

generate_dummy_api_search_results

August 16, 2021

0.1 DOCUMENTATION

AUTHOR: Rebecca Wright

DATE: 6/20/21

PURPOSE

To create a dummy dataset of Google Jobs API search results for use creating dummy trackerlog dataset in TRKR App product development. Program is called upon to run a API query call up to 100 times using a FOR loop. Each iteration of the FOR loop offsets the query results pagination by 10 (as 10 is the default number of results returned by each query call). As a system wait command was not used, the results of each query call are written out as an append to a csv file, to guard against loss of query results due to timing errors.

After the Google Jobs API query calls are finished, the CSV results are imported as a new sample dataframe. A new FOR loop API query call is made to the Google Jobs Listing API using each job result id number. The first apply_option link value from each call is stored in a temporary app_url list. The app_url_list is appended to the sample dataframe as a new column. The sample dataframe then contains job listings joined together with a valid application_url value. The dataframe is saved as CSV file 'job_listing_result_log_with_links.csv' to be imported by generate_dummy_trackerlog.ipynb.

DEPENDANCIES

A valid secret_api_key must be obtained through <https://serpapi.com/>.

DEVELOPMENT NOTES

This program does not protect against timing errors within the API query FOR loop. A call to system sleep would be more efficient, as would reconsidering the explicit writing of each search result to the CSV. Current program iteration documented 34 successful API query calls, resulting in 340 job listing results, which were sufficient for initial development needs.

*** ##### IMPORTS

```
[1]: from serpapi import GoogleSearch
import pandas as pd
from csv import writer
```

*** ##### VARIABLES

0.1.1 Api Key

```
[2]: api_key = api_config_file.api_key
```

```
*** ##### MAIN
```

Execute Google Jobs API query call

```
[55]: for x in range(0,100):
        index = x
        start_var = x * 10
        params = {
            "engine": "google_jobs",
            "q": "UX designer new york",
            "hl": "en",
            "start": start_var,
            "api_key": api_key
        }

        search = GoogleSearch(params)
        results = search.get_dict()
        jobs_results = results['jobs_results']
        df = pd.DataFrame(jobs_results)
        df.drop(columns='thumbnail', inplace=True)

        if index == 0:
            df.to_csv('job_listing_result_log.csv', index=False)
        else:
            with open('job_listing_result_log.csv', 'a') as f:
                df.to_csv(f, header=False, index=False)

        print(x)  # used to monitor FOR loop progress
```

```
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
```

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33

```
-----  
KeyError                                Traceback (most recent call last)  
<ipython-input-55-8edeb36c7793> in <module>  
    12     search = GoogleSearch(params)  
    13     results = search.get_dict()  
----> 14     jobs_results = results['jobs_results']  
    15     df = pd.DataFrame(jobs_results)  
    16     df.drop(columns='thumbnail', inplace=True)  
  
KeyError: 'jobs_results'
```

Import Google Jobs API query results CSV

```
[58]: sample_df = pd.read_csv('job_listing_result_log.csv')  
sample_df
```

```
[58]:
```

	title \
0	UX Designer
1	UX Designer
2	Junior UX Designer
3	Sr. UX Designer
4	UX Designer
..	...
335	UX Designer
336	UI /UX Designer
337	User Experience (UX) Designer
338	UX Designer with UI, Full Time
339	Director of User Experience

	company_name	location \
0	Nielsen	New York, NY
1	Amazon	New York, NY
2	Brilliant Experience	Anywhere
3	Microsoft	New York, NY
4	Tory Burch	New York, NY
..
335	Tablet Hotels - The hotel experts at MICHELIN	New York, NY
336	Inabia Software & Consulting Inc.	New York, NY
337	Robert Half The Creative Group	New York, NY
338	Creative Circle	New York, NY (+1 other)
339	Aquent	New York, NY

	via	description \
0	via Nielsen Careers	UX Designer - 85001\n\nTechnology and Engineer...
1	via LinkedIn	Description\n\nWe are looking for a UX Designe...
2	via Glassdoor	Brilliant Experience's team works directly wit...
3	via Microsoft Careers	Senior UX Designer, PromoteIQ\n\nMicrosoft Adv...
4	via The Muse	JOB SUMMARY\n\nThe Designer, will assist in cr...
..
335	via Lensa	Tablet is all about discovering the world's mo...
336	via Nexxt	The ideal candidate Is detail-oriented with st...
337	via Lensa	Description Do you consider yourself a creativ...
338	via Nexxt	Position UX Designer with UI, Full Time Locati...
339	via Lensa	Our NYC based non-profit higher education clie...

	extensions \
0	['15 days ago', 'Full-time']
1	['10 days ago', 'Full-time', 'No degree mentio...
2	['4 days ago', 'Work from home', 'Full-time']
3	['Over 1 month ago', 'Full-time', 'No degree m...
4	['Over 1 month ago', 'Full-time']
..	...
335	['6 days ago', 'Full-time', 'No degree mention...
336	['6 days ago', 'Full-time']
337	['7 days ago', 'Full-time']
338	['Over 1 month ago', 'Full-time', 'No degree m...
339	['9 days ago', 'Full-time']

	detected_extensions \
0	{'posted_at': '15 days ago', 'schedule_type': ...
1	{'posted_at': '10 days ago', 'schedule_type': ...
2	{'posted_at': '4 days ago', 'schedule_type': '...
3	{'posted_at': 'Over 1 month ago', 'schedule_ty...
4	{'posted_at': 'Over 1 month ago', 'schedule_ty...
..	...

```

335 {'posted_at': '6 days ago', 'schedule_type': '...'
336 {'posted_at': '6 days ago', 'schedule_type': '...'
337 {'posted_at': '7 days ago', 'schedule_type': '...'
338 {'posted_at': 'Over 1 month ago', 'schedule_ty...'
339 {'posted_at': '9 days ago', 'schedule_type': '...'

```

```

                                job_id
0    eyJqb2JfdG10bGUiOiJVWCBEZXNpZ251ciIsImNvbXBhbn...
1    eyJqb2JfdG10bGUiOiJVWCBEZXNpZ251ciIsImNvbXBhbn...
2    eyJqb2JfdG10bGUiOiJKdW5pb3IgVVggRGVzaWduZXIiLC...
3    eyJqb2JfdG10bGUiOiJTci4gVVggRGVzaWduZXIiLCJjb2...
4    eyJqb2JfdG10bGUiOiJVWCBEZXNpZ251ciIsImNvbXBhbn...
..
335 eyJqb2JfdG10bGUiOiJVWCBEZXNpZ251ciIsImNvbXBhbn...
336 eyJqb2JfdG10bGUiOiJVSSAvVVggRGVzaWduZXIiLCJjb2...
337 eyJqb2JfdG10bGUiOiJvc2VyIEV4cGVyaWVuY2UgKFVYKS...
338 eyJqb2JfdG10bGUiOiJVWCBEZXNpZ251ciB3aXRoIFVJLC...
339 eyJqb2JfdG10bGUiOiJEaXJlY3RvcjBvZiBvc2VyIEV4cG...

```

[340 rows x 8 columns]

Temporary app_url_list container

```
[83]: app_url_list = []
```

Temporary dataframe size variable

```
[83]: sample_df_size = sample_df.shape[0]
sample_df_size
```

```
[83]: 340
```

Execute Google Jobs Listing API query call

```
[84]: for x in range(sample_df_size):
    job_id = sample_df.loc[x]['job_id']
    params2 = {
        "engine": "google_jobs_listing",
        "q": job_id,      # q variable holds job_id number
        "hl": "en",
        "api_key": api_key
    }
    search2 = GoogleSearch(params2)
    results2 = search2.get_dict()
    app_url_list.append(results2['apply_options'][0]['link'])

```

```

0
1
2

```

3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98

99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146

147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194

195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242

243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290

291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338

339

Append app_url_list to sample_df as new column

```
[86]: sample_df['app_url'] = app_url_list
```

```
[87]: sample_df
```

```
[87]:
```

	title \		company_name	location \		via	description \
0	UX Designer		Nielsen	New York, NY		via Nielsen Careers	UX Designer - 85001\n\nTechnology and Engineer...
1	UX Designer		Amazon	New York, NY		via LinkedIn	Description\n\nWe are looking for a UX Designe...
2	Junior UX Designer		Brilliant Experience	Anywhere		via Glassdoor	Brilliant Experience's team works directly wit...
3	Sr. UX Designer		Microsoft	New York, NY		via Microsoft Careers	Senior UX Designer, PromoteIQ\n\nMicrosoft Adv...
4	UX Designer		Tory Burch	New York, NY		via The Muse	JOB SUMMARY\n\nThe Designer, will assist in cr...
..
335	UX Designer		Tablet Hotels - The hotel experts at MICHELIN	New York, NY		via Lensa	Tablet is all about discovering the world's mo...
336	UI /UX Designer		Inabia Software & Consulting Inc.	New York, NY		via Nexxt	The ideal candidate Is detail-oriented with st...
337	User Experience (UX) Designer		Robert Half The Creative Group	New York, NY		via Lensa	Description Do you consider yourself a creativ...
338	UX Designer with UI, Full Time		Creative Circle	New York, NY (+1 other)		via Nexxt	Position UX Designer with UI, Full Time Locati...
339	Director of User Experience		Aquent	New York, NY		via Lensa	Our NYC based non-profit higher education clie...

extensions \

```

0          ['15 days ago', 'Full-time']
1  ['10 days ago', 'Full-time', 'No degree mentio...
2      ['4 days ago', 'Work from home', 'Full-time']
3  ['Over 1 month ago', 'Full-time', 'No degree m...
4      ['Over 1 month ago', 'Full-time']
..
335  ['6 days ago', 'Full-time', 'No degree mention...
336      ['6 days ago', 'Full-time']
337      ['7 days ago', 'Full-time']
338  ['Over 1 month ago', 'Full-time', 'No degree m...
339      ['9 days ago', 'Full-time']

```

```

                                detected_extensions \
0  {'posted_at': '15 days ago', 'schedule_type': ...
1  {'posted_at': '10 days ago', 'schedule_type': ...
2  {'posted_at': '4 days ago', 'schedule_type': '...
3  {'posted_at': 'Over 1 month ago', 'schedule_ty...
4  {'posted_at': 'Over 1 month ago', 'schedule_ty...
..
335 {'posted_at': '6 days ago', 'schedule_type': '...
336 {'posted_at': '6 days ago', 'schedule_type': '...
337 {'posted_at': '7 days ago', 'schedule_type': '...
338 {'posted_at': 'Over 1 month ago', 'schedule_ty...
339 {'posted_at': '9 days ago', 'schedule_type': '...

```

```

                                job_id \
0  eyJqb2JfdG10bGUiOiJVWCBEZXNpZ25lciIsImNvbXBhbn...
1  eyJqb2JfdG10bGUiOiJVWCBEZXNpZ25lciIsImNvbXBhbn...
2  eyJqb2JfdG10bGUiOiJKdW5pb3IgVVggRGVzaWduZXIiLC...
3  eyJqb2JfdG10bGUiOiJTci4gVVggRGVzaWduZXIiLCJjb2...
4  eyJqb2JfdG10bGUiOiJVWCBEZXNpZ25lciIsImNvbXBhbn...
..
335 eyJqb2JfdG10bGUiOiJVWCBEZXNpZ25lciIsImNvbXBhbn...
336 eyJqb2JfdG10bGUiOiJVSSAvVVggRGVzaWduZXIiLCJjb2...
337 eyJqb2JfdG10bGUiOiJVc2VyIEV4cGVyaWVuY2UgKFVYKS...
338 eyJqb2JfdG10bGUiOiJVWCBEZXNpZ25lciB3aXRoIFVJLC...
339 eyJqb2JfdG10bGUiOiJEaXJlY3RvcjBvZiBvc2VyIEV4cG...

```

```

                                app_url
0  https://jobs.nielsen.com/job/New-York-NY/74248...
1  https://www.themuse.com/jobs/amazon/ux-designe...
2  https://www.glassdoor.com/job-listing/junior-u...
3  https://careers.microsoft.com/us/en/job/102458...
4  https://www.themuse.com/jobs/toryburch/ux-desi...
..
335 https://lensa.com/ux-designer-jobs/new-york/jd...
336 https://www.nexxt.com/jobs/ui-ux-designer-new-...

```

```
337 https://www.jobilize.com/job/user-experience-u...
338 https://www.nexxt.com/jobs/ux-designer-with-ui...
339 https://lensa.com/director-of-user-experience-...
```

```
[340 rows x 9 columns]
```

Save sample job listing dataframe dummy data to CSV file

```
[88]: sample_df.to_csv('job_listing_result_log_with_links.csv', index=False)
```