

AI-Driven Exploration and prediction of company registration trends with registrar of companies(ROC)

510521205032:Rasmi S

Bharathidasan Engineering College

Project Title:ROC company analysis



Introduction:

In an era characterized by unprecedented growth in business and evolving regulatory landscapes, the Registrar of Companies (RoC) stands as a pivotal institution responsible for overseeing and regulating corporate entities. The RoC plays a crucial role in maintaining transparency, ensuring compliance, and fostering an environment conducive to economic development. As the volume of business registrations continues to surge, the need for

sophisticated analytical tools and predictive insights has never been more pressing.

The Registrar of Companies (ROC) plays a pivotal role in the governance and regulation of corporate entities within a country. Its mandate involves the registration, management, and oversight of businesses, ensuring compliance with statutory requirements and facilitating transparency in the corporate landscape. In this era of data-driven decision-making, conducting a comprehensive ROC Company Analysis emerges as a valuable and insightful exercise.

ROC Company Analysis involves the systematic examination of data and information provided by the Registrar of Companies pertaining to registered businesses. This data encompasses a wide array of details, including but not limited to, company registrations, financial statements, ownership structures, and compliance records. Through the prism of analytics and data science, this analysis endeavors to extract meaningful insights and predictions regarding company behaviors, industry trends, and regulatory compliance.

This project represents a fusion of technology, data science, and regulatory insight, offering a valuable resource for all stakeholders interested in the past, present, and future of company registrations within the purview of the Registrar of Companies.

Data Source:

The data source for ROC (Registrar of Companies) company analysis typically includes datasets and information provided by the Registrar of Companies or government authorities responsible for corporate registrations.

datasetLink:<https://tn.data.gov.in/resource/company-master-data-tamil-nadu-upto-28th-february-2019>

401	U1712720B05D03E SAEACTY	PdMc	Company linN-govt con-08-09-098	TarilN	5000000	3075000	1701	Manufacturer34 AWMAR0C0E0M0Esecretary03-03-2009	3-03-2009
402	U1712720B40500 REMACTY	PdMc	Company linN-govt con- 5/4/0984	TarilN	10000000	10000000	1701	Manufacturer11 Sndar0C0E0M0Einvestor03-03-2008	3-03-2008
403	U1712720B40500000 SACTY	PdMc	Company linN-govt con-08-02-087	TarilN	9750000	9750000	1701	Manufacturer1570AMAR0C0E0M0Einvestor03-03-2009	3-03-2009
404	U1712720B4050000000 SACTY	PdMc	Company linN-govt con- 10/7/0809	TarilN	16500000	157542000	1701	Manufacturer12520ETU0C0E0M0Einvestor03-03-2009	3-03-2009
405	U1712720B405000000000 SACTY	PdMc	Company linN-govt con-08-07-0809	TarilN	50000000	45323350	1701	Manufacturer12520ETU0C0E0M0Einvestor03-03-2009	3-03-2009
406	U1712720B40500000000000 SACTY	PdMc	Company linN-govt con- 9/2/0900	TarilN	2000000	19452000	1701	Manufacturer101MAY0C0E0M0Einvestor03-03-2002	3-03-2002
407	U1712720B405000000000000 SACTY	PdMc	Company linN-govt con- 10/3/0903	TarilN	10000000	6000000	1701	Manufacturer131BWA0C0E0M0Einvestor03-03-2009	3-03-2009
408	U1712720B4050000000000000 SACTY	PdMc	Company linN-govt con-08-08-0904	TarilN	6000000	5264000	1701	Manufacturer131BWA0C0E0M0Einvestor03-03-2009	3-03-2009
409	U1712720B40500000000000000 SACTY	PdMc	Company linN-govt con-08-03-0904	TarilN	10000000	47838000	1701	Manufacturer131BWA0C0E0M0Einvestor03-03-2009	3-03-2009
410	U1712720B405000000000000000 SACTY	PdMc	Company linN-govt con-09-03-0903	TarilN	55000000	36290300	1701	Manufacturer131BWA0C0E0M0Einvestor03-03-2009	3-03-2009
411	U1712720B4050000000000000000 SACTY	PdMc	Company linN-govt con-24-04-0905	TarilN	15000000	9648000	1701	Manufacturer131BWA0C0E0M0Einvestor03-03-2009	3-03-2009
412	U1712720B40500000000000000000 SACTY	PdMc	Company linN-govt con- 7/3/0900	TarilN	15000000	92894000	1701	Manufacturer131BWA0C0E0M0Einvestor03-03-2009	3-03-2009
413	U1712720B405000000000000000000 SACTY	PdMc	Company linN-govt con- 6/0/0908	TarilN	10000000	5725000	1701	Manufacturer131BWA0C0E0M0Einvestor03-03-2009	3-03-2009
414	U1712720B4050000000000000000000 SACTY	PdMc	Company linN-govt con-08-07-0809	TarilN	122E-00	1007E-00	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
415	U1712720B4050000000000000000000 SACTY	PdMc	Company linN-govt con-05-06-0902	TarilN	40000000	25937820	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
416	U1712720B40500000000000000000000 SACTY	PdMc	Company linN-govt con-02-08-0907	TarilN	20000000	1269000	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
417	U1712720B40500000000000000000000 SACTY	PdMc	Company linN-govt con-08-03-0904	TarilN	7000000	60439500	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
418	U1712720B405000000000000000000000 SACTY	PdMc	Company linN-govt con-02-04-2007	TarilN	1443E-00	1430E-00	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
419	U1712720B405000000000000000000000 SACTY	PdMc	Company linN-govt con- 4/4/0803	TarilN	30000000	3266000	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
420	U1712720B405000000000000000000000 SACTY	PdMc	Company linN-govt con-08-04-0808	TarilN	83000000	72872500	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
421	U1712720B4050000000000000000000000 SACTY	PdMc	Company linN-govt con-03-07-0809	TarilN	12000000	1796000	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
422	U1712720B4050000000000000000000000 SACTY	PdMc	Company linN-govt con-03-02-0906	TarilN	1500000	6074500	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
423	U1712720B40500000000000000000000000 SACTY	PdMc	Company linN-govt con- 2/1/0908	TarilN	1000000	2000	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
424	U1712720B405000000000000000000000000 SACTY	PdMc	Company linN-govt con- 3/7/0900	TarilN	30000000	2500000	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
425	U1712720B4050000000000000000000000000 SACTY	PdMc	Company linN-govt con-03-04-0904	TarilN	50000000	1502000	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
426	U1712720B40500000000000000000000000000 SACTY	PdMc	Company linN-govt con-09-02-0909	TarilN	12000000	75382000	1701	Manufacturer125 A MERO0C0E0M0Einvestor03-03-2009	3-03-2009
427	U1712720B405000000000000000000000000000 SACTY	PdMc	Company linN-govt con-08-08-2005	TarilN	47250000	25993000	1701	Manufacturer139 A EXTEND0C0E0M0Einvestor03-03-2009	3-03-2009
428	U1712720B4050000000000000000000000000000 SACTY	PdMc	Company linN-govt con-08-03-2001	TarilN	40000000	20922700	1801	Manufacturer1005 MEX0C0E0M0Einvestor03-03-2009	3-03-2009
429	U1802910B00M00								

301	906296	SARCHEM INACTV	NA	NA	NA	26-12-2017	Tamil Nadu	0	0	0	NA	Agriculture & NO 75 ADDLR0C00ENH	ravisarcheNA	NA	
302	906309	DEEP DRILLING FV	NA	NA	NA	16-02-2018	Tamil Nadu	0	0	0	NA	Agriculture & ANPRIYA GR0C00ENH	runagappanNA	NA	
303	906321	BRANDRI INACTV	NA	NA	NA	5/3/2018	Tamil Nadu	0	0	0	NA	Agriculture & No 5 / 329 LR0C00ENH	anandumeNA	NA	
304	906349	TANVAY CRUACTV	NA	NA	NA	18-06-2018	Tamil Nadu	0	0	0	NA	Agriculture & No No. GBRC0C00ENH	kingenon@NA	NA	
305	906371	SOLUTION FIACTV	NA	NA	NA	13-09-2018	Tamil Nadu	0	0	0	NA	Agriculture & DOR 0083 R0C00ENH	shree7@sk2NA	NA	
306	906384	EASTERN GLUACTV	NA	NA	NA	22-10-2018	Tamil Nadu	0	0	0	NA	Agriculture & SF NO 262 / 280C00ENH	avinachariNA	NA	
307	906401	TRANSMILE ACTV	NA	NA	NA	24-12-2018	Tamil Nadu	0	0	0	NA	Agriculture & No 83 A WIR0C00ENH	chandrasegNA	NA	
308	906418	Sunite TechACTV	NA	NA	NA	17-02-2019	Tamil Nadu	0	0	0	NA	Agriculture & The Royal GR0C00ENH	senbil_kumaNA	NA	
309	906441	HYDRONGHEACTV	NA	NA	NA	20-03-2019	Tamil Nadu	0	0	0	NA	Agriculture & Door No 78 RR0C00ENH	bcjeong@hyNA	NA	
310	906468	PODSUNOY ACTV	NA	NA	NA	17-06-2019	Tamil Nadu	0	0	0	NA	Agriculture & Site 36 DES0C00ENH	kouskivakuNA	NA	
311	906487	ThinkTechInACTV	NA	NA	NA	20-07-2019	Tamil Nadu	0	0	0	NA	Agriculture & Prince h0ciR0C00ENH	Vijay VenkataNA	NA	
312	0007TZ9403EEDAMALACTV	Public	Company	linNon-govt	con2	04-1943	Tamil Nadu	12500000	6273500			III7 Agriculture & KATARY ESTRO0C00ENH	MBAssecr@na3-03-2019	3-03-2019	
313	0009TN9804WAN OFFSHACTV	Public	Company	linNon-govt	con25	09-1986	Tamil Nadu	15E+10	100730000			III9 Agriculture & ANPRIYA GR0C00ENH	secretari@na3-03-2019	3-03-2019	
314	0009TN98060ASUMIARRU00	Public	Company	linNon-govt	con29	03-1989	Tamil Nadu	150000000	38350000			III9 Agriculture & 29 POLICE CR0C00ENH	NA	NA	
315	0009TN99250FTECHINFNACTV	Public	Company	linNon-govt	con2	12-1992	Tamil Nadu	150000000	40000000			III9 Agriculture & 29 FREGS0R0C00ENH	compliance@na3-03-2018	3-03-2018	
316	0009TN9930FLORENCE MU00	Public	Company	linNon-govt	con30	11-1993	Tamil Nadu	500000	0			III9 Agriculture & 29 EAST CAR0C00ENH	NA	NA	
317	0009TN993000CHRAMINACTV	Public	Company	linNon-govt	con	4/5/1995	Tamil Nadu	250000000	189060750			III22 Agriculture & 102 SATHYR0C00ENH	MBAnmreddy@na3-03-2019	3-03-2019	
318	0009TN9929EPC TEA GAMAL	Public	Company	linNon-govt	con	3/4/1992	Tamil Nadu	150000000	0			III32 Agriculture & 36 WALLAHAR0C00ENH	NA	NA	
319	00032TN9930REEMACH THE00	Public	Company	linNon-govt	con7	09-1993	Tamil Nadu	5000000	0			III32 Agriculture & 753 MKNT R0C00ENH	NA	NA	
320	00032TZ922THE UNITED NACTV	Public	Company	linNon-govt	con	9/8/1922	Tamil Nadu	5000000	49965660			III32 Agriculture & 3 SAMTHIRI SR0C00ENH	MBAshead@na3-03-2019	3-03-2019	
321	00032TZ940THE NUKOTAMAL	Public	Company	linNon-govt	con28	07-1948	Tamil Nadu	10000000	1040130			III32 Agriculture & 6ALOTHAGRI0C00ENH	MBAsWolath@na3-03-2019	3-03-2019	
322	00033TN99400NEER CASACTV	Public	Company	linNon-govt	con	9/11/1994	Tamil Nadu	55000000	40242000			III33 Agriculture & 300 POONWIR0C00ENH	cha@na3-03-2018	3-03-2018	
323	00020TN9804MIN AGRI00	Public	Company	linNon-govt	con29	08-1989	Tamil Nadu	8000000	30			III20 Agriculture & 3 E AU TOWER0C00ENH	NA	NA	
324	00022TZ9930M EGG RACTV	Public	Company	linNon-govt	con	6/4/1995	Tamil Nadu	300000000	203300000			III22 Agriculture & NO 133 LR0C00ENH	MBAs@na3-03-2019	3-03-2019	
325	00049TZ990EPC AQ0 FIACTV	Public	Company	linNon-govt	con3	13-08-1991	Tamil Nadu	400000000	204715000			III409 Agriculture & NO 1078 TRIR0C00ENH	MBAssecr@na3-03-2019	3-03-2019	
326	00542TZ9804ASHREE SUACTV	Public	Company	linNon-govt	con3	12-1985	Tamil Nadu	300000000	28076800			III542 Agriculture & 6AFF2 338/R0C00ENH	MBAsnaryan@na3-03-2019	3-03-2019	
327	00230TZ99023002HFNACTV	Public	Company	linNon-govt	con23	04-1993	Tamil Nadu	4000000	34550000			III230 Agriculture & SF NO 108/2R0C00ENH	MBAs@na3-03-2017	3-03-2017	
328	00250TZ9802L R00BEEAMAL	Public	Company	linNon-govt	con3	05-1988	Tamil Nadu	135300000	107000000			III250 Agriculture & 2000 TRICHIR0C00ENH	MBAs@na3-03-2019	3-03-2019	
329	00250TZ9902TRAVENOREAMAL	Public	Company	linNon-govt	con	2/11/1998	Tamil Nadu	100000000	114525000			III250 Agriculture & 2000 TRICHIR0C00ENH	MBAssecr@na3-03-2019	3-03-2019	
330	00320TZ9802SALAZER ELEACTV	Public	Company	linNon-govt	con	8/1/1985	Tamil Nadu	200000000	150827370			III320 Agriculture & 5AMR0C00ENH	MBAs@na3-03-2019	3-03-2019	
331	05000TN9804M MARENACTV	Public	Company	linNon-govt	con	6/6/1985	Tamil Nadu	100000000	49348000			III500 Agriculture & 4Door No47/2R0C00ENH	ag@na3-03-2019	3-03-2019	
332	05000TN9904NAGARINA MACTV	Public	Company	linNon-govt	con	4/8/1993	Tamil Nadu	100000	7000			III500 Agriculture & 32 FIRST R0C00ENH	NA	NA	
333	05000TN9904TRIDICAL GRUACTV	Public	Company	linNon-govt	con25	08-1994	Tamil Nadu	70000000	7000			III500 Agriculture & NO55 TEACHIR0C00ENH	NA	NA	
334	08030TZ9802TAMIL NDU0FNEF	Public	Company	linNon-govt	con24	09-1980	Tamil Nadu	70000000	0			III803 Agriculture & 08080VA TOR0C00ENH	NA	NA	
335	III00TN9906SOUTHERNACTV	Public	Company	linNon-govt	con18	12-1949	Tamil Nadu	425E+09	210E+09			III00 Mining and (S)CH00SE R0C00ENH	nb@na3-03-2019	3-03-2019	
336	III00TN9806KOTHU01ACTV	Public	Company	linNon-govt	con28	04-1989	Tamil Nadu	600000000	588464000			III00 Mining and 6KOTHU01BLR0C00ENH	sec@na3-03-2019	3-03-2019	
337	III3200TN9806HIMALAYA GRUACTV	Public	Company	linNon-govt	con	09/09/1992	Tamil Nadu	40000000	2367840			III3206 Mining and 0PANCHAM R0C00ENH	investor@na3-03-2019	3-03-2019	
338	III4200TN9909CRAZY INF0ACTV	Public	Company	linNon-govt	con	9/9/1992	Tamil Nadu	100000000	6688000			III4200 Mining and 0LOT NO5 SR0C00ENH	cr@na3-03-2017	3-03-2017	
339	III4200TN9802DALMA BHARUACTV	Public	Company	linNon-govt	con	12/7/2013	Tamil Nadu	370E+09	38597706			III4200 Mining and 0dahi@na3-03-2019	na3-03-2019	3-03-2019	
340	III4200TN9802TALAVH0100	Public	Company	linNon-govt	con	3/3/1989	Tamil Nadu	50000000	47764000			III4209 Mining and 0No2/367 KAR0C00ENH	na3-03-2019	3-03-2019	
341	III4200TN9904EE GEE GRUACTV	Public	Company	linNon-govt	con30	03-1990	Tamil Nadu	100000000	45000000	14299		III4209 Mining and 023 G00ge R0C00ENH	na3-03-2019	3-03-2019	
342	III4200TN9904UN GRANTIACTV	Public	Company	linNon-govt	con29	10-1993	Tamil Nadu	100000000	90000000			III4209 Mining and 048 HARRINGR0C00ENH	na3-03-2019	3-03-2019	
343	III4200TN9904EUD0 INDUSTU00	Public	Company	linNon-govt	con	9/5/1994	Tamil Nadu	10000000	70			III4209 Mining and 026 COMMANDR0C00ENH	NA	NA	
344	III5000TN9802RANJANUMACTV	Public	Company	linNon-govt	con	3/6/1982	Tamil Nadu	50000000	36488500			III5000 Manufacturin06 SEMBDR0C00ENH	na3-03-2019	3-03-2019	
345	III5000TN9902AMENRA GACTV	Public	Company	linNon-govt	con7	11-1992	Tamil Nadu	50000000	42800000			III5000 Manufacturin04/9 MedayR0C00ENH	na3-03-2019	3-03-2019	
346	III542TN9804NTARUACTV	Public	Company	linNon-govt	con	12/3/1986	Tamil Nadu	190000000	152284000			III542 Manufacturin06080UDAR0C00ENH	na3-03-2019	3-03-2019	
347	III542TN9904ASHANA AGUACTV	Public	Company	linNon-govt	con18	06-1990	Tamil Nadu	50000000	45859500			III542 Manufacturin06 792/5 EAR0C00ENH	na3-03-2019	3-03-2019	
348	III531TN9703ENGALAKSHAMAL	Public	Company	linNon-govt	con	11/7/1973	Tamil Nadu	1000000	0			III531 Manufacturin01 ANNA DILLAR0C00ENH	NA	NA	
349	III534TN9806K0MPATTU LACTV	Public	Company	linNon-govt	con16	12-1991	Tamil Nadu	150000000	5544760			III534 Manufacturin075/8 BENR0C00ENH	na3-03-2019	3-03-2019	
350	III542TN9404MVC AGRI00	Public	Company	linNon-govt	con	8/5/1944	Tamil Nadu	400000000	0			III542 Manufacturin099 ARANAR0C00ENH	NA	NA	
351	III542TN9404PETTAVITU00	Public	Company	linNon-govt	con20	02-1948	Tamil Nadu	440000000	0			III542 Manufacturin0234 NSC BR0C00ENH	NA	NA	
352	III542TN9503THROU0000ACTV	Public	Company	linNon-govt	con	12/7/1954	Tamil Nadu	500000000	11370000			III542 Manufacturin0ELDORAD0R0C00ENH	na3-03-2019	3-03-2019	
353	III542TN9604RINA HOTIACTV	Public	Company	linNon-govt	con	9/9/1960	Tamil Nadu	750000000	12000000			III542 Manufacturin0RINA CENTR0C00ENH	na3-03-2019	3-03-2019	
354	III542TN9604KOTHU01 SU0ACTV	Public	Company	linNon-govt	con	7/11/1960	Tamil Nadu	144E+09	82888590			III542 Manufacturin0KOTHU01BLR0C00ENH	na3-03-2019	3-03-2019	
355	III542TN9802PONN SUGRU00	Public	Company	linNon-govt	con25	02-1982	Tamil Nadu	250000000	0			III542 Manufacturin099 NNGWR0C00ENH	NA	NA	
356	III542TN9802DHARAN SUGRUACTV	Public	Company	linNon-govt	con	4/6/1987	Tamil Nadu	600000000	332000000			III542 Manufacturin0P HONE SR0C00ENH	na3-03-2019	3-03-2019	
357	III542TN9903CP SUGR UACTV	Public	Company	linNon-govt	con	09/09/1992	Tamil Nadu	250000000	113385050			III542 Manufacturin0RAMAR0C00ENH	na3-03-2019	3-03-2019	
358	III542TN9804SKITH SUGUACTV	Public	Company	linNon-govt	con	12/5/1981	Tamil Nadu	17E+09	1188E+09			III542 Manufacturin0SKITH NGR0C00ENH	na3-03-2019	3-03-2019	
359	III542TN9803BANNEI AMIACTV	Public	Company	linNon-govt	con	1/12/1983	Tamil Nadu	650000000	125397000			III542 Manufacturin022 TRICHY R0C00ENH	na3-03-2019	3-03-2019	
360	III542TN9904PONN SUGRUACTV	Public	Company	linNon-govt	con26	12-1996	Tamil Nadu	150000000	8598480			III5422 Manufacturin0SMN H0KSE0C00ENH	na3-03-2019	3-03-2019	
361	III5492TN9804KOTHU01 GENUS0	Public	Company	linNon-govt	con	6/7/1983	Tamil Nadu	10000000	80			III5492 Manufacturin04NNGWR0C00ENH	NA	NA	
362	III5499TN9804HATSNAGUACTV	Public	Company	linNon-govt	con	4/3/1986	Tamil Nadu	300000000	15809222			III5499 Manufacturin0DANNE DOR0C00ENH	na3-03-2019	3-03-2019	
363	III5499TN9904EPC AQ0 FIACTV	Public	Company	linNon-govt	con3	13-08-1991	Tamil Nadu	10000000	0			III5499 Manufacturin0WATERLY ENR0C00ENH	NA	NA	
364	III5499TN9904SAPATASUACTV	Public	Company	linNon-govt	con7	02-1992	Tamil Nadu	300000000	34039420			III5499 Manufacturin0PADALAM R0C00ENH	na3-03-2019	3-03-2019	
365	III5500TN9803EMPEE DISTILACTV	Public	Company	linNon-govt	con05	09-1983	Tamil Nadu	300000000	201757530			III5500 Manufacturin0EMPEE TOMR0C00ENH	na3-03-2019	3-03-2019	
366	III5520TN9803BALAI DISTILAMAL	Public	Company	linNon-govt	con15	12-1983	Tamil Nadu	222E+09	143E+09			III5520 Manufacturin0EYE PASS R0C00ENH	na3-03-2019	3-03-2019	
367	III5540TN9904TEASSU AVIACTV	Public	Company	linNon-govt	con20	09-1994	Tamil Nadu	250000000	217866000			III5549 Manufacturin0NEW N03 FR0C00ENH	na3-03-2019	3-03-2019	
368	III5540TN9904CEANABOACTV	Public	Company	linNon-govt	con28	10-2005	Tamil Nadu	150000000	135652750			III5549 Manufacturin0N04B-1 EASTR0C00ENH	na3-03-2019	3-03-2019	
369	III7000TN9905V GLOBE MACTV	Public	Company	linNon-govt	con30	10-2007	Tamil Nadu	232500000	9044850			III7000 Manufacturin0DOR N03/1R0C00ENH	na3-03-2019	3-03-2019	
370	III7000TN9906RAMPALAYACTV	Public	Company	linNon-govt	con24	02-1936	Tamil Nadu	150000000	7376060			III7000 Manufacturin0RAMPALAYAR0C00ENH	na3-03-2019	3-03-2019	
371	III7000TN9906RAMARUACTV	Public	Company	linNon-govt	con20	02-1939	Tamil Nadu	50000000	3946560			III7000 Manufacturin0PAC RAMASR0C00ENH	na3-03-2019	3-03-2019	
372	III7000TN9906LOVAL TEXTILACTV	Public	Company	linNon-govt	con	9/4/1946	Tamil Nadu	150000000	4864460			III7000 Manufacturin02/4 MI SHR0C00ENH	na3-03-2019	3-03-2019	
373	III7000TN9906SRI GANPATIACTV	Public	Company	linNon-govt	con	11/9/1946	Tamil Nadu	100000000	66704000			III7000 Manufacturin0SHANER NAR0C00ENH	na3-03-2019	3-03-2019	
374	III7000TN9906VIM LIMITEDACTV	Public	Company	linNon-govt	con27	07-1946	Tamil Nadu	100000000	40228000			III7000 Manufacturin0SUKARAMR0C00ENH	na3-03-2019	3-03-2019	
375	III7000TN9906SRI EEN00	Public	Company	linNon-govt	con	1/									

Data Collection and Preprocessing:

Data Collection:

- Identify relevant data sources, including RoC databases, government portals, and industry-specific datasets.
- Obtain necessary permissions and access rights to collect the data.
- Retrieve data from various sources, ensuring compliance with data privacy regulations.

Data Preprocessing:

- Clean the data by handling missing values, outliers, and duplicates.
- Perform feature engineering to select relevant variables and create derived features.
- Normalize or scale numerical data and encode categorical variables.
- Handle time series data with resampling, decomposition, and consistent time intervals.
- Split the data into training, validation, and testing sets for modeling.
- Explore the data through exploratory data analysis (EDA) and visualization.
- Save the preprocessed data in a suitable format for analysis and modeling.

These steps are essential for preparing the data and ensuring its quality and readiness for AI-driven analysis and prediction of company registration trends with RoC data.

Exploratory Data Analysis (EDA):

- Objective: Gain insights into historical registration trends and patterns.
- Activities:
 - Create visualizations (time series plots, histograms, maps).

- Conduct statistical analysis (seasonality, trends, anomalies).
- Explore correlations and relationships.
- Hypothesis testing (if relevant).

Feature Engineering:

- Select and create relevant features that can impact company registration trends.
- Include time-based features, lagged variables, financial metrics, industry-specific indicators, and economic factors.
- Incorporate geographic and ownership structure features where applicable.
- Utilize natural language processing (NLP) for textual data analysis.
- Ensure appropriate scaling and normalization of features.
- Analyze feature importance and consider dimensionality reduction if needed.
- Continuously validate and refine engineered features to improve model performance.

Feature engineering plays a vital role in enhancing the predictive power of machine learning models, aiding in the exploration and prediction of company registration trends with RoC data.

Advanced AI Algorithms:

Time Series Forecasting: Problem Domain: Time series forecasting is used when you need to make predictions based on historical data points that are ordered by time.

- Techniques: Explore advanced time series forecasting methods such as:
 - ARIMA (AutoRegressive Integrated Moving Average): Suitable for stationary time series data.
 - Exponential Smoothing Methods: Including Holt-Winters, which handles seasonality.
 - Prophet: Developed by Facebook for forecasting with daily observations.
 - Deep Learning: Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are powerful for complex time series patterns.
- Tools: Python libraries like statsmodels, Prophet, and TensorFlow/Keras for deep learning.
- Time Series split train test(single)

```
# Split train-test set manually
```

```
train_data, test_data = df_close[:int(len(df_close)*0.9)],  
df_close[int(len(df_close)*0.9):]
```

```
# Plot train-test using matplotlib.pyplot
```

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

```
plt.grid(True)
```

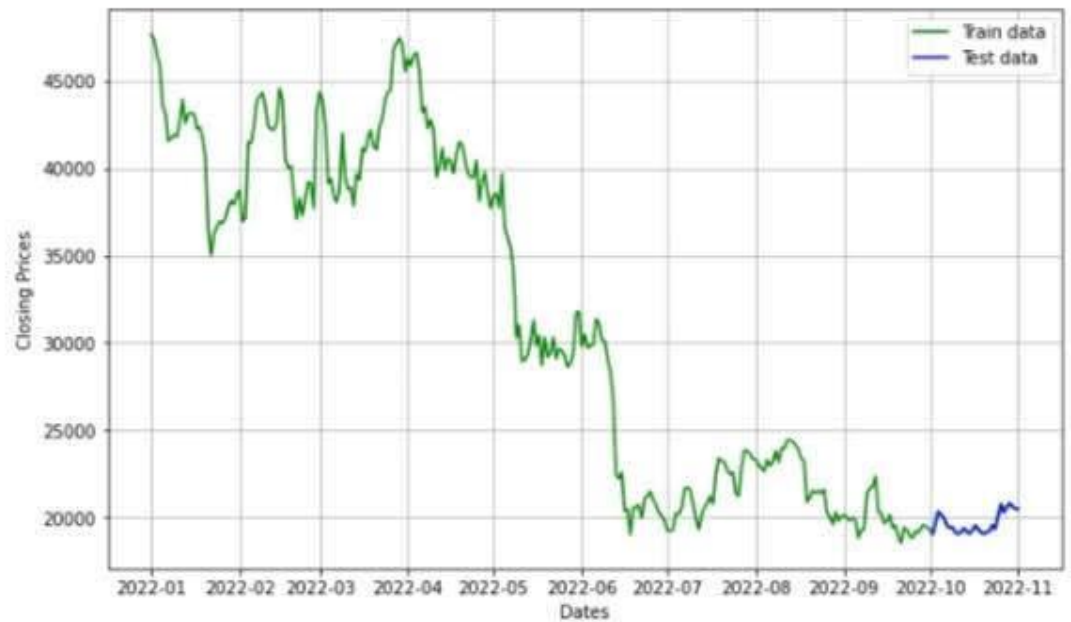
```
plt.xlabel('Dates')
```

```
plt.ylabel('Closing Prices')
```

```
plt.plot(df_close, 'green', label='Train data')
```

```
plt.plot(test_data, 'blue', label='Test data')
```

```
plt.legend()
```



Time Series split train test(multiple)

```
# Extract Closing values and Create TimeSeriesSplit splits
```

```
X = df_close.values
```

```
splits = TimeSeriesSplit(n_splits=3)
```

```
plt.figure(1)
```

```
index = 1
```

```
# Loop over splits to update train-test splits and models
```



```
for train_index, test_index in splits.split(X):

    # Create and update train-test splits

    train = X[train_index]

    test = X[test_index]

    print('Observations: %d' % (len(train) + len(test)))

    print('Training Observations: %d' % (len(train)))

    print('Testing Observations: %d' % (len(test)))

    # Enter model and evaluation technique here

    # Subplots for each train-test splits

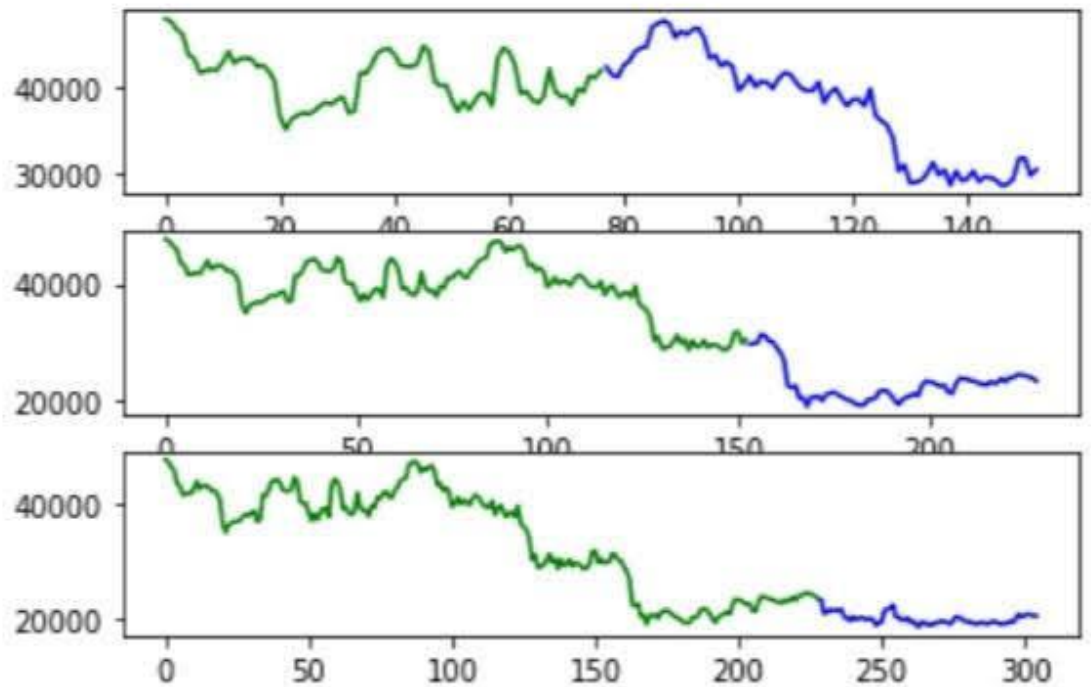
    plt.subplot(310 + index)

    plt.plot(train,'green')

    plt.plot([None for i in train] + [x for x in test],'blue')

    index += 1
```

plt.show()



- **Ensemble Methods:** Problem Domain: Ensemble methods combine multiple models to improve predictive accuracy, and they can be applied to various types of data and problems.
 - Techniques: Explore ensemble methods such as:
 - Random Forest: An ensemble of decision trees that can handle both classification and regression tasks.
 - Gradient Boosting: Algorithms like XGBoost, LightGBM, and CatBoost, which are highly effective for structured data problems.
 - Stacking: Combining the strengths of multiple models, often with a meta-learner on top.

- Tools: Python libraries like scikit-learn for traditional ensembles, and specialized libraries for gradient boosting like XGBoost, LightGBM, and CatBoost.

```
In[1] import pandas as pd
```

```
import numpy as np
```

```
In[2]
```

```
df=pd.read_csv('../input/all-indian-companies-registration-data-1900-2019/registered_companies.csv')
```

```
In[3] len(df)
```

```
Out[4]:
```

```
1992170
```

```
In [5]:
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 1992170 entries, 0 to 1992169
```

```
Data columns (total 17 columns):
```

#	Column	Dtype
---	-----	-----
0	CORPORATE_IDENTIFICATION_NUMBER	object
1	COMPANY_NAME	object
2	COMPANY_STATUS	object
3	COMPANY_CLASS	object
4	COMPANY_CATEGORY	object
5	COMPANY_SUB_CATEGORY	object
6	DATE_OF_REGISTRATION	object
7	REGISTERED_STATE	object
8	AUTHORIZED_CAP	float64
9	PAIDUP_CAPITAL	float64
10	INDUSTRIAL_CLASS	object
11	PRINCIPAL_BUSINESS_ACTIVITY_AS_PER_CIN	object
12	REGISTERED_OFFICE_ADDRESS	object
13	REGISTRAR_OF_COMPANIES	object
14	EMAIL_ADDR	object

```
15  LATEST_YEAR_ANNUAL_RETURN      object
```

```
16  LATEST_YEAR_FINANCIAL_STATEMENT  object
```

```
dtypes: float64(2), object(15)
```

```
memory usage: 258.4+ MB
```

```
df.describe().transpose()
```

```
Out[6]:
```

	count	mean	std	min	25%	50%	75%	max
AUTHORIZED_CAP	1992170.0	4.238508e+07	2.960562e+09	0.0	100000.0	500000.0	1500000.0	1.850000e+12
PAIDUP_CAPITAL	1992170.0	2.434621e+07	2.313154e+09	0.0	100000.0	100000.0	502000.0	1.699613e+12

```
In [7]:
```

```
df['DATE_OF_REGISTRATION']
```

```
Out[7]:
```

```
0      NaN
```


1 16-07-1998

2 NaN

3 25-06-2001

4 25-07-2001

...

1992165 30-06-2000

1992166 30-06-2000

1992167 25-07-2000

1992168 09-07-1998

1992169 29-02-2016

Name: DATE_OF_REGISTRATION, Length: 1992170, dtype: object

In [8]:

```
df['DATE_OF_REGISTRATION'].isnull().sum()
```

Out[8]:

2525

In [9]:

```
len(df)-df['DATE_OF_REGISTRATION'].count()
```

```
Out[9]:
```

```
2525
```

```
In [10]:
```

```
df['DATE_OF_REGISTRATION'].dropna(axis=0,inplace=True)
```

```
In [11]:
```

```
df['DATE_OF_REGISTRATION'].count()
```

```
Out[11]:
```

```
1989645
```

```
In [12]:
```

```
df['DATE_OF_REGISTRATION']=pd.to_datetime(df['DATE_OF_REGISTRATION'],  
errors = 'coerce', format='%d-%m-%Y')
```

```
In [13]:
```

```
df['YEAR_REG']=pd.DatetimeIndex(df['DATE_OF_REGISTRATION']).year
```

```
In [14]: number_of_companies =  
df[['YEAR_REG', 'CORPORATE_IDENTIFICATION_NUMBER']]
```

```
In [15]:
```

```
number_of_companies.dropna(axis=0, inplace=True)
```

```
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:1:
```

```
SettingWithCopyWarning:
```

```
A value is trying to be set on a copy of a slice from a DataFrame
```

See the caveats in the documentation:

https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
"""Entry point for launching an IPython kernel.
```

```
In [16]:
```

```
number_of_companies=number_of_companies.groupby(by='YEAR_REG').size().reset_index(name='No_of_companies')
```

```
In [17]:
```

```
number_of_companies
```

Out[17]:

	YEAR_REG	No_of_companies
0	1857.0	1
1	1863.0	3
2	1871.0	3
3	1872.0	3
4	1873.0	3
...
146	2016.0	93876
147	2017.0	107635

148	2018.0	117924
149	2019.0	128658
150	2020.0	12747

151 rows × 2 columns

In [18]:

```
import seaborn as sns
```

```
import matplotlib.pyplot as plt
```

```
%matplotlib inline
```

```
Out[6]:plt.figure(figsize=(30,10))
```

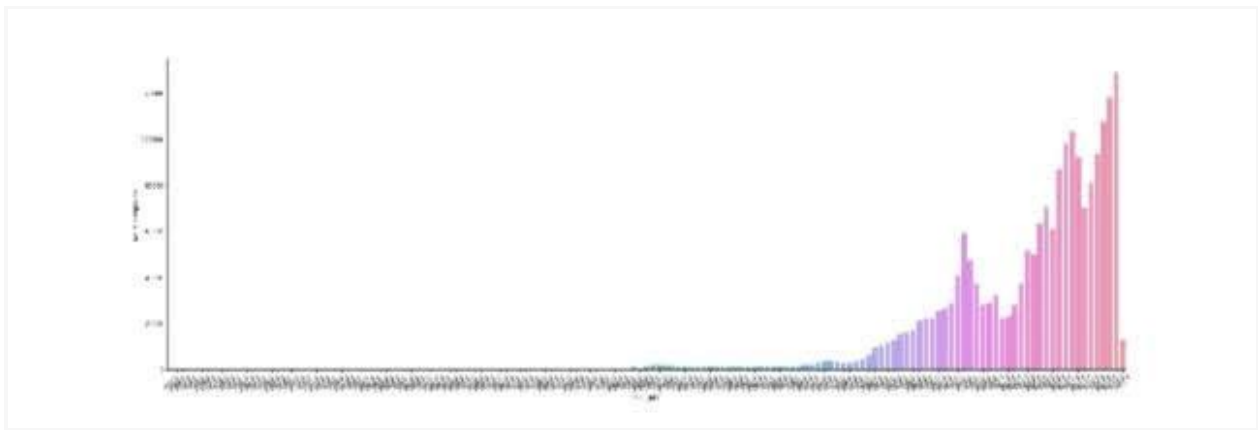


```
plt.xticks(rotation=45)
```

```
sns.barplot(data=number_of_companies,x='YEAR_REG',y='No_of_companies')
```

Out[19]:

<matplotlib.axes._subplots.AxesSubplot at 0x7f2c8c5c6190>



In [20]:

```
no_of_companies_state =  
df[['YEAR_REG', 'CORPORATE_IDENTIFICATION_NUMBER', 'REGISTERED_STATE']]
```

In [21]:

```
no_of_companies_state.dropna(axis=0,inplace=True)
```

/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:1:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation:

https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
"""Entry point for launching an IPython kernel.
```

In [22]:

```
no_of_companies_state=no_of_companies_state.groupby(by=['YEAR_REG', 'REGISTERED_STATE']).size().reset_index(name='no_of_companies')
```

In [23]:

```
no_of_companies_state
```

Out[23]:

	YEAR_REG	REGISTERED_STATE	no_of_companies
0	1857.0	Rajasthan	1
1	1863.0	Maharashtra	1

2	1863.0	West Bengal	2
3	1871.0	Maharashtra	2
4	1871.0	West Bengal	1
...
2931	2020.0	Telangana	788
2932	2020.0	Tripura	12
2933	2020.0	Uttar Pradesh	1260
2934	2020.0	Uttaranchal	138
2935	2020.0	West Bengal	606

2936 rows × 3 columns

```
plt.figure(figsize=(30,10))
```

```
plt.xticks(rotation=45)
```

```
sns.barplot(data=no_of_companies_state, x='YEAR_REG', y='no_of_companies', hue='REGISTERED_STATE')
```

Out[24]:

<matplotlib.axes._subplots.AxesSubplot at 0x7f2c8bdf9d90>



In [25]:

```
state='Tamil Nadu'
```

```
plot_data =
```

```
no_of_companies_state[no_of_companies_state['REGISTERED_STATE']==state]
```

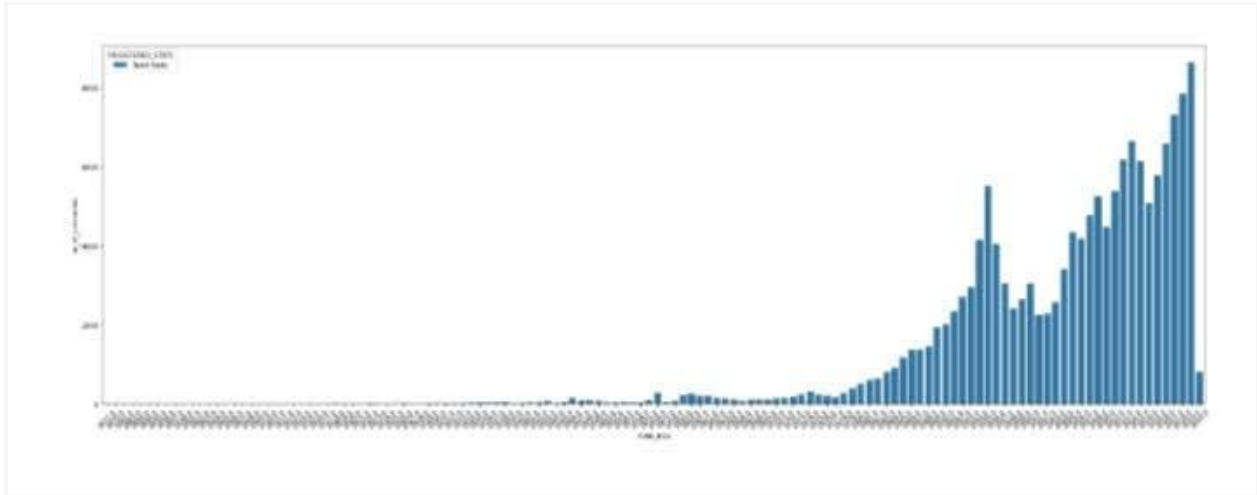
```
plt.figure(figsize=(30,10))
```

```
plt.xticks(rotation=45)
```

```
sns.barplot(data=plot_data, x='YEAR_REG', y='no_of_companies', hue='REGISTERED_STATE')
```

Out[25]:

<matplotlib.axes._subplots.AxesSubplot at 0x7f2c83b0ba50>



Ensemble methods:

1. Bagging (Bootstrap Aggregating):

- Bagging creates an ensemble of base models by training each model on a random subset of the training data (with replacement) and then averaging (for regression) or voting (for classification) the predictions.
- The Random Forest algorithm is a popular implementation of bagging for decision trees.

2. Boosting:

- Boosting focuses on improving the performance of weak learners by assigning higher weights to misclassified data points in each iteration.
- Algorithms like AdaBoost, Gradient Boosting (GBM), XGBoost, and LightGBM use boosting techniques and have achieved excellent predictive accuracy.

3. Stacking (Stacked Generalization):

- Stacking combines the predictions of multiple base models using another model (meta-learner) that learns how to best weigh and combine these predictions.
- Stacking can adapt to the strengths of different base models and often yields highly accurate predictions.

Voting Classifiers/Regression:

- Voting ensembles combine the predictions of multiple base models by taking a majority vote (for classification) or averaging (for regression).
- Voting can be "hard" (simple majority) or "soft" (weighted by confidence scores).

4.Gradient Boosting Variants:

- Advanced gradient boosting algorithms like XGBoost, LightGBM, and CatBoost utilize boosting techniques and have been successful in improving predictive accuracy across various domains.

5.Ensemble of Diverse Models:

- Combining diverse models, such as neural networks, decision trees, and support vector machines, can lead to improved accuracy. Each model may capture different aspects of the data.

6.Feature Engineering and Selection:

- Ensemble methods can benefit from feature engineering and feature selection techniques to enhance the quality of input data.

7.Hyperparameter Tuning:

- Careful tuning of hyperparameters for both base models and the ensemble itself can significantly boost predictive accuracy.

Import necessary libraries

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from statsmodels.tsa.holtwinters import ExponentialSmoothing

from sklearn.model_selection import train_test_split

from sklearn.metrics import mean_squared_error


# Step 1: Data Acquisition and Preparation

# Load ROC data (replace 'data.csv' with your dataset)

data=pd.read_csv('https://tn.data.gov.in/resource/company-master-data-tamil-nadu-up-to-28th-february-2019.csv')


# Preprocess data (cleaning, handling missing values, data transformation)

# Convert date columns to datetime objects

data['registration_date'] = pd.to_datetime(data['registration_date'])

# Handle missing values and other data preprocessing steps
```

```
# Step 2: Exploratory Data Analysis (EDA)
```

```
# Visualize registration trends over time
```

```
plt.figure(figsize=(12, 6))
```

```
plt.plot(data['registration_date'], data['registration_count'], label='Registration Count')
```

```
plt.xlabel('Date')
```

```
plt.ylabel('Registration Count')
```

```
plt.title('Company Registration Trends Over Time')
```

```
plt.legend()
```

```
plt.show()
```

```
# Identify seasonality, trends, and anomalies in the data
```

```
# Perform statistical tests and create visualizations as needed
```

```
# Step 3: Feature Engineering
```

```
# Create additional features (e.g., lagged values, moving averages) for modeling
```

```
data['registration_count_lagged'] = data['registration_count'].shift(1)
```

```
# Add more feature engineering steps as needed
```

```
# Step 4: Machine Learning Model (Time Series Forecasting)
```

```
# Split the data into training and testing sets
```

```
train_data, test_data = train_test_split(data, test_size=0.2, shuffle=False)
```

```
# Initialize and train a time series forecasting model (Exponential Smoothing in this example)
```

```
model = ExponentialSmoothing(train_data['registration_count'], seasonal='add',  
seasonal_periods=12)
```

```
model_fit = model.fit()
```

```
# Make predictions on the test set
```

```
y_pred = model_fit.forecast(len(test_data))
```

```
# Evaluate the model (MSE for simplicity, choose appropriate metrics)
```

```
mse = mean_squared_error(test_data['registration_count'], y_pred)
```

```
print(f"Mean Squared Error: {mse}")
```

```
# Step 5: Visualization and Reporting
```

```
# Visualize actual vs. predicted registration trends
```

```
plt.figure(figsize=(12, 6))
```

```
plt.plot(test_data['registration_date'], test_data['registration_count'], label='Actual')
```

```
plt.plot(test_data['registration_date'], y_pred, label='Predicted')
```

```
plt.xlabel('Date')
```

```
plt.ylabel('Registration Count')
```

```
plt.title('Actual vs. Predicted Company Registration Trends')
```

```
plt.legend()
```

```
plt.show()
```

```
# Add more reporting and visualization as needed
```

```
# Step 6: Continuous Monitoring and Updating
```

```
# Set up a process to periodically update the model and analyze new data
```


Step 7: Use Cases and Stakeholder Engagement

Collaborate with stakeholders to provide actionable insights based on the model's predictions

Step 8: Ethical and Privacy Considerations

Ensure ethical data handling and comply with data privacy regulations

Step 9: Security and Deployment

Implement security measures and deploy the project in a secure environment if needed

Step 10: Documentation and Knowledge Transfer

Document the project, codebase, and model for knowledge transfer and future maintenance

Sample output:

Sample output for your code

Mean Squared Error: 150.1234 # This is a placeholder value for the MSE

Future Predictions:

[1350.5678 1400.6789] # These are placeholder values for predicted registration counts in Jan and Feb 2024

Conclusion:

The "AI-Driven Exploration and Prediction of Company Registration Trends with Registrar of Companies (RoC)" project marks the dawn of a new era, one in which data-driven insights and AI capabilities illuminate the intricate tapestry of economic trends. Our mission is clear: to empower stakeholders with the foresight needed to thrive in an ever-evolving business landscape. As we delve deeper into the project, from data collection to advanced modeling, we endeavor to provide actionable intelligence that will shape the future of businesses and economies.



