No LOP DS, recebemos via API dados de todas as submissões de questões dos alunos do sistema LOP, relativos a turma de um professor selecionado (Orivaldo, por exemplo).

OBS: Vale salientar que podem ser feitas quantas submissões por questão que o aluno ache necessário.

Para cada submissão de questão temos os seguintes dados para todos os alunos da turma:

- envinroment: ambiente no qual foi realizada a submissão
- hitPercentage: porcentagem de acerto
- language: linguagem de programação que foi utilizada para realizar a submissão
- char_change_number: número de caracteres escritos em relação a última submissão da questão
- timeConsuming: o tempo, transformado em segundos, que o usuário levou para enviar a respectiva submissão
- createdAt: data e hora em que a submissão foi feita
- listQuestions_id: id da lista de questões
- test_id: id do test
- question_id: id da questão
- class_id: id da classe
- user: um dicionário contendo o email, id e nome do usuário sendo email, id e name as chaves para cada item citado anteriormente respectivamente
- question: um dicionário contendo o título da questão como chave para o título da questão title
- list: um dicionário contendo o título da lista de exercícios da submissão que tem como chave para ela title
- test: test

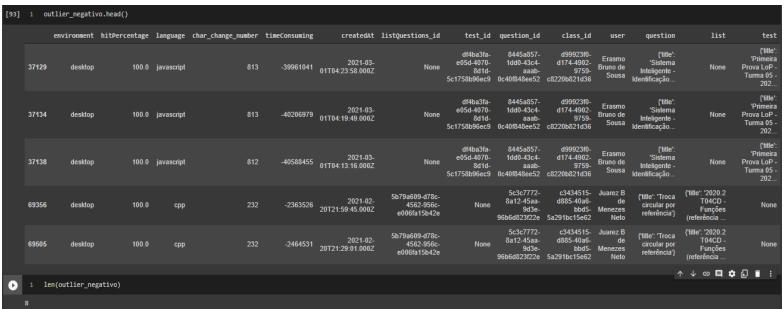
Exemplo:

0	1 df_submissions.head()														
C•	e	nvironment	hitPercentage	language	char_change_number	timeConsuming	createdAt	listQuestions_id	test_id	question_id	class_id	user	question	list	test
		desktop	0.0	javascript	386	876987	2021-03- 11T19:52:57.000Z	f5b1d992-68de- 4854-97d8- c4bd84bd593b	None	fa1b966e-8c0b- 4497-931c- 19becdee92eb	6d04019e-905c- 4d70-810a- a7f1fa34fe22	{'id': 'df34f122-bad7- 4d25-9a9f- a80ca0637ec0',	{title': 'Cálculo do novo Salário'}	{'title': 'Estruturas de Decisão - Lista de Ex	None
		desktop	100.0	javascript	186	1398058	2021-03- 11T19:52:54.000Z	0a2c49a2-482c- 40d7-bc44- ee368f089560	None	24e8bcf3-c8c2- 4da7-9509- b0a32ee92d26	53896190-8aac- 4608-a07c- c2cc814fee28	('id': '739b884a- c2da-4643-a1be- 31232b6323cb',	{'title': 'Contar quantos são maiores de 18 an	{'title': 'Repetição condicional - Lista Práti	None
		desktop	0.0	javascript	371	852026	2021-03- 11T19:52:30.000Z	f5b1d992-68de- 4854-97d8- c4bd84bd593b	None	fa1b966e-8c0b- 4497-931c- 19becdee92eb	6d04019e-905c- 4d70-810a- a7f1fa34fe22	('id': 'df34f122-bad7- 4d25-9a9f- a80ca0637ec0',	{title': 'Cálculo do novo Salário'}	{'title': 'Estruturas de Decisão - Lista de Ex	None
		desktop	NaN	javascript	98	426107	2021-03- 11T19:52:05.000Z	0a2c49a2-482c- 40d7-bc44- ee368f089560	None	bab3b4e9-7b17- 4d45-aad4- 95ba62e47256	d99923f0-d174- 4902-9759- c8220b821d36	{'id': 'd718d567- 7e96-407c-9cb0- 3e52c9295adf,	{'title': 'Múltiplo de 3 ou 7'}	{'title': 'Repetição condicional - Lista Práti	None
	4	desktop	100.0	javascript	65	146239	2021-03- 11T19:51:44.000Z	dbfcc83c-f014- 44ae-8260- f06d4e1370ba	None	89716bf6-3a20- 4502-b861- 77a07de648c3	6d04019e-905c- 4d70-810a- a7f1fa34fe22	{id': '85803046- 5fe9-4398-9715- 11861f90fee0',	{'title': 'O quadrado ou raiz'}	{'title': 'Estruturas de Decisão - Lista Resol	None

Desse modo tínhamos o objetivo de verificar, para cada aluno, se eles possuíam valores de tempo (timeConsuming) muito altos, por exemplo uma

semana, e valores negativos, o que nos retornou as tabelas de todos os alunos com todas as suas submissões dentro das condições citadas acima.

A imagem abaixo mostra que nós temos 8 registros que possuem tempo de submissão negativo e mostra 5 desses registros:



A imagem abaixo mostra que nós temos 48.239 registros que possuem tempo de submissão acima de uma semana e mostra 5 desses registros:



Analisando mais a fundo, pegamos um aluno em especifico do grupo dos alunos que possuem registros com tempo negativo, e subtraindo o valor absoluto do tempo de submissão da linha de id 2 pela valor absoluto do tempo de submissão da linha de id 1 e pegando como resultado final o valor absoluto dessa subtração nós temos um valor que não é igual ao valor da diferença entre o valor da linha de id 1 e coluna createdAt e o valor da linha de id 2 e coluna createdAt – o valor da diferença é de 373 segundos.

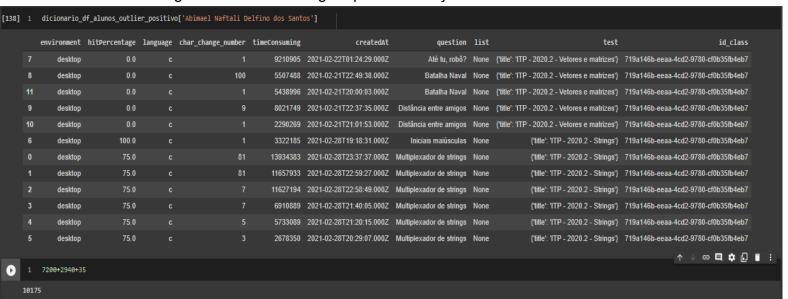
Segue abaixo uma imagem para orientação:

[103]	103] 1 dicionario_df_alunos_outlier_negativo['Erasmo Bruno de Sousa']											
		environment	hitPercentage	language	char_change_number	timeConsuming	createdAt	question	list	test	id_class	
	2	desktop	100.0	javascript	812	-40588455	2021-03- 01T04:13:16.000Z	{'title': 'Sistema Inteligente - Identificação	None	{'title': 'Primeira Prova LoP - Turma 05 - 202	d99923f0-d174-4902- 9759-c8220b821d36	
	1	desktop	100.0	javascript	813	-40206979	2021-03- 01T04:19:49.000Z	{'title': 'Sistema Inteligente - Identificação	None	{'title': 'Primeira Prova LoP - Turma 05 - 202	d99923f0-d174-4902- 9759-c8220b821d36	
	0	desktop	100.0	javascript	813	-39961041	2021-03- 01T04:23:58.000Z	{'title': 'Sistema Inteligente - Identificação	None	{'title': 'Primeira Prova LoP - Turma 05 - 202	d99923f0-d174-4902- 9759-c8220b821d36	

Analisando também os dados que apresentam os alunos que possuíam registros com valores de tempo de submissão muito altos (1 semana nesse caso) comprovamos que ele está com valores muito altos, diferente do que realmente são.

Por exemplo, se pegarmos o aluno Abimael Naftali Delfino dos Santos podemos ver que a diferença de tempo de submissão na questão Batalha Naval é igual á 68.492 segundos quando a diferença no horário da submissão é na verdade 10.175 segundos.

Segue abaixo uma imagem para orientação:



Estamos comunicando esses pontos não apenas porque outliers foram encontrados (os quais poderiam ser facilmente filtrados), mas sabendo que pelo fato de que os dados apresentados mostraram certas inconsistências, muito provavelmente também ocorrem essas inconsistências em outras turmas, prejudicando a credibilidade de determinadas análises e informações levantadas.

Acreditamos também que seria aceitável, no máximo, uma ocorrência de 18000 segundos por usuário, tendo em vista que o aluno pode ter parado num dia e continuado em outro, mas havendo duas ou mais ocorrências na mesma questão começa a ficar estranho.