Exploring datasets – Examples

Workshop at NLU Odisha

CivicDataLab

2022/08/08 (updated: 2022-08-12)

Crime in India

Source: Wikipedia

Objective:

1. Download the table *Absolute number of reported crimes & crime rates across India*
2. Calculate *total reported crimes from 2016-2019* in all states
3. Calculate *percent share of each state* in reported crimes
4. Arrange the results in decreasing order of percent share

Most cited acts and sections

Source: [Justice Hub - Indian Kanoon Statistics](#)

Objective:

1. Import data in google sheets
2. Filter data for IPC sections
3. Find how many sections contribute to 50% of the cases
4. Add a column for the section title

Mortality data from DevDataLab

Source: Dropbox

Link: https://www.dropbox.com/sh/y949ncp39towulf/AACJRhJs6uzbjkB6uM5O9mKga/mortality/csv?dl=0&subfolder_nav_tracking=1

Objective:

1. Import data in google sheets
2. Compare mortality data across states from 2019 - 2021
3. Compare mortality data for Karnataka across years, as shown here

Judges of the Orissa High Court

Source: Justice Hub

Link: https://justicehub.in/dataset/dataset-on-judges-of-the-odisha-high-court/resource/92c89ebb-3612-4ad5-a327-d55f65327577

Objective:

1. Import data in google sheets
2. Find the number of Male and Female judges
3. Find the percentage of judges where *Parent High court* is Odisha High Court
4. Find the total number of judges appointed each year
5. Find the total number of judges retiring each year

India at CommonWealth Games

civic
data
lab

Source: Wikipedia

Link: https://en.wikipedia.org/wiki/India_at_the_Commonwealth_Games

Objective:

1. Find the table titled - *Overall Medal Tally By Sport*
2. Import data in Microsoft Excel
3. Find the percentage share of total medals by sports
4. Arrange the sports in decreasing order of percentage share of total medals
5. Find the percentage share of gold, silver and bronze medals by each sport
6. Remove Grand Totals from Columns
7. Highlight the cells with the largest contribution to gold, silver and bronze medals respectively

Queries and Feedback