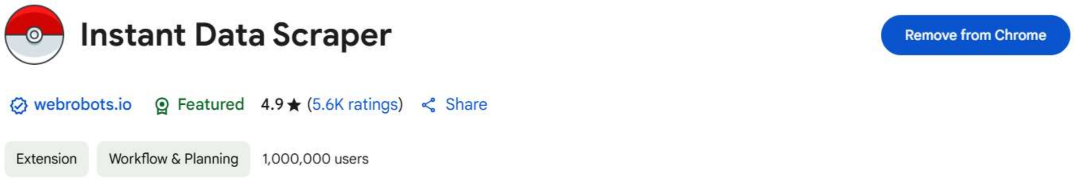
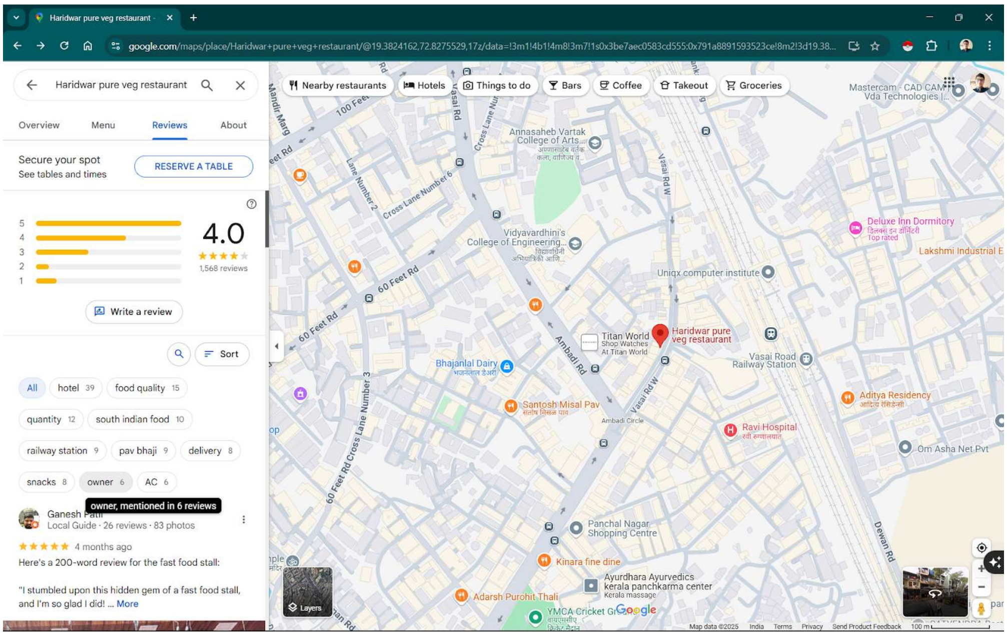
1. To perform web crawling, scraping and parsing using Instant data scraper.
    a. Install Chrome extension



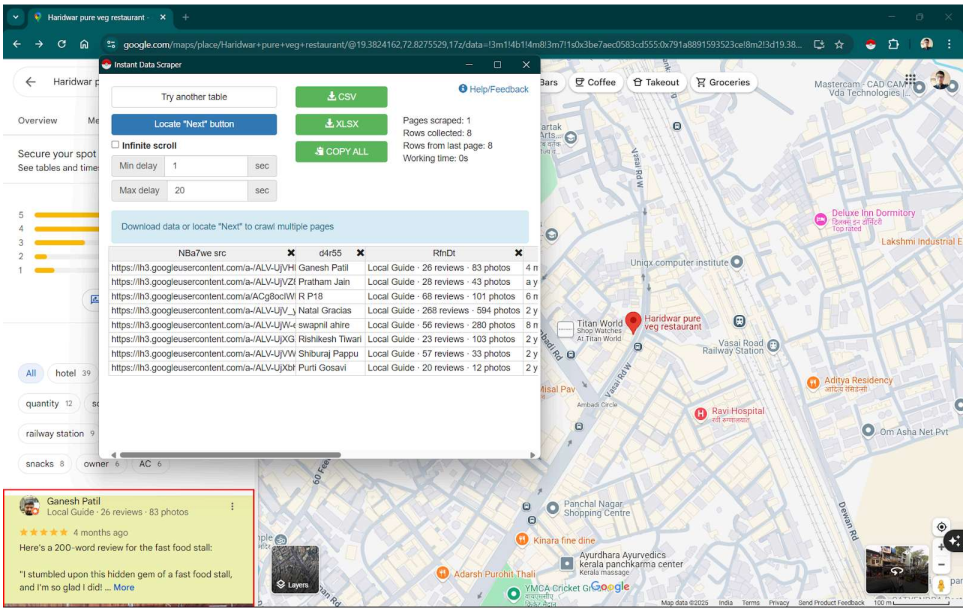
    b. Open maps.google.com
    c. Search for any restaurant for eg Haridwar



Here's a 200-word review for the fast food stall:

"I stumbled upon this hidden gem of a fast food stall, and I'm so glad I did! ... More

    d.
    e. Click pokemon ball icon of instant data scraper
    f. If nahi dikh raha hai, then click on puzzle piece icon it will list all the extensions.

g. 

Instant Data Scraper

Try another table

Locate "Next" button

□ Infinite scroll

Min delay 1 sec

Max delay 20 sec

CSV

XLSX

COPY ALL

Help/Feedback

Pages scraped: 1
Rows collected: 8
Rows from last page: 8
Working time: 0s

Download data or locate "Next" to crawl multiple pages

| NBa7we src | d4r55 | RfnDt | |
|---|---|---|---|
| https://lh3.googleusercontent.com/a-/ALV-UjVHI | Ganesh Patil | Local Guide · 26 reviews · 83 photos | 4 m |
| https://lh3.googleusercontent.com/a-/ALV-UjVZ8 | Pratham Jain | Local Guide · 28 reviews · 43 photos | a y |
| https://lh3.googleusercontent.com/a/ACg8ocIWI | R P18 | Local Guide · 68 reviews · 101 photos | 6 m |
| https://lh3.googleusercontent.com/a-/ALV-UjV_y | Natal Gracias | Local Guide · 268 reviews · 594 photos | 2 y |
| https://lh3.googleusercontent.com/a-/ALV-UjW-e | swapnil ahire | Local Guide · 56 reviews · 280 photos | 8 m |
| https://lh3.googleusercontent.com/a-/ALV-UjXG. | Rishikesh Tiwari | Local Guide · 23 reviews · 103 photos | 2 y |
| https://lh3.googleusercontent.com/a-/ALV-UjVW | Shiburaj Pappu | Local Guide · 57 reviews · 33 photos | 2 y |
| https://lh3.googleusercontent.com/a-/ALV-UjXbh | Purti Gosavi | Local Guide · 20 reviews · 12 photos | 2 y |

Ganesh Patil
Local Guide · 26 reviews · 83 photos
★★★★★ 4 months ago
Here's a 200-word review for the fast food stall:
"I stumbled upon this hidden gem of a fast food stall, and I'm so glad I did! ... More

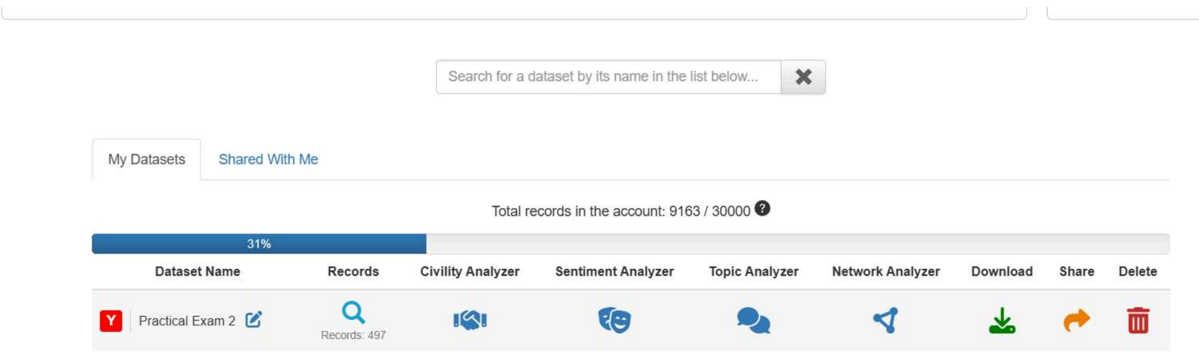
h. Export Xlsx or csv

2. To perform data cleaning on social media data using python or R.

3. To perform exploratory data analysis and visualization of Social media data for business.

```
# Import necessary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import skew, kurtosis
# Load sample social media data (You can replace this with your dataset)
data = pd.DataFrame({
'user_id': [1, 2, 3, 4, 5, 6, 7, 8, 9, 10],
'followers_count': [100, 200, 350, 150, 500, 120, 60, 300, 400, 800],
'likes': [10, 15, 10, 25, 30, 50, 45, 60, 100, 75],
'shares': [5, 3, 7, 8, 6, 12, 10, 9, 15, 18],
'comments': [1, 2, 3, 4, 5, 6, 2, 3, 6, 10],
'engagement_rate': [0.1, 0.15, 0.2, 0.18, 0.3, 0.25, 0.28, 0.35, 0.4,
0.5],
'sentiment_score': [0.8, 0.5, 0.7, 0.9, 0.6, 0.4, 0.6, 0.8, 0.9, 0.7]
})



print("First 5 rows:\n")
print(data.head())
```

```
data.describe()

# Central Tendency - Mean, Median, Mode
print("\nMean of Followers Count:", data['followers_count'].mean())
print("Median of Followers Count:", data['followers_count'].median())
print("Mode of Followers Count:", data['followers_count'].mode()[0])

# Measure of Spread - Standard Deviation, Variance, Range
print("\nStandard Deviation of Followers Count:",
data['followers_count'].std())
print("Variance of Followers Count:", data['followers_count'].var())
print("Range of Followers Count:", data['followers_count'].max() -
data['followers_count'].min())

# Interquartile Range (IQR)
Q1 = data['followers_count'].quantile(0.25)
Q3 = data['followers_count'].quantile(0.75)
IQR = Q3 - Q1
print("\nInterquartile Range (IQR) of Followers Count:", IQR)

# Skewness and Kurtosis
print("\nSkewness of Followers Count:", skew(data['followers_count']))
print("Kurtosis of Followers Count:", kurtosis(data['followers_count']))

print("\nMissing Values in each column:")
print(data.isnull().sum())

# Visualizations
plt.figure(figsize=(14, 8))
# Histogram of Followers Count
plt.subplot(2, 2, 1)
sns.histplot(data['followers_count'], kde=True, color='skyblue', bins=10)
plt.title('Distribution of Followers Count')

# Boxplot of Followers Count
plt.subplot(2, 2, 2)
sns.boxplot(x=data['followers_count'], color='lightgreen')
plt.title('Boxplot of Followers Count')

# Heatmap of correlations
# plt.subplot(2, 2, 3)
plt.figure(figsize=(10, 8))
sns.heatmap(data.corr(), annot=True, cmap='coolwarm', vmin=-1, vmax=1)
plt.title('Correlation Heatmap')
```

4. Develop Content( text, emoticons, image, audio, video) based  social media analytics model for business.
    a. Go to https://communalytic.org/ and create an account or sign in if acc exists already(EDU LOGIN).
    b. After login click Youtube video comments (Youtube api key google se nikalo yaad nahi kaise nikali thi 😨 )
    c. Koi bhi video youtube ki uthao ( ▶ Einstein and the Theory of Relativity | HD |  this in my case)
    d. Click start data collection
    e. 

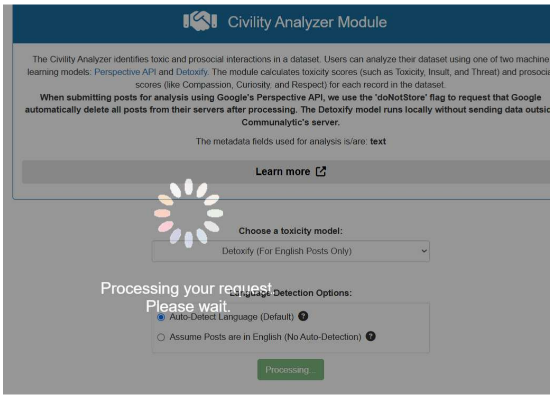
Search for a dataset by its name in the list below... ✖

My Datasets | Shared With Me

Total records in the account: 9163 / 30000 ?

31%

| Dataset Name | Records | Civility Analyzer | Sentiment Analyzer | Topic Analyzer | Network Analyzer | Download | Share | Delete |
|---|---|---|---|---|---|---|---|---|
| Y Practical Exam 2 | Records: 497 | | | | | | | |

    f. Select Civility Analyzer
    g. 

Civility Analyzer Module

The Civility Analyzer identifies toxic and prosocial interactions in a dataset. Users can analyze their dataset using one of two machine learning models: Perspective API and Detoxify. The module calculates toxicity scores (such as Toxicity, Insult, and Threat) and prosocia scores (like Compassion, Curiosity, and Respect) for each record in the dataset.

**When submitting posts for analysis using Google's Perspective API, we use the 'doNotStore' flag to request that Google automatically delete all posts from their servers after processing. The Detoxify model runs locally without sending data outsic Communalytic's server.**

The metadata fields used for analysis is/are: **text**

Choose a toxicity model: Detoxify (For English Posts Only)

Processing your request. Please wait.

Language Detection Options:

● Auto-Detect Language (Default)

○ Assume Posts are in English (No Auto-Detection)

    h. Then sentiment analyzer and baaki sab….

5. Develop Structure based  social media analytics model for any business.

```
import networkx as nx
import matplotlib.pyplot as plt
```

```
# Create a directed graph
G = nx.DiGraph()

# Adding nodes (Users in social media)
users = ["Alice", "Bob", "Charlie", "David", "Eve"]
G.add_nodes_from(users)

# Adding edges (Relationships: Follows, Likes,
Mentions)
edges = [("Alice", "Bob"), ("Bob", "Charlie"),
("Charlie", "David"),
         ("David", "Eve"), ("Eve", "Alice"),
("Alice", "Charlie"),
         ("Bob", "Eve"), ("Charlie", "Alice")]
G.add_edges_from(edges)

# Calculate centrality measures
degree_centrality = nx.degree_centrality(G)
betweenness_centrality =
nx.betweenness_centrality(G)
eigenvector_centrality =
nx.eigenvector_centrality(G)

# Display centrality measures
print("Degree Centrality:", degree_centrality)
print("Betweenness Centrality:",
betweenness_centrality)
print("Eigenvector Centrality:",
eigenvector_centrality)
```

```
# Draw the network
plt.figure(figsize=(8, 6))
pos = nx.spring_layout(G)  # Positioning nodes
visually
nx.draw(G, pos, with_labels=True,
node_color='skyblue', edge_color='gray',
node_size=2000, font_size=10)
nx.draw_networkx_edge_labels(G, pos,
edge_labels={(u, v): 'follows' for u, v in edges})
plt.title("Social Media Network Graph")
plt.show()
```