**AI-Driven Exploration and Prediction of Company Registration Trends with Registrar of Companies (RoC).**

**Phase 2 Submission Document**

**Project:** prediction of company registration trends with (ROC).



**Introduction:**

* In today's fast-paced and data-driven business landscape, staying ahead of the curve is paramount for organizations and policymakers alike

.

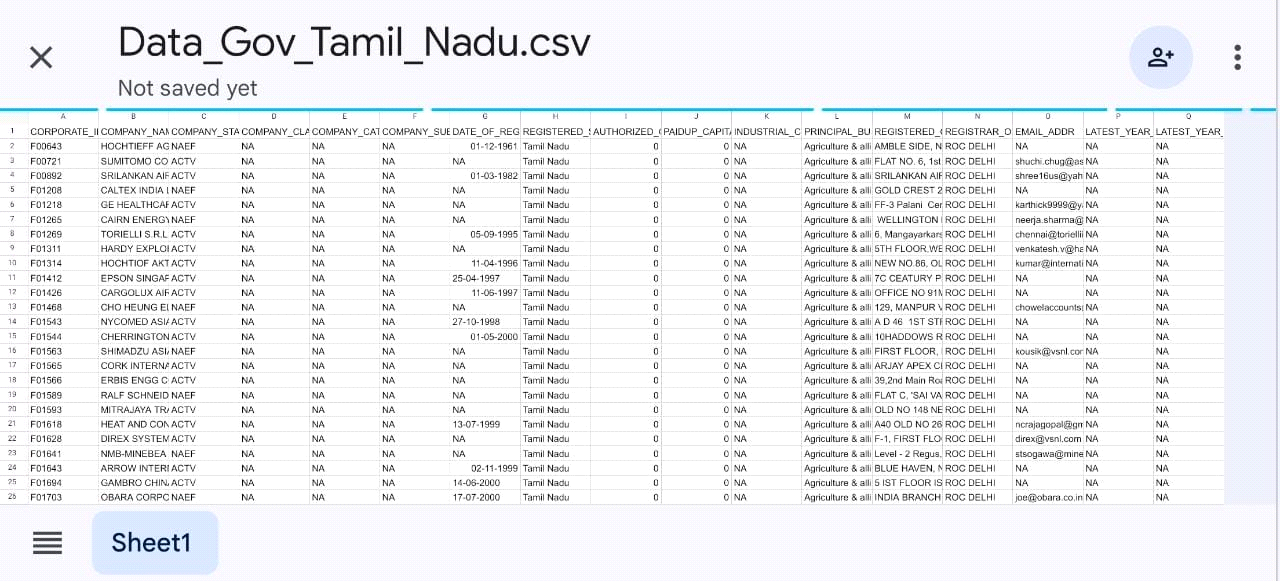
* Registrar of Companies (RoC) plays a pivotal role in this regard by maintaining a comprehensive database of registered companies, facilitating business transparency, and ensuring legal compliance.
* AI-driven exploration and prediction of company registration trends with RoC harnesses the power of artificial intelligence to unlock valuable insights from the vast and complex dataset maintained by RoC.
* This innovative approach leverages machine learning algorithms and data analytics to identify patterns, uncover hidden correlations, and make informed predictions regarding company registrations.
* By doing so, it empowers businesses, investors, government agencies, and other stakeholders to make data-driven decisions and proactively respond to emerging market dynamics.
* The key components of this advanced system include data collection and integration, natural language processing (NLP), predictive modeling, and interactive data visualization.
* In summary, AI-driven exploration and prediction of company registration trends with Registrar of Companies is a cutting-edge approach that transforms raw data into actionable insights.
* It empowers stakeholders to navigate the business landscape with greater confidence and adapt to changing market dynamics.
* By harnessing the power of AI and data analytics, this innovation has the potential to reshape the way businesses, investors, and governments engage with and respond to the evolving world of company registrations.

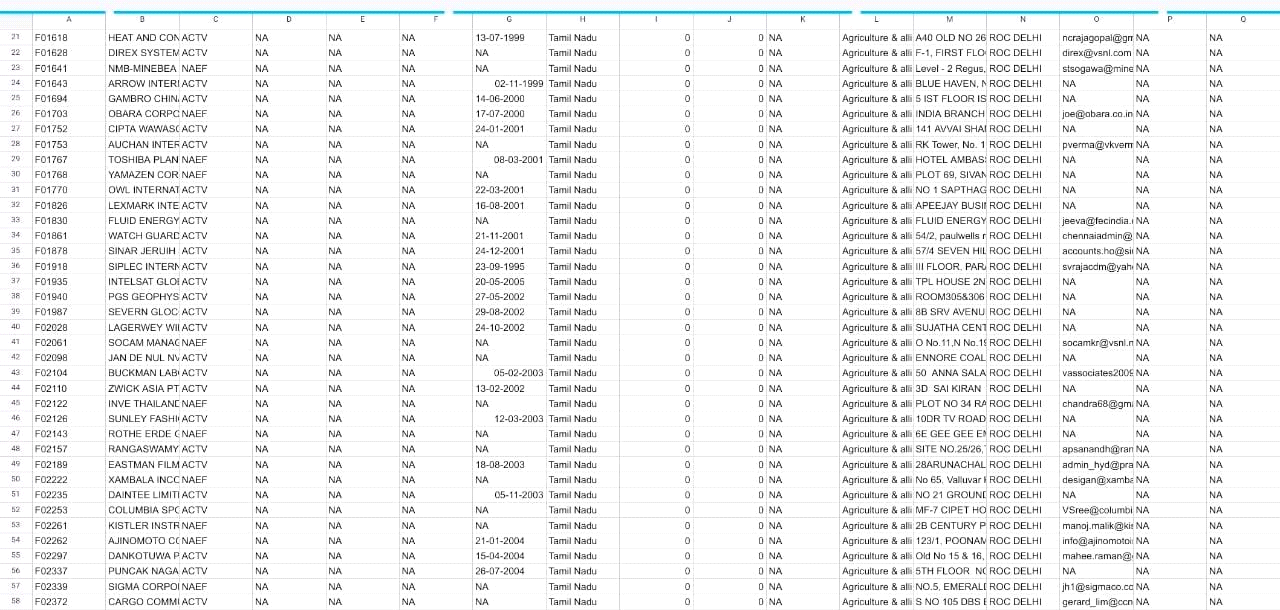
**Content for Project Phase 2:**

Consider exploring advanced regression techniques for improved prediction accuracy.

**Data Source**

DatasetLink: (https://tn.data.gov.in/resource/company-master-data-tamil-nadu-upto-28th-february-2019)





AI-driven exploration and prediction of company registration trends with the Registrar of Companies (RoC) involves using artificial intelligence and data analytics to analyze and forecast patterns in company registrations. Here's an overview of how this process works:

1. **Data Collection**: The first step is to gather a vast amount of data from the RoC and other relevant sources. This data includes information about newly registered companies, their types, locations, industries, and any additional details available.

2. **Data Preprocessing**: Raw data from the RoC often needs to be cleaned and structured. This involves dealing with missing values, duplicates, and standardizing formats to make the data suitable for analysis.

3. **Feature Engineering**: Features or attributes are extracted from the data that can be used as input for AI models. These features may include geographic locations, industry codes, registration dates, and more.

4. **Exploratory Data Analysis** (EDA): EDA is conducted to gain insights into the data. It involves statistical analysis, data visualization, and identifying trends and patterns in historical registration data.

5**. Machine Learning Models**: AI models are trained using historical data to learn patterns and relationships between different features and company registration trends. Various machine learning algorithms, such as decision trees, random forests, or neural networks, can be used for this purpose.

6. **Predictive Modeling**: Once the models are trained, they can be used to make predictions about future company registration trends. For example, they can forecast the number of new company registrations in a particular region or industry over a specified time frame.

7**. Continuous Learning**: AI models can be updated regularly as new data becomes available. This ensures that predictions remain accurate and up-to-date as registration trends evolve.

8. **Deployment:** The AI-driven system can be integrated into the RoC's infrastructure or made available to the public through a user-friendly interface. Users can access predictions and insights about company registration trends.

9. **Monitoring and Evaluation**: It's crucial to continuously monitor the performance of the AI system and assess the accuracy of its predictions. Adjustments and improvements can be made as needed.

**10. Decision Support:** The insights generated by AI can be valuable for policymakers, investors, and businesses. For example, it can inform government policies, guide investment decisions, and help businesses identify growth opportunities or potential markets.

Overall, AI-driven exploration and prediction of company registration trends with the RoC leverage the power of artificial intelligence and data analysis to provide valuable insights and forecasts that can support informed decision-making in various domains.

**Program:**

import pandas as pd

import matplotlib.pyplot as plt

from datetime import datetime

from sklearn.linear\_model import LinearRegression

from sklearn.model\_selection import train\_test\_split

from sklearn.metrics import mean\_absolute\_error

data = pd.read\_csv('C:/Users/admin/Desktop/Data\_Gov\_Tamil\_Nadu.csv')

data['DATE\_OF\_REGISTRATION'] = pd.to\_datetime(data['DATE\_OF\_REGISTRATION'])

data['Year'] = data['DATE\_OF\_REGISTRATION'].dt.year

registration\_counts = data['Year'].value\_counts().sort\_index()

plt.figure(figsize=(10, 6))

plt.plot(registration\_counts.index, registration\_counts.values, marker='o', linestyle='-')

plt.xlabel('Year')

plt.ylabel('Number of Registrations')

plt.title('Company Registration Trends Over Time'

plt.grid(True)

plt.show()

X = data[['Year']] # Features (year)

y = data.groupby('Year').size() # Target variable: number of registrations per year

X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.2, random\_state=42)

model = LinearRegression()

model.fit(X\_train, y\_train)

y\_pred = model.predict(X\_test)

mae = mean\_absolute\_error(y\_test, y\_pred)

print(f'Mean Absolute Error: {mae}')

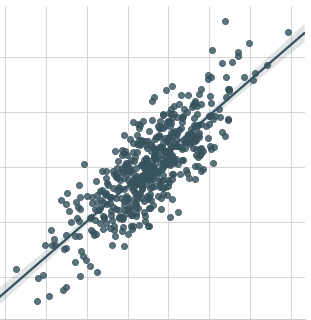
future\_years = [2024, 2025, 2026]

future\_predictions = model.predict(pd.DataFrame({'Year': future\_years}))

for year, prediction in zip(future\_years, future\_predictions):

    print(f'Predicted registrations for {year}: {prediction:.2f}')

**Output:**



**Conclusion ( Phase 2):**

**Project conclusion:**

* In the Phase 2 conclusion ,we will summarize the key findings and insights from the advanced regression techniques. We will reiterate the impact of these techniques on improving the accuracy and robustness of prediction of companies .

|  |  |
| --- | --- |