


From: Michael Booth mjboothaus@gmail.com 
Subject: Fwd: Excel data template
Date: 7 September 2022 at 1:28 pm
To:



Begin forwarded message:

From: Joe <thedatafreelancer@gmail.com>
Subject: RE: Excel data template
Date: 16 July 2017 at 4:31:24 am AEST
To: Michael Booth <mjboothaus@gmail.com>

Hi Michael,

I have used a different scraper and got the character issue sorted I guess. Have a look at the data and let me know.

Joe Abraham
Data Freelancer
0401 368 890

From: Joe
Sent: Sunday, 16 July 2017 12:37 AM
To: Michael Booth
Subject: RE: Excel data template

Hi Michael,

Looks like my scraper couldn't recognise those character. I have no idea why my scraper would have trouble recognising characters. I could try another scraper to check if it really critical for you.

Let me know

Joe Abraham
Data Freelancer
0401 368 890

From: Michael Booth
Sent: Saturday, 15 July 2017 2:40 PM
To: Joe
Subject: Re: Excel data template

Hi Joe

I've only had a chance to take a quick look - however it is looking very promising!

1. I can JOIN the tables successfully using the URL which was my primary concern/objective. Thanks!
2. The only small frustration that is immediately evident is with differences in the character sets in some (non-English) passenger names.

For example, for the first passenger in the "table list" his name is:

ABĪ SA'B, Mr Jirjis Yūsuf (which uses accented non-standard characters)

however in the "Detail list" it shows up as:

Mr Jirjis YÅ«suf AbÄ« Sa'b (Gerios Yousseff)

with the characters in a different character set. However - on the webpages they appear the same (as per the table list).

I don't want you to do any additional scraping work, however, do you have any idea what character set they were in so I might convert them on a consistent basis?

3. I will do some additional consistency spot checks (by hand) on the data to verify the integrity of the automated scraping process -- obviously, being a manual process this takes some time.

Cheers,
Michael

On Sat, Jul 15, 2017 at 12:44 PM, Joe <thedatafreelancer@gmail.com> wrote:

Hi Michael,

Please ignore the file in the last email. Messed up the file as I copied it in the wrong areas. Sleep got the better off me.

Let me know if you have comments.

Joe Abraham
Data Freelancer
0401 368 890

From: [Joe](#)
Sent: Saturday, 15 July 2017 1:27 AM
To: [Michael Booth](#)
Subject: RE: Excel data template

Michael,

As discussed, see the revised scraped data. I had to scrape the page twice to get the URL from the list table.

Let me know if you have comments. I will request payment only after you confirmation.

Thanks

Joe Abraham
Data Freelancer
0401 368 890

From: [Michael Booth](#)
Sent: Friday, 14 July 2017 10:28 PM
To: [Joe](#)
Subject: Re: Excel data template

Hi Joe

Yes -- it is looking good! -- but I didn't want to look too closely at the data until:

1. I have all records and
2. I can join the tables together to cross-check.

I appreciate the work you have done so far.

Cheers,
Michael

On Fri, Jul 14, 2017 at 9:37 PM, Joe <thedatafreelancer@gmail.com> wrote:

Hi Michael,

Will send all together in the next batch. Are you happy with the detail page scrape so far?

Joe Abraham
Data Freelancer
0401 368 890

From: [Michael Booth](#)
Sent: Friday, 14 July 2017 9:26 PM
To: [Joe](#)
Subject: Re: Excel data template

Hi again Joe

I do need to be able to join the two tables (list & details), so I do need the hyperlink also in the list table to do the join -- i.e. this should be available as part of scrapping the list page.

Otherwise it is a bit tricky / time-consuming.

Cheers,
Michael

On Fri, Jul 14, 2017 at 9:16 PM, Michael Booth <mjboothaus@gmail.com> wrote:

Hi Joe

Thanks for the update!

Don't worry about the rearrangement of the data -- I'm good at that - it was the scraping piece that I don't have much experience with (and the expertise I was after).

Let me take a look at the data for quality / completeness and if you could get the remaining ~300+ records that would be fantastic.

Cheers,
Michael

On Fri, Jul 14, 2017 at 9:03 PM, Joe <thedatafreelancer@gmail.com> wrote:

Hi Micahel,

I just checked my data. My scraper ended after scraping about 1000 detail pages. I will reconfigure my scraper to run overnight to give you the results tomorrow.

I had a quick look at the template and realised the data that I scrape and the template excel that you have doesn't match. It would require me to process the data scraped. I lowered my offer assuming I could just give you the data scraped without requiring a scrape. Would you mind if I increase my offer back to \$60 to account for that.

If you agree, I can send a price increase request which becomes effective once you approve. Meanwhile see attached the excel file with the raw data scraped so far in a different tabs.

Note that the data attached has only 1000 records for detail page as my scraper timed out due to some reason. I am reconfiguring and will run the scraper again tonight.

Let me know if you have any questions/comments.

Regards

Joe Abraham
Data Freelancer
0401 368 890

From: [Michael Booth](#)
Sent: Friday, 14 July 2017 10:53 AM
To: thedatafreelancer@gmail.com
Subject: Excel data template

Hi Joe

As attached.

Th...

THANKS,
Michael



Virus-free. www.avast.com



titanica new
detail.xlsx