# Punjab Dataset Survey

**Data set from Punjab survey where citizens shared top two priorities to be addressed. This survey was recorded before the Election result 2022.**
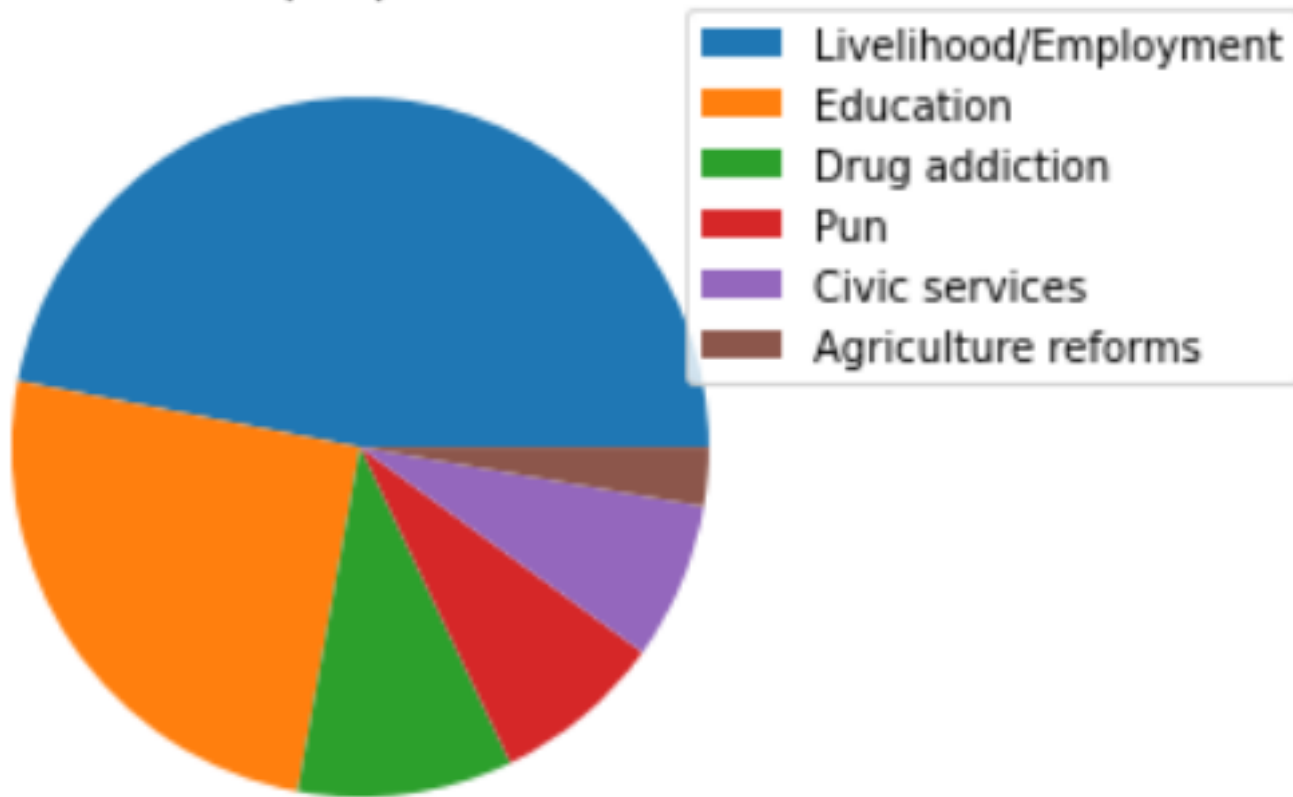
This has a sample set of 886

VOTE

# Participants in this survey who did not write anything in the following sections

| | |
|---|---|
| contact_name | 0 |
| resident_punjab | 26 |
| Pin_constituency | 240 |
| locality_name | 824 |
| priority_1 | 294 |
| Pun_priority_1 | 852 |
| why_priority_1 | 448 |
| priority_2 | 459 |
| Pun_priority_2 | 859 |
| why_priority_ | 600 |

**eg: 26 people did not fill the resident_punjab section**

# Priorities people would like to address



Priorities people selected

- Livelihood/Employment
- Education
- Drug addiction
- Pun
- Civic services
- Agriculture reforms

## Top 4 Locations of the people who filled out this survey

| | |
|---|---:|
| Bathinda | 3 |
| Patiala | 3 |
| Gurdaspur | 3 |
| Dasuya | 3 |

# Personal recommendations after analyzing the data set
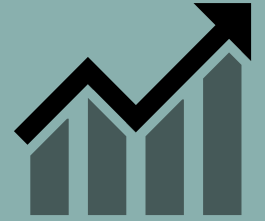
-  A lot of people are unhappy with the goverment school/college infrastructure

-  Unemployment has been increased which leads to crimes

RECOMMEND

-  Civil services in certain areas could be improved

-  Private school and colleges fee is on a higher side and government institutes has lower quality of infrastructure and teachers (As per the data)

RECOMMEND

**Improvements** with the form which could lead to better data next time:

1- Few sections/columns could be made mandatory to fill, in this particular dataset(Out of 886 entries):
 (i) - The resident_punjab column has 26 NaN(no entries) values
(ii) - The Pin_constituency column has 240 NaN values

(iii) - The locality column is not really an important column, we could hence think of skipping it. It however has 824 NaN values, which means only 62 people filled that part.
 (iv) - priority_1 & Priority_2 has 294 & 459 NaN values respectively, which could also be made mandatory.

# 2- Details in few columns has been filled with regional language

(i) - Individuals having regional language knowledge could analyze such datasets
(ii) - English could be made mandatory (Has its cons)

Vedaang Gulhane

Analysis done using python