Detailed Comparison of Data Collected by College Dunia and Siksha

Introduction

College Dunia and Siksha have collected data for **MJ College, Bhilai**, but their datasets differ in structure, depth, and coverage. Below is a detailed qualitative and quantitative comparison of their contributions, highlighting the fields they brought, their advantages, disadvantages, and opportunities for improvement.

Fields Brought by College Dunia

1. **Basic College Information**:

   - **College Name**: MJ College, Bhilai
   - **Ratings**:
     - Academic: 4.0
     - Infrastructure: 3.0
   - **Verified Reviews**: 1 Verified Review
2. **Contact Information**:

   - **Address**: Kohka Junwani Road Kohka Bhilai - 490023 Chhattisgarh, India
   - **Phone/SMS**: 0788-2295044
   - **Website**: mjcit.in
3. **Faculty Details**:

   - Includes names, positions, and departments of 8 faculty members.
4. **Course Information**:

   - **10 Courses** listed with:
     - Course Name
     - 1st Year Fees
     - Eligibility Criteria
5. **Empty Sections**:

   - Facilities, Scholarships, and "What's New" sections are placeholders but lack data.

Fields Brought by Siksha

1. **Basic College Information**:

   - **College Name**: MJ Group of Education
   - **Establishment Year**: 2001
   - **Type**: Private Institution
   - **City and State**: Bhilai, Chhattisgarh
   - **Latitude/Longitude**: 21.13, 81.19
   - **Institute Type**: Private
   - **Mode of Course**: Full Time

2. **Contact Information**:

   - **Address**: Kohka, Junwani Road, Bhilai Near Cseb Sub-Station, Po-Supela
   - **Phone**: 9424114847
   - **Email**: director@mjcit.org
   - **Website**: mjcit.org

3. **Course Information**:

   - **Specific Course**: M.Sc. in Computer Science
   - **Duration**: 2 years
   - **Mode**: Full Time

4. **URLs**:

   - College URL: shiksha.com - MJ Group of Education
   - Course URL: shiksha.com - M.Sc. in Computer Science

5. **Additional Data**:

   - **Unique Identifiers**: College and Course IDs.
   - **Country**: India
   - **Tuition Fee Details**: Fields for fees are present but unpopulated.

6. **Empty Fields**:

   - Many fields related to tuition fees, eligibility, and ratings are placeholders without values.

Structural Comparison :

Red- Siksha Metadata that will not help us in most of the case

Siksha	CD

- CollegeID
- name
- course_level
- Course
- type_of_university
- url_college
- url_course
- city
- code
- contact
 	
- coordinates
  - latitude
  - longitude
 	
- country
- email
- establishment_year
- grade
- institute_type
- locality
- phone
- rating_imp
- reviews_count
- state
- type
- top_recruiters
- currency
- tuition_fees
- total_tuition_fees
- total_fee
	- college_name
	- review_number
	
	- ratings
	  - Academic
	  - Infrastructure
	
	- header_data
	- facilities
	
	- contact_info
	  - address
	  - sms_text
	  - call_details
	  - web_link
	
	- whats_new
	- Scholarship
	
	- Faculty
	  - heading
	  - paragraphs
	  - tables
	
	- courses_fees
	  - course_details
	  - eligibility_section
	  - nirf_details

- post_graduation_obc_number
- other_fees
- other_expenses_1st_year
- one_time_payment
- official_career_service_website
- mode_of_course
- international_students_admission_website
- hostel_fees
- graduation_obc_number
- exam_1
- duration
- 1st_year_tuition_fees
- 12th_general_number
- 10th_general_number
- tuition_fees_description
- rating_pg
- post_graduation_description
- Parent_Code
- other_fees_description
- other_eligibility_criteria
- one_time_payment_description
- number_of_reviews
- international_students_eligibility
- hostel_fees_description
- graduation_description
- fees_components
- Fees_components2
- 12th_description
- 10th_description
- course_college_name
- unique_identifier
- college_unique_identifier
- course_unique_identifier
- InstituteId

- InstituteCourseId - InstituteAddressId - IsProcessed - IsExamProcessed - IsCompanyProcessed	

Qualitative Comparison

Aspect	CollegeDunia	Siksha
Depth of College Info	Basic details with ratings and reviews.	Provides establishment year and geolocation.
Contact Information	Phone, SMS, and website included.	Adds email and a more detailed address.
Faculty Data	Detailed list of faculty names and departments.	No faculty information included.
Courses Offered	Comprehensive list of 10 courses with fees.	Focuses on a single course with URLs.
URLs	Only the official college website.	Provides URLs to Shiksha.com for the college and course.
Unique Identifiers	Absent.	Includes unique identifiers for college and course.
Additional Fields	Facilities, scholarships, and new updates.	Adds geolocation and mode of course.
Data Completeness	More populated but lacks advanced data points.	Many fields are placeholders without values.

Quantitative Comparison

Category	College Dunia	Siksha
College Details	4 fields	8 fields
Contact Information	3 fields	4 fields
Faculty Details	8 entries	0 entries
Courses	10 courses	1 course
URLs	1 URL	2 URLs
Unique Identifiers	0	3
Empty Fields/Placeholders	3 sections	40+ fields

Mapping of the entities

Note-

Entity Name	Fields from CD	Fields from Siksha
Name	- college_name	- name - type_of_university - state - city

Contact Details	- contact_info
	  - address
	  - sms_text
	  - call_details
	  - web_link
		- contact
		- coordinates
		- latitude
		- longitude
		- country
		- email
		- establishment_year
		- grade
		- institute_type
		- locality
		- phone

Rating	- review_number
	- ratings
	  - Academic
	  - Infrastructure
		- rating_imp
		- reviews_count
		- rating_pg
		- number_of_reviews

Course and Fees	- courses_fees
	  - course_details
	  - eligibility_section
	  - nirf_details
		- Course
		- mode_of_course
		- course_level
		- currency
		- tuition_fees
		- total_tuition_fees
		- total_fee
		- other_fees
		- other_expenses_1st_year
		- one_time_payment
		- hostel_fees
		- 1st_year_tuition_fees
		- tuition_fees_description
		- other_fees_description
		- hostel_fees_description
		- one_time_payment_description
		- fees_components
		- international_students_admission_website
		- 10th_general_number
		- 12th_general_number
		- graduation_obc_number
		- post_graduation_obc_number
		- other_eligibility_criteria
		- exam_1
		- duration
		- international_students_eligibility
		- 10th_description
		- 12th_description
		- graduation_description
		- post_graduation_description

Faculty	- Faculty
	  - heading
	  - paragraphs
	  - tables	

Facility	- facilities	
Scholarship	- Scholarship	
Recruiters		- top_recruiters

Advantages and Disadvantages

College Dunia

Advantages:

- Provides more comprehensive course data.
- Includes faculty information, which is useful for prospective students.
- Offers basic ratings and reviews.

Disadvantages:

- Lacks unique identifiers for data standardization.
- Does not provide geolocation or establishment year.
- Some sections like facilities and scholarships are empty.

Siksha

Advantages:

- Offers detailed metadata such as unique identifiers, geolocation, and URLs.
- Provides a more detailed address and includes email information.
- Covers metadata fields for advanced use cases (e.g., currency, tuition breakdown).

Disadvantages:

- Course data is limited to a single course.
- Many fields are placeholders without values.
- No faculty information or ratings included.

Recommendations

1. **Combine Data**: Merge the strengths of both datasets for a comprehensive view:

   - Use College Dunia's detailed course and faculty data.
   - Integrate Siksha's metadata and geolocation information.
2. **Fill Missing Sections**:

   - Populate fields like facilities, scholarships, and ratings.
   - Add detailed tuition fee breakdowns in Siksha's dataset.
3. **Standardize Structure**:

   - Use unique identifiers and geolocation across the combined dataset for consistency.
4. **Validate URLs and Information**:

   - Ensure URLs lead to valid and relevant pages.
   - Cross-check course fees and faculty names with official sources.

Conclusion

Both datasets bring unique strengths and weaknesses to the table. College Dunia focuses on practical information like courses, fees, and faculty, while Siksha provides advanced metadata and infrastructure for more sophisticated analysis. Combining these datasets would create a comprehensive and enriched profile for MJ College, Bhilai.