



Team name

OCTAVIANS CRADLE

Problem statement

UNIFIED DATA VERIFICATION

mock dataset
structure

Unified Data Verification

Akshi Karpe
Officer

Admin Control Panel

CJ KN RK BM DM
RK BK LM +28 more

Selected Dataset select all

All (combined view)

department 1 / Excise

department 2 / Home ...

department 3
department 4
department 5

+ new 1 2

View Control Panel

Unified Data View 42 dataset selections

Visual Analytical View 16 charts created!

saved templates:

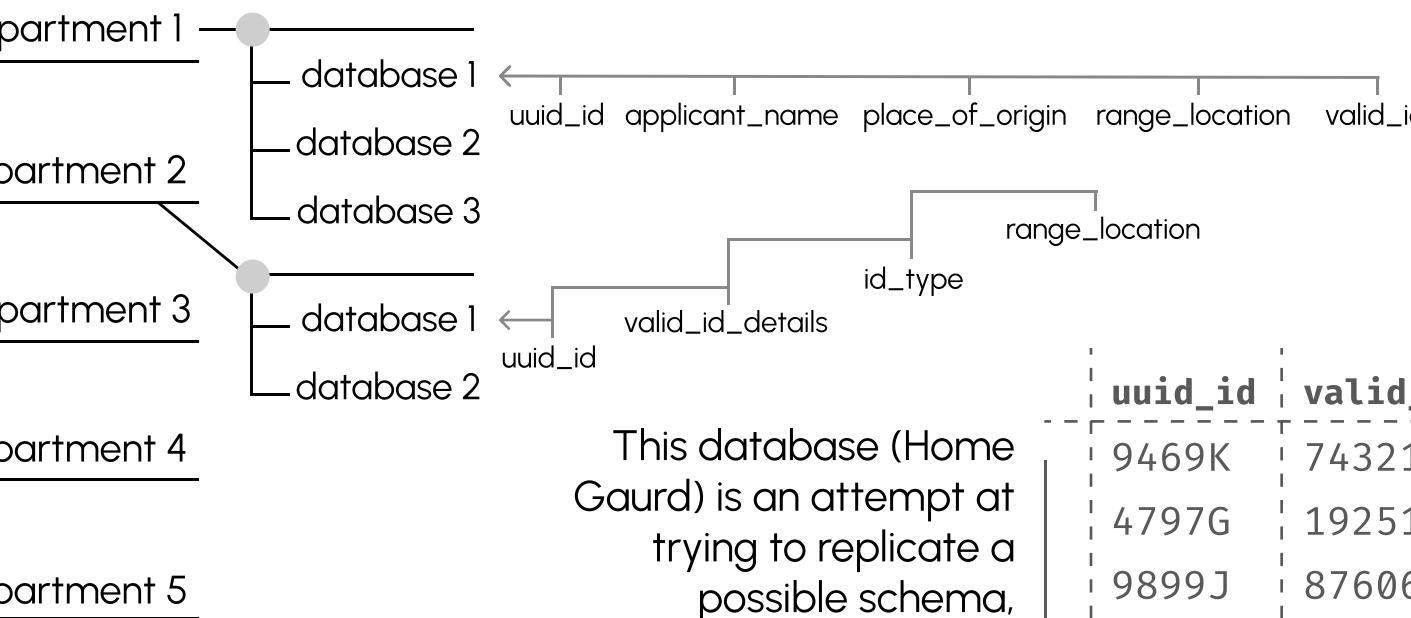
save changes create new +

test template 1 ^

```

graph LR
    UD[Unified Data Verification] --> D1[department 1]
    UD --> D2[department 2]
    UD --> D3[department 3]
    UD --> D4[department 4]
    UD --> D5[department 5]

    D1 --- DB1_1[database 1]
    D2 --- DB2_1[database 1]
    D2 --- DB2_2[database 2]
    D3 --- DB3_1[database 1]
    D3 --- DB3_2[database 2]
  
```

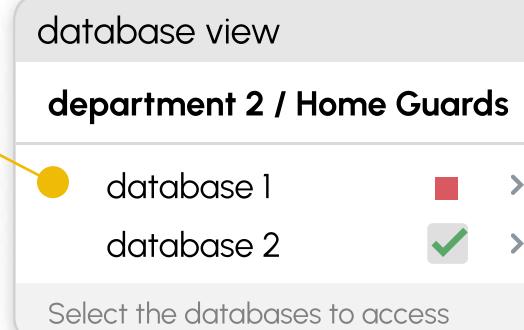


The idea here is to use a government-issued id that is unique to every citizen and hence can be used to GROUP data across different databases.

uid_id	valid_id_details	id_type	updaters_id	nature_of_crime	place_registered
9469K	74321437	aadhar	Anusha	shoplifting	Mysore
4797G	19251693	aadhar	Anagha	trafficViolation	Bengaluru
9899J	87606811	pan_card	Arev	pick_pocketting	Bengaluru
7321N	10892185	driving_li	Gauri	vandalism	Mandy
2347B	65822441	aadhar	Abhishek	shoplifting	Mysore
6369C	42537067	passport	Tanya	shoplifting	Bengaluru
8931D	19251693	aadhar	Anagha	pick_pocketting	Mysore



every user will only have access to the data, as allowed by the admin.



toggle between multiple databases of multiple departments.

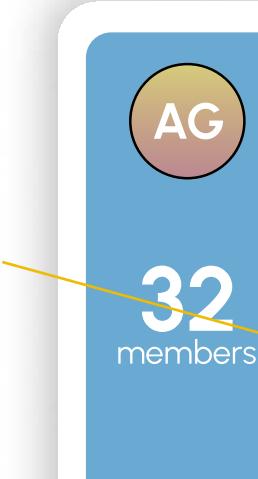
These are the sub-ordinates of the logged in user. The templates (discussed in section 1.3) can be shared among the users, which saves a lot of time.

uid_id	applicant_name	place_of_origin	range_location	valid_id
7249Q	Jayendra	Mandy	Mandy	19251693
6898M	Gaurav	Bengaluru	Bengaluru	67800897
4896C	Vihaan	Bengaluru	Bengaluru	56845905
1176E	Richa	Mandy	Mandy	65822441
3969F	Mehul	Mysore	Mysore	34569085

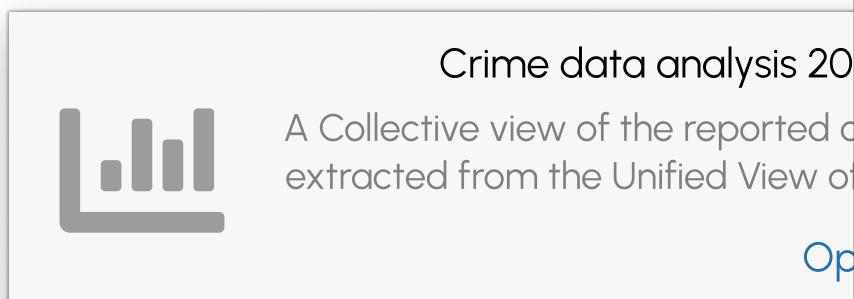
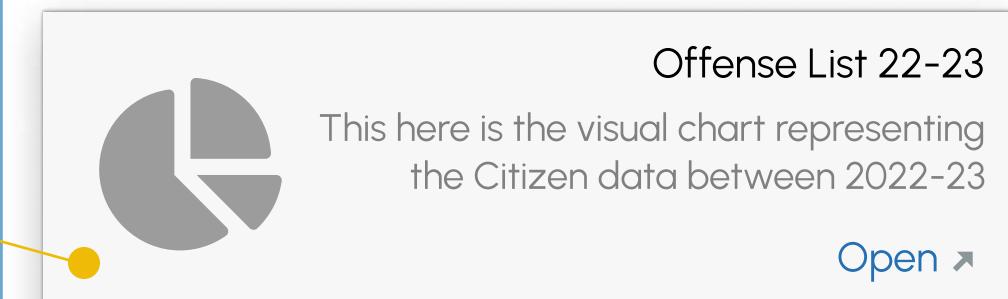
This database (Excise Department) also uses the unique citizen id, Aadhar as a means to register the new data.

Once all the required databases are selected the user clicks on the Unified Data View panel which runs the required complex GROUP-ing algorithms to create a unified view.

The user can get fascinating insights from the selected databases by creating bar, pie chart. These created visualizations can be added to the final report.

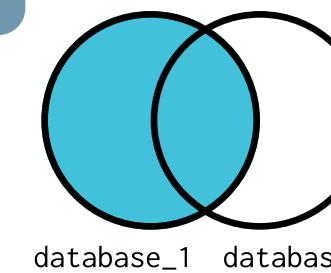


Currently saved visuals!



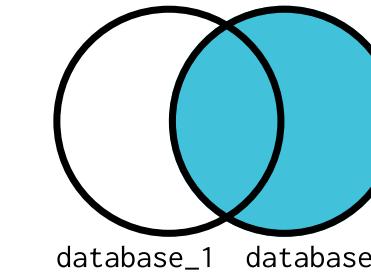
Unified Data View

2 dataset selections



join direction: left full right

by default, the data shown in the Unified Data view are merged together with an **outer join**.



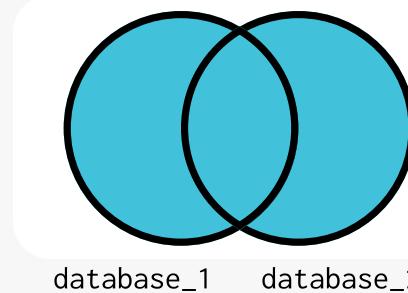
the toggle button switches the order of the join. i.e. **right join** among all the selected databases.

excise (database 1) ➞ home guard (database 2)

represents a **Left Join** between database_1 and database_2.

SQL Query

```
select department_1.applicant_name,
       department_1.place_of_origin,
       department_1.range_location,
       valid_id,
       valid_id_details,
       id_type,
       uploaders_id,
       nature_of_crime,
       place_registered
  from department_2
 full join department_1 on valid_id_details = valid_id;
```



database_1 database_2

uuid_id	applicant_name	place_of_origin	range_location	valid_id	uuid_id_2	valid_id_details	id_type	uploaders_id	nature_of_crime	place_registered
6898M	Gaurav	Bengaluru	Bengaluru	67800897	-	-	-	-	-	-
4896C	Vihaan	Bengaluru	Bengaluru	56845905	-	-	-	-	-	-
3969F	Mehul	Mysore	Mysore	34569085	-	-	-	-	-	-
-	-	-	-	-	9469K	74321437	aadhar	Anusha	shoplifting	Mysore
7249Q	Jayendra	Mandya	Mandya	19251693	4797G	19251693	aadhar	Anagha	trafficViolation	Bengaluru
-	-	-	-	-	9899J	87606811	pan_card	Arev	pick_pocketing	Bengaluru
-	-	-	-	-	7321N	10892185	driving_li	Gauri	vandalism	Mandya
1176E	Richa	Mandya	Mandya	65822441	2347B	65822441	aadhar	Abhishek	shoplifting	Mysore
-	-	-	-	-	6369C	42537067	passport	Tanya	shoplifting	Bengaluru
7249Q	Jayendra	Mandya	Mandya	19251693	8931D	19251693	aadhar	Anagha	pick_pocketing	Mysore

AG

Search: params: applicant_name place_of_origin range_location Filters: + add custom filterscombined view parameters: + add new parametersselection: rows per page: merge direction: excise (database 1) ➞ home guard (database 2)[Download collective PDF report](#)32
members127
datasets

 Unified Data View
2 dataset selections

When a filter related to a specific template is amended, save changes is toggled, which enables the user to save it.

saved templates:

save changes create new +

mysore data-list 2021

test template 1  

mysore data-list 2021  

Mandy locale data-list  

- one of the most defining features of the unified data verification, is the ability to create templates based on the filters selected by the user.
- These templates are stored as a JSON key/pair value in the database, which can be shared between the users of the Organization.

[Download collective PDF report](#)

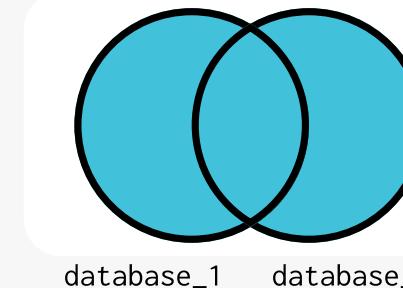
The user can download a PDF format report of the report exactly as seen on screen. Without any change to the order of data.

 Mention the specifics for the PDF report

- Display source on footer
 - Add created charts to the end of report.
 - list sourced databases
- [send to mail](#) [download now](#)

SQL Query

```
select department_1.applicant_name,
       department_1.place_of_origin,
       department_1.range_location,
       valid_id,
       valid_id_details,
       id_type,
       uploaders_id,
       nature_of_crime,
       place_registered
  from department_2
 full join department_1 on valid_id_details = valid_id
 where place_of_origin = 'Mandy'
 and department_2.id_type = 'aadhar';
```



here the user can search across the entire database in selected columns for a specific data

Search: params:

Filters: + add custom filters

combined view parameters: + add new parameters

selection: rows per page: merge direction: excise (database 1)

[Download collective PDF report](#)

valid_id | valid_id_details 

- The entire concept of the joining tables and deriving meaningful data out of it depends a lot on having common key / values that in some way can make the tables relatable to each other.
- This can be created automatically for smaller databases but as the size of the data increases so does computational requirements, so the user might have to select the field.

The filters in the dashboard are auto-generated from selected databases. A group_by SQL query is used to analyze what can be grouped.

SQL Query

```
select count(*) as no_uniques, place_of_origin
from department_1
group by place_of_origin;
```

```
select count(*) as no_uniques, place_registered
from department_2
group by place_registered;
```

nature_of_crime: shoplifting 

pick_pocketing

shoplifting

vandalism

AG

32

members

127

datasets

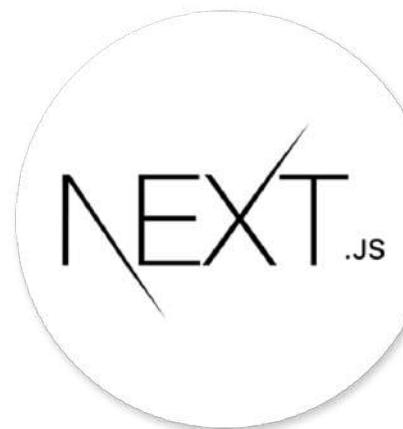
	uuid_id	id_type	uuid_id_2	valid_id	place_of_ori	range_locati	uploaders_id	nature_of_crime	applicant_name
1	-	aadhar	9469K	-	-	-	Anusha	shoplifting	-
2	4797G	aadhar	4797G	19251693	Mandy	Mandy	Anagha	trafficViolation	Jayendra
3	-	pan_card	9899J	87606811	-	-	Arev	pick_pocketing	-
4	1176E	aadhar	2345M	3245093	Mysore	Mysore	Arun	pick_pocketing	Richa

KPIs to measure the solutions / Business logic

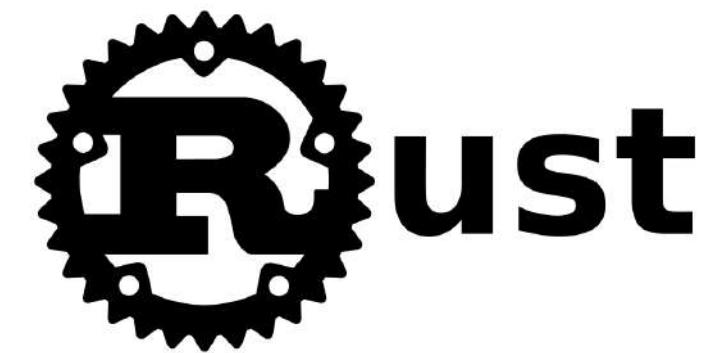
DAU: The ease offered by simplified UI and Integrated View of data reduces the hectic task of having to go through isolated datasets of each department, hence be a major factor in increasing the DAU who use the platform.

Number of User Actions: The number of user creating templates, and sharing it across the organization, is an effective way to measure the effectiveness of this feature.

Technology stack used:



strapi



Related Cost:

The development process by itself shouldn't cost anything as AWS services can be used within their free tier limits, but once deployed an EC2 instance alone cost a minimum of \$ 10 if used beyond its free tier limit (which is very likely) and same can be said for RDS (database) instances running on AWS.