

csv raw data from
main page

Generate Key

Main Web
page

| Companies_Main | |
|----------------|----------------------------|
| PK | <u>company_id NOT NULL</u> |
| | company_name |
| | trustscore |
| | number_of_reviews |
| | domain |

Join Key

| Company details joined | |
|------------------------|----------------------------|
| PK | <u>company_id NOT NULL</u> |
| | company_name |
| | trustscore |
| | domain |
| | rating_class |
| | number_of_reviews |
| | %_star1 |
| | %_star2 |
| | %_star3 |
| | %_star4 |

csv data from sub
pages

Business
page

| Details_Company |
|-------------------------------|
| csv raw data from sub page |

| Reviews_Company |
|-------------------------------|
| csv raw data from sub page |

Generate Key & Join
Key

| Reviews | |
|---------|---------------------------|
| PK | <u>review_id NOT NULL</u> |
| FK1 | company_id |
| | review_star |
| | review_date |
| | reviewer_name |
| | review_title |
| | review_text |
| | experience_date |
| | reply_text |
| | reply_date |

POSTGRESQL

ElasticSearch

reviews csv. from
Postgres

| Reviews | |
|------------------|------|
| <u>review_id</u> | int |
| company_id | int |
| review_star | text |
| review_date | date |
| reviewer_name | text |
| review_title | text |
| review_text | text |
| experience_date | date |
| reply_text | text |
| reply_date | date |