

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:

- environment: 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:

| | environment | hitPercentage | language | char_change_number | timeConsuming | createdAt | listQuestions_id | test_id | question_id | class_id | user | question | list | test |
|---|-------------|---------------|------------|--------------------|---------------|--------------------------|--------------------------------------|---------|--------------------------------------|--------------------------------------|---|---|---|------|
| 0 | 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 |
| 1 | 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 |
| 2 | 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 |
| 3 | 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-f064e1370ba | 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:

[93] 1 outlier_negativo.head()

| | environment | hitPercentage | language | char_change_number | timeConsuming | createdAt | listQuestions_id | test_id | question_id | class_id | user | question | list | test |
|-------|-------------|---------------|------------|--------------------|---------------|--------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------|---|---|---|
| 37129 | desktop | 100.0 | javascript | 813 | -39961041 | 2021-03-01T04:23:58.000Z | None | df4ba3fa-e05d-4070-8d1d-5c1758b96ec9 | 8445a857-1dd0-43c4-aaab-0c40f848ee52 | d99923f0-d174-4902-9759-c8220b821d36 | Erasmus Bruno de Sousa | {'title': 'Sistema Inteligente - Identificação... | None | {'title': 'Primeira Prova LoP - Turma 05 - 202... |
| 37134 | desktop | 100.0 | javascript | 813 | -40206979 | 2021-03-01T04:19:49.000Z | None | df4ba3fa-e05d-4070-8d1d-5c1758b96ec9 | 8445a857-1dd0-43c4-aaab-0c40f848ee52 | d99923f0-d174-4902-9759-c8220b821d36 | Erasmus Bruno de Sousa | {'title': 'Sistema Inteligente - Identificação... | None | {'title': 'Primeira Prova LoP - Turma 05 - 202... |
| 37138 | desktop | 100.0 | javascript | 812 | -40588455 | 2021-03-01T04:13:16.000Z | None | df4ba3fa-e05d-4070-8d1d-5c1758b96ec9 | 8445a857-1dd0-43c4-aaab-0c40f848ee52 | d99923f0-d174-4902-9759-c8220b821d36 | Erasmus Bruno de Sousa | {'title': 'Sistema Inteligente - Identificação... | None | {'title': 'Primeira Prova LoP - Turma 05 - 202... |
| 69356 | desktop | 100.0 | cpp | 232 | -2363526 | 2021-02-20T21:59:45.000Z | 5b79a609-d78c-4562-956c-e006fa15b42e | None | 5c3c7772-8a12-45aa-9d3e-96b6d823f22e | c3434515-d885-40a6-bbd5-5a291bc15e62 | Juarez B de Menezes Neto | {'title': 'Troca circular por referência'} | {'title': '2020.2 T04CD - Funções (referência ... | None |
| 69505 | desktop | 100.0 | cpp | 232 | -2464531 | 2021-02-20T21:29:01.000Z | 5b79a609-d78c-4562-956c-e006fa15b42e | None | 5c3c7772-8a12-45aa-9d3e-96b6d823f22e | c3434515-d885-40a6-bbd5-5a291bc15e62 | Juarez B de Menezes Neto | {'title': 'Troca circular por referência'} | {'title': '2020.2 T04CD - Funções (referência ... | None |

1 len(outlier_negativo)
8

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

1 outlier_positivo.head()

| | environment | hitPercentage | language | char_change_number | timeConsuming | createdAt | listQuestions_id | test_id | question_id | class_id | user | question | list | test |
|---|-------------|---------------|------------|--------------------|---------------|--------------------------|--------------------------------------|---------|--------------------------------------|--------------------------------------|----------------------------|---|---|------|
| 0 | desktop | 75.0 | c | 313 | 1310024 | 2021-03-11T17:46:28.000Z | 6893c2d2-49b5-4476-8d0d-cf393130aacc | None | e23cd857-2be5-4ffa-a245-6f36441f3099 | 52e39ed5-e297-4fa5-953b-763d119b7af6 | Ana Caroline Dantas | {'title': 'Numeros Fracionários'} | {'title': 'TTP - Structs e Enumerações - Lista'} | None |
| 1 | desktop | 80.0 | javascript | 299 | 2806722 | 2021-03-11T17:46:20.000Z | aca84206-87b2-4478-afa4-6193119cc06a | None | d6ce9371-80bb-4506-9a7e-84ab32cd7709 | 53896190-8aac-4608-a07c-c2cc814fee28 | José Davi | {'title': 'Tempo de vida de um material radioa... | {'title': 'Repetição condicional - Lista de Ex... | None |
| 4 | desktop | 0.0 | python | 874 | 3852003 | 2021-03-11T17:44:39.000Z | 887ac98d-84c8-46c1-a9e3-1598e25621ba | None | 4a66d56f-d8d0-4670-a9be-62ca90756de9 | 149b9597-cca2-4274-9bc8-754bed4068de | Vinicius Lamas von Sohsten | {'title': 'Python: Gauss-Seidel sistemas'} | {'title': 'CN T3T4 lista 3 (sistemas)'} | None |
| 5 | desktop | 0.0 | python | 873 | 3842741 | 2021-03-11T17:44:27.000Z | 887ac98d-84c8-46c1-a9e3-1598e25621ba | None | 4a66d56f-d8d0-4670-a9be-62ca90756de9 | 149b9597-cca2-4274-9bc8-754bed4068de | Vinicius Lamas von Sohsten | {'title': 'Python: Gauss-Seidel sistemas'} | {'title': 'CN T3T4 lista 3 (sistemas)'} | None |
| 6 | desktop | 100.0 | cpp | 110 | 32203036 | 2021-03-11T17:44:02.000Z | 498acd7b-f6b5-4727-b5c0-2b9800e7ecc1 | None | b016fcbd-5043-4ac0-9cf8-c65feab76557 | 74c79e20-9e33-4fb9-8876-5aaa9173b123 | Felipe Melo | {'title': 'ES2020 Est - Jogadores'} | {'title': 'Lista IP C 03'} | None |

1 len(outlier_positivo)
48239

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] 1  dicionario_df_alunos_outlier_negativo['Erasmus 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:

```
[138] 1  dicionario_df_alunos_outlier_positivo['Abimael Naftali Delfino dos Santos']
```

| | environment | hitPercentage | language | char_change_number | timeConsuming | createdAt | question | list | test | id_class |
|----|-------------|---------------|----------|--------------------|---------------|--------------------------|--------------------------|------|--|--------------------------------------|
| 7 | desktop | 0.0 | c | 1 | 9210905 | 2021-02-22T01:24:29.000Z | Alé tu, robô? | None | {'title': 'TTP - 2020.2 - Vetores e matrizes'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 8 | desktop | 0.0 | c | 100 | 5507488 | 2021-02-21T22:49:38.000Z | Batalha Naval | None | {'title': 'TTP - 2020.2 - Vetores e matrizes'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 11 | desktop | 0.0 | c | 1 | 5438996 | 2021-02-21T20:00:03.000Z | Batalha Naval | None | {'title': 'TTP - 2020.2 - Vetores e matrizes'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 9 | desktop | 0.0 | c | 9 | 8021749 | 2021-02-21T22:37:35.000Z | Distância entre amigos | None | {'title': 'TTP - 2020.2 - Vetores e matrizes'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 10 | desktop | 0.0 | c | 1 | 2290269 | 2021-02-21T21:01:53.000Z | Distância entre amigos | None | {'title': 'TTP - 2020.2 - Vetores e matrizes'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 6 | desktop | 100.0 | c | 1 | 3322185 | 2021-02-28T19:18:31.000Z | Iniciais maiúsculas | None | {'title': 'TTP - 2020.2 - Strings'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 0 | desktop | 75.0 | c | 81 | 13934383 | 2021-02-28T23:37:37.000Z | Multiplexador de strings | None | {'title': 'TTP - 2020.2 - Strings'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 1 | desktop | 75.0 | c | 81 | 11657933 | 2021-02-28T22:59:27.000Z | Multiplexador de strings | None | {'title': 'TTP - 2020.2 - Strings'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 2 | desktop | 75.0 | c | 7 | 11627194 | 2021-02-28T22:58:49.000Z | Multiplexador de strings | None | {'title': 'TTP - 2020.2 - Strings'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 3 | desktop | 75.0 | c | 7 | 6910889 | 2021-02-28T21:40:05.000Z | Multiplexador de strings | None | {'title': 'TTP - 2020.2 - Strings'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 4 | desktop | 75.0 | c | 5 | 5733089 | 2021-02-28T21:20:15.000Z | Multiplexador de strings | None | {'title': 'TTP - 2020.2 - Strings'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |
| 5 | desktop | 75.0 | c | 3 | 2678350 | 2021-02-28T20:29:07.000Z | Multiplexador de strings | None | {'title': 'TTP - 2020.2 - Strings'} | 719a146b-eeaa-4cd2-9780-cf0b35fb4eb7 |

1 7200+2940+35

10175

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.