|  |
| --- |
| AI-DRIVEN EXPLORATION AND PREDICTION OF COMPANY REGISTRATION TRENDS WITH (ROC) |

PHASE 5

**ABSTRAC:**

This project report outlines the development and implementation of an AI- driven system for the exploration and prediction of company registration trends using data from the Register of Companies. The primary objective of this project is to leverage artificial intelligence and machine learning

Technique to analyze historical company registration data, identify patterns and trends, and make predictive forecasts to assist government authorities, business analysts, and policymakers in making informed decisions regarding economic development and regulatory changes.

# INTRODUCTION:

The Register of Companies (RoC) maintains a database of company

Registration and related information. Analyzing this data can provide valuable insights into economic trends, industrial growth, and regulatory compliance.

This project aims to develop an AI-driven system that can automatically process, analyze, and predict company registration trends based on historical RoC data.

# PROJECT SCOPE AND OBJECTIVE:

# *SCOPE*:

* Collect historical company registration data from RoC.
* Preprocess and clean the data for analysis.
* Explore the data to identify patterns, correlations, and anomalies.
* Develop AI and machine learning models for trend analysis and

Prediction.

* Create a user-friendly interface for stakeholders to access the insights.

# *OBJECTIVE*:

* Build a data pipeline for regular data updates.
* Develop predictive models for company registration trends.
* Provide interactive visualizations to convey insights.
* Evaluate model performance and validate predictions.

# REGISTRATION OF A COMPANY IN INDIA:

The present article aims to briefly describe the procedure for registration of a “Company” in India and procedure followed for establishment of Liaison and Branch Office of a foreign company in India.

The Companies Act, 1956 governs the formation, regulation and winding up of companies in India and the Registrar of Companies appointed under the Companies Act registers companies and ensure all statutory requirements are

Compile by the Company Objective of Registration of a Company

The Companies Act defines the procedure that governs the registration or incorporation of a Company which leads to formation or birth of a Company. It is highly significant to register a Company.

**HOW TO REGISTER A COMPANY IN INDIA**

* **Step 1:**

Digital Signature Certificate

* **Step 2:**

Director Identification Number

* **Step 3:**

Registration on the MCA Portal

* **Step 4:**

Certificate of Incorporation

|  |  |
| --- | --- |
| **DOCUMENTS REQUIRED FOR COMPANY** |  |
|  |  |
| **REGISTRATION :** |  |

|  |  |
| --- | --- |
| * Proof of identification of all the company’s directors and shareholders (partners in case of LLP). Any one of the below documents can be submitted as proof of identification: |  |

* Pan card
* Aadhar card
* Driving license
* Passport

* Proof of address
* Bank account statement having address

**BENEFITS OF COMPANY REGISTRATION IN INDIA**

A company registration provides many advantages. A licensed company makes it genuine and enhances the business’ credibility.

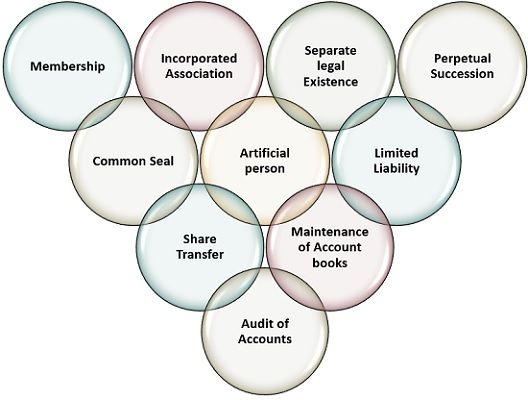
|  |
| --- |
| * Protects against personal obligation, and defends against other threats and losses. |
| * Builds goodwill and also supports more customer attraction |
| * Gives reliable investors bank credits and good investment with ease. |
| * Provides cover of the responsibility to protect the company’s assets |
| * Bigger commitment to wealth and greater stability |
| * Increases the ability to develop and grow large |

# TYPES OF COMPANIES IN INDIA :

There are six types of companies, each of them taking a different approach to the company registration process in India.

* Private Limited Company
* Public Limited Company
* One Person Company
* Producer Company
* Nidhi Company
* Section 8 Company

# FEATURES COMMON TO EVERY TYPE OF COMPANY

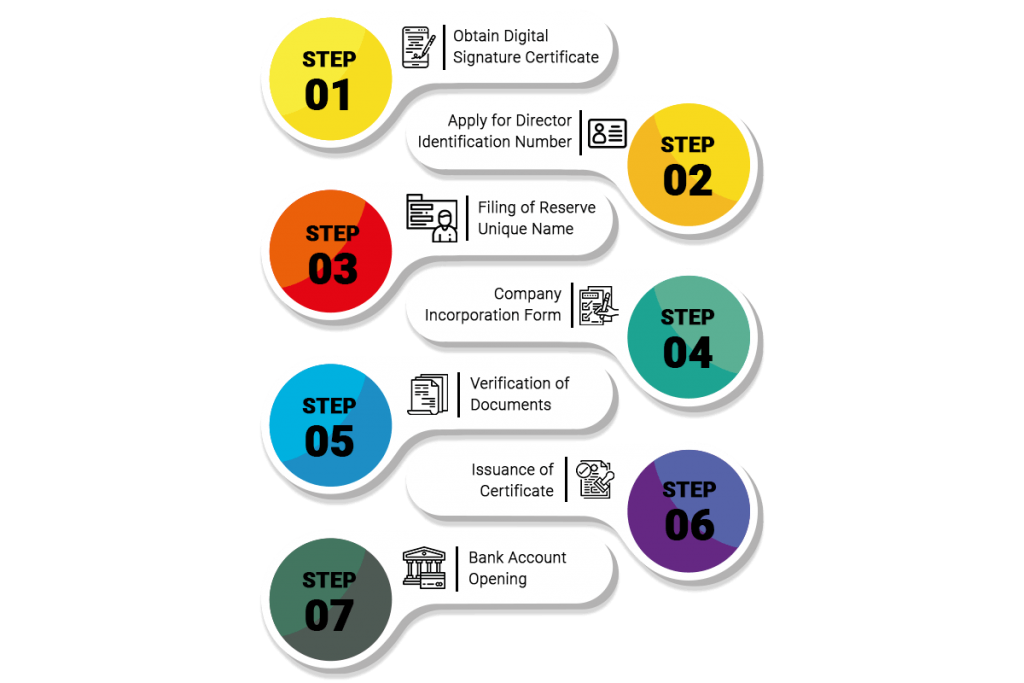


**FEATURE ENGINEERING:**

The ROCs also ensure that LLPs comply with the statutory requirements under the LLP Act. The office of the ROC maintains a registry of records related to Company registered with them, and permits the general public to access this data on payment of a fee. The Union Government maintains administrative Control over ROCs through Regional Directors. There are 7 Regional Directors, and they supervise the functioning of ROCs within their respective regions.

The Registrar of Companies takes care of company registration (also known as incorporation) in India, completes reporting and regulation of companies and their directors and shareholders, and also oversees government reporting of various matters including the annual filling of various documents.

**TYPE OF COMPANY REGISTRATION IN INDIA**



* **Step 1:** **Project Setup**

Before starting the analysis, you need to set up your project. This involves creating a directory structure and installing necessary libraries. You can use a virtual environment to manage dependencies.

```bash

# Create a virtual environment (optional but recommended)

python -m venv roc-analysis-env

# Activate the virtual environment (replace 'source' with 'activate' on Windows)

source roc-analysis-env/bin/activate

# Install required libraries

pip install pandas matplotlib

```

* **Step 2**: **Data Loading**

Assuming you have a CSV file (`roc\_data.csv`) containing RoC data, you can load it into your Python script:

```python

import pandas as pd

# Load the RoC data from a CSV file

data = pd.read\_csv('roc\_data.csv')

# Display the first few rows of the dataset

print(data.head())

```

* **Step 3: Basic Data Exploration (EDA)**

Let's perform some basic data exploration to get a feel for the data. This includes examining data statistics, understanding data types, and visualizing some basic information.

```python

# Get an overview of the data

print(data.info())

# Summary statistics

print(data.describe())

# Count the number of unique companies

num\_unique\_companies = data['Company Name'].nunique()

print(f"Number of unique companies: {num\_unique\_companies}")

# Visualize a histogram of company sizes

import matplotlib.pyplot as plt

plt.hist(data['Company Size'], bins=20)

plt.xlabel('Company Size')

plt.ylabel('Count')

plt.title('Distribution of Company Sizes')

plt.show()

```

At this point, you've set up your project, loaded the RoC data, and performed some basic data exploration. Next, you can continue to Part 2, where you'll dive deeper into data analysis, including feature engineering, more advanced visualizations, and specific analyses based on your research goals.

Remember to replace `'roc\_data.csv'` with the actual path to your RoC dataset, and adapt the analysis to your specific data and objectives.

* **Step 4: Feature Engineering**

Feature engineering involves creating new features or transforming existing ones to extract more valuable information. Depending on your specific goals, you might want to calculate financial ratios, create date-related features, or encode categorical variables.

Here's an example of calculating a simple financial ratio, Return on Assets (ROA):

```python

# Calculate Return on Assets (ROA) and add it as a new feature

data['ROA'] = data['Net Income'] / data['Total Assets']

**```**

* **Step 5: Advanced Data Visualization**

Create more advanced visualizations to gain deeper insights into the data. Libraries like Seaborn or Plotly can help you create visually appealing plots.

For example, you can create a pair plot to visualize relationships between different numerical variables:

```python

import seaborn as sns

sns.pairplot(data, vars=['Revenue', 'Net Income', 'Total Assets'])

plt.show()

```

* **Step 6**: **Specific Analyses**

Perform analyses tailored to your research objectives. Here are a few examples:

**- Company Demographics:** Analyze the distribution of companies by industry or location.

```python

# Plot a bar chart of the top 10 industries in the dataset

industry\_counts = data['Industry'].value\_counts().head(10)

industry\_counts.plot(kind='bar')

plt.xlabel('Industry')

plt.ylabel('Count')

plt.title('Top 10 Industries')

plt.show()

```

- **Financial Analysis:** Calculate financial metrics and visualize them.

```python

# Calculate and visualize average ROA by industry

average\_roa\_by\_industry = data.groupby('Industry')['ROA'].mean().sort\_values(ascending=False)

average\_roa\_by\_industry.plot(kind='bar')

plt.xlabel('Industry')

plt.ylabel('Average ROA')

plt.title('Average ROA by Industry')

plt.show()

```

**Predictive Modeling :** If your goal is to build predictive models (e.g., predicting bankruptcy), you can use machine learning libraries like scikit-learn to train and evaluate models.

* **Step 7: Interpretation and Reporting**

Once you've conducted your analyses, it's crucial to interpret your findings and create a report or presentation to communicate your results effectively. This might involve summarizing key insights, creating dashboards, or generating reports using tools like Jupiter notebooks, RMarkdown, or data visualization libraries.

Remember to adapt your analysis to your specific ROC data and research questions. Each analysis will be unique based on your objectives. Continue refining your analysis by exploring more advanced statistical methods or machine learning techniques if needed.

Lastly, ensure that your analysis follows ethical and legal considerations, especially when working with sensitive data. Data privacy and security are essential when dealing with real-world data, especially from government agencies like the Registrar of Companies.

## **Exploratory data analysis for a three-class ROC marker**

### 

### **Description**

A function that investigates data that arose from a single marker and contains the reference standard of the three classes "healthy", "intermediate" and "diseased".

### **Usage**

roc.eda(

x,

y,

z,

dat = NULL,

type = c("empirical", "trinormal"),

plotVUS = FALSE,

saveVUS = FALSE,

sep.dens = FALSE,

scatter = FALSE,

conf.level = 0.95,

n.boot = 1000,

verbose = TRUE,

alternative = c("two.sided", "less", "greater")

)

### **Examples**

### data(krebs)

# empirical EDA:

roc.eda(dat = krebs[,c(1,5)], type = "e", plotVUS = FALSE)

# equal data input via:

x <- with(krebs, krebs[trueClass=="healthy", 5])

y <- with(krebs, krebs[trueClass=="intermediate", 5])

z <- with(krebs, krebs[trueClass=="diseased", 5])

roc.eda(x, y, z, type = "e", sep.dens = TRUE)

data(cancer)

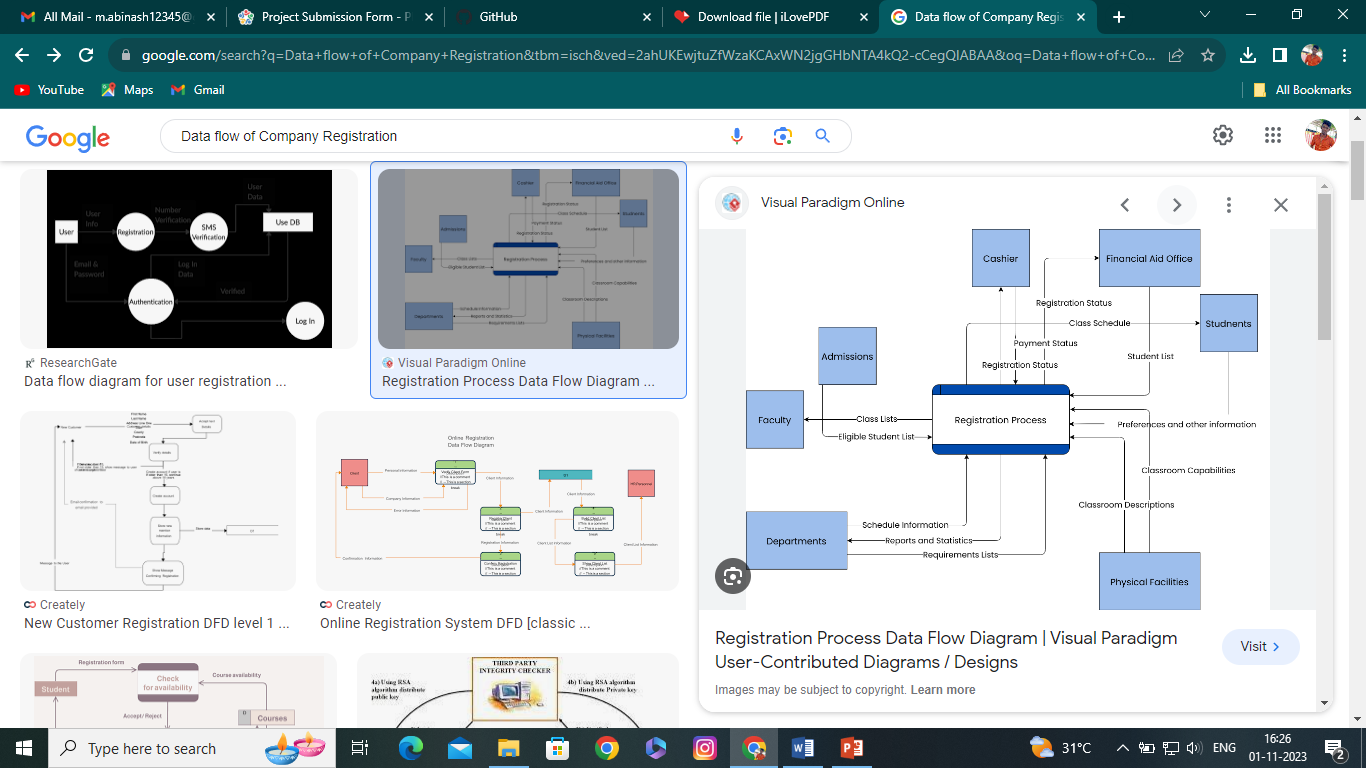
# trinormal EDA:

roc.eda(dat = cancer[,c(1,10)], type = "trin", plotVUS = FALSE)

# trinormal EDA with different plots:

roc.eda(dat = cancer[,c(1,5)], type = "t", sep.dens = TRUE, scatter = TRUE)

**Architecture of Roc**



**YEAR : III YEAR   
DEP : CSE/ECE  
KARAIKUDI INSTITUTE OF TECHNOLOGY & KARAIKUDI INSTITUTE OF MANAGEMENT,  
KIRANIPPATTI , SIVAGANGAI (DT),  
PIN : 630 307.**

**GUIDED BY :**

**Mr. MUTHURAMU , M.E.,**

**HEAD OF THE DEPARTMENT,**

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING,**

**KARAIKUDI INSTITUTE OF TECHNOLOGY & KARAIKUDI INSTITUTE OF MANAGEMENT**

**KIRANIPPATTI , SIVAGANGAI DISTRICT,**

**PIN : 630 307.**

**SUBMITTED BY**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.NO** | **NAME** | **REGISTER NUMBER** | **DEPARTMENT** |
| 1. | PL.KARUPPIAH | 910821106001 | ECE |
| 2. | M.ABINASH | 910821104001 | CSE |
| 3. | R.KISHORE KUMAR | 910821104005 | CSE |
| 4. | K.SARAN KUMAR | 910821104009 | CSE |
| 5. | T.VEERAMANI | 910821104014 | CSE |