Assignment

On Statistics

ES1101: COMPUTATIONAL DATA ANALYSIS

Industries, Research and Power in India: The road ahead

Sustainable Development Goal 9: Industrial Innovation and Infrastructure

SUBMITTED BY: TEAM 13

Priyanshu Joshi (2021BTech127)

Mayank Bohra (2021BTech072)

Steven Moses (2021BTech110)

Vaibhav Nalotia (2021BTech118)

SUBMITTED TO:

Dr. Richa Sharma

Mr. Santosh Verma

Institute of Engineering and Technology (IET)

JK LAKSHMIPAT UNIVERSITY, JAIPUR

February 2022

INDEX

LIST OF FIGURES

LIST OF TABLES

ABSRACT

The 2030 agenda for Sustainable development, is an initiative that was introduced by the UN in the year 2015, its main aim is show all nations of the world to have bright and sustainable future ahead. At its heart, there are 17 sustainable development goals (SDGs), which are an acute and important call for action by all nations, whether it is developed or developing nation, in a worldwide collaboration. Today our country runs on power generated by industries and our GDP depends a lot on the funds generated by industries and funds that are invested in industries as well in different forms such as R&D. This report focuses on analyzing the monetary inputs and outcomes by industries, their growth and analyze power generation by industries and their sustainability as well.

Keywords: Capital, Growth, India, SDGs, Power, R&D

1. INTRODUCTION

The Sustainable Development Goals (SDGs) or Global Goals are a collection of 17 interlinked global goals designed to be a "blueprint to achieve a better and more sustainable future for all". The SDG's were set up in 2015 and intend to be achieved by the year 2030. The SDG's are universally applicable to all countries for their future growth and development of the country which is sustainable in the long run as well. Although, they are applicable to India as well. There are many areas in India where we can focus in by these SDG's like life on land, Industry Innovation and Infrastructure etc.

Fig i

Let's take a deep dive into the present Industrial situation of India by taking a survey of the

current situation of India involving different factors that are responsible for development of Industrial growth, manufacturing and financial situation of the Industries. Finally, and also going to have a look the power generation situation in India which is responsible to power these Industries and are actually generated by Industries as well and the environmental consequences that can cause by the said development in power situation as well.

A deep statistical analysis is going to be seen of the efforts by India in terms of investments in research and development sectors along with developments in different Industries, comparing the large industries to small businesses in India (MSMEs) and analyze the financial situation of these Industries. Along with the overall power situation in India. And also, analysis of the areas in which there can be a room of improvement.

2. Survey of Literature:

Arrow K. in his paper stated that a few issues related with the regional financial disparities in India got recent consideration. The greater part of these issues is connected with or inspired by the fluctuation and unpredictability in development rates experienced by the Indian States. Productivity analysis helps to understand the degree of financial prosperity, way of life and the level of seriousness of a nation or an area. The two main theories that explain the productivity growth in countries are Neoclassical Growth Theory and Endogenous Growth Theory. Neoclassical Growth Theory: This hypothesis found its beginning in the work of Solow (1956). The minor result of capital is low in developing countries and high in developed countries. This is because of the high development pace of pay in developing countries and this reality is known as convergence theory in financial development compositions. The convergence hypothesis yields in low-income countries tends to converge of high-income countries. (Baumol, 1986; Baumol et al., 1989). The soul of the convergence hypothesis lies in the declining returns to capital. Endogenous Growth Theory: This hypothesis originated from the work of Arrow (1962). It states that distinction in productivity remains constant or even diverges over time, between developing and developed countries. As indicated by the Endogenous Growth Theory, the declining gets back to financial scale vanishes and the development path of developing economies separates from the developed countries.(2.1)

Small and medium scale industries are an essential part of any economy. In a country like India, its role is much more significant, giving employment to over 80 million people in nearly 36 million wide-spread MSMEs across nation: contributing for 45 percent of manufactured products, 41 percent of India’s total export, also producing more than eight thousand products. Furthermore, these enterprises are the hub for innovation and entrepreneurship, which is key of India’s future growth. It is also a well-known fact that this sector can help us to reach the target of the proposed National Manufacturing Policy to enhance the share of manufacturing in GDP to 25 percent and to create more than 100 million jobs by the end of the financial year 2022, as well as to foster growth and to take India from its present 2 trillion-dollar economy to a 20 trillion-dollar economy. Despite the sector’s high potential and inherent potential to grow, its growth is still facing several challenges. In this context, this paper also analyses the role of Indian MSMEs in India’s financial growth and explores various challenges which are faced by the sector. The paper also attempts to analyze various policy measures that can be undertaken by the Government of India to strengthen MSMEs of India. (2.2)

S. Chinnammai in her paper, sheds light on the fact that energy is something that is essential for each and every country in the world and India is no different, according to an article by Business Standards it has been ranked 3rd in the world when it comes to renewable sources of energy though the facto of the matter is that India is still heavily dependent on coal to meet its maximum energy demands as it has one of the biggest coal reserves in the world. Though this fact remains true, the fact of the matter remains that India has been not able to meet its energy demands in the recent years. Though there is a significant rise in electric power, there has been an issue when it comes to distribution and Generation of power. Moreover, many of the technologies like water heaters that are used for industries are made or designed to work on good quality of coal that is in minority which leads to again have a need to import coal for around the world. With all this kept in mind it’s time for our nation to come up and use its resources to improve its renewable energy infrastructure as many of the renewable energy sources remain underutilized to this day. Resources like solar energy have a very wide scope in the field. Though the country does face some challenges to use it, the factors can be such as lack of skilled labor or in many cases the upfront cost of installation can be a turning factor for many industries as well. Now acknowledging that targets achieved by our country. The future seems bright as by 2020, India has been the 4th most attractive market in the world when it comes to wind power, and 5th when it came to solar power (though more potential is there as mentioned earlier).(2.3)

Acknowledging some government initiatives, ministry of power in 2021 in India has announced some new rules that point at reducing financial thoughts that stakeholders have when investing in electric generation. Also, Indian government has proposed new rules for purchasing clean and green energy that benefit the purchases in many ways. In March of 2021 Union Cabinet had approved a memo of understanding or MOU in renewable power between India and France and many such initiatives can be seen. India is clearly serious to use the energy that it has potential to do, but the road is long way ahead. Research and technological development (R&D) are the set of innovative activities undertaken by corporations or governments in developing new services or products and improving existing ones.R&D in the field of Information and Communication Technology (ICT) is one the most important factor for a better equitable and sustainable world. According to research, there are certain gaps to fill regarding digitalization and creating a disturbance free communication chain. The trends in the variables indicate the need for more planned expenditure on ICT and R&D to build a robust ICT and to prepare for digitalization. It was also found that the ICT development indicator is correlated to the amount of money invested. Therefore, R&D on ICT is also one of the leading indicators of SDG Industry, Innovation, and Infrastructure. Globally, R&D has continued to grow at a brisk pace, reaching $2.2 trillion (purchasing power parity) in 2017, up from $1.4 trillion in 2010 and $741 billion in 2000. After COVID-19, the importance of more investment in R&D and at a faster rate has been observed. More investment is needed in the pharmaceutical industry and in emerging technologies such as artificial intelligence, which can assist in developing drugs and vaccines and in managing related services and resources.(2.4)

3. Problem Statement:

The aim of this report is to study and analyze the growth of industries in recent years based on parameters like capital, impact of MSME's on Industries, distribution of R&D finances on different Industries and power generated by the said Industries.

Objective 1

1) To study and analyze the capital approach for sustainability of the manufacturing industries.

1.1) To determine the rank among top 10 big states of India which highly contributes to the growth and profitability of the manufacturing industries

1.2) To test the hypothesis that there is no significant mean difference between productive and physical working capital of the manufacturing industries.

1.3) To analyze the relation between the productive and physical working capital of the manufacturing industries.

Objective 2

2) To study the access of medium and small-scale manufacturing industries in financial services and their integration into the value chain

2.1) To determine the relation between growth rate of overall Industry and MSME sector

2.2) To test the hypothesis that there is a significant difference between the growth rate of overall Industry and MSME sector

2.3) To analyze and predict the contribution of MSMEs in GDP of India

Objective 3

3) To study the power generated by industries in India and analyze different factors that affect generation of power

3.1) To test the hypotheses that there is significant difference between solar and thermal energies contribution to the power generation by industries in India.

3.2) To determine the relation between production of coal and power requirement.

3.3) To analyze the country's power requirement trends and hence, predict the country's power requirements in future.

Objective 4

4) To study the expenditure on Research and Development Sector of different industries of India by several Govt. organization and private industrials.

4.1) To analyze the approximate range of Rupees spent in Research and Development sector.

4.2) To Test the hypothesis that there is no significant difference in the mean of the expenditure in R&D in public and private sector of different industries of India.

4.3) To analyze and determine which state of India has invested more money in R&D.

4. DATA COLLECTION

Table 1.1 Table showing data for physical working capital of manufacturing industries

State/ UT	PHYSICAL WORKING CAPITAL (in crores)					
	2014-15	**2015-16**	**2016-17**	**2017-18**	**2018-19**	**Total**
Andhra Pradesh	4766145	4548320	5212273	5549586	6558372	26634696
Arunachal Pradesh	23868	25906	15744	13770	10690	89978
Assam	965936	755150	933890	926654	890570	4472200
Bihar	595204	516081	567831	644944	750568	3074628
Chandigarh	53181	64075	93945	58820	66776	336797
Chhattisgarh	1916121	1556111	1779409	1956706	2487520	9695867
Delhi	1256594	1084808	1130609	662156	674568	4808735
Goa	631387	573317	608996	784783	852002	3450485
Gujarat	16885736	15663667	18149946	19047793	21415721	91162863
Haryana	4795509	4956393	6402659	6171379	8026644	30352584
Himachal Pradesh	1469568	1293515	1248458	1314289	1427573	6753403
Jammu and Kashmir	328139	327468	387954	441728	439832	1925121
Jharkhand	2743019	2114327	2400861	2135694	2301800	11695701
Karnataka	6459882	6560601	7065744	7838556	8518634	36443417
Kerala	1430486	1601924	1709895	1823124	2117473	8682902
Madhya Pradesh	2630506	2755882	3044069	3858564	4034740	16323761
Maharashtra	16083653	18455255	16482040	17880963	20054541	88956452
Manipur	6569	6918	4891	6696	7429	32503
Meghalaya	80893	102844	103309	133103	129194	549343
Nagaland	10967	11016	11742	11369	11309	56403
Odisha	3547298	3625816	3177714	3806745	4011614	18169187
Puducherry	283287	288124	258177	338170	428711	1596469
Punjab	4234762	3758829	3695024	3935475	4085185	19709275
Rajasthan	3169469	3414137	3932695	4349174	4436998	19302473
Sikkim	96226	100542	158874	185651	228520	769813
Tamil Nadu	10785180	11205520	12011779	12693893	13956396	60652768
Telangana	3020495	3247806	3842310	4022345	4608648	18741604
Tripura	17489	19827	23252	31702	31895	124165
Uttar Pradesh	6630483	6767153	7542633	7847833	8803984	37592086
Uttarakhand	2107937	2197074	2253138	2212010	2455267	11225426
West Bengal	4586860	4156151	4562639	5180424	5473662	23959736

Table 1.2 Table showing data for productive capital of manufacturing industries.

State/ UT	Productive Capital (in crores)					
	2014-15	**2015-16**	**2016-17**	**2017-18**	**2018-19**	**Total**
Andhra Pradesh	15392577	18487203	18433656	19878183	22855973	95047592
Arunachal Pradesh	55218	60073	50996	35976	38600	240863
Assam	2730729	3927370	3563851	3639827	4034434	17896211
Bihar	1343628	1136371	2012130	2298576	2701644	9492349
Chandigarh	124731	150624	166455	150624	124731	717165
Chhattisgarh	12775998	11650781	13511018	13653864	14815817	66407478
Delhi	1937171	1553116	1493598	1151644	1362987	7498516
Goa	1509802	1731829	2108973	2452079	2158500	9961183
Gujarat	50095434	58579854	65966700	71894406	78886302	3.25E+08
Haryana	11273430	12829502	14783527	15369734	19640581	73896774
Himachal Pradesh	6515812	7164591	5015491	5828268	5716371	30240533
Jammu and Kashmir	896382	983419	1519672	1438118	1373641	6211232
Jharkhand	7686462	8208819	10757729	11348741	12201593	50203344
Karnataka	21476053	22558219	24023003	24563385	26860472	1.19E+08
Kerala	3787340	4746403	5816758	6290710	6359288	27000499
Madhya Pradesh	11022723	14994248	15710742	16577799	16499919	74805431
Maharashtra	44711432	57850848	54053943	47559377	55961858	2.6E+08
Manipur	28287	16584	19598	17968	16752	99189
Meghalaya	446537	388810	359998	431938	463516	2090799
Nagaland	36724	39879	39738	39428	38053	193822
Odisha	22650016	26928696	31499704	30373705	28917201	1.4E+08
Puducherry	738789	826008	746126	878077	931145	4120145
Punjab	7354482	8168199	7512885	8180831	8959933	40176330
Rajasthan	10499980	11504805	13693270	14396163	16323309	66417527
Sikkim	325733	449508	847909	1010273	1037943	3671366
Tamil Nadu	28844947	28964334	36272274	38986490	38730422	1.72E+08
Telangana	8537175	10279884	10467449	12681281	15182594	57148383
Tripura	34600	47807	54337	63498	53591	253833
Uttar Pradesh	13349782	14968405	17335309	17634483	20315978	83603957
Uttarakhand	6421631	7145513	7658688	7804143	8332374	37362349
West Bengal	10868074	10946934	13923847	14026860	12629982	62395697

Table 2.1 Data for Growth rate in MSME Sector and Overall Industry Sector

Year	Growth rate of overall Industry Sector	Growth rate of MSME Sector
2000-01	5	8
2001-02	2.7	6.1
2002-03	5.7	8.68
2003-04	7	9.64
2004-05	8.4	10.8
2005-06	8.2	12.32
2006-07	11.6	12.6
2007-08	8.5	13
2008-09	12.8	10.3
2009-10	10.4	6.7
2010-11	7.4	4.4
2011-12	7.8	4.4
2012-13	6.85	5.6
2013-14	6.76	9.1
2014-15	6.54	8.56
2015-16	6.6	7.62

Table 2.2 Data for State-wise Net value Added by Industry

Year	Andaman & Nicobar Islands	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chandigarh	Chhattisgarh	Dadra & Nagar Haveli	Daman & Diu
2004-05	449	1593777	-	372082	91016	20800	870540	489707	274618
2005-06	633	1766116	-	357617	42223	22733	697579	324635	447397
2006-07	823	2707610	-	364199	31610	32256	1191892	512559	450920
2007-08	1861	3145778	-	327473	115947	37104	1456960	525186	677392
2008-09	830	3547084	-	301389	318374	44150	1661367	778462	439760
2009-10	864	3920987	-	506212	232130	51953	1328067	662291	568985
2010-11	1190	5777901	-	667349	441499	71618	1286739	715170	632621
2011-12	1339	6705275	-	654789	564392	69272	1260536	777326	651306
2012-13	121	2382242	-	629229	130356	64938	1521724	967248	1248760
2013-14	761	1862780	-	801755	515457	68185	2125353	1022199	616659
2014-15	1081	2808114	33964	818613	582379	65235	1815125	1241146	570915
2015-16	523	3119878	24789	994758	523963	61011	939603	1335392	911687
2016-17	296	2611022	20203	1282165	536902	73922	1138267	870751	694636
2017-18	659	3536269	12206	1364876	640408	61891	1449542	1473519	765817
2018-19	889	3112374	12463	1415078	595922	42473	1874773	791313	698431

Year	Delhi	Goa	Gujarat	Haryana	Himachal Pradesh	Jammu and Kashmir	Jharkhand	Karnataka
2004-05	243083	297092	3601558	1169560	218516	57597	1677796	2053182
2005-06	273798	329525	4787158	1358920	569280	101110	1251802	2089464
2006-07	312753	361251	4795240	1530711	709902	176630	1076971	3129507
2007-08	362468	390527	6210755	1827000	1125555	274702	2062175	3465815
2008-09	347958	490164	6041722	2035386	1333237	320759	1430000	4253250
2009-10	375807	564038	9002801	2797454	1375547	263062	1510298	3669992
2010-11	587940	582752	8944778	2467880	1820876	277880	1951782	4086077
2011-12	612153	1178219	8769124	2985298	2161155	405195	1678877	4620311
2012-13	788079	1059957	11887604	3883175	2224673	397106	2091303	5347936
2013-14	787841	763987	12842005	3709582	2276625	414922	1987246	5533356
2014-15	795462	1368214	16966808	4879568	3102355	518554	2120095	5699577

2015-16	769950	1518226	18000465	5087135	3093347	585663	1353950	6426541
2016-17	620589	1367219	16543326	6690746	2832143	695502	2120637	8276614
2017-18	555781	1411722	18304059	6007400	2964376	640797	2341836	8721906
2018-19	530066	1373321	18762157	7889276	2836230	650387	3125397	8847067

Year	**Kerala**	**Madhya Pradesh**	**Maharas htra**	**Man ipur**	**Meghalaya**	**Nagaland**	**Odisha**	**Puduch erry**	**Punjab**
2004-05	406601	555351	5130921	575	9212	1995	604542	151776	580741
2005-06	446784	714211	7435401	992	23132	3498	628749	304956	661710
2006-07	355895	1112914	9537109	642	45035	3763	902270	251374	1042899
2007-08	592824	1422273	11053762	1096	53969	3716	1351150	260724	1416746
2008-09	770511	1502123	11287800	1243	52352	2823	1667360	299159	1256840
2009-10	719733	1535751	11787242	1486	49303	13602	1479974	384544	1473619
2010-11	873426	1745095	14969569	3759	66047	5366	1693137	359152	2034953
2011-12	926623	1895071	15676558	5920	80962	6813	1820476	365281	1775220
2012-13	1171460	2146115	17872932	4378	76692	8821	1805725	170457	1984470
2013-14	1345969	2124725	20051648	5208	47451	7629	2133980	407272	2006476
2014-15	1197535	2147740	20883081	6976	62697	12225	1668493	297709	2052371
2015-16	1597014	2740594	21725857	7127	60160	11817	1444942	467149	2189438
2016-17	1991866	3050925	20880019	7383	28578	12540	1847663	496782	2489373
2017-18	1887210	3991861	22372401	8376	122400	13167	2698489	500796	2679446
2018-19	1978951	4699683	19621798	9131	120267	14690	3722012	694099	2939632

Year	Rajasthan	Sikkim	Tamil Nadu	Telangana	Tripura	Uttar Pradesh	Uttarakhand	West Bengal	All-India
2004-05	658264	-	2156743	.	23124	1429776	194801	1054892	25990686
2005-06	802700	-	2798101	.	10878	1642591	345668	947059	31186419
2006-07	1215276	-	3778425	.	12637	2282767	497901	1148784	39572526
2007-08	1224435	-	3934122	.	14348	2574292	831520	1417593	48159268
2008-09	1688436	-	3971350	.	15515	2413888	2843285	1659980	52776558
2009-10	2027772	169641	5914310	.	22420	3097665	1771875	1931961	59211387
2010-11	1722902	277803	7199321	.	34129	4438129	2634767	2085971	70457581
2011-12	3926264	332717	7695571	.	18716	3714441	3064360	2055832	76455389
2012-13	3013961	364809	9136446	2810175	34809	4111251	3629866	2228051	85194869
2013-14	2681144	411195	8721268	3012612	31661	5187064	3786410	2243761	89534187
2014-15	3261992	442439	8786379	2883168	47161	4374723	4161647	1842630	97516172
2015-16	3599480	528194	10900073	3909569	28542	5485378	5552657	2253636	107248509
2016-17	4166552	965582	11788606	3551504	33840	9049549	4951666	2904543	114591911
2017-18	4248614	1064105	13701555	3948741	30480	7285531	4445813	3715370	122967418
2018-19	4669013	1084827	14079539	4868162	26736	7100261	5004840	4178172	127369428

Table 3.1 Data for Thermal Energy Solar Energy generation in India (in MW) for past 10 years

State Wise Solar and Thermal Energy Output For years (2020 – 2021)			
S. No.	States/Uts	Thermal energy (In MW)	Solar Energy Output (In MW)
1	Andaman and Nicobar Islands	40.5	29.22
2	Andhra Pradesh	14714.46	4203
3	Arunachal Pradesh	83.87	5.61
4	Assam	1167.44	42.99
5	Bihar	6528.21	159.51
6	Chhattisgarh	12221.89	45.16
7	Chandigarh	59.86	252.48
8	Dadra & Nagar Haveli	488.78	5.46
9	Daman & Diu	208.08	40.55
10	Delhi	6520.93	192.97
11	Goa	559.94	7.44
12	Gujarat	24289.09	4430.82
13	Haryana	9322.19	407.83
14	Himachal Pradesh	213.7	42.73
15	Jammu and Kashmir & Ladakh	881.22	20.73
16	Jharkhand	2426.5	52.06
17	Karnataka	10343.4	7355.17
18	Kerala	3066.66	257
19	Lakshadweep	0	0.75
20	Madhya Pradesh	16419.48	2463.22
21	Maharashtra	28766.91	2289.97
22	Manipur	154.67	6.36
23	Meghalaya	161.29	0.12
24	Mizoram	71.51	1.53
25	Nagaland	81.03	1
26	NLC	166	0
27	Odisha	5027.21	401.72
28	Puducherry	285.1	9.33
29	Punjab	8729.51	959.5
30	Rajasthan	14004.49	5732.58
31	Sikkim	50.27	0.07
32	Tamil Nadu	14839.17	4475.21
33	Telangana	10211.19	3953.12
34	Tripura	630.05	9.41
35	Uttar Pradesh	20303.33	1712.5
36	Uttarakhand	1011.26	368.41
37	West Bengal	12444.36	149.84

Table 3.2 Data for coal generation (in Million Tones) and Power Requirements (in MW) for the past 10 years

Year wise coal generated and power requirement in India		
Years	Coal generated (In Million Tonnes)	Power requirement (In MW)
2010	532.042	83059
2011	532.694	86159
2012	539.95	93720
2013	556.402	99811
2014	565.765	100226
2015	609.179	106892
2016	639.23	111441
2017	657.868	114293
2018	675.4	119215
2019	728.718	127456
2020	730.874	129101
2021	716.084	127553

Table 4.1: Data for money spent by different states of India for the advancement in Research and Development of the nation.(in crores)

S. No.	State	2015-16	2016-17	2017-18	Total Sum
1	Andhra Pradesh	493.69	533.43	543.49	1570.61
2	Assam	276.59	289.58	455.13	1021.3
3	Bihar	132.79	129.71	53.68	316.18
4	Chhattisgarh	169.99	208.57	257.99	636.55
5	Gujarat	728.01	755.44	789.52	2272.97
6	Haryana	285.62	313.3	374.58	973.5
7	Himachal Pradesh	140.16	141.15	158.97	440.28
8	Jammu & Kashmir	238.17	260.58	290.68	789.43
9	Jharkhand	70.88	106.2	91.54	268.62
10	Karnataka	369.17	344.57	373.48	1087.22
11	Kerala	138.9	123.82	160.41	423.13
12	Madhya Pradesh	417.93	466.98	488.51	1373.42
13	Maharashtra	190.09	234.1	221.77	645.96
14	Manipur	200.19	113.19	102	415.38
15	Meghalaya	8.76	8.58	9.01	26.35
16	Odisha	130.97	147.26	154.85	433.08
17	Punjab	513.68	534.38	551.44	1599.5
18	Rajasthan	203.55	233.11	192.56	629.22
19	Tamil Nadu	589.14	670.51	688.34	1947.99
20	Telangana	269.89	284.07	276.02	829.98
21	Andhra Pradesh	371.87	448.39	471.02	1291.28
22	Assam	255.53	253.98	281.41	790.92
23	Bihar	252.36	266.1	278.41	796.87

Table 4.2: Data for money spent by the Government of India regarding the Research and Development of different Industries of the nation.

Investment made by Government of India in R&D (in crores)					
S. No.	**Industry Group**	**2015-16**	**2016-17**	**2017-18**	**Average Govt. Expenditure**
1	Agricultural Machinery	0.76	1.23	0.59	0.86
2	Biotechnology	0.12	0.12	0.07	0.103
3	Boilers & Steam Generating Plants	129.68	162.28	129.68	140.546
4	Chemicals (other than fertilizers)	17.54	13.15	15.64	15.443
5	Consultancy Services	42.97	41.66	90.52	58.383
6	Defense Industries	2021.49	2199	2712.22	2310.903
7	Drugs & Pharmaceuticals	2.09	2.85	3.15	2.69
8	Earth Moving Machinery	66.61	78.07	102.05	82.243
9	Electricals & Electronics	95.66	94.97	102.3	97.643
10	Fertilizers	19.86	17.13	14.34	17.11
11	Food Processing Industries	0.05	0.05	0.05	0.05
12	Fuels	1023.25	1123.63	1163.99	1103.623
13	Industrial Equipment	2.09	1.23	2.41	1.91
14	Industrial Machinery	749.1	648.86	646.64	681.533
15	Information Technology	0	0	0	0
16	Machine Tools	1.65	1.93	1.73	1.77
17	Medical & Surgical Appliances	11.69	9.02	6.57	9.0933
18	Metallurgical Industries	198.04	246.81	208.64	217.83
19	Paper & Pulp	9.72	10.94	11.15	10.603
20	Scientific Instruments	4.21	3.57	3.74	3.84
21	Soaps,Cosmetics,Toilet Preparations	0.54	0.48	0.65	0.5577
22	Sugar	0.68	0.67	0.57	0.64
23	Telecommunications	33.61	33.25	22.76	29.873
24	Textiles(Dyed,Printed,Processed)	0.39	0.46	0.28	0.377
25	Transportation	9.59	6.65	0.85	5.697
26	Miscellaneous Industries	12.49	8.28	12.66	11.143

Table 4.3: Data for money spent by different Private Industrial Group regarding Research and Development of their respective Industry.

Investment made by Private Industries in R&D (in crores)					
S.No.	**Industry Group**	**2015-16**	**2016-17**	**2017-18**	**Average Private Expenditure**
1	Agricultural Machinery	697.03	777.57	831.69	768.763
2	Biotechnology	1052.67	1075.28	1071.3	1066.416
3	Boilers & Steam Generating Plants	50.66	51.07	68.21	56.646
4	Chemicals (other than fertilizers)	2437.86	2813.85	3004.56	2752.09
5	Consultancy Services	27.66	28.74	26.35	27.583
6	Defense Industries	112.58	117.86	140.25	123.563
7	Drugs & Pharmaceuticals	8832.96	10294.44	10159.11	9762.17
8	Earth Moving Machinery	53.13	94.3	149.41	98.947
9	Electricals & Electronics	1737.6	1768.45	1935.47	1813.84
10	Fertilizers	37.79	38.38	44.03	40.067
11	Food Processing Industries	248	280.92	245.86	258.26
12	Fuels	39.91	45.1	51.32	45.443
13	Industrial Equipment	362.82	402.92	470.49	412.077
14	Industrial Machinery	186.27	171.39	201.52	186.393
15	Information Technology	3209.52	3210.62	3625.17	3348.437
16	Machine Tools	121.41	125.68	142.16	129.75
17	Medical & Surgical Appliances	983.2	1176.57	1277.97	1145.913
18	Metallurgical Industries	173	216.09	263.9	217.663
19	Paper & Pulp	16	21.25	32.22	23.157
20	Scientific Instruments	13	13.13	13.77	13.3
21	Soaps, Cosmetics, Toilet Preparations	13	278.17	308.91	200.027
22	Sugar	11	93.63	87.58	64.07
23	Telecommunications	13	179.13	187.76	126.63
24	Textiles(Dyed, Printed, Processed)	51	228.43	242.8	174.077
25	Transportation	100	6096.48	6061.18	4085.887
26	Miscellaneous Industries	132	604.68	596.79	444.49

5. METHODOLOGY

DESCRIPTIVE STATS:

Descriptive stats help us to get a basic idea about information by variables like mean, mode and median present in a dataset and helps us to determine some sort of relationship between the values of the dataset with variables like Standard Deviation and Variance. The variables used in this report are as follows.

1.) Mean: Mean is the most basic operation in mathematics and statistics. The mean is the average value in a given dataset.

   Formula: $$\overline{X} = \frac{\sum X}{N}$$

   Where X denotes discrete values and N denotes number of values of the dataset.

2.) Standard Deviation: In statistics, the standard deviation is a measure of the amount of variation or dispersion of a set of values.

   Steps for determining Standard Deviation:

   Step 1: Find mean of the dataset. (μ)

   Step 2: Find distance of every data value to the mean and square it. $| x - \mu |^2$

   Step 3: Add up all the values found I Step 2. $\sum| x - \mu|^2$

   Step 4: Divide the result by N and Sqrt it.

   $$S = \sqrt{\frac{\sum |x - \mu|^2}{N}}$$

3.) Variance: It's simply the square of Standard Deviation.

$$\sigma^2 = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{n}$$

Inferential Statistics:

Inferential statistics is the tool that statisticians use to draw conclusions about the characteristics of a population. It is used to make generalization about large groups, also it uses probability to determine how right our conclusion is that assume while analyzing the data.

Hypothesis Testing

Hypothesis Testing is a form of statistical inference that uses data from a sample to draw conclusions about the dataset. It is basically an assumption which we generally make by considering the population parameter. First, a tentative assumption is made about the parameter, this assumption is called null hypothesis and is denoted by Ho. Then we define, alternate hypothesis which is the opposite of what is stated in the null hypothesis.

There are two types of hypotheses:

1) One-tailed Test

Left-tailed

Right-tailed

2) Two-Tailed Test

Left-tailed

If the alternate hypothesis contains the less-than inequality symbol (<), the hypothesis test is left-tailed test.

$$H_0: \mu \geq k$$
$$H_a: \mu < k$$

Fig ii

Right-tailed

If the alternative hypothesis contains the greater than symbol (>), the hypothesis test is a right-tailed test.

$$H_0: \mu \leq k$$
$$H_a: \mu > k$$

Fig iii

Two-tailed

If the alternative hypothesis contains the not-equal-to symbol, the hypothesis test is a two-tailed test.

$$H_0: \mu = k$$
$$H_a: \mu \neq k$$

Fig iv

Steps For Hypothesis Testing:

Step 1: Determine null and alternate hypothesis. They should by contradictory.

Step 2: Determine the random variable and the distribution for the test.

Step 3: Draw a graph, calculate the test statistics, and use statistics to calculate the p-value.

Step 4: Compare the preconceived level of significance with the p-value and make the decision and write the clear conclusion using English sentences.

Z-Test for Two Means:

Z-test for two samples is used to compare the means of two samples to see if the two proportions are same or not.

Steps For Z-Test for two means:

Step 1: Determine the null and alternate hypothesis.

Step 2: Find the means and standard deviations for the two samples.

Step 3: Then z-score can be calculated using the following formula:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$$

Step 4: Determine the z-score from the table at a certain level of significance. The critical Z value for level of significance, $\alpha = 0.05$ is 1.96

Step 5: Compare the calculated z-score from step 3 with the tabled z-score, determined in step If the calculated z-score is larger, then one can reject the null hypothesis.

Rejection Criteria

Criteria to reject the null hypothesis for z-test.

1) $|Z| > Z_{\alpha/2}$ (In Case of Two-Tailed Test)

2) $Z < -Z_\alpha$ (In Case of Left-Tailed Test)

3) $Z > Z_\alpha$ (In Case of Right-Tailed Test)

T-Test for Two Means

T-Test for two samples is used to identify, how the means of two data sets differ from one another when variance is not given.

Steps For T-Test for two means:

Step 1: Determine the null and alternate hypothesis.

Step 2: Find the means and standard deviations for the two samples.

Step 3: Then t-score can be calculated using the following formula:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Step 4: Determine the t-score from the table at a certain level of significance. The critical t-value for level of significance, $\alpha = 0.05$ depends on the Degree of freedom () of the dataset.

Step 5: Compare the calculated t-score from step 3 with the tabled t-score, determined in step 4. If the calculated t-score is larger, then one can reject the null hypothesis.

Rejection Criteria

Criteria to reject the null hypothesis for t-test.

1) $|t| > t_{\alpha/2}$ (In Case of Two-Tailed Test)

2) $t < -t_\alpha$ (In Case of Left-Tailed Test)

3) $t > t_\alpha$ (In Case of Right-Tailed Test)

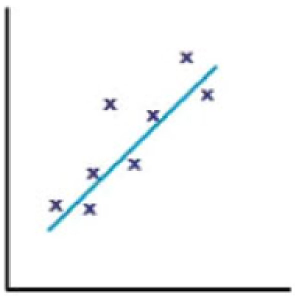
Correlation

Correlation is a statistic that measures the degree to which two parameters move in relation to each other. It shows the strength of a relationship between two variables and is expressed numerically by the correlation coefficient. The correlation coefficient's values range between –1.0 and 1.0 lower the value is, weaker is the correlation and vice versa.

Direction and Magnitude:

The sign of the correlation coefficient represents the direction of the relationship.

If there is a positive correlation the value of first variable has strong chance to increase on the increase of second variable. They have a up slope



Strong +ve Correlation: $0.7 > r > 1$

Moderate +ve Correlation: $0.3 > r > 0.7$

Weak +ve Correlation: $0 > r > 0.3$

Fig v

Negative coefficients represent cases when the value of one variable increases, the value of the other variable has a tendency to decreased. Negative relationships produce a downward slope.

Strong -ve Correlation: $-1 < r < -0.7$

Moderate -ve Correlation: $-0.3 < r < -0.5$

Weak -ve Correlation: $0 < r < -0.3$

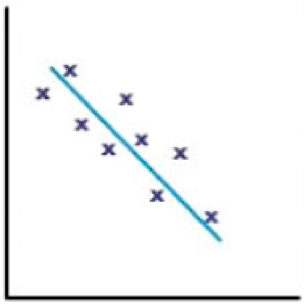


Fig vi

Method and formula for calculation of correlation coefficient.

Step 1 Find means for the 1st and 2nd variable dataset, My and Mx.

Step 2 Find Sum of squares (SS_X) and (SS_y), these are sum of squares of corresponding difference of mean and 1st and 2nd variable values.

$\sum(X - Mx)^2$ and $\sum(Y - My)^2$

Step 3 Find products of sum of corresponding difference of mean and 1st and 2nd variable values.

$\sum(X - Mx)(Y - My)$

Plug these values in the equation below for finding Pearson correlation coefficient:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Where,

x_i = Numerical data for variable 1

y_i = Numerical data for variable 2

$\bar{x}$ = Mean of variable 1

$\bar{y}$ = Mean of variable 2

Regression:

Simple linear regression is used to model the relationship between two continuous variables. Linear regression is often used to predict the value of a dependent variable based on values from independent variable.

The value so, are obtained by plugging the independent variable in a Linear Regression Equation.

Y = m*X + c

Slope and y-intercept can be found as follows:

Step 1 -> Find means for the dependent and independent variable dataset, My and Mx.

Step 2 -> Find Sum of squares (SS_X), these are sum of squares of corresponding difference of mean and independent variable values.

$\sum(X - Mx)^2$

Step 3 -> Obtain Sum of products (SP) which is the sum of product of difference of independent variable values and dependent variable values for their means. $\sum (X - Mx) (Y - My)$

Step 4 -> Obtain the slop and y-intercept by the following relation.

Slope (**m**)= SP/SS_X

y-intercept (**c**) = $M_Y - m*M_X$

Power Method:

This method is used to find out the dominant eigen value and the corresponding dominant eigen vector of a given matrix.

To apply this method on a square matrix A, guess an initial eigenvector for the corresponding dominant eigen value. Multiply that eigen vector to matrix A keeping it on the right side. Standardize the result and keep repeating the process until getting the desired eigenvector or until the results aren't similar to each other. If it occurs, then the norm of that eigenvector will be considered as the absolute value of the dominant eigenvalue.

Finding Rank using Power Method:

- First form a square matrix of corresponding order
- Find the required dominant eigenvalue and vector using the power method explained above.
- Maybe for matrix of higher order (as of in our case) it can done with the help of Python code as well.
- After the dominant eigenvalue and corresponding eigenvector is known we can rank the of the corresponding data with the help of the eigenvector.

6. Analysis:

6.1) For finding rank among top ten states of India based on their contribution in productive capital for manufacturing industries, power method has been implemented on table 1.2.

Modelling for Ranking: (Refer table 1.2)

- To find the rank, matrix contains elements 0,1,2
- Based on data of table 1.2 of Productive Capital of manufacturing industries, modelling has been done as written below:
- If both values are the same, element 0 is considered.
- If the first value is less than the second value, element 1 is considered.
- If the first value is greater than the second value, element 2 is considered.
  The following Matrix has obtained

	Madhya Pradesh	Maharashtra	Punjab	Rajasthan	Telangana	Uttar Pradesh	West Bengal	Karnataka	Delhi	Goa
Madhya Pradesh	0	1	2	2	2	1	2	1	1	1
Maharashtra	2	0	2	2	2	2	2	2	2	1
Punjab	1	1	0	1	1	1	1	1	1	1
Rajasthan	1	1	2	0	2	1	2	1	1	1
Telangana	1	1	2	2	0	1	1	1	1	1
Uttar Pradesh	2	1	2	2	2	0	2	1	2	1
West Bengal	1	1	2	2	2	1	0	1	1	1
Karnataka	2	1	2	2	2	2	2	0	2	1
Delhi	2	1	2	2	2	1	2	1	0	1
Goa	2	2	2	2	2	2	2	2	2	0

Python is used to apply power method, dominant eigenvalue and eigenvector are calculated as follows:

Dominant eigenvalue is: 12.93

Dominant eigenvector is:

$$\begin{bmatrix} 0.71 \\ 0.93 \\ 0.54 \\ 0.66 \\ 0.57 \\ 0.81 \\ 0.62 \\ 0.87 \\ 0.76 \\ 1 \end{bmatrix}$$

6.2) For determining if there is a statistically significant difference between Physical working capital and Productive capital, z test for difference of means has been implemented (as the sample size is greater than 30) on table 1.1 and 1.2

Claim: There is no significant difference between the means of Physical working capital and Productive capital.

Step 1:

Null hypothesis (H_o): There is no significant difference between the means of Physical working capital and Productive capital.

$\mu_2 - \mu_1 = 0$

Alternate Hypothesis (H_a): There is a significant difference between the means of Physical working capital and Productive capital.

$\mu_2 - \mu_1 \neq 0$

Step 2:

The level of significance is taken as 5%: $\alpha = 0.05$

Step 3:

The critical Z value for $\alpha = 0.05$ is 1.96

Step 4:

Plugging values into the equation

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{{\sigma_1}^2}{n_1} + \frac{{\sigma_2}^2}{n_2}}}$$

We get $|Z| = 3.85$, by putting the values of the corresponding variables in the python code.

6.3) For finding if there is a relation between the Physical Working Capital and productive capital of manufacturing industries, Pearson's Correlation method has been applied on table 1.1 and 1.2

Step 1: Calculation for X Values: Values for Physical Working Capital
$\sum X = 557340841$

$M_x = 17978736.806$

$\sum(X - Mx)^2 = SS_x = 1.708E+16$

Step 2: Calculation for Y Values: values for Productive Capital
$\sum Y = 1844361642$

M_y = 59495536.839
$\sum(Y - My)^2 = SS_y$ = 1.782E+17

Step 3: Calculation for X and Y Combined
N = 31
$\sum(X - Mx)(Y - My)$ = 5.304E+16

Step 4: Calculation for r

$$r = \sum((X - M_y)(Y - M_x)) / \sqrt{((SS_x)(SS_y))}$$

Plug in these values in the formula above and obtain the Pearson correlation coefficient (r).

Here, r = 0.961

6.4) Applying Descriptive Statistics on Table 2.1 for analyzing relation between growth of Industries and MSMEs

Count = 15

Mean = 7.654882e+07

Minimum value = 25990686

Max = 127369428

Standard Deviation(s)

$s^2 = \sum (x - \bar{x})^2 / n-1$

=1.5500076148948e+16/14

=1.1071482963534e+15

s =√1.1071482963534e+15 = 33273838.0166

6.5) For determining if there is a statistically significant difference between the growth rate of overall industry and MSME sector, T - test for difference of means has been applied (as the sample size is less than 30) on Table 2.1.

Here, X determines growth rate of overall Industry sector.

And Y determines growth rate of overall MSME sector.

Claim: There is not a significant difference between the growth rate of overall Industry and MSME sector.

[Hypothesis Test]

Step 1:

Null Hypothesis (H_0): There is no significant difference between the growth rate of overall Industry and MSME sector.

H_0: $\mu_{1=}$ μ_2

Step 2:

Alternate Hypotheses (H_a): There is no significant difference between the growth rate of overall Industry and MSME sector.

H_a: $\mu_1 \neq \mu_2$

Step 3:

Level of significance is 5%, i.e., $\alpha = 0.05$

Step 4:

Evaluating t value

N_1: 16

$df_1 = N - 1 = 16 - 1 = 15$

M_1: 7.64

SS_1: 90.88

$s^2_1 = SS_1/ (N - 1) = 90.88/ (16-1) = 6.06$

N_2: 16

$df_2= N - 1 = 16 - 1 = 15$

M_2: 8.61

SS_2: 113.73

$S^2_2 = SS_2/ (N - 1) = 113.73/ (16-1) = 7.58$

$s^2_p = ((df_1/ (df_1 + df_2)) * s^2_1) + ((df_2/ (df_2 + df_2)) * s^2_2)$

$= ((15/30) * 6.06) + ((15/30) * 7.58)$

$= 6.82$

$s^2_{M_1}= s^2_p /N_1$

$= 6.82/16 = 0.43$ S

$s^2_{M_2}= s^2_p /N_2$

$= 6.82/16 = 0.43$

$t = (M_1 - M_2)/\sqrt{} (s^2_{M_1} + s^2_{M_2})$

$= -0.97/\sqrt{0.85} = -1.05$

6.6) For predicting the contribution of MSMEs in GDP in India, Linear Regression has been implemented. (Refer Table 2.2)

For the Equation:

Y = m*X + c

Calculation

Here X denotes years from 2005 to 2019

And Y denotes Net value added (in Crores)

Sum of X = 30180

Sum of Y = 11482323.08

Mean X = 2012

Mean Y = 765488.2053

Sum of squares (SS_X) = 280

Sum of products (SP) = 20813928.07

Regression Equation = $\hat{y} = mX + c$

$m = SP/SS_X$

$= 20813928.07/280$

$= 74335.45739$

$c = M_Y - M_X$

$= 765488.21 - (74335.46*2012)$

$= -148797452.0691$

$\hat{y} = 74335.45739X - 148797452.0691$

6.7) For determining if there is a statistically significant difference between solar and thermal energies contribution to the power generation by industries in India, z test for difference of means has been implemented on Table 3.1.

Claim: There is not a significant difference between the solar energy and thermal energy generated by industries.

[Hypothesis Test]

Step 1:

Null Hypothesis (H_0): There is no significant difference between the means solar energy and thermal energy generated by industries.

H_0: $\mu_1 = \mu_2$

Step 2:

Alternate Hypotheses (H_a): There is a significant difference between the means solar energy and thermal energy generated by industries.

H_a: $\mu_1 \neq \mu_2$

Calculation has been done through Python Code.

we obtain $|Z| = 3.845$, which is greater than 1.96.

6.8) For determining if there is any relation between the coal production and power generated for the power requirement in the country, Pearson's Correlation has been implemented.

Calculation:

For values of year wise coal generated(X):
$\sum X=7484.206$
Mean= 623.684
$\sum(X - Mx)^2 = SSx = 66925.058$

For values of year wise Power Generated (Y):
$\sum Y= 1298926$
Mean = 108243.833
$\sum(Y - My)^2 = SSy = 2814336667.667$

X and Y Combined:
N = 12
$\sum(X - Mx)(Y - My) = 13446647.396$

R Calculation:
$r = \sum((X - My)(Y - Mx)) / \sqrt{((SSx)(SSy))}$

$r = 13446647.396 / \sqrt{((66925.058)(2814336667.667))} = 0.9798$

6.9) For analyzing and predicting the country's power requirement trends, Linear Regression has been applied on Table 3.1

For the regression equation

Y = m*X + c

Where X denotes the years and Y denotes Coal production in million tonnes.

Sum of X = 24186
Sum of Y = 1298926
Mean X = 2015.5
Mean Y = 108243.8333
Sum of squares (SS_x) = 143
Sum of products (SP) = 627917

m = SP/SS_x = 627917/143 = 4391.02797

c = $M_Y - m*M_X$ = 108243.83 - (4391.03*2015.5) = -8741873.04429
The slope comes out to be = 4391.027
and y intercept = - 8741873.044

And the regression equation is
Y = 4391.028X - 8741873.044

6.10) For analysis of contribution of different states by expenditure on R&D, descriptive statistics have been applied on Table 4.1

Count = 23

Mean (μ) = 894.771

For Standard Deviation (σ)

$$\sigma^2 = \frac{\Sigma(x_i - \mu)^2}{N}$$

$= (1570.61 - 894.77)^2 + ... + (796.87 - 894.77130434783)^2/23$

$= 304148.200$

σ = 551.496

6.11) For determining if there is a significant difference in the mean of the expenditure in R&D by govt. and private sector of different industries T-test has been applied on Table 4.2 and 4.3 as the sample size is less than 30.

Claim: There is not a significant difference between the money spent by the govt. and Private Industries in the field of R&D of the nation.

[Hypothesis Test]

Step 1:

Null Hypothesis (H_0): There is no significant difference between the money spent by Govt. of India and Private Industries in the field of R&D of the nation.

H_0: $\mu_1 = \mu_2$

Step 2:

Alternate Hypotheses (H_a) : There is a significant difference between the money spent by Govt. of India and Private Industries in the field of R&D of the nation.

Step 3:

Level of significance is 5%, i.e., $\alpha = 0.05$

Step 4:

Test Statistic (Data in Table 4.2 and 4.3)

N1: 26
df1 = N - 1 = 26 - 1 = 25
M1: 184.79
SS1: 6223655.95
s21 = SS1/(N - 1) = 6223655.95/(26-1) = 248946.24

N2: 26
df2 = N - 1 = 26 - 1 = 25
M2: 1053.29
SS2: 108923563.4
s22 = SS2/(N - 1) = 108923563.4/(26-1) = 4356942.54

s2p = ((df1/(df1 + df2)) * s21) + ((df2/(df2 + df2)) * s22) = ((25/50) * 248946.24) + ((25/50) * 4356942.54) = 2302944.39

s2M1 = s2p/N1 = 2302944.39/26 = 88574.78
s2M2 = s2p/N2 = 2302944.39/26 = 88574.78

t = (M1 - M2)/√(s2M1 + s2M2) = -868.51/√177149.57 = -2.06

6.12) For determining the rank among top 10 states of India which has spent more amount of money in the field of Research & Development, Power Method has been used on Table 4.1.

Modelling for Ranking: (Refer table 4.1)

The Modelling has been done as follows:

- To find the rank, matrix contains elements 0,1,2
- Based on data of Research and Development Expenditure compared for each state the modelling has been done as follows:
- If both values are the same, element 0 is considered.
- If the first value is less than the second value, element 1 is considered.
- If the first value is greater than the second value, element 2 is considered.

The following Matrix has obtained

	Andhra Pradesh	Gujarat	Assam	Karnataka	Madhya Pradesh	Punjab	Tamil Nadu	Uttar Pradesh	Haryana	Telangana
Andhra Pradesh	0	1	2	2	2	1	1	2	2	2
Gujarat	2	0	2	2	2	2	2	2	2	2
Assam	1	1	0	1	1	1	1	1	2	2
Karnataka	1	1	2	0	1	1	1	1	2	2
Madhya Pradesh	1	1	2	2	0	1	1	2	2	2
Punjab	2	1	2	2	2	0	1	2	2	2
Tamil Nadu	2	1	2	2	2	2	0	2	2	2
Uttar Pradesh	1	1	2	2	1	1	1	0	2	2
Haryana	1	1	1	1	1	1	1	1	0	2
Telangana	1	1	1	1	1	1	1	1	1	0

Applying power method using python code, the following Eigen values and Eigen Vectors are obtained:

The code went for 6 iterations and the dominant Eigen value = 12.93

And Eigen vector is :

$$\begin{bmatrix} 0.81 \\ 1 \\ 0.62 \\ 0.66 \\ 0.76 \\ 0.87 \\ 0.93 \\ 0.71 \\ 0.57 \\ 0.54 \end{bmatrix}$$

7. Result and Discussion

The Rank of different states (based on their contributions in productive capital for the manufacturing industries) obtained using power method are as follows:

1st Rank: Gujarat

2nd Rank: Maharashtra

3rd Rank: Karnataka

4th Rank: Uttar Pradesh

5th Rank: Delhi

6th Rank: Madhya Pradesh

7th Rank: Rajasthan

8th Rank: West Bengal

9th Rank: Telangana

10th Rank: Punjab

From the above result, one can observe that Gujarat contributes highest in the total productive capital for the growth and profitability of the manufacturing industries, whereas Punjab has the lowest contribution in the productive capital shares.

As calculated in section 6.2, the |Z| value is 3.85 and critical Z value is 1.96

Hence, as the |Z| is greater than critical Z value, so null hypothesis is rejected, and alternate hypothesis is accepted. Therefore, the claim that there is no significant difference between the means of Physical working capital and Productive capital is false.

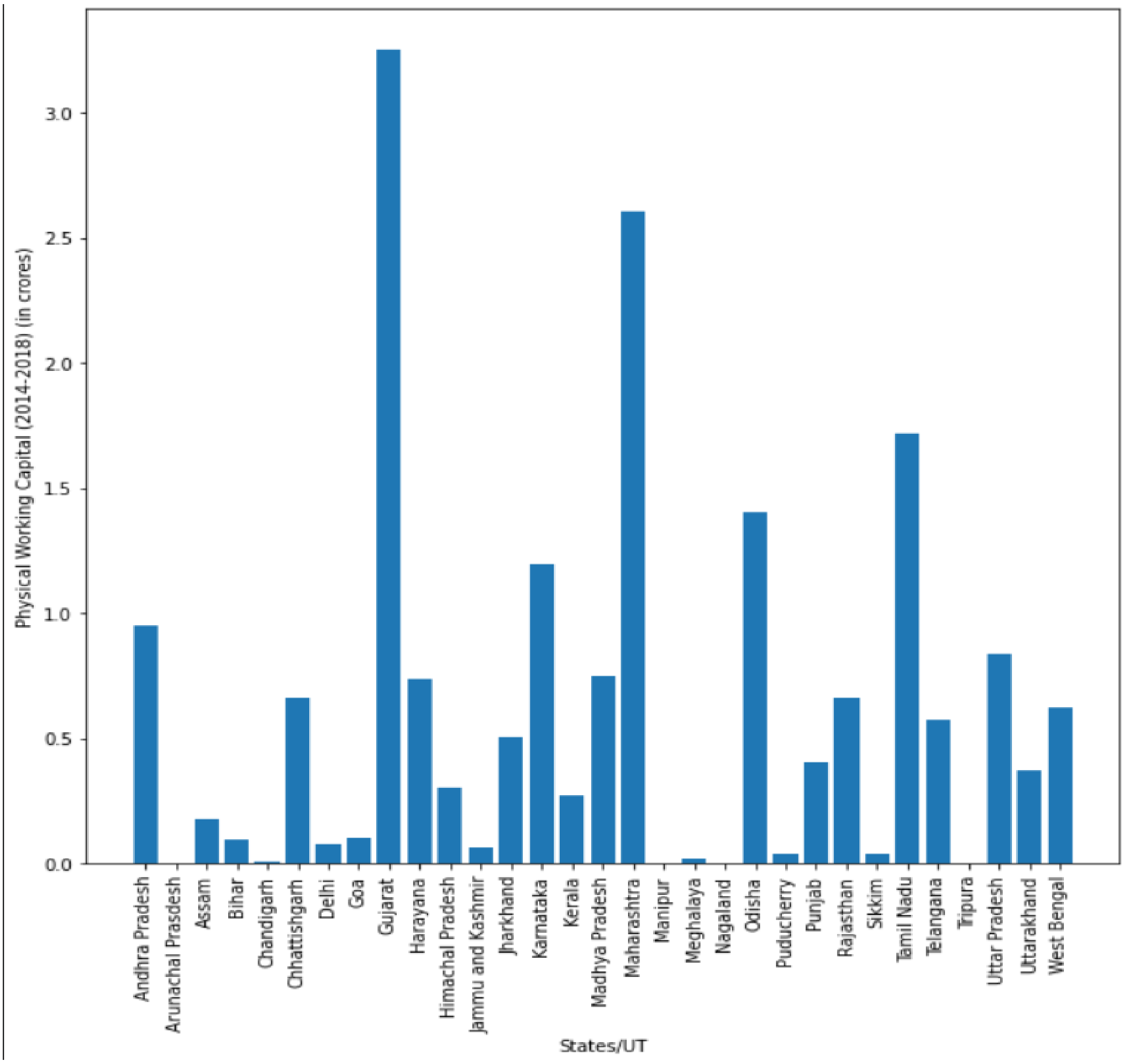


Figure 1: Bar Plot graph State wise representing Total Productive Capital of the manufacturing industries.

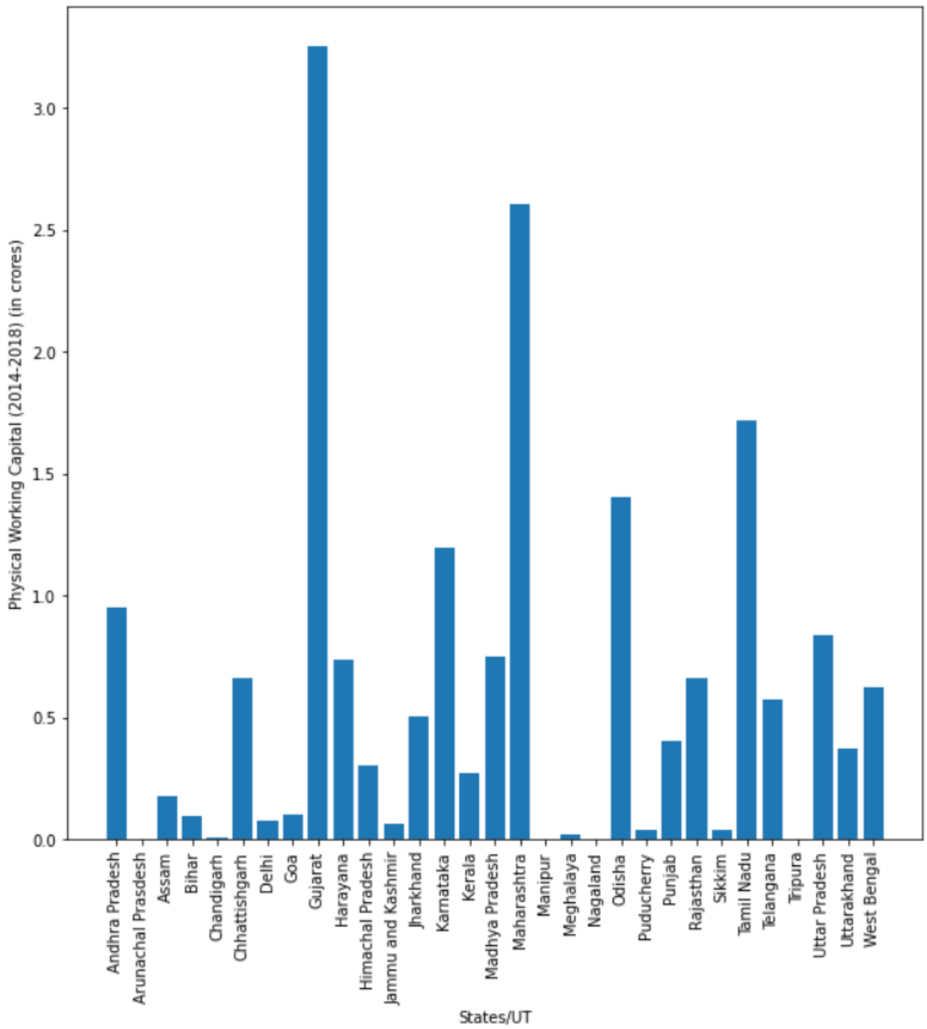


Figure 2: Bar Plot graph State wise representing the money invested in Physical Working Capital by the manufacturing industries.

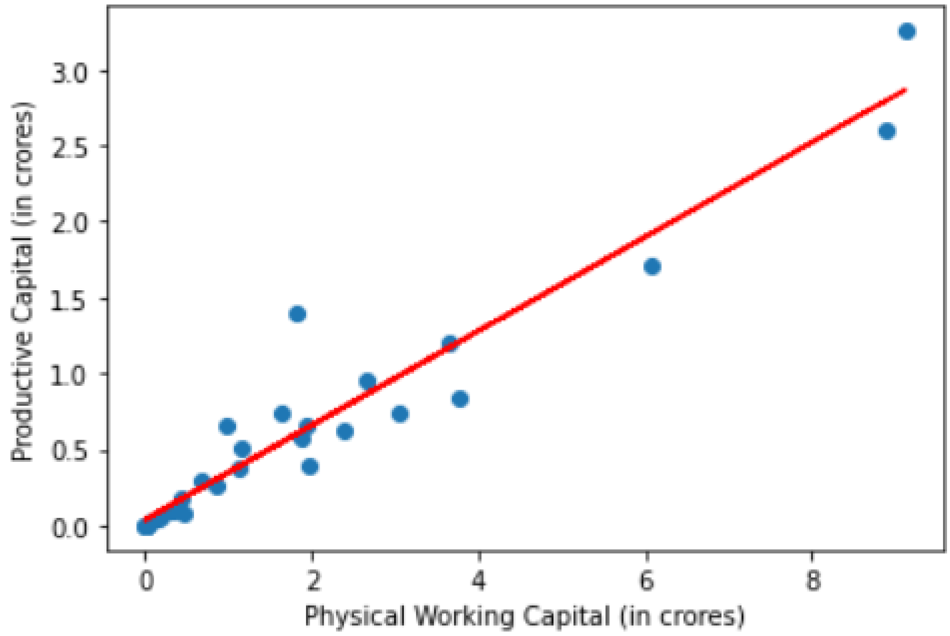


Figure 3: Scatter cum Line Plot graph representing the relation between the Physical Working Capital (in crores) and Productive Capital (in crores) for the manufacturing Industries.

As calculated in section 6.3, r = 0.961, and hence from figure 3 as there is a strong positive correlation between the two variables. So, years in which there is a high Physical Working Capital also tend to the years with high Productive Capital by industries.

From figure 1 and 2 one can see that there is a significant difference between physical working capital and productive capital, and also one can infer that with more productive capital Gujarat ranks first in the contribution of productive capital for the manufacturing industries.

To analyze the growth rate of MSME contribution in GDP

In previous years MSME sector shown amazing development and its contribution to GDP in also increasing every year (Fig.6) making India the fastest growing economy.

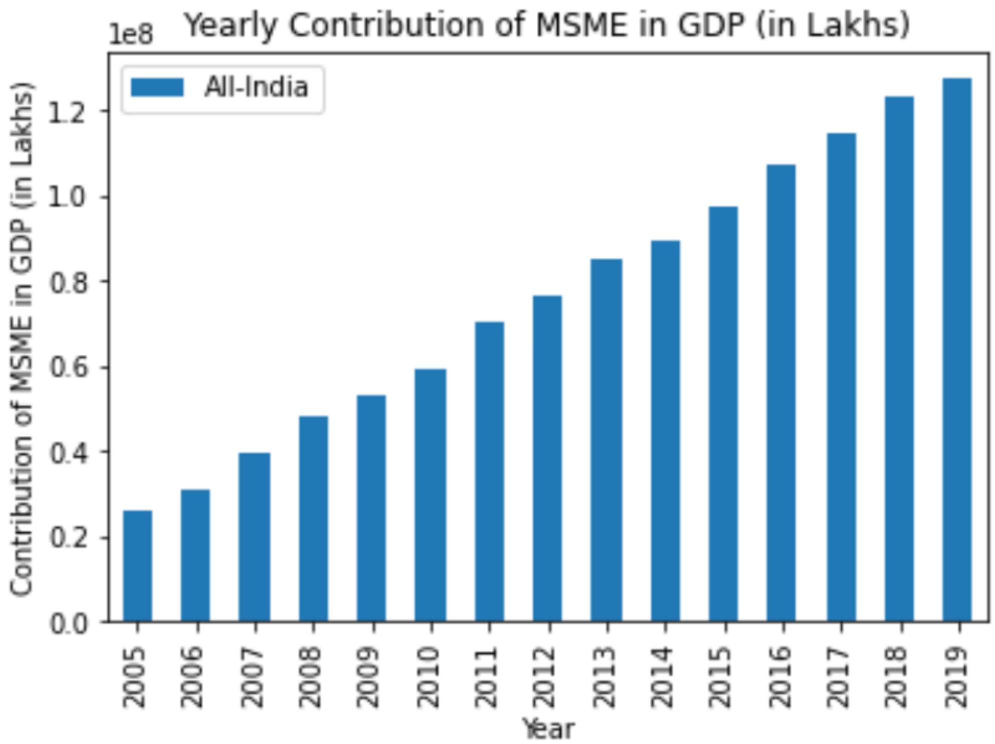


Figure 4: Yearly Contribution of MSME in GDP (in Lakhs)

To test the hypothesis that there is a significant difference between the growth rate of overall Industry and MSME sector.

The t-value is -1.05395.

The tabled value of t at 5% of level of significance is 1.706

Therefore, we have enough evidence to state that there is not a significant difference between the growth rate of overall Industry and MSME sector.

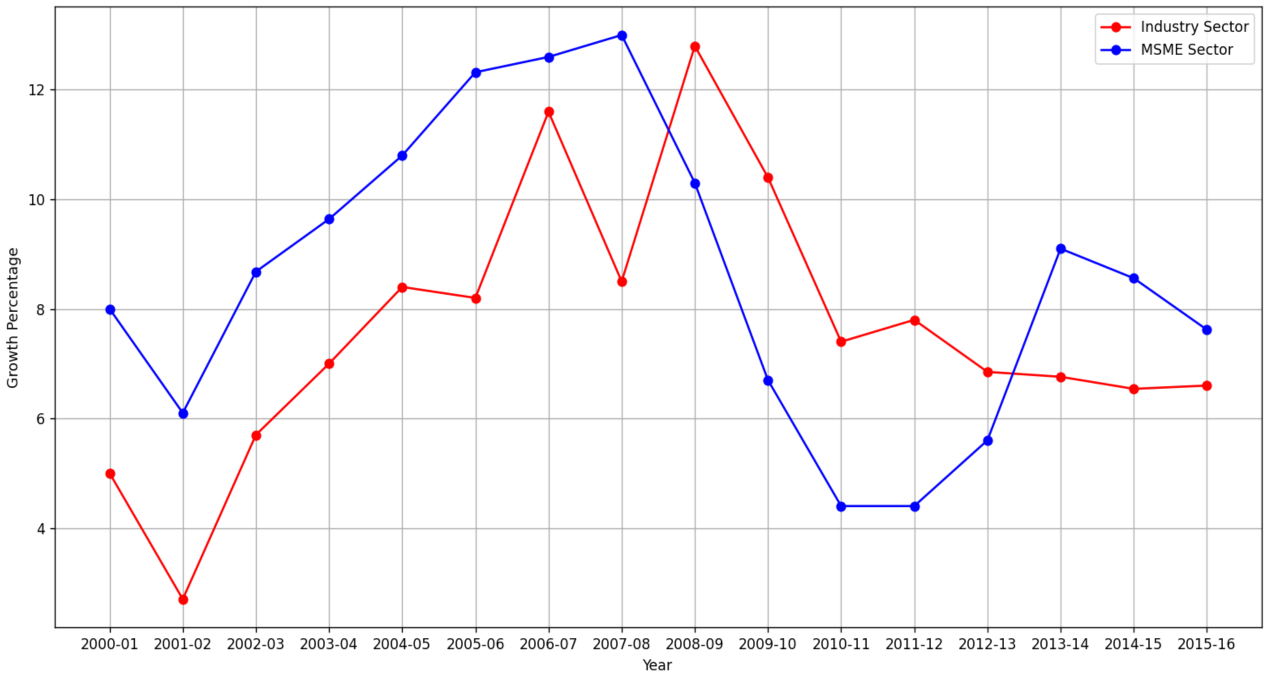


Figure 5: Previous year trend of Growth between MSME Sector and Industry Sector

To analyze and predict the contribution of MSMEs in GDP of India.

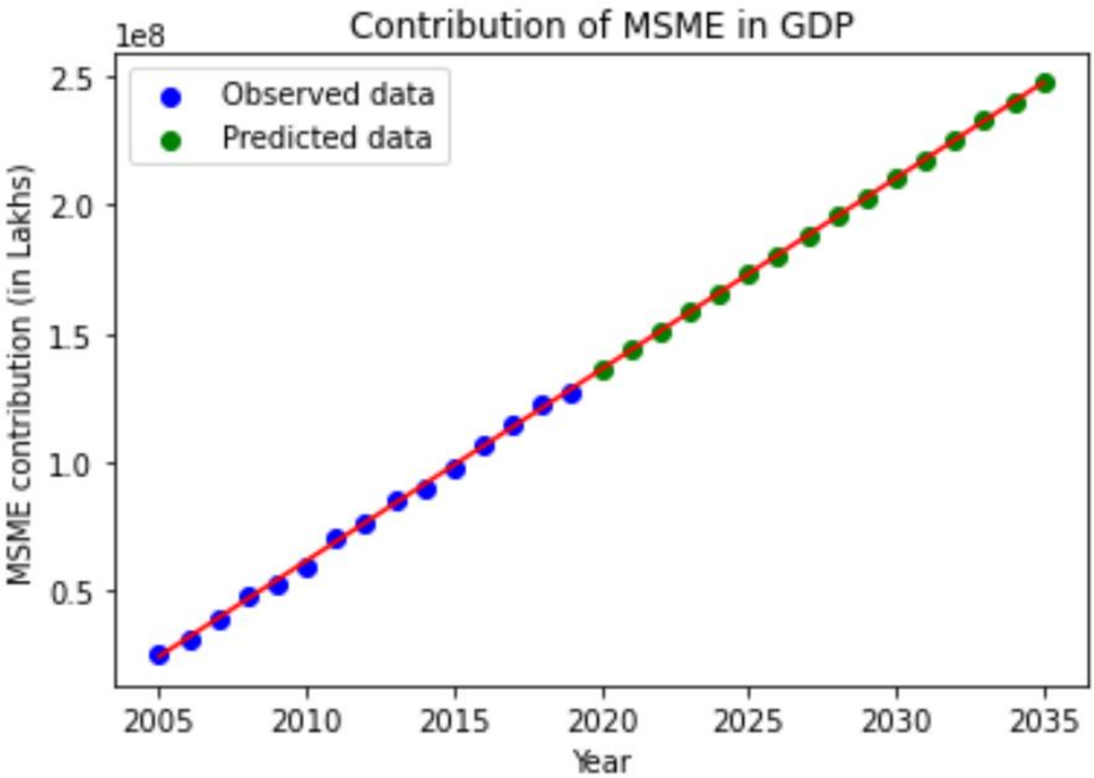


Figure 6: Scattered plot of contribution of MSME sector in GDP along with predicted data until 2035

In 2005, the contribution of MSME sector in GDP was 2.45 Lakh Crores. It is expected that contribution of MSME sector in 2035 will be around ten times from 2005 being 24.75 Lakh Crores (Fig.7)

For determining if there is a statistically significant difference between solar and thermal energies contribution to the power generation by industries in India (as calculated on section 6.7)

we obtain $|Z| = 3.85$, which is greater than $Z_c = 1.96$.

Hence, we can say that null hypothesis is rejected and the alternate hypothesis is accepted and therefore we can say that there is a significant difference between the contribution of solar and thermal energies produced by the industries.

And this Graph (figure 7) further Supports this claim.

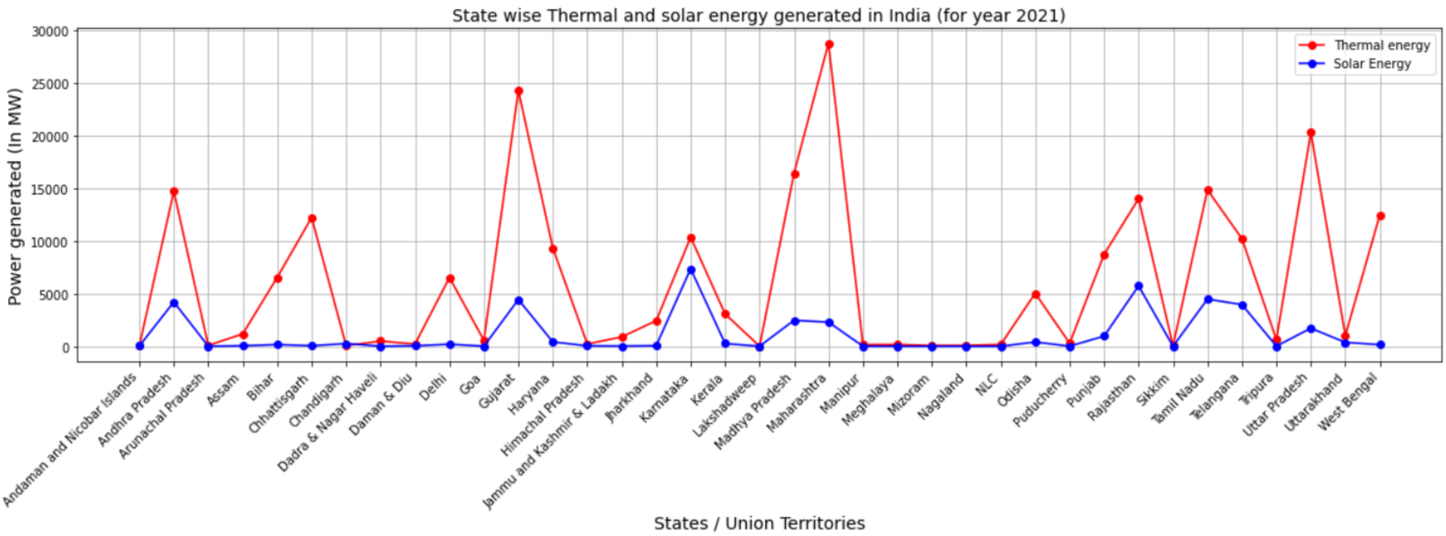


Figure 7: Line chart showing thermal and solar energies for respective states in India for the year 2021

The Correlation Coefficient between coal demands and power requirements by the industries is as, (r) = 0.9798. Hence there is a strong positive correlation between the two variables.

So, years in which there is a high coal demand also tend to the years with high power requirements by the industries. (Fig 8)

And with a significant difference between the solar and thermal energy differences, we can infer that coal is majorly responsible for power generation in India.

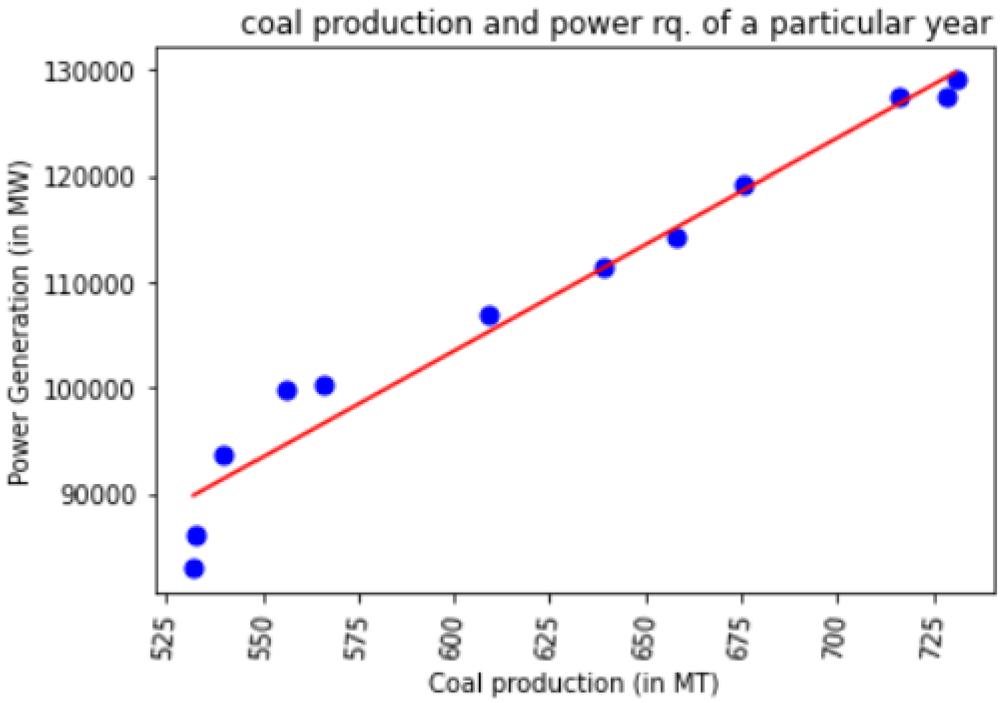


Figure 8: Scattered plot for coal production and power generation in India of past 11 years

There is also a year wise increase in power requirement in India (Figure 9)

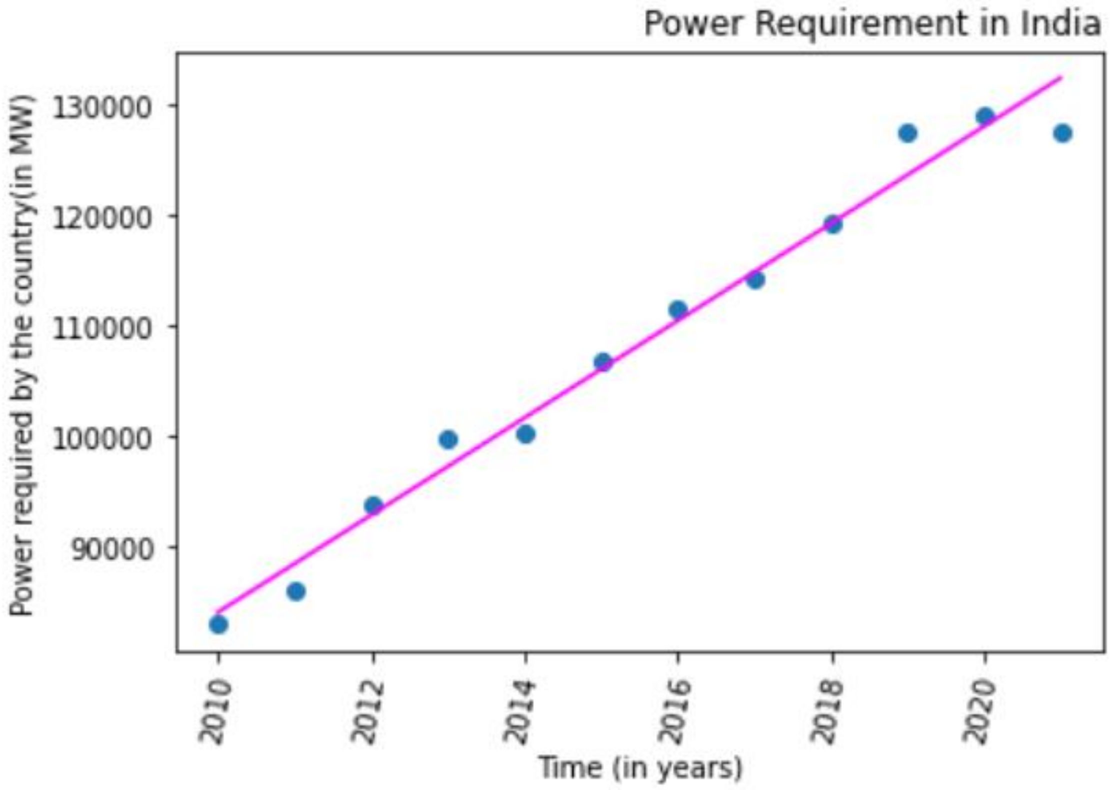


Figure 9: Future power requirements of India based on predicted values

Applying Regression on table 3.2 as calculated in section 6.9 the regression equation obtained is:

Y = 4391.028X - 8741873.044

This equation can be used to predict the Power generated in India for a particular year and the equation has been implemented in python to predict the future power requirements in India (Fig 10)

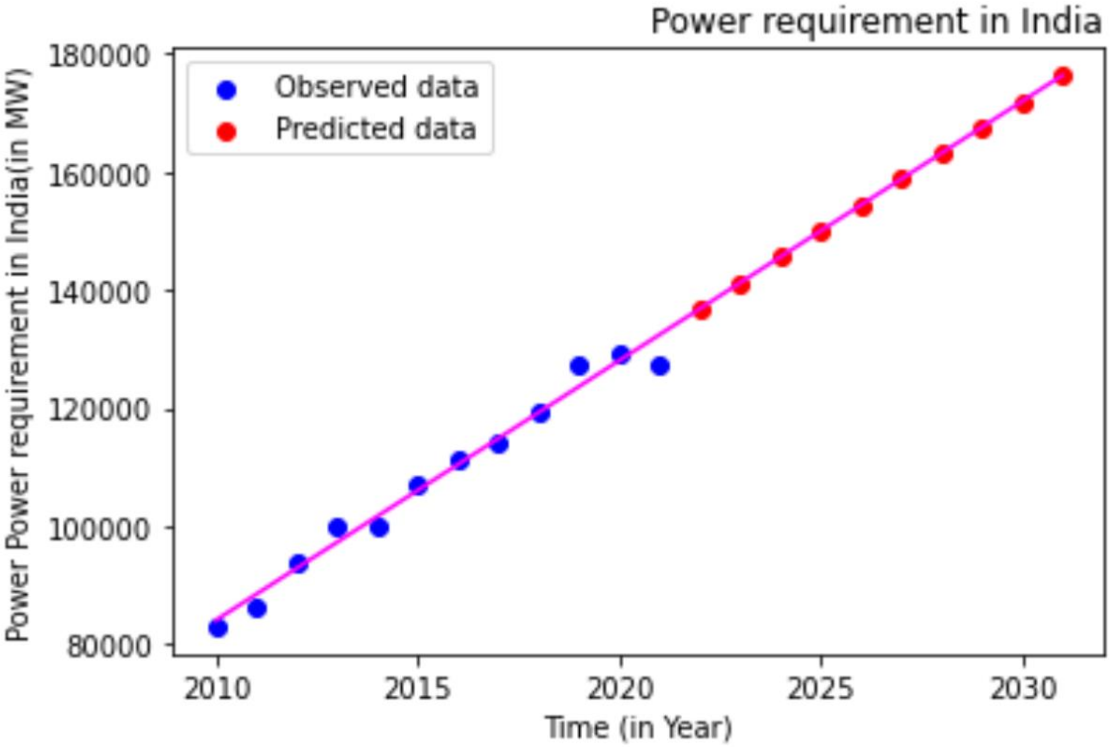


Figure 10: Future power requirements of India based on predicted values upto 2030

And according to the trends in the recent year, the coal production recently has seen a dip. Due to rising prices of coal all round the world. (Fig 11)

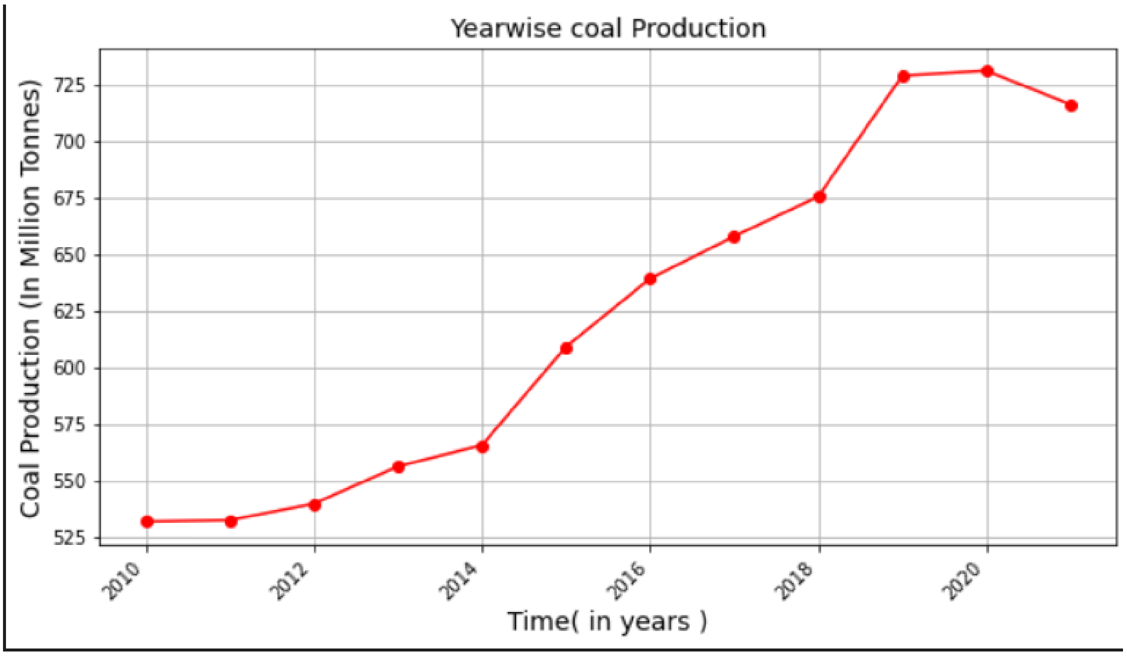


Figure 11: Year wise coal products in India for the past 11 year

To analyze the approximate range of Rupees spent in Research and Development by various states application of Descriptive stats is applied and the following results are obtained.

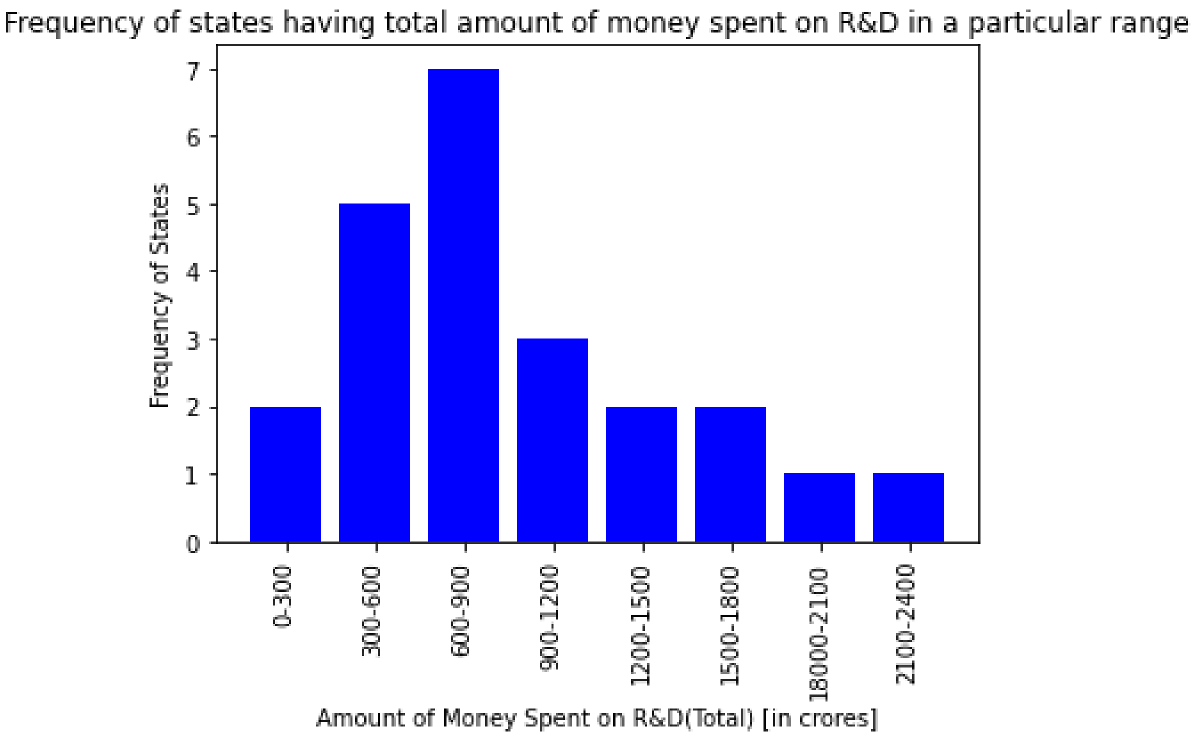


Figure 12: Frequency Distribution graph between no. of states and range of amount of money spent on R&D by states of India. (Refer Table 4.1)

From the above frequency distribution graph one can conclude that maximum states have made total expenditure in R&D is in the range of Rs.600 - Rs.900 crores.(Fig. 12)

Since the standard deviation and the mean is 551.49 and 894.77 (in crores) respectively, one can estimate from the above statement that approximately 95% of the total expenditure in R&D fall in the range of 208.2 – 1997.7 (in crores).

To test the hypothesis that there is a significant difference in the mean of the expenditure in R&D by Govt. and Private sector of different industries.

As calculated in section 6.11, the tabled value of t at 5% of level of significance is 1.706

Therefore, as calculated value of t is less than tabled value of t, hence, the null hypothesis is accepted. So, our claim that there is no significant difference between the money spent by the govt. and Private Industries in the field of R&D of the nation is true.(Fig. 13)

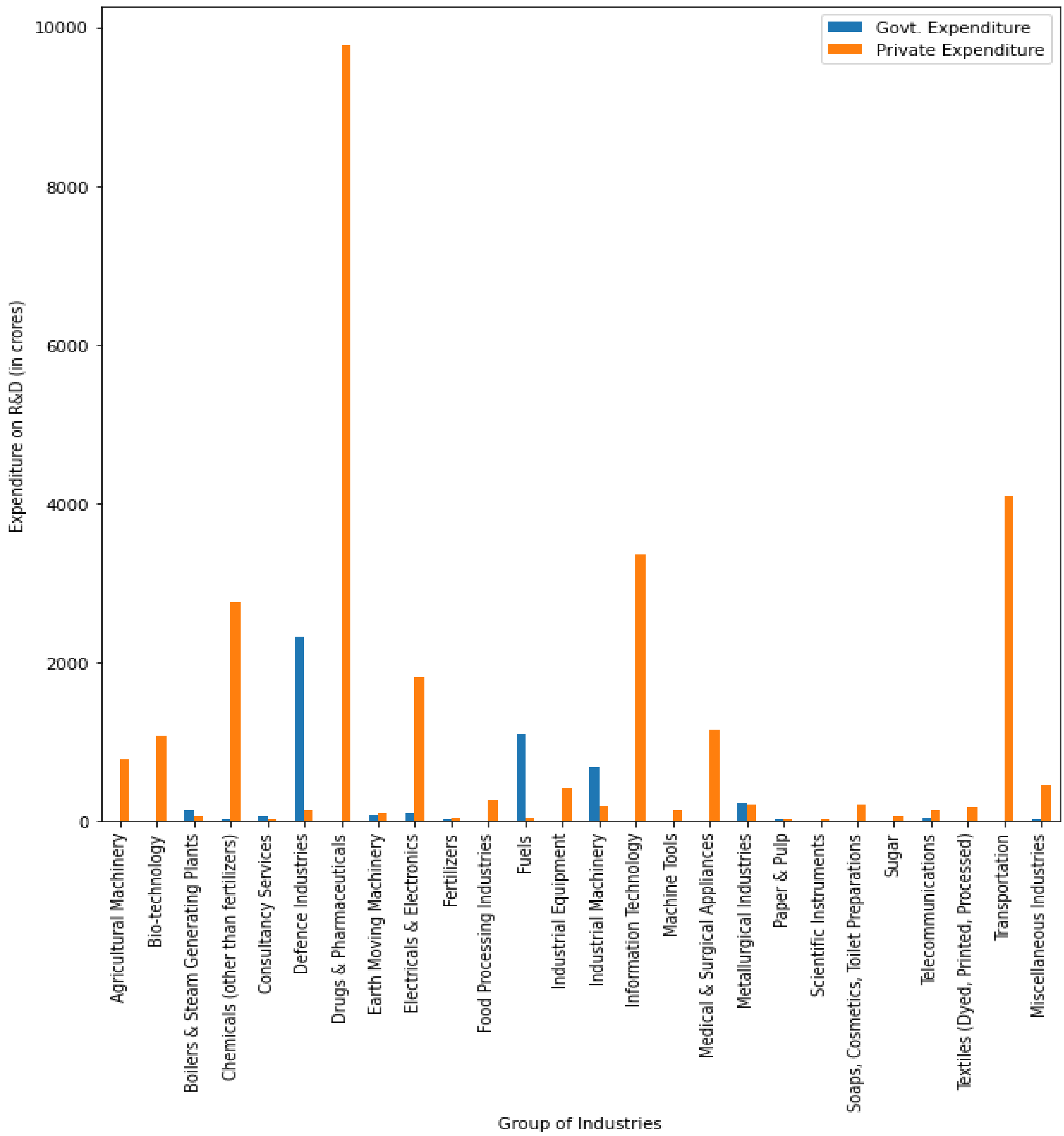


Figure 13: Bar plot graph of expenditure made by various Govt. and Private Industries of India in R&D sector of the industry.(Refer Table 4.2 and 4.3)

To determine the rank of top 10 states who has contributed more to R&D of nation applying power method to get dominant eigen vector.

Rank of the States are as:

1st Rank: Gujarat

2nd Rank: Tamil Nadu

3rd Rank: Punjab

4th Rank: Andhra Pradesh

5th Rank: Madhya Pradesh

6th Rank: Uttar Pradesh

7th Rank: Karnataka

8th Rank: Assam

9th Rank: Haryana

10th Rank: Telangana

From the above-mentioned result, one can conclude that Gujarat contributes the highest in Research and Development, whereas Telangana has the lowest contribution in Research and Development.

7. Conclusion:

- The physical working capital and productive capital can play a substantial role in formulating profitability and productivity and accordingly affect industry's financial substantiality. Therefore, there is the impact of physical working capital on industry's productivity with state-wise benchmarks.
- The MSME sector is the key to India's future economic growth. Growth rate of MSMEs and overall industry are statistically not significant, thus stating the importance of the sector. Further, analyzing the contribution of MSME in GDP. It will ten times the contribution of MSME in 2035 compared to 2005.
- Due to the fall in coal production, strong correlation between coal production and power requirement in India and future power demands seemingly rising by our predictions rising, there is strong need to increase solar energy generation by the industries. Else India can face a sever power crisis in future.
- So far in all those years, Indian State Govt.'s, private industries and many Indian govt. agencies have invested heavily in the Advancement of Research and Development Sector of the nation.

  Private Industries spends the most in Pharma, while Govt. spend the most on defense security of the nation. In the coming days of digital era, there is a requirement to fulfill by investing on Information and Technology sector as well to make India a superpower nation.

The capitals of our country are mostly invested on Drug and pharmaceuticals that is a positive trend as it can lead to more development in the country's health sector (SDG 3). The capital that is invested in different industries also leads better productive output as well and hence, it's a safe bet for our country to invest more on R&D of Industries. MSMEs can also learn a lot from these industries that will further develop more employment opportunities which will go a great mile to reduce poverty in our country (SDG 1). The power generation scenario of India, that's responsible to run the industries as well have plenty of room for improvement as supply of coal decreases and need for power increases. Further the report can also be linked to SDG 7 (affordable and clean energy) as better use of solar resources will lead to clean and sustainable energy and improve general population's health as well due to reduced carbon emissions generated by coal.

Overall, industries in India have progress in recent times, though there is always room for improvement and these factors if kept in mind can lead to a prosperous nation.

8. References:

Literature review:

2.1) Arrow K. (1962) The Economic implications of learning by doing. Review of Economic Studies 29(3): 155–173.

2.1) Baumol W. J. (1986) Productivity growth, convergence and welfare: what the long-run data show. American Economic Review 76(5): 1072–1085.

2.2) Solow, Robert M. (1956) ‘A Contribution to the Theory of Economic Growth.’ Quarterly Journal of Economics, 70, 65–94.

2.2) Singh, S., & Paliwal, M. (2017). Unleashing The Growth Potential of Indian MSME Sector. Comparative Economic Research. Central And Eastern Europe, 20(2), 35-52. https://doi.org/10.1515/cer-2017-0011

2.2) MSME Annual Report by Government of India

https://msme.gov.in/sites/default/files/MSME-ANNUAL-REPORT-ENGLISH%202020-21.pdf

2.3) S. Chinnammai. (2014, August 4). *Srinivasan Chinnammai - IJESD*. http://www.ijesd.org Retrieved February 3, 2022, from http://www.ijesd.org/papers/518-A1001.pdf

2.3) India, P. T. of. (2021, December 1). Business Standard. Retrieved February 3, 2022, from https://www.business-standard.com/article/economy-policy/india-s-power-consumption-rises-3-6-to-100-42-bu-in-november-121120100496_1.html

2.3) Central Electricity Authority, Ministry of New and Renewable Energy. (2021, November). *Brand India*. IBEF. Retrieved February 3, 2022, from https://www.ibef.org/industry/renewable-energy.aspx

2.4) Mavuri, Sudha & Chavali, Kavita & Kumar, Ajith. (2020). SDG 9 Imperatives.

https://www.researchgate.net/publication/341509895_SDG_9_Imperatives

Bangera, Sharyn & Gandhi, Usha. (2021). A STUDY ON PROGRESS OF SUSTAINABLE DEVELOPMENT GOALS BY INDIA.

For Data:

Table 1.1: https://m.rbi.org.in/scripts/PublicationsView.aspx?id=20762

Table 1 1.2: https://m.rbi.org.in/scripts/PublicationsView.aspx?id=20763

Table 2.1 and 2.2

https://www.rbi.org.in/Scripts/PublicationReportDetails.aspx?UrlPage=&ID=924#CH2

Table 2.2 Source: Annual Survey of Industries (ASI), Ministry of Statistics and Programme Implementation, Government of India

Table 3.1 Reserve Bank of India - Publications (rbi.org.in)

Table:3.2https://en.wikipedia.org/wiki/States_of_India_by_installed_power_capacity

Table 4.1

https://dst.gov.in/sites/default/files/S%26T%20Indicators%20Tables%202019-20.pdf

Table 4.2 and 4.3

https://dst.gov.in/sites/default/files/Research%20and%20Deveopment%20Statistics%202019-20_0.pdf

For Figures:

Fig i

Sustainable Development Goals (gavi.org)

Fig ii/iii/iv

(1) How do you know whether the hypothesis is left tailed, right tailed or two tailed? - Quora

Fig v/vi

Positive and Negative Correlation | eMathZone

Appendix:

Code for all graphs and plots:

For objective 1 [Power Method]

```
import pandas as pd
matrix = pd.read_csv("Rank.csv")

mat = matrix.iloc[:,1:].to_numpy()
print(mat)
```

```
[[0 1 2 2 2 1 2 1 1 1]
 [2 0 2 2 2 2 2 2 2 1]
 [1 1 0 1 1 1 1 1 1 1]
 [1 1 2 0 2 1 2 1 1 1]
 [1 1 2 1 0 1 1 1 1 1]
 [2 1 2 2 2 0 2 1 2 1]
 [1 1 2 1 2 1 0 1 1 1]
 [2 1 2 2 2 2 2 0 2 1]
 [2 1 2 2 2 1 2 1 0 1]
 [2 2 2 2 2 2 2 2 2 0]]
```

```
import numpy as np

x = []
eigen = [0]

x.append(np.array([[1],[1],[1],[1],[1],[1],[1],[1],[1],[1]]))
print(mat,'\n\n',x)
print('-----------------------------------------------')

i=0
while len(eigen) == 1 or np.round(eigen[i],4)!=np.round(eigen[i-1],4):
    x.append(np.matmul(mat,x[i]))
    eigen.append(max(x[i+1]))
    x[i+1] = np.round(x[i+1]/eigen[i+1],4)
    print('Iteration',i+1,':')
    print(x[i+1],'\t',eigen[i+1])
    print()
    i+=1
```

For [Z-Test for Difference of Means]

```python
import pandas as pd
from numpy import sqrt, abs, round
from scipy.stats import norm

data = pd.read_csv("Total.csv")
data
```

	State/ UT	Physical Working Capital	Productive Capital
0	Andhra Pradesh	26634696	95047592
1	Arunachal Prasdesh	89978	240863
2	Assam	4472200	17896211
3	Bihar	3074628	9492349
4	Chandigarh	336797	717165
5	Chhattishgarh	9695867	66407478
6	Delhi	4808735	7498516
7	Goa	3450485	9961183
8	Gujarat	91162863	325422696
9	Harayana	30352584	73896774
10	Himachal Pradesh	6753403	30240533
11	Jammu and Kashmir	1925121	6211232
12	Jharkhand	11695701	50203344
13	Karnataka	36443417	119481132
14	Kerala	8682902	27000499
15	Madhya Pradesh	16323761	74805431
16	Maharashtra	88956452	260137458
17	Manipur	32503	99189
18	Meghalaya	549343	2090799
19	Nagaland	56403	193822
20	Odisha	18169187	140369322
21	Puducherry	1596469	4120145
22	Punjab	19709275	40176330
23	Rajasthan	19302473	66417527
24	Sikkim	769813	3671366
25	Tamil Nadu	60652768	171798467
26	Telangana	18741604	57148383
27	Tripura	124165	253833
28	Uttar Pradesh	37592086	83603957
29	Uttarakhand	11225426	37362349
30	West Bengal	23959736	62395697

```python
physical = data["Physical Working Capital"]
productive = data["Productive Capital"]

mean1 = physical.mean()
mean2 = productive.mean()

std1 = physical.std()
std2 = productive.std()

print("The mean and standard deviation of Physical working capital is",round(mean1,3),',', round(std1,3), "respectively.")
print("The mean and standard deviation of Productive capital is",round(mean2,3),',',round(std2,3), "respectively.")
```

```
The mean and standard deviation of Physical working capital is 17978736.806 , 23865521.024 respectively.
The mean and standard deviation of Productive capital is 59495536.839 , 77073275.371 respectively.
```

```python
n1 = physical.count()
n2 = productive.count()

Sp = sqrt(std1*std1/n1 + std2*std2/n2)
z = (mean2-mean1)/Sp

print('z =',abs(round(z,3)))
```

✓ 0.1s

```
z = 2.865
```

```python
if abs(z)< 1.96:
    print('z =',round(z, 3), "\nThe null hypothesis is accepted, there is no significant difference between the means of the two variables" )
elif abs(z)> 1.96:
    print('z =',round(z, 3), "\nThe null hypothesis is rejected, therefore, there is a significant difference between the means of the two variables")
```

✓ 0.1s

```
z = 2.865
The null hypothesis is rejected, therefore, there is a significant difference between the means of the two variables
```

```python
import pandas as pd

data1 = pd.read_csv("physical_working_capital.csv")
data1
```

	State/ UT	2014-15	2015-16	2016-17	2017-18	2018-19	Total
0	Andhra Pradesh	4766145	4548320	5212273	5549586	6558372	26634696
1	Arunachal Prasdesh	23868	25906	15744	13770	10690	89978
2	Assam	965936	755150	933890	926654	890570	4472200
3	Bihar	595204	516081	567831	644944	750568	3074628
4	Chandigarh	53181	64075	93945	58820	66776	336797
5	Chhattishgarh	1916121	1556111	1779409	1956706	2487520	9695867
6	Delhi	1256594	1084808	1130609	662156	674568	4808735
7	Goa	631387	573317	608996	784783	852002	3450485
8	Gujarat	16885736	15663667	18149946	19047793	21415721	91162863
9	Harayana	4795509	4956393	6402659	6171379	8026644	30352584
10	Himachal Pradesh	1469568	1293515	1248458	1314289	1427573	6753403
11	Jammu and Kashmir	328139	327468	387954	441728	439832	1925121
12	Jharkhand	2743019	2114327	2400861	2135694	2301800	11695701
13	Karnataka	6459882	6560601	7065744	7838556	8518634	36443417
14	Kerala	1430486	1601924	1709895	1823124	2117473	8682902
15	Madhya Pradesh	2630506	2755882	3044069	3858564	4034740	16323761
16	Maharashtra	16083653	18455255	16482040	17880963	20054541	88956452
17	Manipur	6569	6918	4891	6696	7429	32503
18	Meghalaya	80893	102844	103309	133103	129194	549343
19	Nagaland	10967	11016	11742	11369	11309	56403
20	Odisha	3547298	3625816	3177714	3806745	4011614	18169187
21	Puducherry	283287	288124	258177	338170	428711	1596469
22	Punjab	4234762	3758829	3695024	3935475	4085185	19709275
23	Rajasthan	3169469	3414137	3932695	4349174	4436998	19302473
24	Sikkim	96226	100542	158874	185651	228520	769813
25	Tamil Nadu	10785180	11205520	12011779	12693893	13956396	60652768
26	Telangana	3020495	3247806	3842310	4022345	4608648	18741604
27	Tripura	17489	19827	23252	31702	31895	124165
28	Uttar Pradesh	6630483	6767153	7542633	7847833	8803984	37592086
29	Uttarakhand	2107937	2197074	2253138	2212010	2455267	11225426
30	West Bengal	4586860	4156151	4562639	5180424	5473662	23959736

```
plt.figure(figsize = (10,10))
plt.bar(x=data1["State/ UT"],height=data1["Total"])

labels = ['200']

plt.xticks(rotation = 90)

plt.xlabel("States/UT")
plt.ylabel("Physical Working Capital (2014-2018) (in crores)")
plt.show()
```

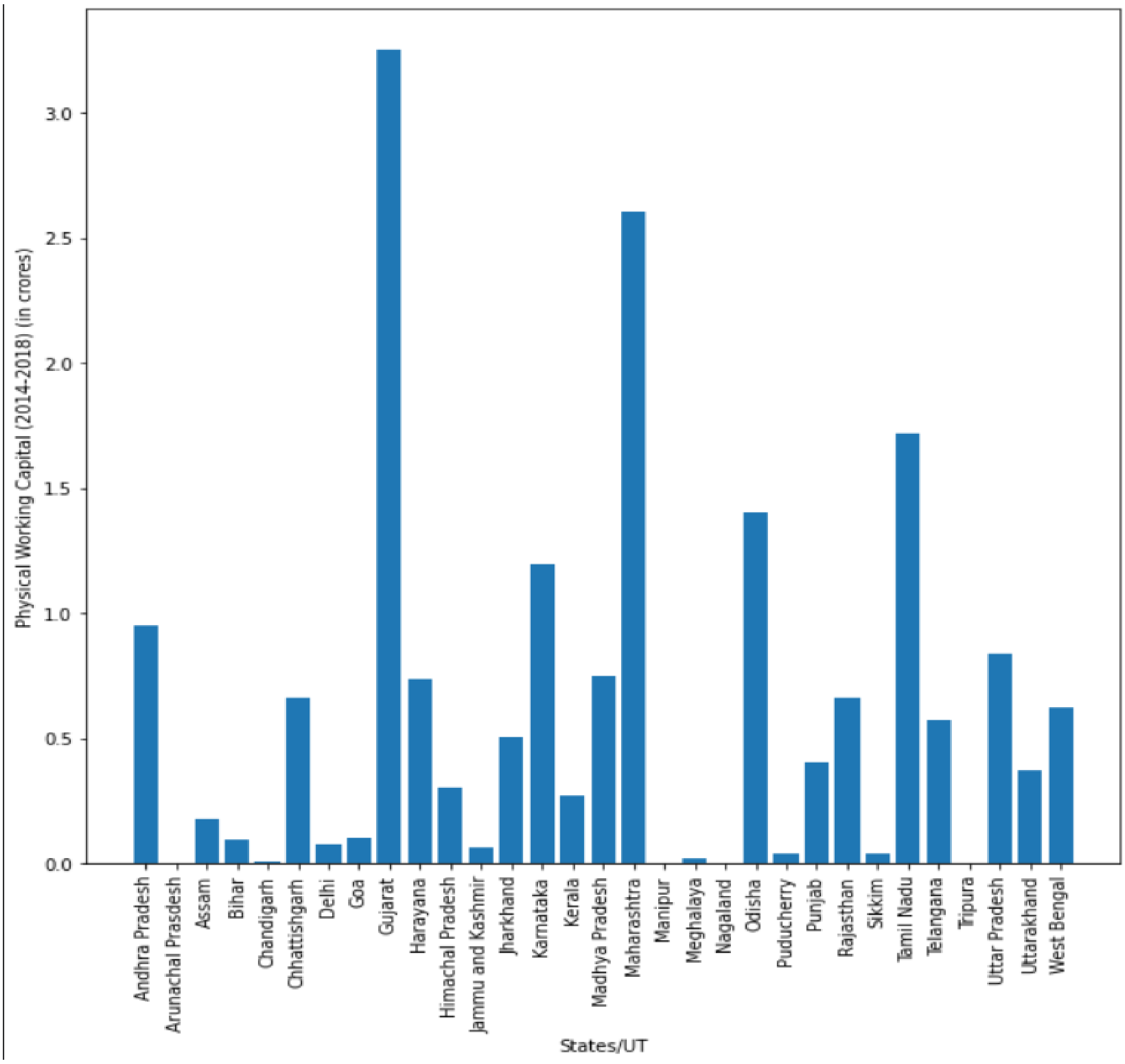


Figure 1: Bar Plot graph State wise representing Total Productive Capital of the manufacturing industries.

```
import pandas as pd

data2 = pd.read_csv("productive_capital.csv")
data2
```

	State/ UT	2014-15	2015-16	2016-17	2017-18	2018-19	Total
0	Andhra Pradesh	15392577	18487203	18433656	19878183	22855973	95047592
1	Arunachal Prasdesh	55218	60073	50996	35976	38600	240863
2	Assam	2730729	3927370	3563851	3639827	4034434	17896211
3	Bihar	1343628	1136371	2012130	2298576	2701644	9492349
4	Chandigarh	124731	150624	166455	150624	124731	717165
5	Chhattishgarh	12775998	11650781	13511018	13653864	14815817	66407478
6	Delhi	1937171	1553116	1493598	1151644	1362987	7498516
7	Goa	1509802	1731829	2108973	2452079	2158500	9961183
8	Gujarat	50095434	58579854	65966700	71894406	78886302	325422696
9	Harayana	11273430	12829502	14783527	15369734	19640581	73896774
10	Himachal Pradesh	6515812	7164591	5015491	5828268	5716371	30240533
11	Jammu and Kashmir	896382	983419	1519672	1438118	1373641	6211232
12	Jharkhand	7686462	8208819	10757729	11348741	12201593	50203344
13	Karnataka	21476053	22558219	24023003	24563385	26860472	119481132
14	Kerala	3787340	4746403	5816758	6290710	6359288	27000499
15	Madhya Pradesh	11022723	14994248	15710742	16577799	16499919	74805431
16	Maharashtra	44711432	57850848	54053943	47559377	55961858	260137458
17	Manipur	28287	16584	19598	17968	16752	99189
18	Meghalaya	446537	388810	359998	431938	463516	2090799
19	Nagaland	36724	39879	39738	39428	38053	193822
20	Odisha	22650016	26928696	31499704	30373705	28917201	140369322
21	Puducherry	738789	826008	746126	878077	931145	4120145
22	Punjab	7354482	8168199	7512885	8180831	8959933	40176330
23	Rajasthan	10499980	11504805	13693270	14396163	16323309	66417527
24	Sikkim	325733	449508	847909	1010273	1037943	3671366
25	Tamil Nadu	28844947	28964334	36272274	38986490	38730422	171798467
26	Telangana	8537175	10279884	10467449	12681281	15182594	57148383
27	Tripura	34600	47807	54337	63498	53591	253833
28	Uttar Pradesh	13349782	14968405	17335309	17634483	20315978	83603957
29	Uttarakhand	6421631	7145513	7658688	7804143	8332374	37362349
30	West Bengal	10868074	10946934	13923847	14026860	12629982	62395697

```
plt.figure(figsize = (10,10))
plt.bar(x=data2["State/ UT"],height=data2["Total"])

labels = ['200']

plt.xticks(rotation = 90)

plt.xlabel("States/UT")
plt.ylabel("Physical Working Capital (2014-2018) (in crores)")
plt.show()
```

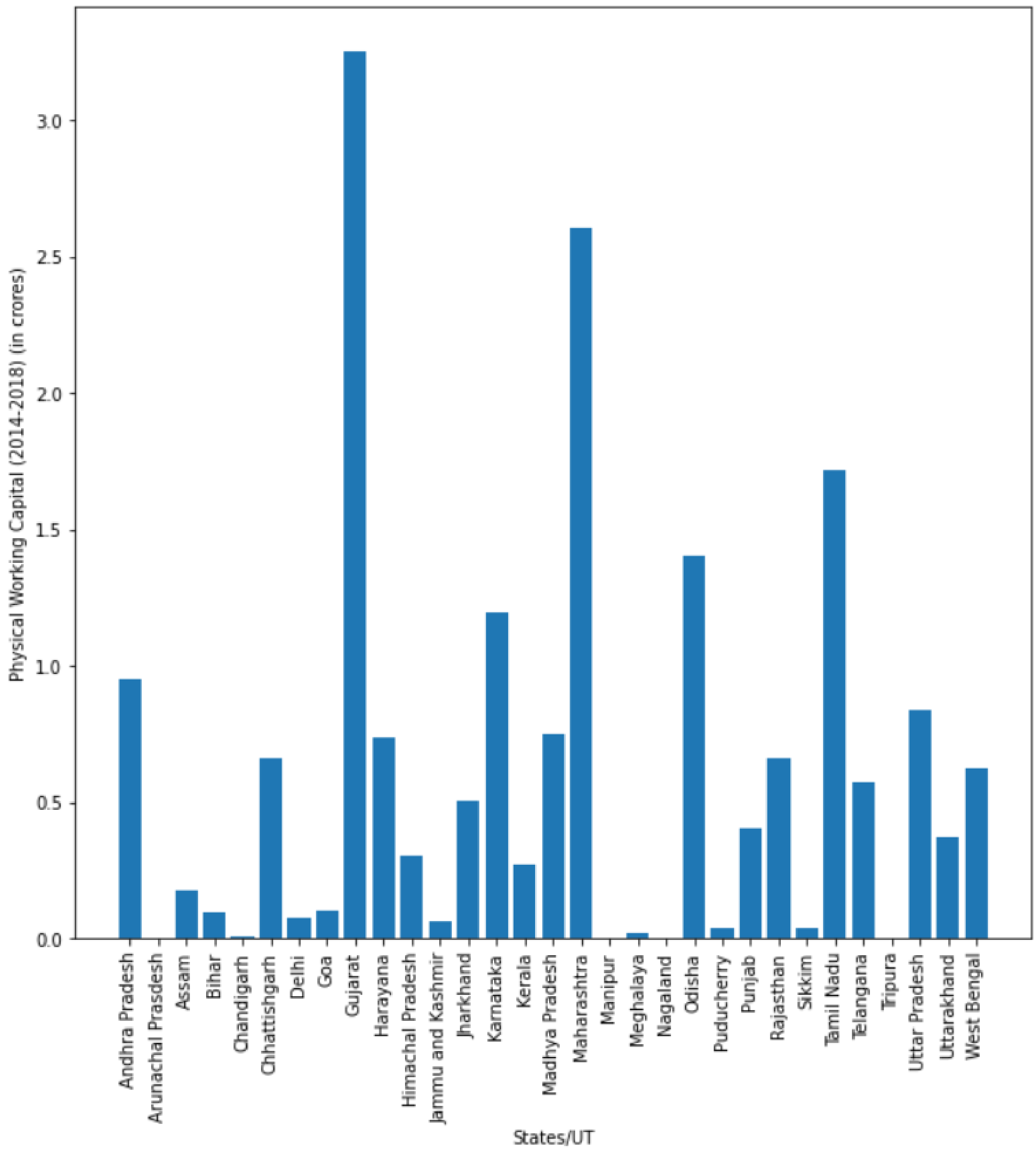


Figure 2: Bar Plot graph State wise representing the money invested in Physical Working Capital by the manufacturing industries.

```
import pandas as pd

df = pd.read_csv("Total.csv")
df
```

	State/ UT	Physical Working Capital	Productive Capital
0	Andhra Pradesh	26634696	95047592
1	Arunachal Prasdesh	89978	240863
2	Assam	4472200	17896211
3	Bihar	3074628	9492349
4	Chandigarh	336797	717165
5	Chhattishgarh	9695867	66407478
6	Delhi	4808735	7498516
7	Goa	3450485	9961183
8	Gujarat	91162863	325422696
9	Harayana	30352584	73896774
10	Himachal Pradesh	6753403	30240533
11	Jammu and Kashmir	1925121	6211232
12	Jharkhand	11695701	50203344
13	Karnataka	36443417	119481132
14	Kerala	8682902	27000499
15	Madhya Pradesh	16323761	74805431
16	Maharashtra	88956452	260137458
17	Manipur	32503	99189
18	Meghalaya	549343	2090799
19	Nagaland	56403	193822
20	Odisha	18169187	140369322
21	Puducherry	1596469	4120145
22	Punjab	19709275	40176330
23	Rajasthan	19302473	66417527
24	Sikkim	769813	3671366
25	Tamil Nadu	60652768	171798467
26	Telangana	18741604	57148383
27	Tripura	124165	253833
28	Uttar Pradesh	37592086	83603957
29	Uttarakhand	11225426	37362349
30	West Bengal	23959736	62395697

For [Pearson Correlation]

```
import numpy as np
import matplotlib.pyplot as plt

x = df['Physical Working Capital']
y = df['Productive Capital']

n = np.size(x)
x_mean = np.mean(x)
y_mean = np.mean(y)

Sxy = np.sum(x*y)- n*x_mean*y_mean
Sxx = np.sum(x*x)-n*x_mean*x_mean
slope = Sxy/Sxx
y_intercept = y_mean-slope*x_mean

trend_line = [(slope * x) + y_intercept for x in x]  # List of predictions on basis of

plt.scatter(x,y)
plt.plot(x, trend_line, color='r')
plt.xlabel('Physical Working Capital (in crores)')
plt.ylabel('Productive Capital (in crores)')
plt.show()
```

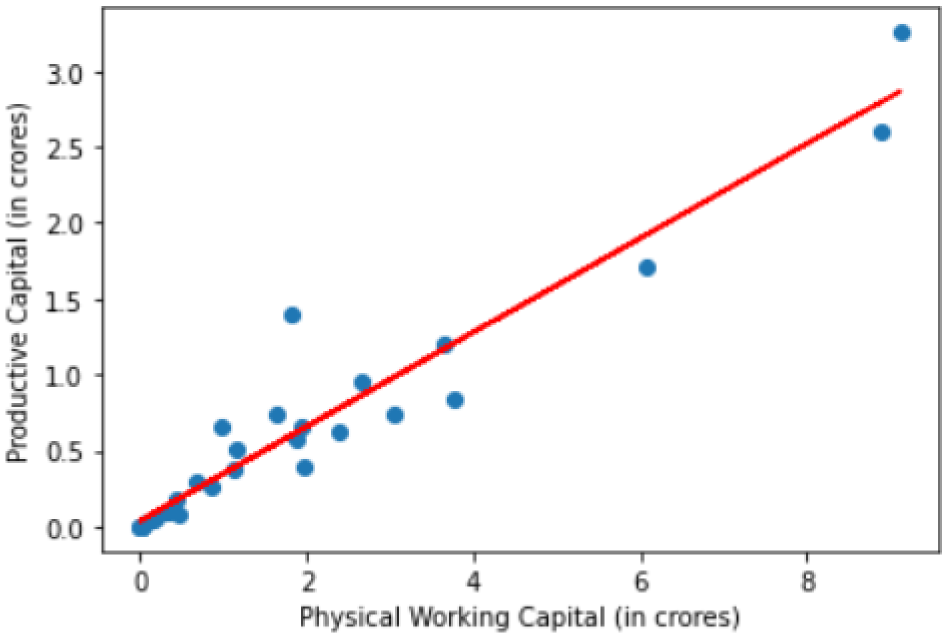


Figure 3: Scatter Plot graph representing the relation between the Physical Working Capital and Productive Capital for the manufacturing Industries for the year 2014-18

For objective 2

Tables used

```python
import pandas as pd
table2_1 = pd.read_csv("Contribution of MSME.csv")
table2_1
```

	Year	All-India
0	2005	25990686
1	2006	31186419
2	2007	39572526
3	2008	48159268
4	2009	52776558
5	2010	59211387
6	2011	70457581
7	2012	76455389
8	2013	85194869
9	2014	89534187
10	2015	97516172
11	2016	107248509
12	2017	114591911
13	2018	122967418
14	2019	127369428

```python
import pandas as pd
table2_2 = pd.read_csv("Growth of MSME.csv")
table2_2
```

	Year	Industry Sector	MSME Sector
0	2000-01	5.00	8.00
1	2001-02	2.70	6.10
2	2002-03	5.70	8.68
3	2003-04	7.00	9.64
4	2004-05	8.40	10.80
5	2005-06	8.20	12.32
6	2006-07	11.60	12.60
7	2007-08	8.50	13.00
8	2008-09	12.80	10.30
9	2009-10	10.40	6.70
10	2010-11	7.40	4.40
11	2011-12	7.80	4.40
12	2012-13	6.85	5.60
13	2013-14	6.76	9.10
14	2014-15	6.54	8.56
15	2015-16	6.60	7.62

[Descriptive Statistics]

```
import pandas as pd
data = pd.read_csv("Contribution of MSME.csv")

data.describe()
```

	Year	All-India
count	15.000000	1.500000e+01
mean	2012.000000	7.654882e+07
std	4.472136	3.327384e+07
min	2005.000000	2.599069e+07
25%	2008.500000	5.046791e+07
50%	2012.000000	7.645539e+07
75%	2015.500000	1.023823e+08
max	2019.000000	1.273694e+08

```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
data = pd.read_csv("Contribution of MSME.csv")
data.plot.bar(x='Year',y='All-India')
plt.xlabel('Year')
plt.ylabel('Contribution of MSME in GDP (in Lakhs)')
plt.title('Yearly Contribution of MSME in GDP (in Lakhs)')
```

```
Text(0.5, 1.0, 'Yearly Contribution of MSME in GDP (in Lakhs)')
```

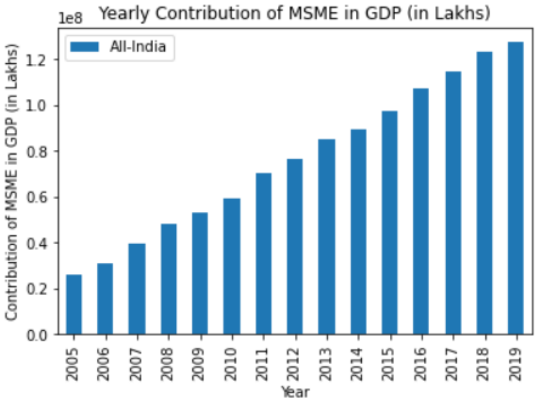


Figure 4: Yearly Contribution of MSME in GDP (in Lakhs)

For T-test

```
import pandas as pd
from statsmodels.stats.weightstats import ttest_ind

df= pd.read_csv("Growth of MSME.csv")

x= df["Industry Sector"]
y= df["MSME Sector"]

ttest_ind(x,y,value=0)
```

```
(-1.0539517427824379, 0.30031991935120755, 30.0)
```

```
import pandas as pd
import matplotlib.pyplot as plt
fig,ax=plt.subplots(figsize=(20,5))
df=pd.read_csv("Growth of MSME.csv")

plt.plot(df["Year"],df["Industry Sector"],color="red", marker="o" ,label="Industry Sector")
plt.plot(df["Year"],df["MSME Sector"],color="blue", marker="o", label="MSME Sector")

labels= ax.get_xticklabels()
plt.xlabel("Year")
plt.ylabel("Growth Percentage")
plt.grid(True)
plt.legend()
plt.show()
```

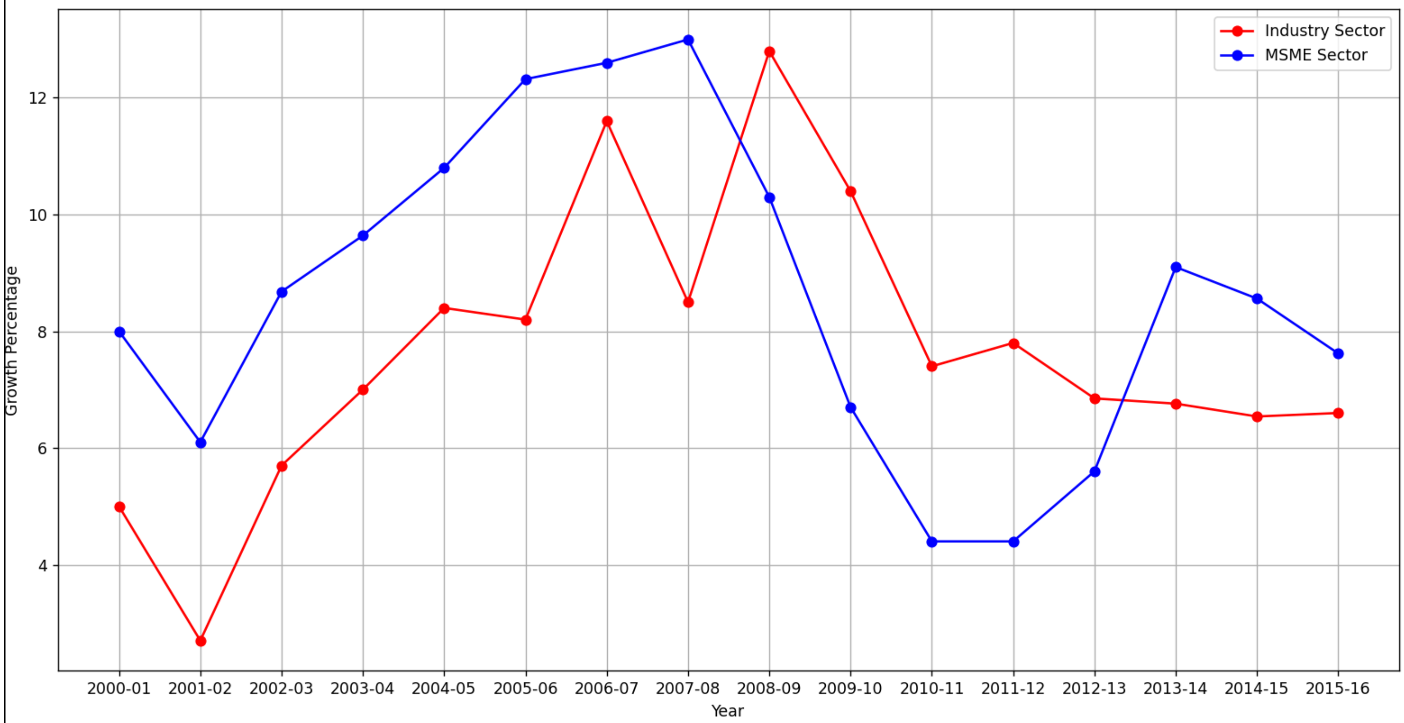


Figure 5: Previous year trend of Growth between MSME Sector and Industry Sector

For Linear Regression

Showing Data

```
import numpy as np
import matplotlib.pyplot as plt  # To visualize
import pandas as pd  # To read data
from sklearn.linear_model import LinearRegression

data = pd.read_csv("Contribution of MSME.csv")
data
```

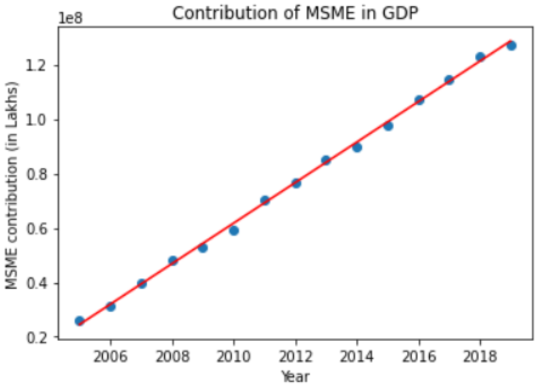
	Year	All-India
0	2005	25990686
1	2006	31186419
2	2007	39572526
3	2008	48159268
4	2009	52776558
5	2010	59211387
6	2011	70457581
7	2012	76455389
8	2013	85194869
9	2014	89534187
10	2015	97516172
11	2016	107248509
12	2017	114591911
13	2018	122967418
14	2019	127369428

Linear Regression Graph Before Prediction

```
X = data.iloc[:, 0].values.reshape(-1, 1)  # values converts it into a numpy array
Y = data.iloc[:, 1].values.reshape(-1, 1)  # -1 means that calculate the dimension of rows, but have 1 column

reg= LinearRegression()  # create object for the class
reg.fit(X, Y)  # perform linear regression
Y_pred = reg.predict(X)  # make predictions

plt.xlabel('Year')
plt.ylabel('MSME contribution (in Lakhs)')
plt.title('Contribution of MSME in GDP')
plt.scatter(X, Y)
plt.plot(X, Y_pred, color='red')
plt.show()
```



Prediction Data

```
next_X= np.array([[i for i in range(2020,2036)]]).T
next_Y = reg.predict(next_X)
cont_X=np.vstack((X,next_X))
pred_Y=reg.predict(cont_X)

plt.xlabel('Year')
plt.ylabel('MSME contribution (in Lakhs)')
plt.title('Contribution of MSME in GDP')
plt.scatter(X, Y, c ="blue",label="Observed data")
plt.scatter(next_X, next_Y, c ="green",label="Predicted data")
plt.plot(cont_X, pred_Y, color='red')
plt.legend()
plt.show()
```

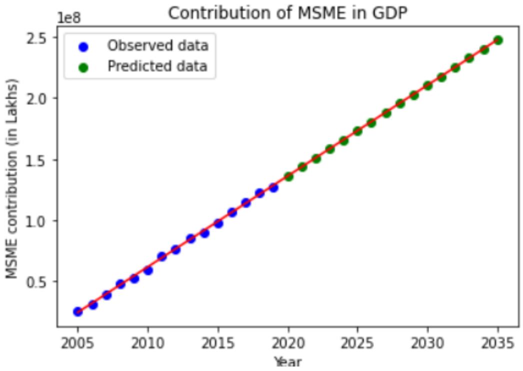


Figure 6: Scattered plot of contribution of MSME sector in GDP along with predicted data until 2035

Objective 3:

Table used:

```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
data  = pd.read_csv('ThermalvSolar.csv')
data
```

✓ 0.5s

	States/Uts	Thermal energy	Solar Energy Output
0	Andaman and Nicobar Islands	40.50	29.22
1	Andhra Pradesh	14714.46	4203.00
2	Arunachal Pradesh	83.87	5.61
3	Assam	1167.44	42.99
4	Bihar	6528.21	159.51
5	Chhattisgarh	12221.89	45.16
6	Chandigarh	59.86	252.48
7	Dadra & Nagar Haveli	488.78	5.46
8	Daman & Diu	208.08	40.55
9	Delhi	6520.93	192.97
10	Goa	559.94	7.44
11	Gujarat	24289.09	4430.82
12	Haryana	9322.19	407.83
13	Himachal Pradesh	213.70	42.73
14	Jammu and Kashmir & Ladakh	881.22	20.73
15	Jharkhand	2426.50	52.06
16	Karnataka	10343.40	7355.17
17	Kerala	3066.66	257.00
18	Lakshadweep	0.00	0.75
19	Madhya Pradesh	16419.48	2463.22
20	Maharashtra	28766.91	2289.97
21	Manipur	154.67	6.36
22	Meghalaya	161.29	0.12
23	Mizoram	71.51	1.53
24	Nagaland	81.03	1.00
25	NLC	166.00	0.00
26	Odisha	5027.21	401.72
27	Puducherry	285.10	9.33
28	Punjab	8729.51	959.50
29	Rajasthan	14004.49	5732.58
30	Sikkim	50.27	0.07
31	Tamil Nadu	14839.17	4475.21
32	Telangana	10211.19	3953.12
33	Tripura	630.05	9.41
34	Uttar Pradesh	20303.33	1712.50
35	Uttarakhand	1011.26	368.41
36	West Bengal	12444.36	149.84

```python
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
regData  = pd.read_csv('Regression.csv')
regData
```

✓ 0.6s

	Years	coal yearwise	power yearwise
0	2010	532.042	83059
1	2011	532.694	86159
2	2012	539.950	93720
3	2013	556.402	99811
4	2014	565.765	100226
5	2015	609.179	106892
6	2016	639.230	111441
7	2017	657.868	114293
8	2018	675.400	119215
9	2019	728.718	127456
10	2020	730.874	129101
11	2021	716.084	127553

Figure 7: Line chart showing thermal and solar energies for respective states in India for the year 2021

```python
import pandas as pd
import matplotlib.pyplot as plt

X = data["Thermal energy"]
Y = data["Solar Energy Output"]
St = data["States/Uts"]

fig, ax = plt.subplots(figsize = (20,5))  # Don't forget fig is the reso and ax is axes subplot
plt.plot(St,X, color= 'red' ,marker = 'o',label='Thermal Energy' )
plt.plot(St,Y, color= 'blue',marker = 'o',label='Solar Energy' )
plt.title("Yearwise coal Production", fontsize=14)
labels = ax.get_xticklabels( )
plt.setp(labels, rotation=45, horizontalalignment= 'right' )
plt.xlabel("Time( in years )"  , fontsize=14)
plt.ylabel("Coal Production (In Million Tonnes)" , fontsize=14)
plt.grid (True)
plt.legend()
plt.show ()
```

✓ 0.3s

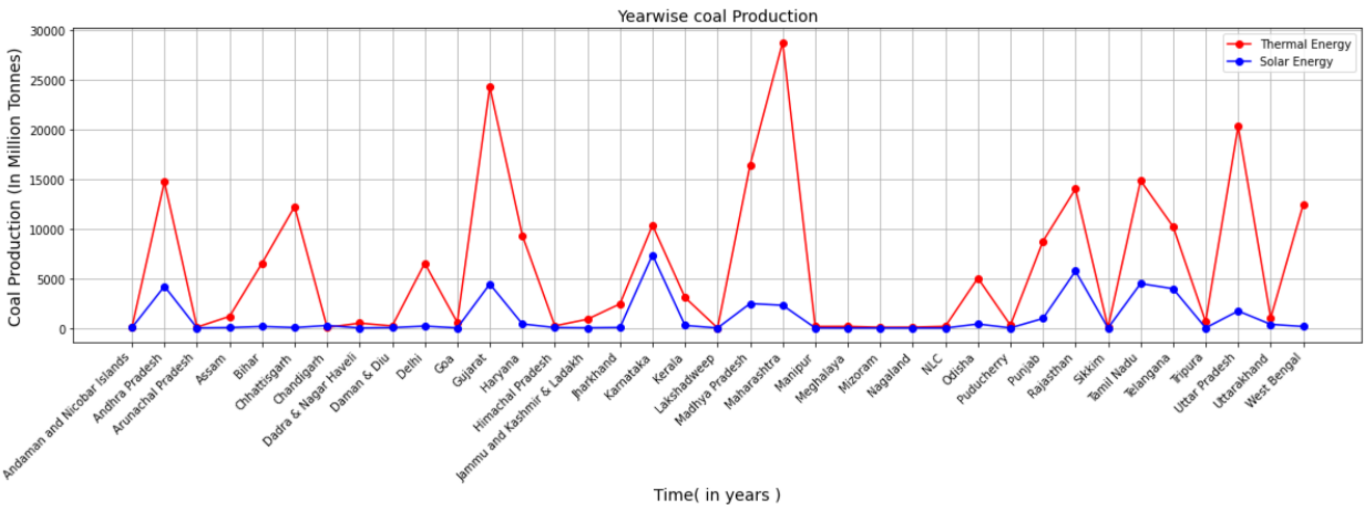


Figure 8: Scattered plot for coal production and power generation in India of past 11 years

```
plt.scatter(b,c , color="b",s=50)
n = np.size(b)
x_mean = np.mean(b)
y_mean = np.mean(c)
Sxy = np.sum(b*c)- n*x_mean*y_mean
Sxx = np.sum(b*b)-n*x_mean*x_mean
M = Sxy/Sxx
C = y_mean-M*x_mean

regline = [(M * x) + C for x in b]  # List of predictions on basis of years....

plt.scatter(b,c)
plt.plot(b,regline,color = 'r')
plt.xticks(rotation=85)
plt.title("coal production and power rq. of a particular year", loc="right")
plt.xlabel("Coal production (in MT)")
plt.ylabel("Power Generation (in MW)")

plt.show()
```

✓ 0.5s

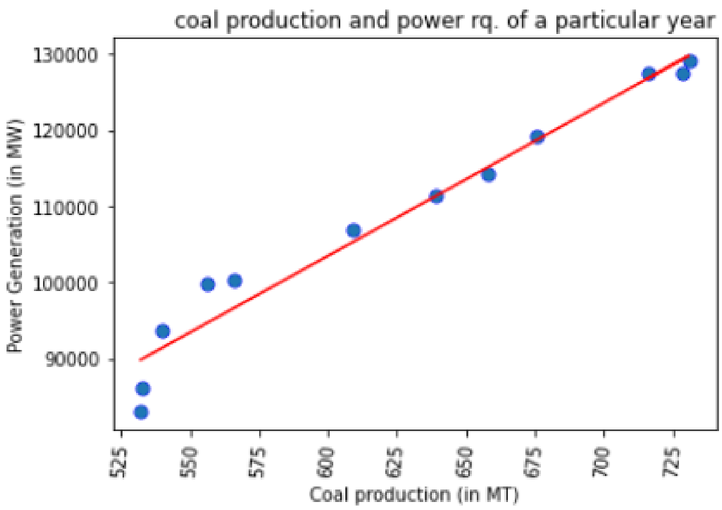


Figure 9: Year wise power requirements of India past 11 years

```
import numpy as np
import matplotlib.pyplot as plt  # To visualize
import pandas as pd  # To read data
regData  = pd.read_csv('Regression.csv')
a = regData["Years"]              #years
b = regData["coal yearwise"]      # coal in million tonnes
c = regData["power yearwise"]     # power yearwise

n = np.size(a)
x_mean = np.mean(a)
y_mean = np.mean(c)
Sxy = np.sum(a*c)- n*x_mean*y_mean
Sxx = np.sum(a*a)-n*x_mean*x_mean
M = Sxy/Sxx
C = y_mean-M*x_mean

regline = [(M * x) + C for x in a]  # List of predictions on basis of years....

print('slope M is', round(M,3))
print('intercept C is', round(C,3))
print("And the required Equation is Y = ",M,"* X +",C)

plt.xticks( rotation = 80)
plt.xlabel("Time (in years)")
plt.ylabel("Power required by the country(in MW)")
plt.title("Power Requirement in India", loc="right")
plt.scatter(a, c)

plt.plot(a, regline, color='Magenta')
plt.show()
```

```
slope M is 4391.028
intercept C is -8741873.044
And the required Equation is Y =  4391.027972027972 * X + -8741873.044289045
```

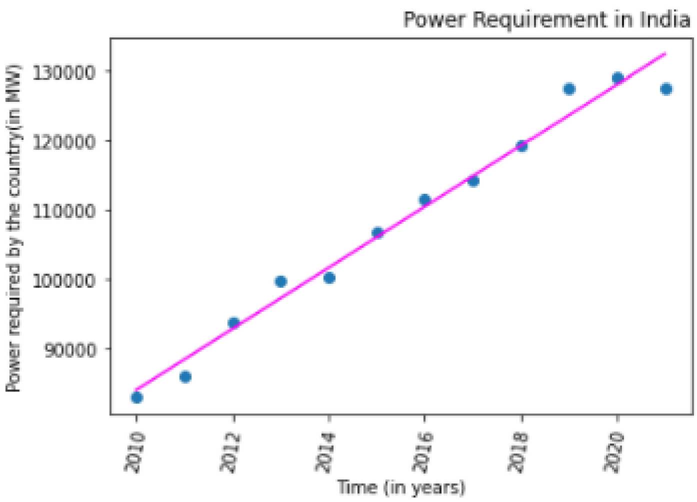


Figure 10: Year wise power requirement in India predicted values

```
# Predction code..
import numpy as np
import matplotlib.pyplot as plt  # To visualize
import pandas as pd  # To read data
regData  = pd.read_csv('Regression.csv')
a = regData["Years"]              #years
b = regData["coal yearwise"]      # coal in million tonnes
c = regData["power yearwise"]     # power yearwise

n = np.size(a)
x_mean = np.mean(a)
y_mean = np.mean(c)
Sxy = np.sum(a*c)- n*x_mean*y_mean
Sxx = np.sum(a*a)-n*x_mean*x_mean
M = Sxy/Sxx
C = y_mean-M*x_mean

futureYears = np.array([i for i in range(2022,2034)])
allYears = np.append(a,futureYears)
regline = [(M * x) + C for x in futureYears]   # List of predictions on basis of years....
predvalsForRegline = [(M * x) + C for x in allYears]
allVals = np.append(c,regline)

plt.xlabel('Time (in Year)')
plt.ylabel('Power Power requirement in India(in MW)')
plt.title('Power requirement in India',loc="right")

plt.scatter(a, c, c ="blue",label="Observed data")
plt.scatter(futureYears,regline, c ="red",label="Predicted data")
plt.plot(allYears,predvalsForRegline , color='magenta')   # Reg is applied on all values from 2010 to
plt.legend()
plt.show()
```

✓ 0.2s

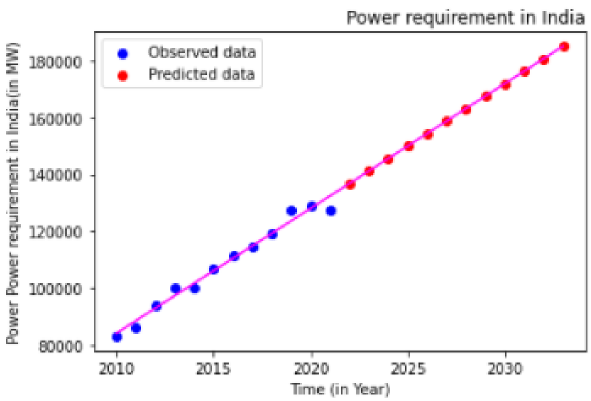
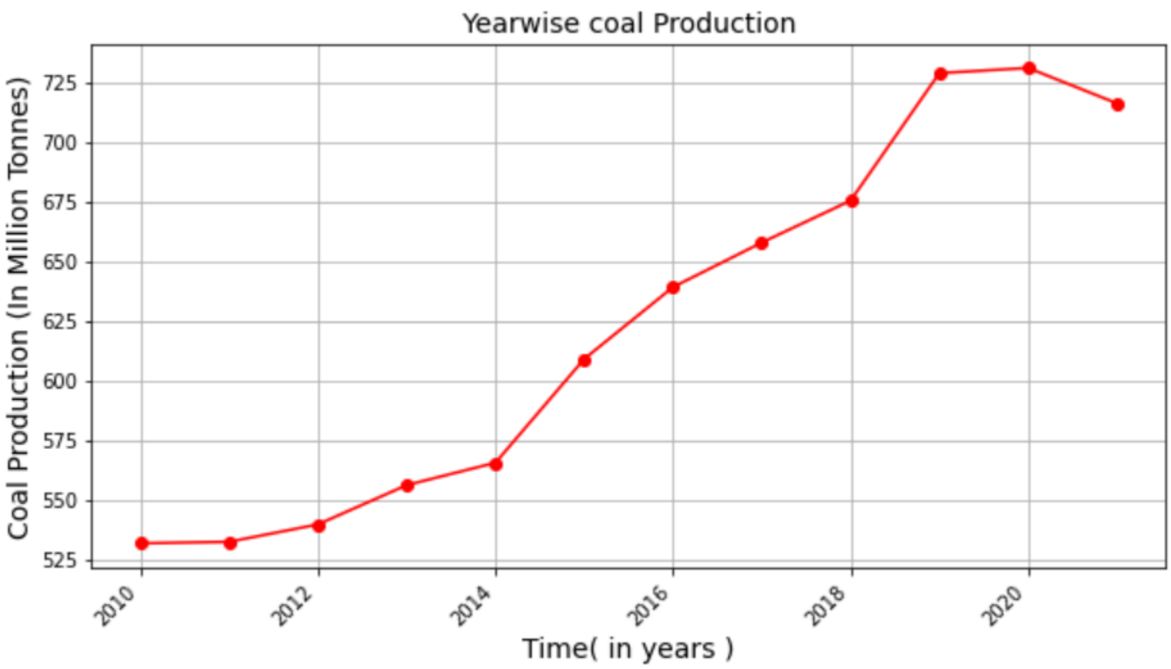


Figure 11: Year wise coal production in past 11 years

```
"And as we can see that there is a recent fall in coal production hence "
import pandas as pd
import matplotlib.pyplot as plt
fig, ax = plt.subplots(figsize = (10,5))
plt.plot(a,b, color= 'red',marker = 'o' )
plt.title("Yearwise coal Production", fontsize=14)
labels = ax.get_xticklabels( )
plt.setp(labels, rotation=45, horizontalalignment= 'right' )
plt.xlabel( "Time( in years )"  , fontsize=14)
plt.ylabel( "Coal Production (In Million Tonnes)" , fontsize=14)
plt.grid ( True)
plt.show ( )
```



Calculation of Z test (section 6.7)

```python
def twoSampz(X,Y):
    from numpy import sqrt, abs, round

    X1 = X.mean()
    X2 = Y.mean()
    sd1 = X.std()
    sd2 = Y.std()
    n1 = X.count()
    n2 = Y.count()
    Denom = sqrt(sd1*sd1/n1 + sd2*sd2/n2)
    z = (X1 - X2)/Denom
    if abs(z)< 1.96:
        return round(z, 3), "The null hypothesis is accepted, there is no significant difference between the means of the two variables"
    elif abs(z)> 1.96:
        return round(z, 3), "The null hypothesis is rejected,there is a significant difference between the means of the two variables"
print(twoSampz(X,Y))
```

```
(3.845, 'The null hypothesis is rejected,there is a significant difference between the means of the two variables')
```

For Objective 4 [Descriptive stats]

```python
import matplotlib.pyplot as plt
import numpy as np

total = [2,5,7,3,2,2,1,1]

money_range = ['0-300','300-600','600-900','900-1200','1200-1500','1500-1800','18000-2100','2100-2400']

indices = np.arange(len(money_range))
# width = 0.4

plt.bar(indices, total, color = 'b')
plt.xticks(ticks=indices, labels = money_range, rotation='vertical')

plt.xlabel('Amount of Money Spent on R&D(Total) [in crores]')
plt.ylabel('Frequency of States')
plt.title('Frequency of states having total amount of money spent on R&D in a particular range')

plt.show()
```

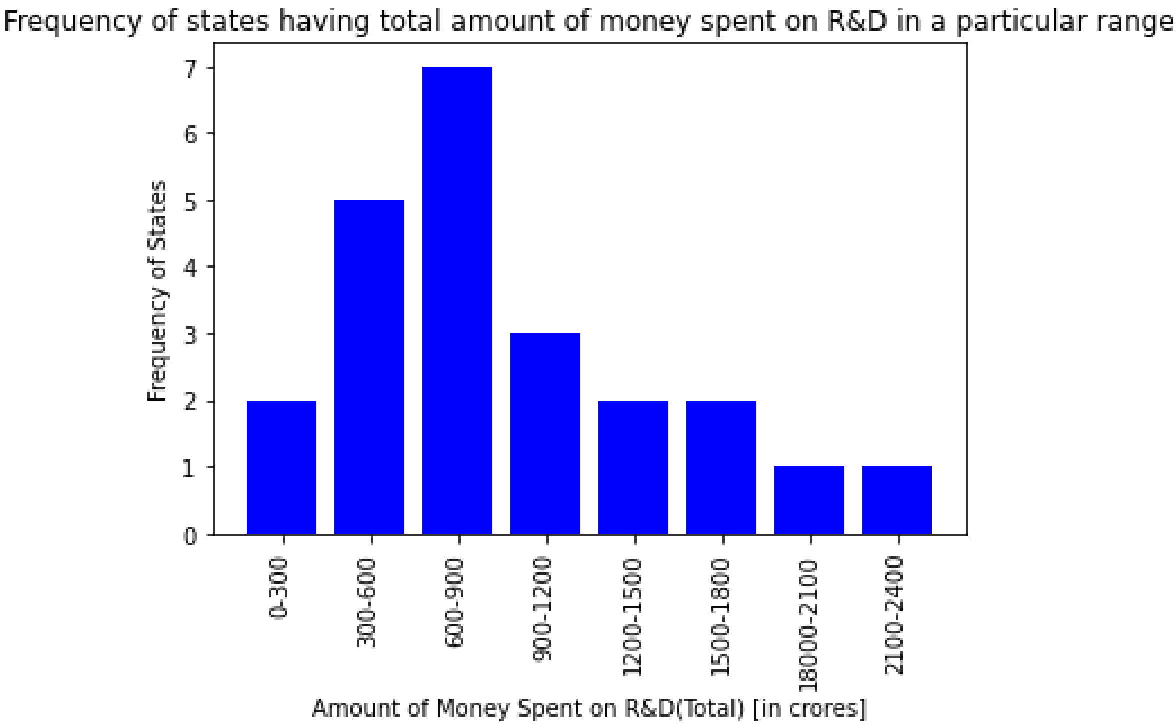
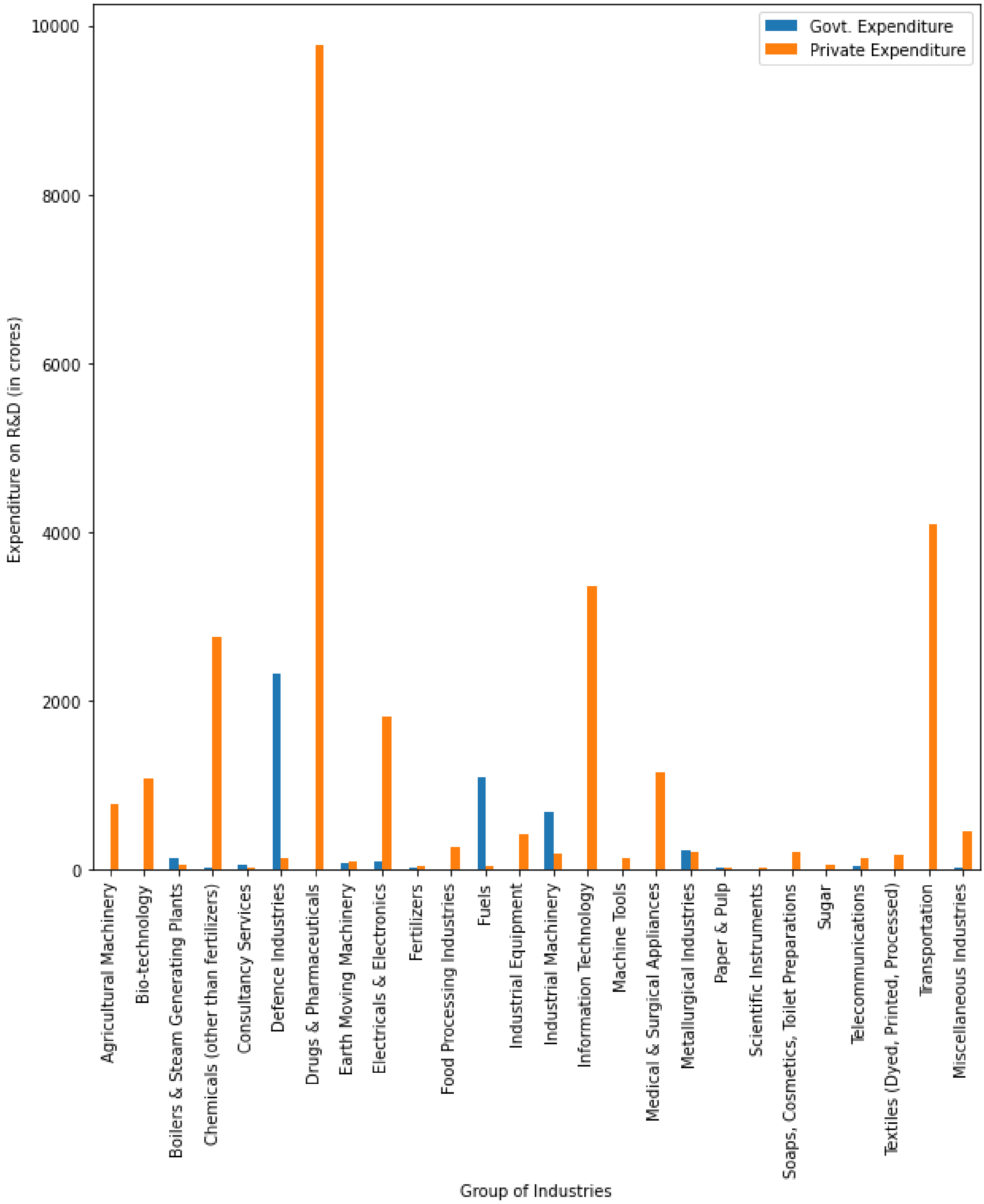


Figure 12: Frequency Distribution graph between no. of states and range of amount of money spent on R&D by states of India.

Figure 13: Bar plot graph of expenditure made by various Govt. and Private Industries of India in R&D sector of the industry.

[T-Test for Difference of Means]

```
import matplotlib as plt
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
data1.plot.bar(x='Industry Group',y=['Govt. Expenditure','Private Expenditure'],figsize=(
10,10))
plt.xlabel("Group of Industries")
plt.ylabel("Expenditure on R&D (in crores)")
```



For [Power Method]

```
import pandas as pd
matrix = pd.read_csv("Rank Table.csv")

mat = matrix.iloc[:,1:].to_numpy()
print(mat)
```

```
[[0 1 2 2 2 1 1 2 2 2]
 [2 0 2 2 2 2 2 2 2 2]
 [1 1 0 1 1 1 1 1 2 2]
 [1 1 2 0 1 1 1 1 2 2]
 [1 1 2 2 0 1 1 2 2 2]
 [2 1 2 2 2 0 1 2 2 2]
 [2 1 2 2 2 2 0 2 2 2]
 [1 1 2 2 1 1 1 0 2 2]
 [1 1 1 1 1 1 1 1 0 2]
 [1 1 1 1 1 1 1 1 1 0]]
```

```python
import numpy as np

x = []
eigen = [0]

x.append(np.array([[1],[1],[1],[1],[1],[1],[1],[1],[1],[1]]))
print(mat,'\n\n',x)
print('-----------------------------------------------')

i=0
while len(eigen) == 1 or np.round(eigen[i],4)!=np.round(eigen[i-1],4):
    x.append(np.matmul(mat,x[i]))
    eigen.append(max(x[i+1]))
    x[i+1] = np.round(x[i+1]/eigen[i+1],4)
    print('Iteration',i+1,':')
    print(x[i+1],'\t',eigen[i+1])
    print()
    i+=1
```

Computational Data Analysis –Project
Individual Contribution Report

Team No. 13 Section B Project Title: Industries, Research and Power in India: The road ahead

SDG Goal No. Planned 9 Other SDG Goals Covered project 1,11,8,7

Student ID	Name	Objective(s)	Summary on Contribution (Brief method, result summary and conclusion)	Signature
2021BTECH127	Priyanshu Joshi	To study the access of medium and small-scale manufacturing industries in financial services and their integration into the value chain	Used descriptive stats to analyze contribution of MSME in GDP, Linear Reg to predict future contribution of MSME, T-Test to compare rate of growth in Industry and MSME	
2021BTECH072	Mayank Bohra	To study and analyze the capital approach for sustainability of the manufacturing industries	Used t-test and correlation to compare the physical working and productive capital for the manufacturing industries and to determine which state contributes highest in terms of capital for the manufacturing industries, using power method.	
2021BTECH118	Vaibhav Nalotia	To study the expenditure on Research and Development Sector of different industries of India by several Govt. organization and private industrials.	Used descriptive stats to determine state-wise contribution in R&D, Used T-Test to compare mean of expenditure between private and public sector in R&D, Used Power Method to find Rank of top 10 states of India investing the more in R&D of the nation	
2021BTECH110	Steven Moses	To study the power generated by industries in India and analyze different factors that affect generation of power	Used Z-Test and correlation to determine the major factor in India's power generation and predicted future power requirements in India using linear regression.