

```
#-----  
-----  
# Centro Universitário Unicarioca  
# Pós-Graduação em Ciência de Dados - 2024  
# Disciplina: P241A - VISUALIZAÇÃO DE DADOS 20240 - Prof. Dr. Leonardo Bezerra  
# Aluno: Paulo Cesar dos Santos  
# Case Study 01 - Instituto Imbuhy  
#-----  
-----  
# 1. Prezados, tomo a liberdade de apresentar um caso que ocorreu recentemente e  
mostra  
# como a tarefa de extração e envio de dados pode ser "marota", se não tomarmos  
cuidado.  
# Aqui, o problema é que os dados foram testados apenas com os primeiros  
registros,  
# quando deveriam abranger um "range" mais aleatório. Talvez aí o problema  
aparecesse.  
# A falta de uniformidade de preparação também pode ter ocasionado o problema.  
# Lembrando que nunca podemos criticar o trabalho de alguém se não sabemos em  
que  
# condições esse trabalho se deu.  
#  
# Agora, leia os comentários de baixo para cima, desde o fim para cá.  
#-----  
-----
```

```
#-----  
-----  
# 8. Cavalheiros atuam assim !!!  
#-----  
-----
```

Tranquilo, Fernanda.
Muitas vezes nos defrontamos com problemas assim por aqui, mas é difícil de
achar, dado o volume de registros.
Não haverá problema em repassá-los. Estamos aqui para isso mesmo ... :-D

Cordial abraço.

Peter

```
#-----  
-----  
# 7. Pegou mal para a Fernanda, que descobriu seu erro por terceiros e teve de  
dar uma  
# explicação que, sinceramente, era desnecessária.  
# Bem, recuperação de danos ! Ele poderia ter agradecido a minha dica :-D  
# Espero que esteja bem.
```

#-----

De: Fernanda Abreu [mailto:Fernanda.Abreu@CompanhiaB.gov.br]
Enviada em: quarta-feira, 28 de fevereiro de 2024 17:23
Para: Peter <Peter@imbuhy.com>
Cc: suporte CompanhiaB <suporte@CompanhiaB.gov.br>; sac@imbuhy.com
Assunto: Re: Problemas na carga de atualizações nos arquivos do RJ, mês de Janeiro.

Peter,

Para retirar os caracteres que estavam impedindo o carregamento, passei uma macro e acho que isso causou a transformação dos números. Estou reenviando um novo arquivo "entrada.cvs", nesse a coluna de número do processo está preservada.

Se der certo, me avise, terei que refazer as extrações de 02.2024. Provavelmente estará com o mesmo problema.

Dr. Agamenon,

Gerei mais um relatório e salvei no DRIVE2, poderia, por gentileza, compartilhar com o colega.

Att.,

Fernada Abreu

Dept de Contas/CompanhiaB
(61) 22222-33333
(61) 99999-8888

#-----

6. Aqui, eu dou a dica à operadora: o número da linha onde os erros começam.
Certamente ela ficou mal com o chefe. Melhro ajudá-la, antes que perca a função !
Meu chefe foi claro: Não faça o serviço dos outros - Apenas informe o problema !
#-----

Em qua., 28 de fev. de 2024 às 17:01, Peter <Peter@imbuhy.com> escreveu:
Olá Fernanda,

Se ajudar, verifiquei que alguns valores numéricos estão em notação científica. Veja, por exemplo, a linha 254713 do arquivo mencionado anteriormente. Enquanto isso, vou limpar essa última carga de nosso sistema, passar para o outro arquivo e ver o que acontece.

Cordial abraço,

Peter

```
#-----  
-----  
# 5. A operadora da CompanhiaB tinha total confiança em seus dados. Ela testou  
os "N"  
# primeiros registros e tudo estava OK. Mas "N" não foi grande o suficiente. O  
problema  
# aconteceu muito tarde, com N > 200 mil ! Ela foi peremptória: Não há  
problema.  
# Uma abordagem mais conveniente seria pesquisar partes da mensagem de erro,  
como eu  
# mesmo fiz.  
#-----  
-----
```

De: Fernanda Abreu [mailto:Fernanda.Abreu@CompanhiaB.gov.br]
Enviada em: quarta-feira, 28 de fevereiro de 2024 14:04
Para: Peter <Peter@imbuhy.com>
Cc: suporte CompanhiaB <suporte@CompanhiaB.gov.br>; sac@imbuhy.com
Assunto: Re: Problemas na carga de atualizações nos arquivos do RJ, mês de Janeiro.

Prezado Peter, boa tarde!

O arquivo [janeiro.csv] possui 265.350 inscrições, nenhuma repetida. Já as colunas de CPF/CNPJ do Devedor e de Processo possuem registros repetidos visto que um devedor pode ser responsável por mais de um processo e por mais de uma inscrição.

Porém isso sempre ocorreu nos arquivos enviados anteriormente.

Seria possível identificar especificamente qual registro está gerando o problema?

Transformei o arquivo .csv em excel para as contagens.

Att.,

Fernada Abreu
DICAD
Dept de Contas/CompanhiaB
(61) 2222-33333
(61) 99999-8888

```
#-----  
-----  
# 4. Um novo problema ocorreu em nossa aplicação, que não "engoliu" os dados.  
#-----  
-----
```

Em qua., 28 de fev. de 2024 às 13:37, Peter <Peter@imbuhy.com> escreveu:

Dr. Agamenon boa tarde.

Tivemos o seguinte problema ao tentar processar o arquivo [entrada.csv]:

Duplicate entry '138924758724872937238723487349-9' for key 'PRIMARY' [294350745]

Parece que há registro duplicado. Isso pode estar ocorrendo também nos outros arquivos ?

Cordial abraço.

Peter

```
#-----  
-----  
# 3. A operadora da CompanhiaB responde, enviando as correções.  
# Isso provavelmente resolveria os problemas ocorridos na hora de decodificar os  
# registros e inseri-los no banco MySQL. Note que a nossa aplicação devia ser  
# mais  
# rigorosa e explicativa na crítica de entrada.  
#-----  
-----
```

De: suporte CompanhiaB [mailto:suporte@CompanhiaB.gov.br]
Enviada em: sexta-feira, 16 de fevereiro de 2024 10:55
Para: Peter <Peter@imbuhy.com>; Fernanda Abreu <Fernanda.Abreu@CompanhiaB.gov.br>
Assunto: Re: Problemas na carga de atualizações nos arquivos do RJ, mês de Janeiro.

Prezado Peter, bom dia!

Fizemos um ajuste na extração retirando os caracteres especiais para que o erro, em tese, não volte a acontecer.

Qualquer problema, por gentileza, nos avise.

Segue o link: <https://drive.google.com/drive/folders/drive2>

Fernanda, muito obrigado!

Forte abraço!

Em ter., 6 de fev. de 2024 às 13:50, Peter <Peter@imbuhy.com> escreveu:

#-----

2. Aqui, após tentar carregar sem sucesso o arquivo enviado pela CompanhiaB,
mandei um recado para a chefia, que encaminhou o problema ao fornecedor dos
dapds da fornecedora.

#-----

De: Peter

Enviada em: quarta-feira, 31 de janeiro de 2024 19:10

Para: Chefia (chefia@imbuhy.com) <chefia@imbuhy.com>

Cc: sac@imbuhy.com

Assunto: Problemas na carga de atualizações nos arquivos do RJ, mês de Janeiro.

Prezado Supervisor,

Alguns arquivos de atualização de situação do Rio, para o mês de janeiro/2024 vieram com caracteres estranhos (ex.: "@" e plique simples).

Os arquivos foram obtidos no site do Google Drive deste mês.

Não foi possível processar a carga destes arquivos.

Em anexo, segue o relatório com os nomes dos arquivos e linhas com problemas.

Sugiro, s.m.j, que a CompanhiaB produza novos arquivos, do Rio efetuando a remoção dos caracteres mencionados.

Peter