EDGAR Form D Screener

Problem: Manual screening of Form D filings is slow and error-prone.
Goal: Surface only CA LLCs in Residential claiming 506(b), one per company.

Sources: SEC EDGAR master index + primary_doc.xml (no API key; respectful rate limiting)

Approach: Index → resolve XML → verify 4 (or more) criteria → de-dupe by CIK → keep latest per company

Funnel:

Stage	Count
Total filings (D + D/A)	[N_total]
"LLC" in company name	[N_llc]
XML reachable	[N_xml]
Principal office state = CA	[N_ca]
Industry contains "Residential"	[N_res]
Exemptions include 506(b)/06b	[N_506b]
Unique companies kept (1 per CIK)	[N_final]

Results:
- Kept [N_final] companies from [N_total] filings (de-duped by CIK; latest only).
- Median fetch latency <[ms; retry rate x%

Samples:
BLDG Pioneer LLC   2023-01-05   CIK 0001960560
XML:https://www.sec.gov/Archives/edgar/data/1960560/000168316823000028/primary_doc.xml

Stoakes Group LLC   2023-10-12   CIK 0001996708
XML:https://www.sec.gov/Archives/edgar/data/1996708/000199670823000001/primary_doc.xml

Repo:
GitHub Link

Screenshots:

```
$ python filter-lists.py
Loaded 16155 rows; 13930 unique CIKs. Processing newest filings per CIK…
Processed 25/13930 CIKs… kept 0 | reqs:31, 429/503:0, retries:0, avg_req_ms(last31):200
Processed 50/13930 CIKs… kept 0 | reqs:66, 429/503:0, retries:0, avg_req_ms(last97):196
Processed 75/13930 CIKs… kept 0 | reqs:26, 429/503:0, retries:0, avg_req_ms(last123):194
Processed 100/13930 CIKs… kept 0 | reqs:33, 429/503:0, retries:0, avg_req_ms(last156):195
Processed 125/13930 CIKs… kept 0 | reqs:27, 429/503:0, retries:0, avg_req_ms(last183):195
Processed 150/13930 CIKs… kept 0 | reqs:40, 429/503:0, retries:0, avg_req_ms(last200):197

Processed 13825/13930 CIKs… kept 64 | reqs:27, 429/503:0, retries:0, avg_req_ms(last200):191
Processed 13850/13930 CIKs… kept 65 | reqs:25, 429/503:0, retries:0, avg_req_ms(last200):191
Processed 13875/13930 CIKs… kept 65 | reqs:25, 429/503:0, retries:0, avg_req_ms(last200):191
Processed 13900/13930 CIKs… kept 65 | reqs:26, 429/503:0, retries:0, avg_req_ms(last200):191
Processed 13925/13930 CIKs… kept 65 | reqs:25, 429/503:0, retries:0, avg_req_ms(last200):191
Wrote 65 rows to D:\.Programs\VisualStudioProjects\SEC-feeds\edgar-search\data\filtered_list.csv
```