

CloudWalk

Business Monitoring Intelligence Intern / Daniel Campos
França / Universidade Estácio de Sá / 23.07.2024



cloudwalk

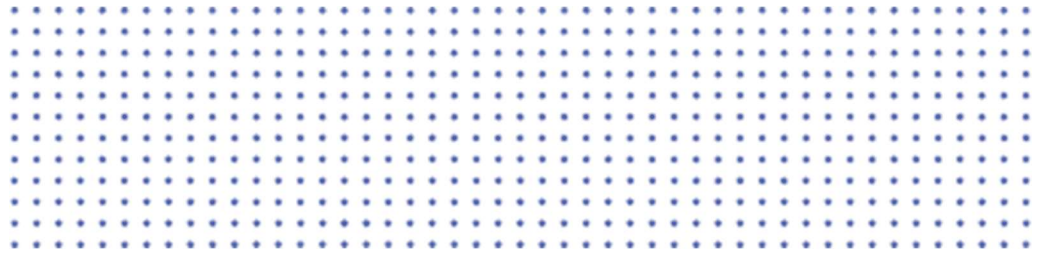


TABLE OF CONTENTS

SUMÁRIO

1. INTRODUÇÃO	3
2. DESAFIO 1	3
3. DESAFIO 2	15
4. CONCLUSÃO FINAL	23



1. INTRODUÇÃO

Este relatório técnico descreve como o candidato ao cargo de estagiário de analista de monitoramento na CloudWalk, desenvolveu as soluções para os dois desafios propostos, sendo um, a análise de anomalias em dados hipotéticos de checkout e outro, o desenvolvimento de solução de problema real relacionado com alerta de incidentes nas transações do cotidiano da empresa.

Descrição dos Desafios

O **primeiro desafio** envolve a análise de um conjunto de dados hipotéticos de checkout fornecidos num arquivo CSV. O candidato deve desenvolver um modelo de análise para identificar comportamentos anômalos nas operações, apresentar os resultados obtidos e suas conclusões.

O **segundo desafio** envolve o desenvolvimento de uma solução de alerta de incidentes de transações, utilizando arquivos CSV fornecidos. O candidato deve desenvolver um sistema que monitore transações em tempo real, detectando e alertando às equipes sobre anomalias nas transações com falha, revertidas ou negadas.

Objetivos

- **Análise de Dados:** Realizar uma análise detalhada dos dados fornecidos, identificando padrões e anomalias.
- **Desenvolvimento de Sistema:** Implementar um sistema de monitoramento que detecte anomalias em tempo real e notifique as equipes relevantes.
- **Apresentação de Resultados:** Comunicar de forma clara e concisa os resultados da análise e o funcionamento do sistema desenvolvido, através de uma apresentação estruturada e detalhada.

2. DESAFIO 1 – Get your hand dirty

O primeiro desafio consiste em realizar uma análise de dados hipotéticos de checkout, utilizando um arquivo CSV (Comma-Separated Values) fornecido, com o objetivo de identificar comportamentos anômalos em um POS (Point of Sale), apresentar os resultados obtidos e as conclusões. Esta tarefa exigiu habilidades de



interpretação de dados, criação de algoritmos para calcular e reconhecer dados anômalos, elaboração de consultas SQL e criação de representações gráficas para comunicar os insights encontrados.

Métodos matemáticos utilizados

1. Média Ponderada (Weighted Mean): A média ponderada foi calculada combinando a média dos últimos 7 dias e do último mês:
$$\text{weighted_mean} = \frac{(avg_{last_week} \times 7) + (avg_{last_month} \times 30)}{(7 + 30)}$$
2. Média Simples (Mean): A média aritmética dos valores diários e das médias semanais e mensais:
$$\text{mean} = \frac{today + yesterd + day_{last_week} + avg_{last_week} + avg_{last_month}}{5}$$
3. Variância: A variância mede a dispersão dos dados em relação à média. Foi calculada usando:
$$\text{variance} = \frac{\sum (x - \text{mean})^2}{n - 1}$$
4. Desvio Padrão: O desvio padrão é a raiz quadrada da variância e quantifica a dispersão dos dados:
$$\text{std_dev} = \sqrt{\text{variance}}$$
5. Limites Superior e Inferior: Determinados com base na média ponderada e no desvio padrão para identificar anomalias.
 - **Limite Superior:** $\text{upper_limit} = \text{weighted_mean} + (2 \times \text{std_dev})$
 - **Limite Inferior:** $\text{lower_limit} = \text{weighte_mean} - (2 \times \text{std_dev})$
6. Detecção de Anomalias dentro dos limites superiores e inferiores: Identificação de valores que estão abaixo de 50% da média ponderada como possíveis anomalias.

Métodos utilizados

O **primeiro método** de análise utiliza o cálculo do desvio padrão e dos limites superiores e inferiores, onde o valor de K é uma constante definida como 2 ($K = 2$). Cabe observar, que o valor de K determina a largura do intervalo de controle, onde K poderia eventualmente assumir os seguintes valores: K=1 abrange 68% dos dados, K=2 abrange 95% dos dados e K=3 abrange 99,7% dos dados. Este método permite identificar valores que se encontram fora dos limites calculados, classificando-os como anomalias. Valores acima do limite podem ser considerados fraudulentos, como no caso de um cartão roubado sendo usado para múltiplos pagamentos por aproximação no mesmo ponto de venda (POS).



Relatório Técnico

O **segundo método**, complementa-se o primeiro método, que se limita a identificar valores acima ou abaixo dos limites estabelecidos. No entanto, existem também valores dentro desses limites que podem ser considerados anômalos devido a desvios significativos da média. Isso pode indicar problemas com o sistema de pagamentos ou falhas de sinal ao efetuar transações. Para identificar essas anomalias, utilizou-se a média ponderada como valor de referência, comparando-a com os valores de "Today", "Yesterday" e "Same Day Last Week". Um valor é considerado anômalo se for 50% menor que a média ponderada.

Análise do checkout 1

Ao processar os dados conforme explicado anteriormente, foram identificadas as anomalias descritas pelos gráficos e tabelas dos Métodos 1 e 2.

O primeiro gráfico e tabela representam as anomalias encontradas utilizando o Método 1, onde foram detectadas 17 anomalias. Na tabela, as células pintadas de verde correspondem às anomalias na coluna "Today", as células amarelas correspondem às anomalias na coluna "Yesterday", e as células azuis correspondem às anomalias na coluna "Same Day Last Week". A coluna "upper_limit" indica o limite superior para considerar um valor como anômalo e está pintada com a cor correspondente ao dia para facilitar a visualização. Esses valores encontrados estão sendo representados por suas respectivas cores no gráfico.

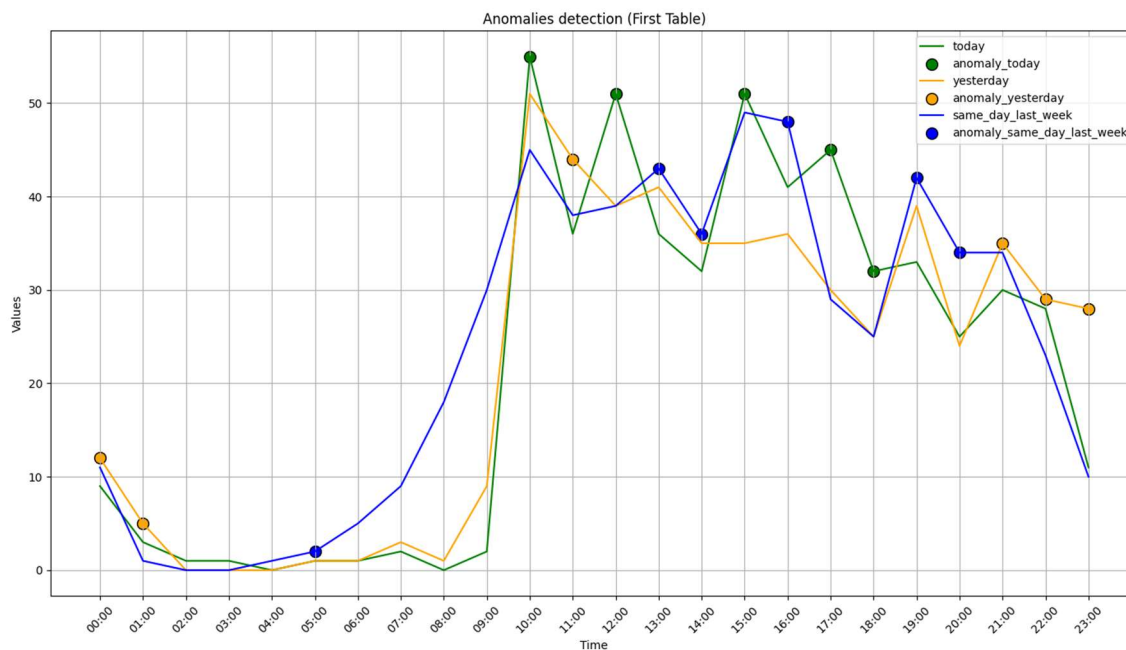
No segundo gráfico e tabela, foram identificadas 13 anomalias. As células possuem as mesmas cores da primeira tabela: verde para "Today", amarelo para "Yesterday", e azul para "Same Day Last Week". As colunas "today_vs_weighted_mean", "yesterday_vs_weighted_mean" e "same_day_last_week_vs_weighted_mean" são colunas do tipo booleano, ou seja, contêm apenas valores verdadeiros ou falsos, com o valor verdadeiro indicando uma anomalia encontrada.



Relatório Técnico

Método de análise 1

Gráfico





Relatório Técnico

Tabela

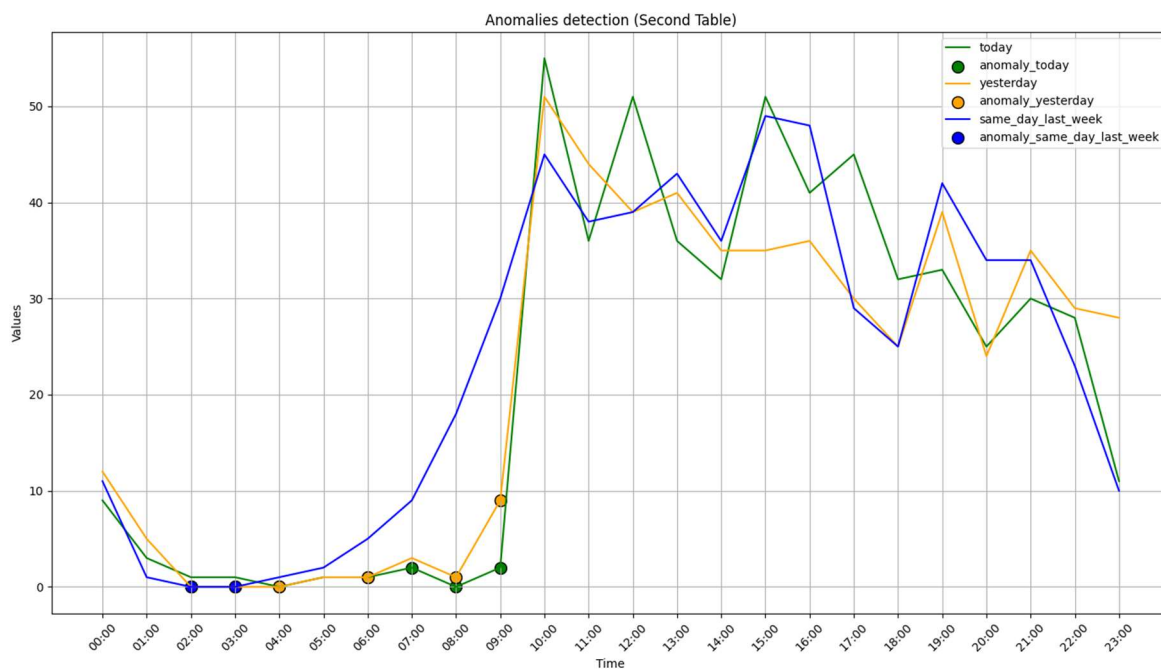
time	today	yesterday	same_day_last_week	avg_last_week	avg_last_month	weighted_mean	mean	variance	std_dev	upper_limit	lower_limit
0:00	9	12	11	6.42	4.85	5.15	8.65	9.07	3.01	11.17	-0.88
1:00	3	5	1	1.85	1.92	1.91	2.55	2.37	1.54	4.99	-1.17
2:00	1	0	0	0.28	0.82	0.72	0.42	0.22	0.47	1.65	-0.21
3:00	1	0	0	0.42	0.46	0.45	0.38	0.17	0.41	1.28	-0.37
4:00	0	0	1	0.42	0.21	0.25	0.33	0.17	0.42	1.08	-0.58
5:00	1	1	2	1.28	0.75	0.85	1.21	0.23	0.48	1.81	-0.11
6:00	1	1	5	2.85	2.28	2.39	2.43	2.72	1.65	5.69	-0.91
7:00	2	3	9	5.57	5.21	5.28	4.96	7.34	2.71	10.7	-0.14
8:00	0	1	18	8.71	10.42	10.1	7.63	54.67	7.39	24.88	-4.69
9:00	2	9	30	20	19.07	19.25	16.01	116.61	10.8	40.84	-2.35
10:00	55	51	45	29.42	28.35	28.55	41.75	150.82	12.28	53.11	3.99
11:00	36	44	38	33.71	28.5	29.49	36.04	32.37	5.69	40.86	18.11
12:00	51	39	39	27.57	25.42	25.83	36.4	106.3	10.31	46.45	5.21
13:00	36	41	43	25.85	24.21	24.52	34.01	74.07	8.61	41.73	7.31
14:00	32	35	36	26.14	25.21	25.39	30.87	24.76	4.98	35.34	15.43
15:00	51	35	49	28.14	27.71	27.79	38.17	125.49	11.2	50.2	5.39
16:00	41	36	48	27.71	25.64	26.03	35.67	86.13	9.28	44.59	7.47
17:00	45	30	29	20.42	22.28	21.93	29.34	93.8	9.68	41.3	2.56
18:00	32	25	25	21.57	18.28	18.9	24.37	25.98	5.1	29.1	8.71
19:00	33	39	42	22.14	18.67	19.33	30.96	104.88	10.24	39.81	-1.16
20:00	25	24	34	17.42	18.92	18.64	23.87	42.5	6.52	31.68	5.6
21:00	30	35	34	18.71	17.57	17.79	27.06	69.91	8.36	34.51	1.06
22:00	28	29	23	15.42	15.64	15.6	22.21	42.38	6.51	28.62	2.58
23:00	11	28	10	9.57	8.75	8.91	13.46	66.69	8.17	25.24	-7.43



Relatório Técnico

Método de análise 2

Gráfico





Relatório Técnico

Tabela

time	today	yesterday	same_day_last_week	weighted_mean	today_vs_weighted_meany	yesterday_vs_weighted_mean	same_day_last_week_vs_weighted_mean
0:00	9	12	11	5.15	FALSE	FALSE	FALSE
1:00	3	5	1	1.91	FALSE	FALSE	FALSE
2:00	1	0	0	0.72	FALSE	TRUE	TRUE
3:00	1	0	0	0.45	FALSE	TRUE	TRUE
4:00	0	0	1	0.25	TRUE	TRUE	FALSE
5:00	1	1	2	0.85	FALSE	FALSE	FALSE
6:00	1	1	5	2.39	TRUE	TRUE	FALSE
7:00	2	3	9	5.28	TRUE	FALSE	FALSE
8:00	0	1	18	10.1	TRUE	TRUE	FALSE
9:00	2	9	30	19.25	TRUE	TRUE	FALSE
10:00	55	51	45	28.55	FALSE	FALSE	FALSE
11:00	36	44	38	29.49	FALSE	FALSE	FALSE
12:00	51	39	39	25.83	FALSE	FALSE	FALSE
13:00	36	41	43	24.52	FALSE	FALSE	FALSE
14:00	32	35	36	25.39	FALSE	FALSE	FALSE
15:00	51	35	49	27.79	FALSE	FALSE	FALSE
16:00	41	36	48	26.03	FALSE	FALSE	FALSE
17:00	45	30	29	21.93	FALSE	FALSE	FALSE
18:00	32	25	25	18.9	FALSE	FALSE	FALSE
19:00	33	39	42	19.33	FALSE	FALSE	FALSE
20:00	25	24	34	18.64	FALSE	FALSE	FALSE
21:00	30	35	34	17.79	FALSE	FALSE	FALSE
22:00	28	29	23	15.6	FALSE	FALSE	FALSE
23:00	11	28	10	8.91	FALSE	FALSE	FALSE



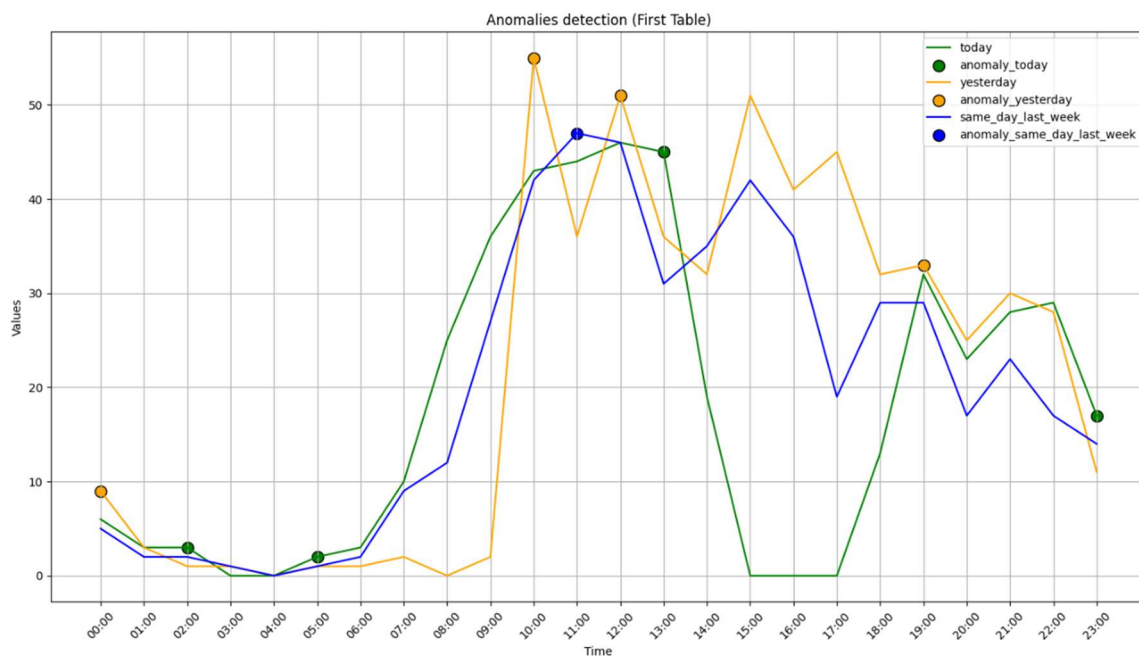
Relatório Técnico

Análise do checkout 2

Ao processar os dados do checkout 2 da mesma forma como processado os dados no checkout 1, se pode observar que no primeiro processo foram identificadas 09 anomalias e no segundo processo foram identificadas 10 anomalias.

Método de análise 1

Gráfico





Relatório Técnico

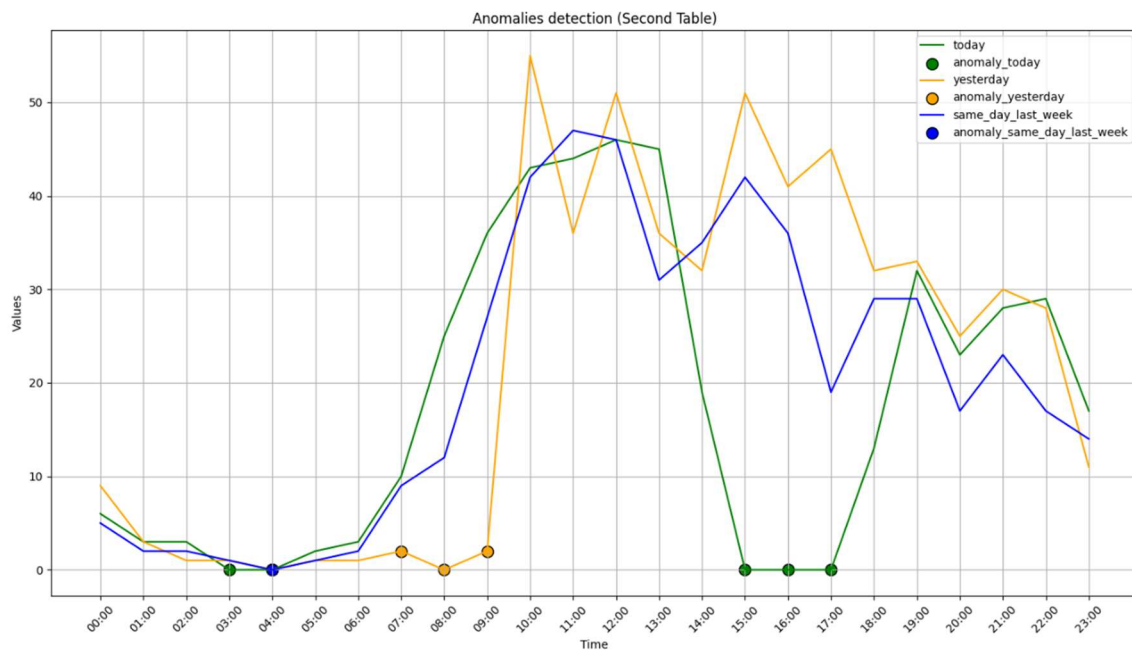
Tabela

time	today	yesterday	same_day_last_week	avg_last_week	avg_last_month	weighted_mean	mean	variance	std_dev	upper_limit	lower_limit
0:00	6	9	5	5	4.92	4.94	5.98	3.04	1.74	8.42	1.45
1:00	3	3	2	2	1.92	1.94	2.38	0.32	0.56	3.06	0.81
2:00	3	1	2	0.42	0.75	0.69	1.43	1.11	1.06	2.8	-1.42
3:00	0	1	1	0.42	0.46	0.45	0.58	0.18	0.43	1.31	-0.4
4:00	0	0	0	0.14	0.21	0.2	0.07	0.01	0.1	0.39	0
5:00	2	1	1	0.71	0.71	0.71	1.08	0.28	0.53	1.77	-0.35
6:00	3	1	2	1.42	2.1	1.97	1.9	0.58	0.76	3.49	0.45
7:00	10	2	9	3	5.03	4.65	5.81	12.69	3.56	11.77	-2.48
8:00	25	0	12	3.71	9.82	8.66	10.11	92.14	9.6	27.86	-10.53
9:00	36	2	27	10.14	17.64	16.22	18.56	180.34	13.43	43.08	-10.64
10:00	43	55	42	26.14	28.57	28.11	38.94	138.79	11.78	51.67	4.55
11:00	44	36	47	25	28.28	27.66	36.06	91.4	9.56	46.78	8.54
12:00	46	51	46	24	25.89	25.53	38.58	159.5	12.63	50.79	0.27
13:00	45	36	31	20.28	24.17	23.43	31.29	95.54	9.77	42.98	3.89
14:00	19	32	35	19.57	24.89	23.88	26.09	52.13	7.22	38.32	9.44
15:00	0	51	42	22.427	27.78	26.77	28.64	384.51	19.61	65.99	-12.45
16:00	0	41	36	21.57	25.53	24.78	24.82	253.47	15.92	56.62	-7.06
17:00	0	45	19	17.71	22.67	21.73	20.88	258.63	16.08	53.9	-10.43
18:00	13	32	29	16.85	18.46	18.16	21.86	67.24	8.2	34.56	1.76
19:00	32	33	29	18	18.21	18.17	26.04	54.67	7.39	32.96	3.38
20:00	23	25	17	12.14	18.53	17.32	19.13	25.8	5.08	27.48	7.16
21:00	28	30	23	14.85	17.82	17.26	22.73	41.73	6.46	30.18	4.34
22:00	29	28	17	12.71	15.5	14.97	20.44	56.6	7.52	30.02	-0.08
23:00	17	11	14	8.28	8.75	8.66	11.81	13.55	3.68	16.02	1.3



Método de análise 2

Gráfico





Relatório Técnico

Tabela

time	today	yesterday	same_day_last_week	weighted_mean	today_vs_weighted_mean	yesterday_vs_weighted_mean	same_day_last_week_vs_weighted_mean
0:00	6	9	5	4.94	FALSE	FALSE	FALSE
1:00	3	3	2	1.94	FALSE	FALSE	FALSE
2:00	3	1	2	0.69	FALSE	FALSE	FALSE
3:00	0	1	1	0.45	TRUE	FALSE	FALSE
4:00	0	0	0	0.2	TRUE	TRUE	TRUE
5:00	2	1	1	0.71	FALSE	FALSE	FALSE
6:00	3	1	2	1.97	FALSE	FALSE	FALSE
7:00	10	2	9	4.65	FALSE	TRUE	FALSE
8:00	25	0	12	8.66	FALSE	TRUE	FALSE
9:00	36	2	27	16.22	FALSE	TRUE	FALSE
10:00	43	55	42	28.11	FALSE	FALSE	FALSE
11:00	44	36	47	27.66	FALSE	FALSE	FALSE
12:00	46	51	46	25.53	FALSE	FALSE	FALSE
13:00	45	36	31	23.43	FALSE	FALSE	FALSE
14:00	19	32	35	23.88	FALSE	FALSE	FALSE
15:00	0	51	42	26.77	TRUE	FALSE	FALSE
16:00	0	41	36	24.78	TRUE	FALSE	FALSE
17:00	0	45	19	21.73	TRUE	FALSE	FALSE
18:00	13	32	29	18.16	FALSE	FALSE	FALSE
19:00	32	33	29	18.17	FALSE	FALSE	FALSE
20:00	23	25	17	17.32	FALSE	FALSE	FALSE
21:00	28	30	23	17.26	FALSE	FALSE	FALSE
22:00	29	28	17	14.97	FALSE	FALSE	FALSE
23:00	17	11	14	8.66	FALSE	FALSE	FALSE



Conclusão – Desafio 1 Get your hands dirty

A análise dos dados, conforme detalhado nos métodos apresentados, revelou informações significativas sobre as anomalias dos dados analisados, conforme relatado a seguir:

1. Análise das Anomalias Identificadas:

- Valores Acima dos Limites Superiores: Valores que excedem os limites superiores podem sugerir eventos locais significativos, como um aumento abrupto no número de transações. Deve ser avaliado se tais picos estão associados a eventos especiais, promoções ou períodos de alta demanda.
- Valores Anômalos que tendem a Zero: Valores anômalos iguais a zero podem indicar possíveis problemas operacionais, como falta de energia no local, o que resultaria no fechamento temporário do estabelecimento, ou imprevistos que não estão relacionados ao sistema de pagamentos.

2. A análise das tabelas de checkout 1 e 2 forneceu insights adicionais que podem ser úteis para a investigação:

- Transações 24 Horas: O local em questão registra transações ao longo de 24 horas, sugerindo que é uma loja de conveniência aberta 24 horas. O maior número de clientes frequenta o estabelecimento entre aproximadamente 8h e 22h.
- Investigação de Transações Fraudulentas: Para verificar a possibilidade de transações fraudulentas, é recomendável examinar o local onde o POS (ponto de venda) está instalado. Algumas localidades podem apresentar taxas de fraude mais elevadas do que outras, e a análise do histórico de transações pode ajudar a identificar padrões suspeitos.

Finalizando, a análise inicial fornece uma base sólida para compreender as anomalias detectadas. As investigações subsequentes devem focar em correlacionar essas anomalias com eventos específicos e avaliar possíveis causas operacionais ou de segurança.



3. DESAFIO 2 – Solve the problem

O desafio apresentado envolve a implementação de um sistema de monitoramento de alertas em tempo real para transações financeiras, utilizando dados de um arquivo CSV contendo informações sobre transações aprovadas, negadas, falhas e revertidas, com objetivo de identificar comportamentos anômalos e gerar notificações automáticas para as equipes responsáveis.

Esta tarefa exigiu habilidades de interpretação de dados, criação de algoritmos para calcular e reconhecer anomalias, elaboração de consultas SQL e desenvolvimento de representações gráficas para comunicar os insights encontrados. Além disso, foi necessário implementar um endpoint para receber dados de transações e retornar recomendações de alertas, garantindo que o sistema pudesse operar de maneira eficiente e eficaz.

Métodos matemáticos utilizados

1. Média Simples (Mean)
2. Desvio Padrão
3. Z-score : é uma forma padronizada de medir o quanto um valor se desvia da média de um conjunto de dados: $z = \frac{x - \mu}{\sigma}$
 - x: O valor individual do conjunto de dados que está sendo analisado.
 - μ : A média dos valores do conjunto de dados.
 - σ : O desvio padrão dos valores do conjunto de dados.

Método utilizado

O método utilizado para identificar as anomalias no arquivo "transactions" consistiu em somar todos os valores por status em cada uma das 24 horas apresentadas. Após essa soma, foi calculada a média simples dos valores para os status "denied", "reversed" e "failed", resultando em 24 métricas para cada status. Definiu-se um z-score de 0,7, pois, após análise dos dados processados, verificou-se que valores acima desse limiar não são reconhecidos como anomalias, pois são perceptíveis ao se inspecionar a tabela.

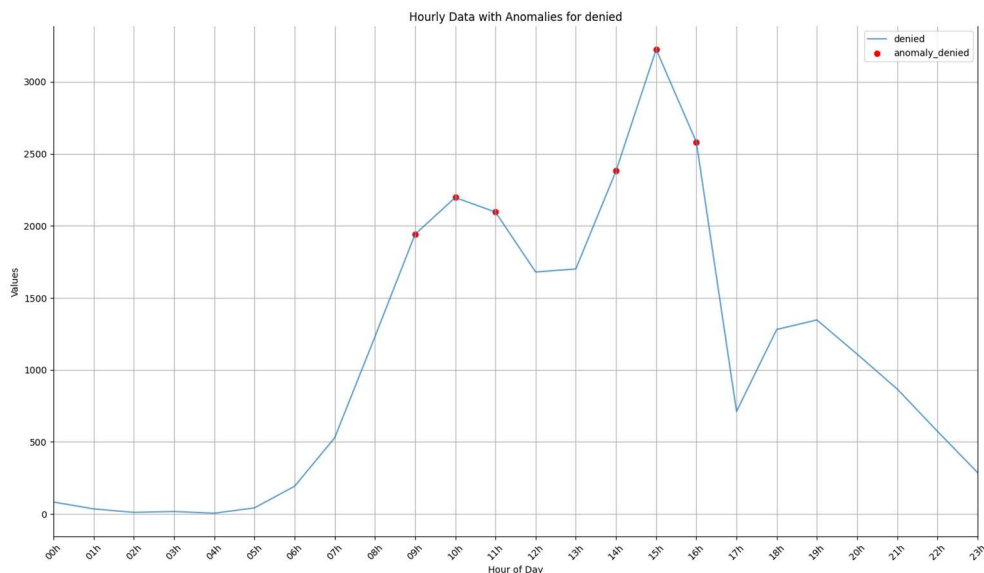


Relatório Técnico

Análise do transaction 1

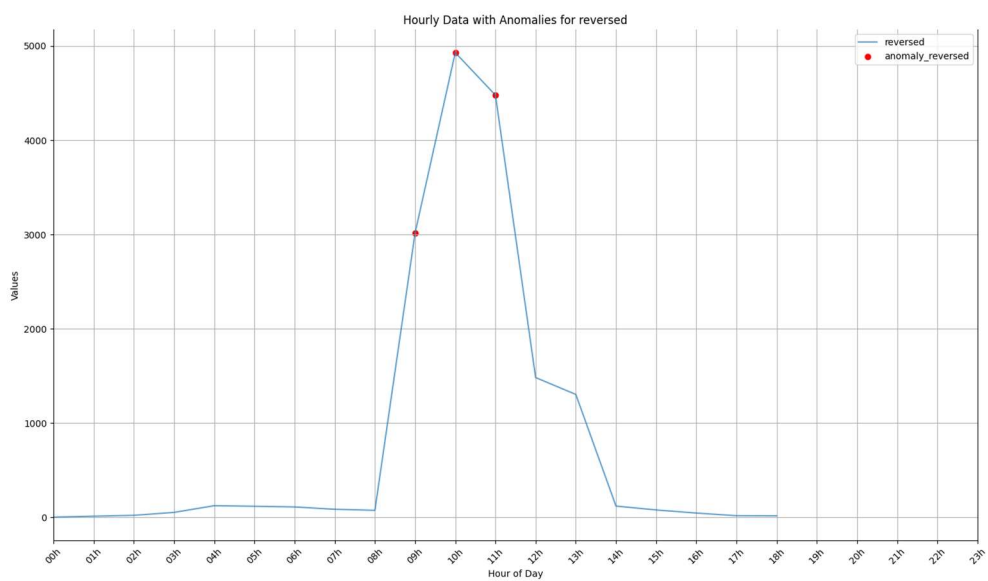
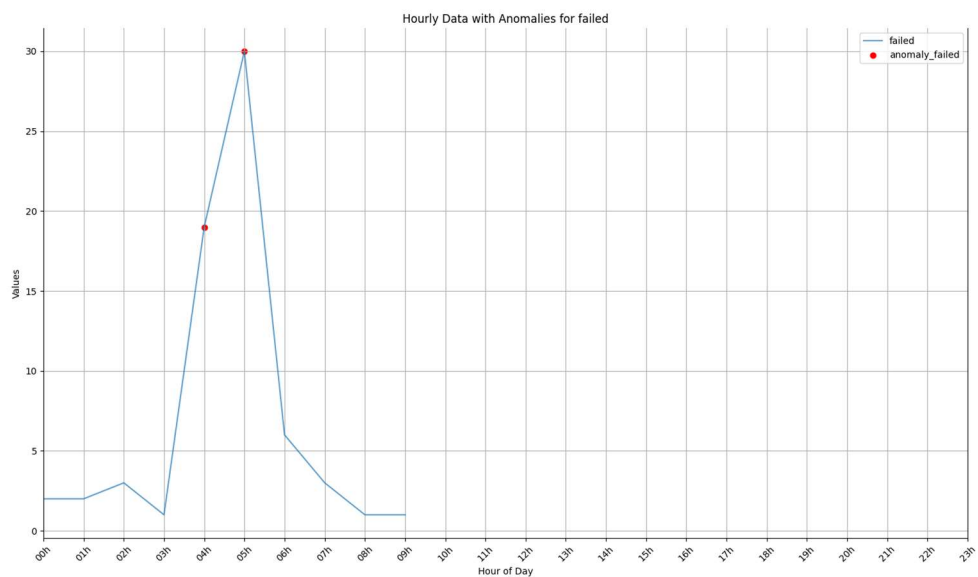
Ao processar os dados do arquivo "transactions 1", foram identificados 6 valores anômalos com status "denied", 2 com status "failed" e 3 com status "reversed". Na tabela, os valores anômalos foram destacados com cores específicas para facilitar a visualização: amarelo para "denied", verde para "failed" e azul para "reversed". Além disso, verificou-se que todos os valores anômalos apresentavam um z-score superior a 0,7.

Gráficos





Relatório Técnico





Relatório Técnico

Tabela

		Z- SCORE		Z- SCORE		Z- SCORE
	denied	denied	failed	failed	reversed	reversed
0h	83.00	-1.05	0.00	-0.40	3.00	-0.47
1h	35.00	-1.10	0.00	-0.40	0.00	-0.47
2h	11.00	-1.12	0.00	-0.40	0.00	-0.47
3h	17.00	-1.12	0.00	-0.40	0.00	-0.47
4h	5.00	-1.13	0.00	-0.40	0.00	-0.47
5h	42.00	-1.09	0.00	-0.40	0.00	-0.47
6h	192.00	-0.93	0.00	-0.40	13.00	-0.46
7h	529.00	-0.58	0.00	-0.40	22.00	-0.46
8h	1230.00	0.15	2.00	-0.12	53.00	-0.43
9h	1942.00	0.89	0.00	-0.40	124.00	-0.38
10h	2197.00	1.15	2.00	-0.12	118.00	-0.39
11h	2096.00	1.05	0.00	-0.40	111.00	-0.39
12h	1680.00	0.62	3.00	0.02	86.00	-0.41
13h	1701.00	0.64	1.00	-0.26	75.00	-0.42
14h	2381.00	1.35	19.00	2.30	3017.00	1.65
15h	3225.00	2.23	30.00	3.87	4929.00	2.99
16h	2582.00	1.56	6.00	0.45	4478.00	2.67
17h	712.00	-0.39	0.00	-0.40	1483.00	0.57
18h	1281.00	0.20	3.00	0.02	1304.00	0.44
19h	1347.00	0.27	0.00	-0.40	120.00	-0.39
20h	1110.00	0.02	1.00	-0.26	79.00	-0.42
21h	867.00	-0.23	0.00	-0.40	46.00	-0.44
22h	574.00	-0.54	0.00	-0.40	18.00	-0.46
23h	287.00	-0.83	1.00	-0.26	17.00	-0.46

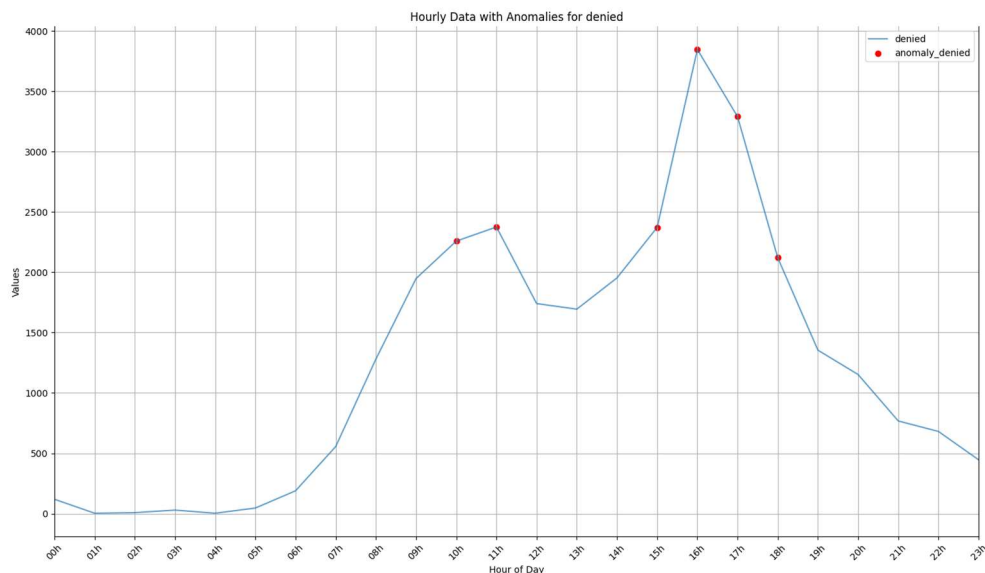


Relatório Técnico

Análise do transaction 2

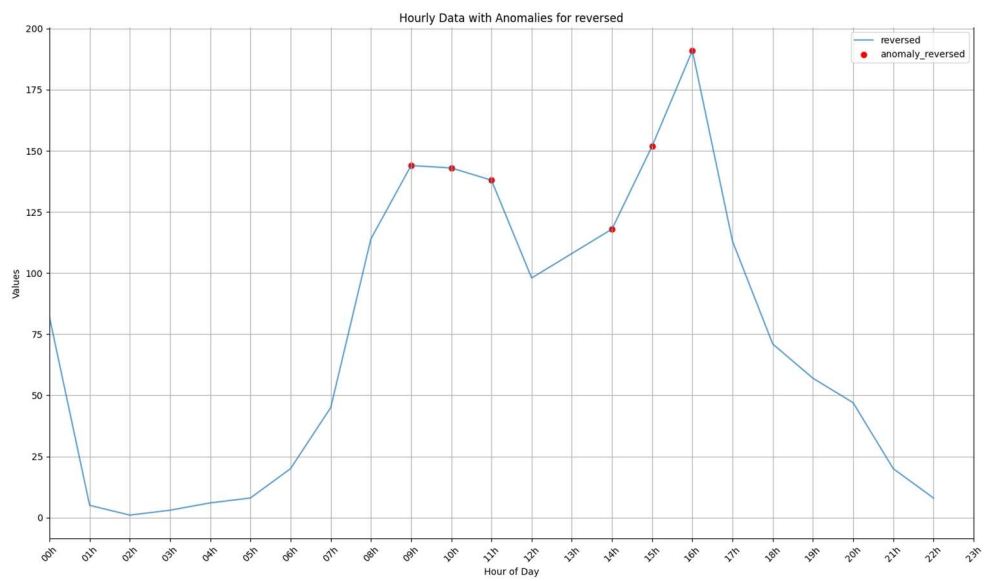
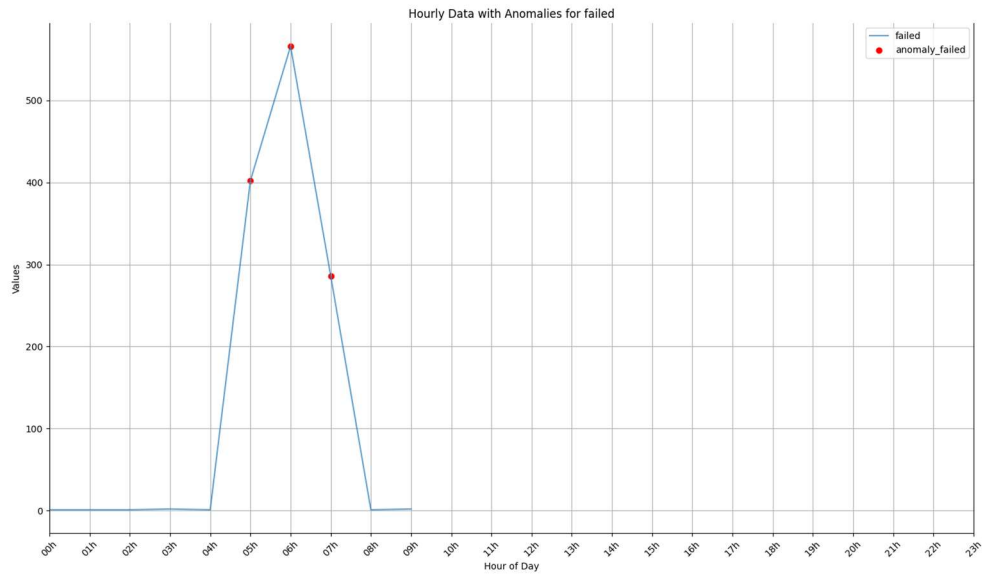
Ao processar os dados do arquivo "transactions 1", foram identificados 6 valores anômalos com status "denied", 3 com status "failed" e 6 com status "reversed". Na tabela, os valores anômalos foram destacados com cores específicas para facilitar a visualização: amarelo para "denied", verde para "failed" e azul para "reversed". Além disso, verificou-se que todos os valores anômalos apresentavam um z-score superior a 0,7.

Gráficos





Relatório Técnico



Tabela



Relatório Técnico

Tabela

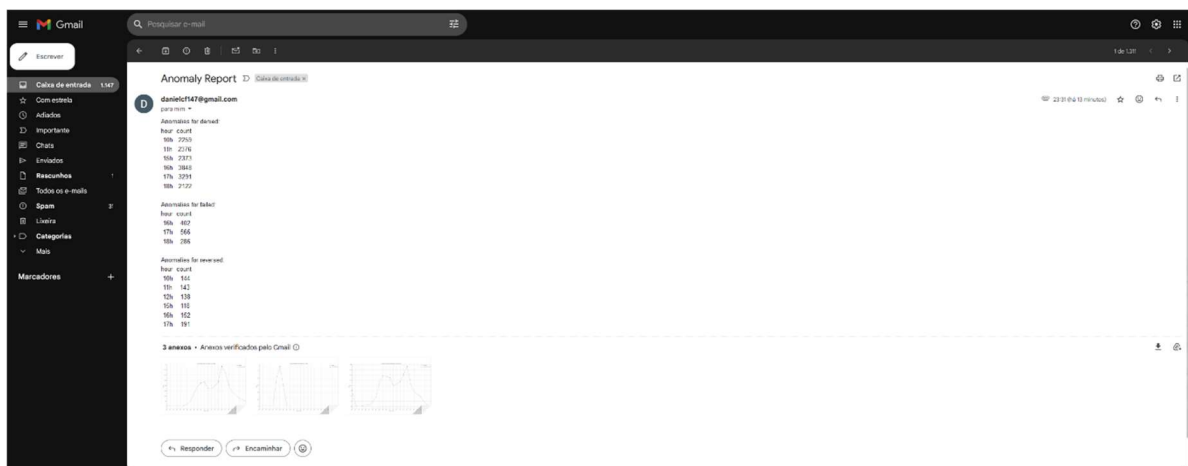
		Z- SCORE		Z- SCORE		Z- SCORE
	denied	denied	failed	failed	reversed	reversed
0h	119.00	-1.03	0.00	-0.36	82.00	0.19
1h	3.00	-1.13	0.00	-0.36	5.00	-1.10
2h	8.00	-1.13	0.00	-0.36	1.00	-1.17
3h	29.00	-1.11	0.00	-0.36	3.00	-1.13
4h	3.00	-1.13	0.00	-0.36	0.00	-1.18
5h	46.00	-1.09	0.00	-0.36	6.00	-1.08
6h	189.00	-0.96	0.00	-0.36	8.00	-1.05
7h	556.00	-0.63	0.00	-0.36	20.00	-0.85
8h	1281.00	0.02	1.00	-0.35	45.00	-0.43
9h	1949.00	0.62	0.00	-0.36	114.00	0.73
10h	2259.00	0.90	1.00	-0.35	144.00	1.23
11h	2376.00	1.01	1.00	-0.35	143.00	1.22
12h	1741.00	0.43	2.00	-0.34	138.00	1.13
13h	1695.00	0.39	1.00	-0.35	98.00	0.46
14h	1954.00	0.62	0.00	-0.36	108.00	0.63
15h	2373.00	1.00	0.00	-0.36	118.00	0.80
16h	3848.00	2.33	402.00	2.38	152.00	1.37
17h	3291.00	1.83	566.00	3.49	191.00	2.02
18h	2122.00	0.78	286.00	1.59	113.00	0.71
19h	1355.00	0.09	1.00	-0.35	71.00	0.01
20h	1153.00	-0.10	0.00	-0.36	57.00	-0.23
21h	768.00	-0.44	0.00	-0.36	47.00	-0.39
22h	681.00	-0.52	2.00	-0.34	20.00	-0.85
23h	446.00	-0.73	0.00	-0.36	8.00	-1.05



Relatório Técnico

Alert system

Foi desenvolvido um sistema de alerta para notificar a equipe responsável sobre possíveis anomalias utilizando os arquivos CSV fornecidos. Este sistema envia um e-mail sempre que um arquivo é processado, podendo conter até três imagens, cada uma correspondente a um status específico (denied, failed e reversed). O e-mail também inclui uma lista detalhada, organizada por status, que informa os horários identificados como anômalos.





Conclusão – Desafio 2 Solve the problem

Os dados analisados destacam os horários em que foram identificadas anomalias em ambos os arquivos. No entanto, é essencial que a equipe que recebeu o alerta realize uma investigação mais aprofundada, examinando todas as transações desses horários para identificar a origem do problema.

Além disso, foi desenvolvido um endpoint “/metrics” para integração com Prometheus e Grafana, com o objetivo de permitir o monitoramento e a atualização dos gráficos em tempo real. No entanto, surgiram problemas relacionados à parametrização dos dados exportados e à configuração do dashboard no Grafana, o que impediu a implementação completa dessas soluções.

4. CONCLUSÃO FINAL

Este relatório técnico apresenta as análises, desenvolvimentos e conclusões inerentes a cada um dos desafios propostos pela CloudWalk para o cargo de estagiário de analista de monitoramento, que envolveu a análise de anomalias em dados hipotéticos de checkout e o desenvolvimento de um sistema de monitoramento de transações financeiras para detecção de anomalias.

Para tanto, foram utilizadas técnicas de análise de dados e desenvolvimento de sistemas, onde foi possível identificar padrões suspeitos e gerar alertas em tempo real para a equipe responsável, proporcionando uma maior segurança nas operações financeiras.

Numa análise inicial, observa-se que caso o sistema desenvolvido fosse colocado em produção, resultaria na análise de milhões de transações diárias, resultando num alto custo de processamento. Uma solução potencial é implementar um sistema de classificação prévia de risco para cada POS, que fosse escalável e baseada em fatores como análise do perfil do cliente, localização do empreendimento (estado, município e bairro), tipo de negócio e histórico de transações.

Por exemplo, uma joalheria em um shopping de alto padrão em São Paulo teria uma classificação de risco diferente de um supermercado em um bairro periférico do Rio de Janeiro. A joalheria, com poucas transações e alto controle de segurança, teria um risco menor de fraudes, enquanto o supermercado, com muitas transações e um maior número de funcionários, teria um risco mais elevado.



Relatório Técnico

Com base nessa classificação, POS de baixo risco poderiam ter uma análise menos criteriosa, focada apenas em transações discrepantes, enquanto POS de alto risco necessitariam de uma investigação mais detalhada de todas as transações. Essa abordagem permitiria uma análise mais eficiente e escalável, otimizando os recursos técnicos e financeiros e mantendo a segurança das transações financeiras.