

O impacto de defeitos observados na biometria do mesário sobre o resultado do segundo turno das eleições 2022

A importância de se analisar os votos sem mesário, com falha dupla, é que estes votos são uma fraude largamente conhecida e de fácil implementação, conhecida como voto fantasma em que um terceiro vota por alguém, este voto em um sistema com controle por reconhecimento biométrico em que o mesário não quiser se comprometer, ou ainda em que se quiser medir a efetividade da fraude para algum outro processo do esquema, resultaria na mensagem de log que será descrita abaixo.

Definição do defeito considerado, o voto sem mesário identificado.

Considerando o Voto Padrão nas urnas disponibilizadas em 2022 para as eleições ocorrem em alguns passos considerando os fluxos de exceção e a nomenclatura que serão dados neste documento.

Os Passos de votação com os fluxos de exceção de interesse:

1. O “eleitor” entrega seu título de eleitor e um documento de identificação a um “primeiro” mesário.
2. O primeiro lê em voz alta o número do título de eleitor.
3. O “segundo” mesário digita o título de eleitor lido pelo primeiro, uma vez o título do eleitor digitado o nome do eleitor é exibido, o nome é lido em voz alta pelo mesmo.
4. O “primeiro” mesário confirma o nome e pede ao eleitor que ele se identifique biometricamente no terminal do mesário.
5. O “eleitor” oferece um de seus dedos que tiveram a biometria coletada anteriormente.
6. A biometria é avaliada pela máquina por similaridade em quatro tentativas.
 - 6.1. Caso a biometria não identifique o “eleitor” o “segundo” mesário exige um documento de identificação.
 - 6.2. O “segundo” mesário confirma a identificação do “eleitor” e oferece um de seus dedos para que o equipamento o reconheça.
 - 6.2.1. O mesário tem a sua identificação confirmada e a votação continua no passo 7.
 - 6.2.2. O mesário não tem sua identificação reconhecida e algumas logs são possivelmente informadas neste momento:
 - 6.2.2.1. Não encontrou digital coletada em nenhum dos arquivos, vai salvar a digital em novo arquivo.
 - 6.2.2.2. Digital coletada bate com uma digital gravada após o registro inicial de mesário. Não é possível associar a habilitação a um mesário.
 - 6.2.2.3. Biometria do mesário \$ID encontrada em eleitores \$ID2.
 - 6.2.2.4. Biometria do mesário \$ID encontrada nos arquivos coletados \$ID4.
7. O terminal do eleitor é habilitado.
8. O “eleitor” vota.
9. O voto do eleitor é computado.
10. Reinicia-se o processo.

O trabalho abaixo irá se concentrar nas seguintes falhas encadeadas:

Durante a identificação do eleitor:

1. Eleitor não identificado por biometria.

2. Mesário não identificado por biometria com as seguintes mensagens que avaliei como falha catastrófica sem possível controle imediato daqui em diante chamada de “dupla falha”:

2.1. Não encontrou digital coletada em nenhum dos arquivos, vai salvar a digital em novo arquivo.

2.2. Digital coletada bate com uma digital gravada após o registro inicial de mesário. Não é possível associar a habilitação a um mesário.

Comparando a existência e a não existência de falhas:

Estatística Descritiva básica:

Realizando um levantamento apenas comparando a existência e não existência dos votos a partir dos boletins de urna foram estratificados três grupos:

Chamaremos a amostra que contém o público completo de: **TODOS**

Chamaremos a amostra **que não contém** votos com “dupla falha” de: **ZERO**

Chamaremos a amostra **que contém** votos com “dupla falha” de: **> 0**

Agrupados em CD_MUNICIPIO + NR_VOTAVEL, resultando em 22773 categorias observável em Contagem abaixo:

TODOS		ZERO		>0	
Média	5456,14526	Média	1183,162824	Média	4272,982435
Erro padrão	291,7718119	Erro padrão	63,96301901	Erro padrão	231,554264
Mediana	816	Mediana	106	Mediana	607
Modo	33	Modo	0	Modo	0
Desvio padrão	44030,4836	Desvio padrão	9652,483702	Desvio padrão	34943,21866
Variância da amostra	1938683486	Variância da amostra	93170441,62	Variância da amostra	1221028530
Curtose	3654,431293	Curtose	1983,936497	Curtose	4261,273015
Assimetria	52,11767086	Assimetria	38,55353295	Assimetria	56,45231665
Intervalo	3677921	Intervalo	613502	Intervalo	3064419
Mínimo	0	Mínimo	0	Mínimo	0
Máximo	3677921	Máximo	613502	Máximo	3064419
Soma	124252796	Soma	26944167	Soma	97308629
Contagem	22773	Contagem	22773	Contagem	22773
Maior(10)	720322	Maior(10)	183491	Maior(10)	545804
Menor(10)	1	Menor(10)	0	Menor(10)	0
Nível de confiança(571,89264	Nível de confiança(125,3718773	Nível de confiança(453,8621413

Considerando apenas a correlação entre estes três grupos agrupados por cidade e nr de voto e temos o seguinte resultado:

	TODOS	ZERO	>0
TODOS	1		
ZERO	0,953904	1	
>0	0,996558	0,925741	1

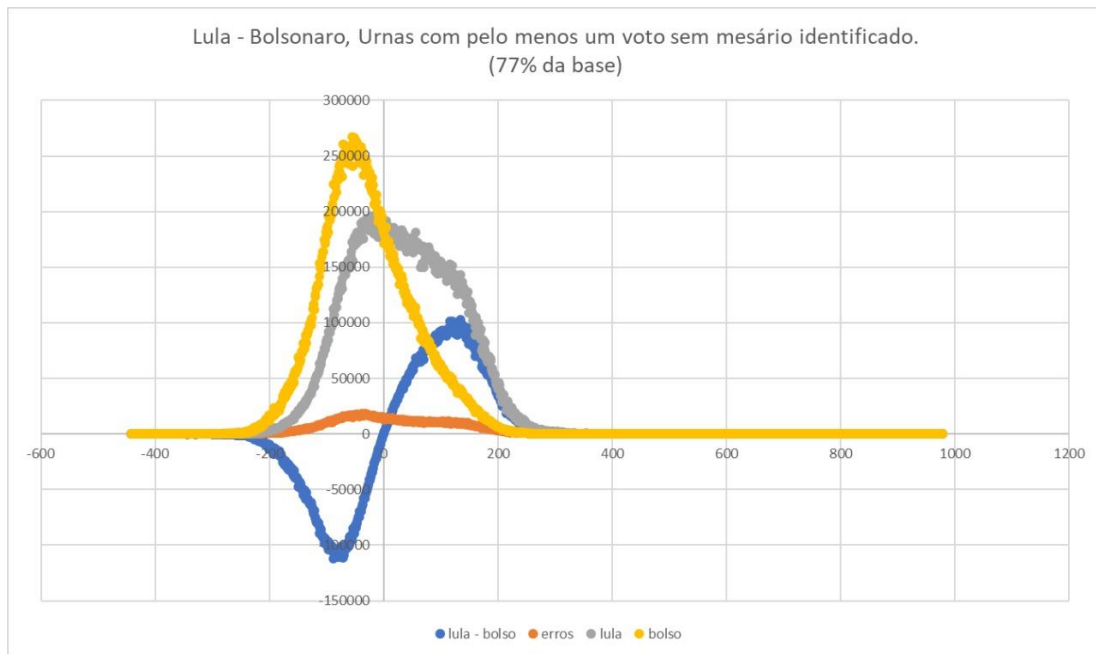
Onde podemos enxergar uma alta correlação (> 0,95) de **TODOS** com os outros 2 grupos e uma baixa correlação de **ZERO** com **>0**

Analisando ainda os tipos de urnas em cada grupo temos:

Votos agrupados por urna			
>0		ZERO	
Tamanho		Tamanho	
Média	263,7133	Média	508,3049
Erro padrão	0,093759	Erro padrão	0,431023
Mediana	270	Mediana	530
Modo	296	Modo	598
Desvio padrão	56,95384	Desvio padrão	137,5896
Variância da amostra	3243,74	Variância da amostra	18930,88
Curtose	0,606413	Curtose	0,28928
Assimetria	-0,49826	Assimetria	-0,60705
Intervalo	497	Intervalo	966
Mínimo	11	Mínimo	8
Máximo	508	Máximo	974
Soma	97308629	Soma	51795764
Contagem	368994	Contagem	101899
Maior(1)	508	Maior(1)	974
Menor(1)	11	Menor(1)	8
Nível de confiança(95,0%)	0,183765	Nível de confiança(95,0%)	0,8448

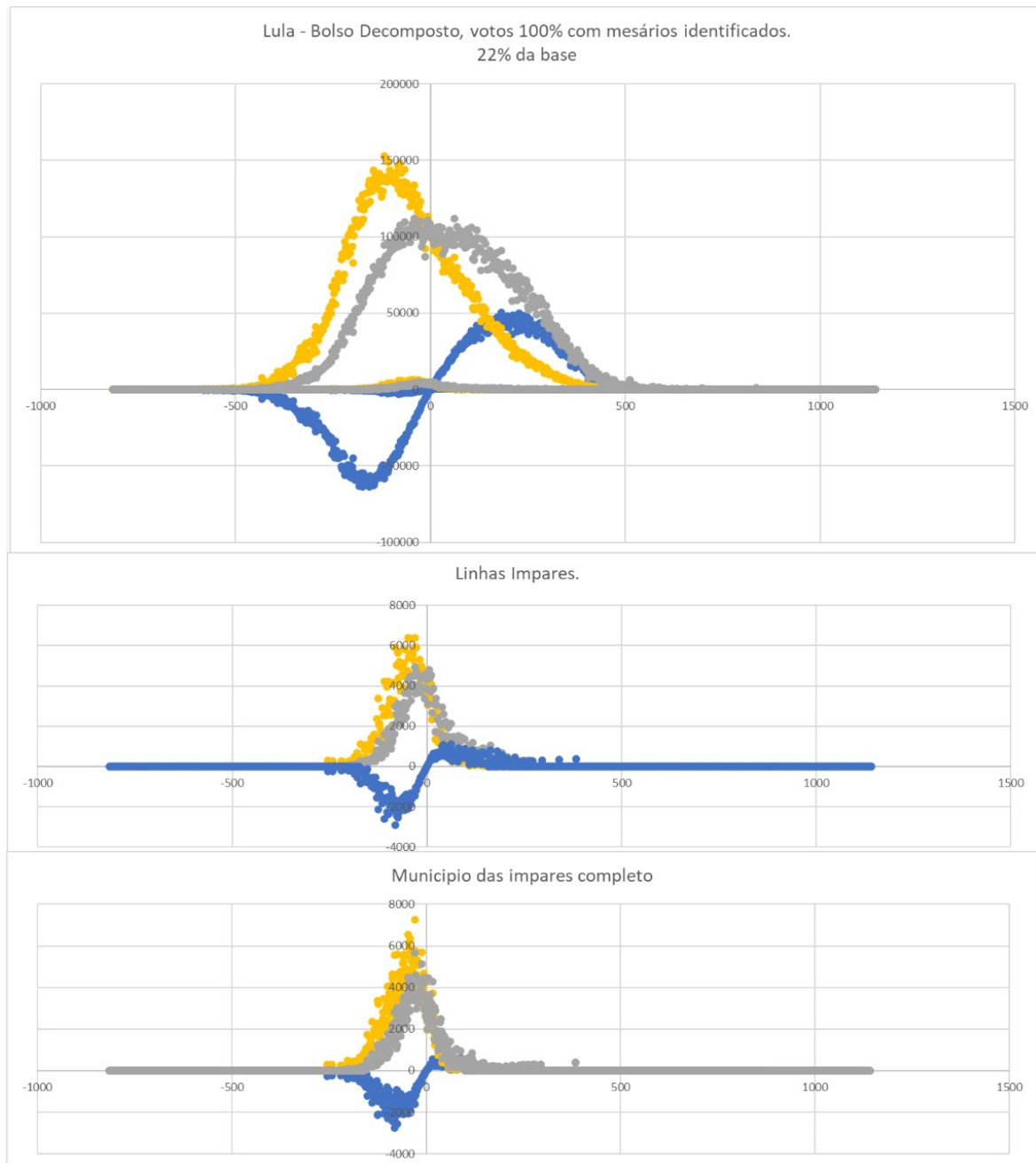
Podemos observar acima que por algum motivo urnas com a dupla falha são **96% maiores** do que urnas em que a falha não ocorreu.

Para aprofundar a análise sobre estes três grupos sobre a mesma base de boletins de urnas separados nos grupos acima em relação a uma curva normal os dados foram agrupados pelo resultado da subtração entre o voto 13 e o voto 22 como x e sumarizados os votos de Lula e Bolsonaro, as curvas resultantes são as seguintes:



Em laranja a quantidade de votos com dupla falha, podemos observar que ela não escalou na mesma proporção que os votos apurados nas urnas, o que por si só já demonstra uma distribuição exclusiva aparentemente mais relacionada a distribuição espacial do que a ocorrência de votos, possivelmente descartando o mesmo como falha, dado que como falha deveria ser mais frequente quanto mais um fenômeno é observado.

Ao obter a mesma curva para a base de dados de urnas **ZERO** foi possível observar:



No primeiro gráfico os votos de Lula e Bolsonaro aparentemente encaixam em normais, mas foi possível observar que os resultados pares e resultados ímpares tinham tendências diferentes perfazendo normais diferentes, a partir desta observação foi possível estratificar mais um público:

Exemplo de observações de dados ímpares e pares em tendências diferentes:

	lula - bols	bolsonaro	lula
-195	0	0	0
-194	-59364	111068	51704
-193	0	0	0
-192	-58368	110624	52256
-191	-191	259	68
-190	-57570	107968	50398
-189	-189	227	38
-188	-57904	111111	53207
-187	-374	456	82
-186	-55428	106405	50977
-185	0	0	0
-184	-56488	109139	52651
-183	-366	467	101
-182	-63154	123698	60544
-181	0	0	0

Sumarizando a situação até o momento:

1 – As urnas foram separadas da ocorrência ou não de duas determinadas linhas no log.

2 – As urnas que não tinham esta linha no log ao sumarizar seus dados pelo RESULTADO de Votos 13 – Votos 22 apresentou DUAS NORMAIS DIFERENTES uma considerando os números PARES e outra considerando os números ímpares que daqui em diante chamaremos de **IMPAR**.

Então sem o encadeamento positivo das duas variáveis as seguintes observações não seriam possíveis:

Os votos das urnas ímpares cobrem todos os estados da federação e, totalizando todos os votos e agrupando as duas massas pela cidade e comparando o total de urnas temos que:

cd_municipio	n_municipio	sg_uf	total	sem falha	mini	com falha	sem falha	sem falha mini	/ total	sem falha	/ total
56014	AFONSO CLAUDIO	ES	74	37	0	74		50,0%		100,0%	
56065	ALTO RIO NOVO	ES	23	12	0	23		52,2%		100,0%	
56111	ARACRUZ	ES	217	103	0	217		47,5%		100,0%	
56138	ATILIO VIVACQUA	ES	28	10	3	25		35,7%		89,3%	
56146	BREJETUBA	ES	32	10	6	26		31,3%		81,3%	
56251	CARIACICA	ES	802	405	0	802		50,5%		100,0%	
56316	CONCEIÇÃO DA BARRA	ES	68	37	0	68		54,4%		100,0%	
56332	CONCEIÇÃO DO CASTELO	ES	30	14	0	30		46,7%		100,0%	
56308	GOVERNADOR LINDENBERG	ES	27	15	0	27		55,6%		100,0%	
56456	GUACUÍ	ES	72	16	19	53		22,2%		73,6%	
56049	IRUPI	ES	33	7	12	21		21,2%		63,6%	
56596	ITUNA	ES	74	15	22	52		20,3%		70,3%	
57134	JAGUARÉ	ES	67	31	0	67		46,3%		100,0%	
56618	JERÔNIMO MONTEIRO	ES	31	8	8	23		25,8%		74,2%	
56634	LINHARES	ES	335	173	0	335		51,6%		100,0%	
56650	MANTENÓPOLIS	ES	33	16	0	33		48,5%		100,0%	
56758	MUQUÍ	ES	37	12	7	30		32,4%		81,1%	
56774	NOVA VENÉCIA	ES	137	74	0	137		54,0%		100,0%	
56790	PANCAS	ES	50	23	0	50		46,0%		100,0%	
57150	PEDRO CANÁRIO	ES	63	24	0	63		38,1%		100,0%	
57118	RIO BANANAL	ES	50	21	0	50		42,0%		100,0%	
57932	SANTA MARIA DE JETIBÁ	ES	83	22	19	64		26,5%		77,1%	
56979	SÃO MATEUS	ES	268	136	0	268		50,7%		100,0%	
56995	SERRA	ES	993	487	0	993		49,0%		100,0%	
56189	SOORETAMA	ES	59	31	0	59		52,5%		100,0%	
57959	VARGEM ALTA	ES	50	16	13	37		32,0%		74,0%	
57975	VENDA NOVA DO IMIGRANTE	ES	54	26	0	54		48,1%		100,0%	
90263	APIACÁS	MT	21	8	0	21		38,1%		100,0%	
90905	BOM JESUS DO ARAGUAIA	MT	17	12	0	17		70,6%		100,0%	
98655	CAMPINÁPOLIS	MT	32	13	0	32		40,6%		100,0%	
90085	CANABRAVA DO NORTE	MT	12	8	0	12		66,7%		100,0%	
90760	CARLINDA	MT	27	15	0	27		55,6%		100,0%	
90549	CASTANHEIRA	MT	17	9	0	17		52,9%		100,0%	
89788	COLNIZA	MT	59	32	0	59		54,2%		100,0%	
90280	CONFRESA	MT	64	33	0	64		51,6%		100,0%	
91006	CONQUISTA D'OESTE	MT	9	5	0	9		55,6%		100,0%	
90786	FELIZ NATAL	MT	22	12	0	22		54,5%		100,0%	
90824	GAÚCHA DO NORTE	MT	18	11	0	18		61,1%		100,0%	
90778	GENERAL CARNEIRO	MT	14	7	0	14		50,0%		100,0%	
89915	JAUURU	MT	24	13	0	24		54,2%		100,0%	
90581	JURUENA	MT	21	11	0	21		52,4%		100,0%	
98710	MARCELÂNIA	MT	32	17	0	32		53,1%		100,0%	
90387	NOVA BANDEIRANTES	MT	29	19	0	29		65,5%		100,0%	
98752	NOVA CANAÃ DO NORTE	MT	36	16	0	36		44,4%		100,0%	
90140	NOVA GUARITA	MT	11	4	0	11		36,4%		100,0%	
90867	NOVA LACERDA	MT	15	7	0	15		46,7%		100,0%	
90166	NOVA MARILÂNDIA	MT	9	3	0	9		33,3%		100,0%	
90123	NOVA MARINGÁ	MT	14	7	0	14		50,0%		100,0%	
90964	NOVA NAZARÉ	MT	8	4	0	8		50,0%		100,0%	
90883	NOVA UBIATÁ	MT	22	9	0	22		40,9%		100,0%	
98795	NOVO HORIZONTE DO NORTE	MT	10	6	0	10		60,0%		100,0%	
90808	NOVO MUNDO	MT	16	10	0	16		62,5%		100,0%	
90948	NOVO SANTO ANTONIO	MT	6	2	0	6		33,3%		100,0%	
98779	NOVO SÃO JOAQUIM	MT	17	11	0	17		64,7%		100,0%	
91359	PORTO DOS GAÚCHOS	MT	12	7	0	12		58,3%		100,0%	
98876	PORTO ESPERIDIÃO	MT	25	14	0	25		56,0%		100,0%	
90182	PORTO ESTRELA	MT	11	6	0	11		54,5%		100,0%	
98914	RESERVA DO CABO AÇUL	MT	6	4	0	6		66,7%		100,0%	
90662	RIBEIRÃO CASCAIS	MT	26	16	0	26		61,5%		100,0%	
89958	RIO BRANCO	MT	13	6	0	13		46,2%		100,0%	
89960	RONDOLÂNDIA	MT	12	4	0	12		33,3%		100,0%	
89974	SALTO DO CÉU	MT	9	2	0	9		22,2%		100,0%	
90980	SANTA CRUZ DO XINGU	MT	6	3	0	6		50,0%		100,0%	
91979	SANTA TEREZINHA	MT	22	8	0	22		36,4%		100,0%	
90328	SANTO AFONSO	MT	7	5	0	7		71,4%		100,0%	
90026	SÃO JOSÉ DO POVO	MT	10	6	0	10		60,0%		100,0%	
90247	SÃO JOSÉ DO XINGU	MT	14	7	0	14		50,0%		100,0%	
90409	SÃO PEDRO DA CIPA	MT	10	6	0	10		60,0%		100,0%	
90220	TABAPORÃ	MT	20	12	0	20		60,0%		100,0%	
98957	TERRA NOVA DO NORTE	MT	34	13	0	34		38,2%		100,0%	
90840	UNIÃO DO SUL	MT	9	6	0	9		66,7%		100,0%	
25011	PALMEIRINA	PE	23	5	8	15		21,7%		65,2%	

Urnas agrupadas por cidade com maior cobertura entre o público **TODOS vs ZERO** e **TODOS vs ÍMPAR**

Podemos observar que as urnas ÍMPAR em uma grande frequência cobrem 50%+ das urnas de um determinado município e que com uma frequência muito grande as urnas **ZERO** cobrem 100% das urnas observadas.

Essa observação me levou a cogitar se não seriam propositalmente agrupadas para observar o processamento das mesmas, o que me levou a analisar o modelo das urnas conforme base urnas para as cidade sumarizadas por modelo:

UF	CD	MUNICIPIO	NM	MUNICIPIO	sem falha impar	sem falha cidade	Modelos Urnas	quantidade
ES	56014	AFONSO CLÁUDIO			50%	100%	UE2010	74
ES	56049	IRUPI			21%	64%	UE2015	33
ES	56065	ALTO RIO NOVO			52%	100%	UE2010	23
ES	56111	ARACRUZ			47%	100%	UE2010	217
ES	56138	ATÍLIO VIVÁQUA			36%	89%	UE2020	28
ES	56146	BREJETUBA			31%	81%	UE2013	32
ES	56189	SOORETAMA			53%	100%	UE2010	59
ES	56251	CARIACICA			50%	100%	UE2010	390
ES	56251	CARIACICA			50%	100%	UE2009	412
ES	56308	GOVERNADOR LINDENBERG			56%	100%	UE2010	27
ES	56316	CONCEIÇÃO DA BARRA			54%	100%	UE2010	68
ES	56332	CONCEIÇÃO DO CASTELO			47%	100%	UE2010	30
ES	56456	GUACUÍ			22%	74%	UE2013	72
ES	56596	IÚNA			20%	70%	UE2013	74
ES	56618	JERÔNIMO MONTEIRO			26%	74%	UE2015	31
ES	56634	LINHARES			52%	100%	UE2009	335
ES	56650	MANTENÓPOLIS			48%	100%	UE2010	33
ES	56758	MUQUI			32%	81%	UE2015	37
ES	56774	NOVA VENÉCIA			54%	100%	UE2010	137
ES	56790	PANCAS			46%	100%	UE2010	50
ES	56979	SÃO MATEUS			51%	100%	UE2009	268
ES	56995	SERRA			49%	100%	UE2010	345
ES	56995	SERRA			49%	100%	UE2009	648
ES	57118	RIO BANANAL			42%	100%	UE2010	50
ES	57134	JAGUARÉ			46%	100%	UE2010	67
ES	57150	PEDRO CANÁRIO			38%	100%	UE2010	63
ES	57932	SANTA MARIA DE JETIBÁ			27%	77%	UE2015	83
ES	57959	VARGEM ALTA			32%	74%	UE2020	50
ES	57975	VENDA NOVA DO IMIGRANTE			48%	100%	UE2010	54
MT	89788	COLNIZA			54%	100%	UE2015	57
MT	89788	COLNIZA			54%	100%	UE2009	3
MT	89915	JAURU			54%	100%	UE2010	23
MT	89915	JAURU			54%	100%	UE2009	1
MT	89958	RIO BRANCO			46%	100%	UE2010	13
MT	89960	RONDOLÂNDIA			33%	100%	UE2011	1
MT	89960	RONDOLÂNDIA			33%	100%	UE2015	11
MT	89974	SALTO DO CÉU			22%	100%	UE2009	4
MT	89974	SALTO DO CÉU			22%	100%	UE2011	2
MT	89974	SALTO DO CÉU			22%	100%	UE2010	3
MT	90026	SÃO JOSÉ DO POVO			60%	100%	UE2020	10
MT	90085	CANABRAVA DO NORTE			67%	100%	UE2009	12
MT	90123	NOVA MARINGÁ			50%	100%	UE2009	14
MT	90140	NOVA GUARITA			36%	100%	UE2015	11
MT	90166	NOVA MARILÂNDIA			33%	100%	UE2011	9
MT	90182	PORTO ESTRELA			55%	100%	UE2013	11
MT	90220	TABAPORÃ			60%	100%	UE2009	20
MT	90247	SÃO JOSÉ DO XINGU			50%	100%	UE2009	14
MT	90263	APIACÁS			38%	100%	UE2015	21
MT	90280	CONFRESA			52%	100%	UE2009	64
MT	90328	SANTO AFONSO			71%	100%	UE2011	7
MT	90387	NOVA BANDEIRANTES			66%	100%	UE2015	29
MT	90409	SÃO PEDRO DA CIPA			60%	100%	UE2020	10
MT	90549	CASTANHEIRA			53%	100%	UE2015	6
MT	90549	CASTANHEIRA			53%	100%	UE2011	10
MT	90549	CASTANHEIRA			53%	100%	UE2009	1
MT	90581	JURUENA			52%	100%	UE2015	21
MT	90662	RIBEIRÃO CASCAIS			62%	100%	UE2009	26
MT	90760	CARLINDA			56%	100%	UE2009	27
MT	90778	GENERAL CARNEIRO			50%	100%	UE2009	1
MT	90778	GENERAL CARNEIRO			50%	100%	UE2010	10
MT	90778	GENERAL CARNEIRO			50%	100%	UE2011	3
MT	90786	FELIZ NATAL			55%	100%	UE2009	18
MT	90786	FELIZ NATAL			55%	100%	UE2011	4
MT	90808	NOVO MUNDO			63%	100%	UE2015	16
MT	90824	GAÚCHA DO NORTE			61%	100%	UE2015	18
MT	90840	UNIÃO DO SUL			67%	100%	UE2015	9
MT	90867	NOVA LACERDA			47%	100%	UE2015	15
MT	90883	NOVA UBIRATÃ			41%	100%	UE2013	22
MT	90905	BOM JESUS DO ARAGUAIA			71%	100%	UE2015	17
MT	90948	NOVO SANTO ANTÔNIO			33%	100%	UE2015	4
MT	90948	NOVO SANTO ANTÔNIO			33%	100%	UE2009	2
MT	90964	NOVA NAZARÉ			50%	100%	UE2009	9
MT	90980	SANTA CRUZ DO XINGU			50%	100%	UE2015	6
MT	91006	CONQUISTA D'OESTE			56%	100%	UE2011	1
MT	91006	CONQUISTA D'OESTE			56%	100%	UE2015	8
MT	91359	PORTO DOS GAÚCHOS			58%	100%	UE2009	12
MT	91979	SANTA TEREZINHA			36%	100%	UE2015	22
MT	98655	CAMPINÁPOLIS			41%	100%	UE2009	32
MT	98710	MARCELÂNDIA			53%	100%	UE2009	11
MT	98710	MARCELÂNDIA			53%	100%	UE2015	21
MT	98752	NOVA CANAÃ DO NORTE			44%	100%	UE2015	36
MT	98779	NOVO SÃO JOAQUIM			65%	100%	UE2009	17
MT	98795	NOVO HORIZONTE DO NORTE			60%	100%	UE2009	10
MT	98876	PORTO ESPERIDIÃO			56%	100%	UE2011	23
MT	98876	PORTO ESPERIDIÃO			56%	100%	UE2009	2
MT	98914	RESERVA DO CABAÇAL			67%	100%	UE2010	6
MT	98957	TERRA NOVA DO NORTE			38%	100%	UE2015	34

O que nos levou a observar que os modelos de urnas para cada cidade são quase sempre exclusivos em comparação ao resto da massa em que foram empregadas um número misto de urnas.

Como engenheiro de sistemas me parece a implementação de uma massa de testes para um piloto controlado facilmente observável através de gráficos.

Surpreende, no entanto, a necessidade de manipular o resultado de mais de 101 mil urnas em que as mensagens de LOG objetivo do estudo pode potencialmente ter sido desativado propositalmente, assim como essa manipulação necessitar de código fonte específico que pode ser implementado como parâmetro ou integrado como um novo código fonte, o que faz mais sentido, dado que como estamos falando de hardware, uma massa de testes desse tipo deveria envolver inclusive o próprio código fonte.

A implementação disto aparentemente foi feita sem camadas de controle e até forjando dados, estes mesmos artifícios podem ter sido utilizados por um agente externo com um simples acesso ao código fonte e conhecimento dos procedimentos e das chaves, manipulando o resultado da própria eleição.

Distribuição de falhas:

Voltando ao objetivo original do estudo, fizemos análise de coberturas por categoria, estruturei a nossa análise com o objetivo de comparar através da fórmula de correlação os diversos municípios do Brasil e, como fator de correlação o voto dado, desta forma os dados foram estratificados em `cd_municipio` e `nr_votável` do boletim de urna dos dados informados pelos dados abertos do TSE, resultando em 12986 categorias a serem comparadas nas análises seguintes.

Cobertura de base.

A primeira análise é uma análise de cobertura das categorias das bases conforme evoluem as quantidades de “dupla falha” nas seções:

Cobertura / categoria categoria = cd_municipio + nr_votavel: total 12986							
	TODOS	ZERO	>0	>8	>13	>21	>34
Amostra / Base	100,0%	22%	78%	31,1%	20,0%	10,4%	3,9%
Categorias preenchidas cidade / voto	99,996%	92,306%	96,637%	94,934%	92,540%	77,400%	24,594%

Nas colunas:

TODOS: todos os votos

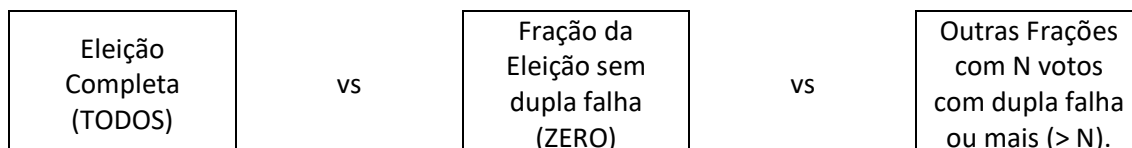
ZERO: seções **sem** votos com “dupla falha”.

>N : seções **com** mais que N votos com “dupla falha”, estratificadas conforme uma sequência numérica Fibonacci a partir de 8 com o objetivo de obter números mais significativos.

Podemos observar que com **ZERO** e com **>13** votos com “dupla falha” temos **20%** ou mais dos votos o que já tornaria uma amostra válida pela regra de Pareto, considerando uma divisão aleatória, em ambos extratos temos mais de 92% das categorias preenchidas.

Correlação entre os públicos estratificado por quantidade de defeitos

Se fossemos considerar falha simples e aleatória de procedimento, considerando uma cobertura de municípios x voto de 92% em uma classificação de 12986 categorias para estratificar os dados para identificar se as bases não mudam a correlação com a ocorrência de “dupla falha” fizemos uma correlação entre grupos da seguinte forma:



Correlação de amostras: TODOS e ZERO vs >N "duplas falhas"							
	TODOS	ZERO	>0	>8	>13	>21	>34
TODOS	-	0,95390432	0,996557817	0,959650261	0,917135099	0,794577436	0,563954693
ZERO	-0,95390432	-	0,925741418	0,840652189	0,774486455	0,61937406	0,396253637

Foi utilizada a fórmula de correlação CORREL(AmostraA, AmostraB) do excel, considerando as 12986 categorias da amostra TODOS preenchidas zeros os dados não encontrados em outras amostras.

Por padrões estatísticos vamos utilizar o grau de confiança de 95% como padrão para as afirmações do resultado das nossas estatísticas, desta forma podemos afirmar:

1: pertencem a um público compatível a amostra que contém **TODOS** os votos e **ZERO** votos com “dupla falha” (95,3 > 95 é verdadeiro).

2: pertencem a um público compatível a amostra que contém **TODOS** os votos e **>0** votos com “dupla falha” (99,6 > 95 é verdadeiro).

3: não pertencem a um público compatível amostras com **ZERO** votos com “dupla falha” e **>0** votos com “dupla falha” (92,5 > 95 é falso).

E considerando que o fato da cobertura de 20% ser um fator relevante na correlação, a mesma amostra com **>13** votos com dupla falha apresentam falsa correlação com o **TODOS** (79,4 > 95 é falso) os votos e com a amostra de **ZERO** (61,9 > 95 é falso) votos com “dupla falha”, assim:

4: não pertence a um público compatível a amostra com **>13** votos com dupla falha incompatível com a amostra que contém **TODOS** os votos e com a amostra que contém **ZERO** votos com dupla falha.

Desta a afirmativa de que as amostras que a “dupla falha” acima descrita tenha ocorrência puramente aleatória é falsa com grau de certeza de 95%.

Resultado da apuração do segundo turno nas amostras:

Ok, a comparação entre amostras mostra que a dupla falha não é aleatória, e, como depende da operação de um usuário pode indicar ação intencional e não intencional, a ação não intencional tende a ser aleatória, já a ação intencional não é aleatória e visa atingir um determinado resultado.

Vamos ver qual é o impacto dos votos com “dupla falha” sobre os votos dos dois candidatos:

	Apuração						
	TODOS	ZERO	>0	>8	>13	>21	>34
% 22 / validos	49,10%	50,20%	48,79%	47,63%	46,88%	45,60%	43,20%
% 13 / validos	50,90%	49,80%	51,21%	52,37%	53,12%	54,40%	56,80%
% brancos / total	1,47%	1,47%	1,47%	1,44%	1,42%	1,36%	1,29%
% nulos / total	3,27%	3,22%	3,28%	3,27%	3,25%	3,20%	3,17%

É possível ver que quanto maior a concentração de duplas falhas uma urna tem maior o resultado percentual para um dos candidatos, observamos também que os votos brancos e nulos têm uma distorção negativa, começando em um percentual que dá a vitória ao candidato com menos votos no resultado que considera o público como um todo.

O que nos leva a não descartar a intencionalidade da dupla falha.

Resultado de uma análise de falha sigma sobre e o resultado numérica quantitativa sobre o resultado.

Vamos considerar a modelagem de defeito industrial como padronizado pela Fundação Seis Sigma para processos industriais no levantamento inicial de oportunidades do fluxo DMAIC:

Em uma abordagem pragmática considerando que a formula de:

Fórmula DPMO	Tabela DPMO x Sigma
$DPMO = \frac{\text{número defeitos} \times 10^6}{\text{número oportunidades} \times \text{número peças}}$	ppm/ DPMO
	Nível Sigma
	0,002
	2
	96
	2700
	71860
	134000

Número de defeitos: Cada defeito identificado na verdade consiste de dois defeitos: um defeito no reconhecimento da biometria do eleitor, e um segundo defeito na biometria do mesário.

Número de oportunidades: Duas oportunidades por peça.

Número de peças: Número Total de Votos.

Estou chamando de pragmático pois o número de defeitos poderia ser ainda maior se considerarmos a falha única que seria a falha apenas da biometria do eleitor, nestes casos uma vez que estamos considerando que ambas as observações estão sendo feitos para ambas as leituras (mesário + eleitor) o tempo todo o que não é um fato, mas que otimizaria o número apresentando um DPMO menor e consequentemente o sigma maior, mesmo assim observamos:

Cálculo sigma de defeitos para cada amostra							
	TODOS	ZERO	>0	>8	>13	>21	>34
Votos	124.252.796	26.944.167	97.308.629	40.518.421	26.447.507	14.167.346	5.529.899
Falhas	4.001.032	-	4.001.032	3.241.959	2.667.628	1.876.477	986.060
% Falhas	3,2%	0,0%	4,1%	8,0%	10,1%	13,2%	17,8%
Defeitos	8.002.064	-	8.002.064	6.483.918	5.335.256	3.752.954	1.972.120
Observações	2	2	2	2	2	2	2
Unidades	124.252.796	26.944.167	97.308.629	40.518.421	26.447.507	14.167.346	5.529.899
Oportunidades	248.505.592	53.888.334	194.617.258	81.036.842	52.895.014	28.334.692	11.059.798
DPMO	32.201	-	41.117	80.012	100.865	132.451	178.314
Sigmas	2,4	3,4	2,0	1,8	1,6	1,5	1,4

O padrão de um processo otimizado em sigma é de **3** para um processo industrial bem construído, entretanto podemos observar que considerando a amostra integral, o sigma está em 2,4 o que é abaixo do ideal (sigma, quanto maior melhor) e, se considerar apenas a massa em que ocorreu pelo menos uma vez por urna, **2**, o que é quase inaceitável em um processo industrial, especialmente se consideramos que o novo padrão da indústria tende a ser o 6.

Efeito potencial sobre o resultado da eleição:

Votos							
Diferença efetiva sobre o resultado							
	TODOS	ZERO	>0	>8	>13	>21	>34
22	58.206.354	12.911.237	45.295.117	18.418.951	11.837.702	6.175.107	2.285.669
13	60.345.999	12.809.312	47.536.687	20.250.509	13.414.503	7.366.049	3.005.059
95	1.769.678	383.863	1.385.815	565.009	363.935	187.355	69.186
96	3.930.765	839.755	3.091.010	1.283.952	831.367	438.835	169.985
Dif 22 - 13	- 2.139.645	101.925	- 2.241.570	- 1.831.558	- 1.576.801	- 1.190.942	- 719.390
% Dif / Eleito	100,00%	-4,76%	104,76%	85,60%	73,69%	55,66%	33,62%
Duplas falhas:	4.001.032	-	4.001.032	3.241.959	2.667.628	1.876.477	986.060

Montando apenas os dados quantitativos podemos observar que: o número da diferença entre os candidatos sempre é maior que o número de defeitos independente da quantidade de erros por urna, a diferença entre os votos apurados para o candidato 13 e candidato 22, é positiva para 22 com **ZERO** duplas falhas.

Conclusão:

Em uma análise técnica podemos observar que considerando apenas as urnas que tiveram erros catastróficos, que seria o erro no reconhecimento biométrico simultâneo do eleitor e do mesário, sem a qual não é possível avaliar a lisura completa da votação, é muito abaixo de um padrão industrialmente aceito, calculado à 2.4 Sigma, considerando ser este um projeto moderno de processo está bastante fora do padrão, é possível afirmar ainda que o resultado da votação está sendo distorcido para um dos candidatos na proporção da existência da ocorrência do defeito na votação da urna, criando públicos entre urnas que não se correlacionam considerando a estratificação por quantidade de defeitos, reforçando a impressão de que uma mudança relevante no resultado teve a origem em um defeito na votação.

A análise quantitativa indica que, se descartadas as urnas com defeito catastrófico, o resultado final da eleição seria mudado drasticamente dando a vitória a outro candidato.