 e de nº 11, em que a proporção inicial de cada um é de 9,4% (Licitações anteriores do próprio órgão) e de 17% (licitações anteriores de outros órgãos), enquanto a proporção de fontes de valor mais próximos dos resultados alcança 12,5% (Licitações anteriores do próprio órgão) e 18,8% (licitações anteriores de outros órgãos). Destarte, apesar das propostas comerciais serem mais numerosas parece haver mais distorção nos valores constantes delas, corroborando a preferência de órgãos de controle no uso de valores obtidos em licitações anteriores, seja do próprio órgão ou de terceiros;
- Entretanto, não se pode desprezar o conjunto das fontes de valores, especialmente porque a grande maioria dos valores obtidos tinham como fonte as propostas comerciais, tema que inclusive pode ser objeto de estudo específico;
- Outro achado digno de nota é que a fórmula utilizada em algumas licitações pelo Poder Executivo Federal em que se busca eliminar valores outliers desempenhou de forma satisfatória, uma vez que os resultados das análises revelam um melhor desempenho da estimativa feita quando os valores estavam saneados, mesmo que de forma discreta (especialmente na mediana saneada e na média saneada, uma vez que os menores valores não foram descartados quando considerada a base de dados do ano de 2023);

- Por fim, essas afirmações e achados permitem concluir que pode-se recomendar a utilização do script (fórmula do Poder Executivo Federal) em Python, para remoção de outliers, propondo-se que o valor mais baixo obtido em fase preliminar da licitação no mais das vezes será um norte para a formulação da estimativa de custo, ressalvada a hipótese de utilização dos demais tipos (média e mediana, saneadas ou não) em casos específicos, ou ainda de forma combinada, conforme mencionado anteriormente.

Grupo

O grupo é formado por:

- | | |
|---------------------------------------|---------------|
| - Andre Gustavo Monteiro Dos Santos F | – RA 10424359 |
| - Felipe Bernardi Da Silva | – RA 10328856 |
| - Milan Mirco Moraes Mazur | – RA 10363757 |
| - Raul Santos Lages | – RA 10424621 |

Github do projeto

<https://github.com/milanmazur/Projeto-Aplicado-I---Mackenzie>

Arquivos constantes do Github:

- Analisa Estimativa por Objeto – v1
- MAPA_PRECOS Design Gráfico 2022 antes PGFN 2022 → arquivo referência para descarte de valores outliers, conforme Poder Executivo Federal
- SEI_ME - 26677408 - Anexo - Design Grafico Metodologia Pesquisa de Preços Media-Mediana Saneada → arquivo referência para descarte de valores outliers, conforme Poder Executivo Federal
- Planilha_valores_obtidos_antes_da_licitação → Tabela 1
- Resultados_de_Licitacoes_e_Conjecturas_v2 → Tabela 2
- soma_das_discrepancias → Tabela 3
- Dif_e_discrep_resultado_media.py
- Dif_e_discrep_resultado_mediana.py
- Dif_e_discrep_resultado_mediana_saneada.py
- Dif_e_discrep_resultado_mediasaneada.py
- Dif_e_discrep_resultado_menorvalor.py
- Dif_e_discrep_resultado_menorvalorsaneado.py

- Distribuicao_por_fonte.py
- soma_das_discrepancias.py
- Comparar_fontes_e_resultados_v2.py

Referências bibliográficas e sites consultados:

<https://ww2.trt2.jus.br/transparencia/planejamento-e-gestao/planejamento-estrategico/plano-estrategico-institucional> – Acesso em 17/03/2024

<https://www.cnj.jus.br/pesquisas-judiciarias/paineis-cnj/> - Acesso em 17/03/2024