Scraper Quick Guide (CRM Import)

Goal
Collect public lead data and prepare a clean CSV for import into the CRM.

Rules (must follow)
- Respect robots.txt and website Terms of Service.
- Scrape only public business information.
- Use a clear user agent and add delays between requests (rate limit).
- Do not scrape personal/private data that is not explicitly public.

Minimum CSV columns accepted by the CRM
- business_name
- contact

Recommended CSV columns
- business_name, contact, website_url, notes, source_type, source_detail

Data quality checklist before import
- business_name is not empty.
- contact is email, phone, handle, or clear contact channel.
- website_url is a valid domain or URL when available.
- Remove duplicate rows (same domain/contact).
- Keep notes short and actionable.

Import workflow
1) Export scraped leads to CSV.
2) Open CRM table -> Import CSV.
3) Upload file and review duplicate warnings.
4) After import: claim leads, contact, update stage, set follow-up.

Safe defaults for scraping jobs
- Concurrency: 1-3 requests at a time.
- Delay: 500-1500ms between requests.
- Retries: max 2 with backoff.
- Always support dry-run mode before writing output files.

Template header
business_name,contact,website_url,notes,source_type,source_detail