

Exploratory Data Analysis on GDP

by
Group 10



Mihir Paija
ID: 202101205
Course: BTech ICT



Dhruv Shah
ID: 202101208
Course: BTech ICT



Nikita Shah
ID: 202101209
Course: BTech ICT



Rohan Mistry
ID: 202101231
Course: BTech ICT

Course Code: IT 462
Semester: Winter 2023

Under the guidance of

Dr. Gopinath Panda



Dhirubhai Ambani Institute of Information and Communication Technology

April 29, 2024

ACKNOWLEDGMENT

I am writing this letter to express my heartfelt gratitude for your guidance and support throughout my project titled 'Analysing and Predicting the Contribution of Various Sectors and States Towards National GDP'. Your invaluable assistance has played a pivotal role in shaping the successful completion of this endeavor.

I am incredibly fortunate to have had the opportunity to work under your mentorship. Your expertise, encouragement, and willingness to share your knowledge have been instrumental in elevating the quality and scope of my project. Your constructive feedback and insightful suggestions have helped me overcome challenges and develop a deeper understanding of the subject matter.

Furthermore, I would like to thank the entire team at Dhirubhai Ambani Institute of Information and Communication Technology for fostering an environment of collaboration and innovation. The resources and facilities provided have been crucial in conducting comprehensive research and analysis.

I would also like to express my gratitude to my peers and colleagues who have been supportive throughout this journey. Their valuable input and camaraderie have been a constant source of motivation.

Completing this project has been a tremendous learning experience. I am confident that the knowledge and skills acquired during this endeavor will be a solid foundation for my future endeavors.

Sincerely,

Mihir Paija, 202101205

Dhruv Shah, 202101208

Nikita Shah, 202101209

Rohan Mistry, 202101231

DECLARATION

We, 202101205, 202101208, 202101209, 202101231 now declare that the EDA project work presented in this report is our original work and has not been submitted for any other academic degree. All the sources cited in this report have been appropriately referenced.

We acknowledge that the data used in this project is obtained from the mospi.gov.in and worldbank.org. We also declare that we have adhered to the terms and conditions mentioned on the website for using the dataset. We confirm that the dataset used in this project is true and accurate to the best of our knowledge.

We acknowledge that we have received no external help or assistance in conducting this project except for the guidance provided by our mentor, Prof. Gopinath Panda. We declare no conflict of interest in conducting this EDA project.

We have now signed the declaration statement and confirmed the submission of this report on 29th April 2024.



Mihir Paija
ID: 202101205
Course: BTech ICT



Dhruv Shah
ID: 202101208
Course: BTech ICT



Nikita Shah
ID: 202101209
Course: BTech ICT



Rohan Mistry
ID: 202101231
Course: BTech ICT

CERTIFICATE

This is to certify that Group 10 comprising of Mihir Pajja, Dhruv Shah, Nikita Shah and Rohan Mistry has completed an exploratory data analysis (EDA) project on GDP, which was obtained from mospi.gove.in and worldbank.org

The EDA project presented by Group 10 is their original work. It was completed under the guidance of the course instructor, Prof. Gopinath Panda, who provided support and guidance throughout the project. The project is based on a thorough analysis of the GDP datasets, and the results presented in the report are based on the data obtained from the dataset.

This certificate is issued to recognize the successful completion of the EDA project on the Analysis and Prediction of Contribution of Various Sectors and States Towards National GDP, which demonstrates the analytical skills and knowledge of the students of Group 10 in the field of data analysis.



Signed,
Dr. Gopinath Panda,
IT 462 Course Instructor
Dhirubhai Ambani Institute of Information and Communication Technology
Gandhinagar, Gujarat, INDIA.

April 29, 2024

Contents

List of Figures	6
1 Introduction	1
1.1 Project idea	1
1.2 Data Collection	1
1.3 Dataset Description	1
1.3.1 World	2
1.3.1.1 GDP of G20 Countries	2
1.3.2 Contribution of Sectors	2
1.3.2.1 GVA	2
1.3.2.2 Growth Rate of GVA	3
1.3.3 Contribution of States	4
1.3.3.1 GSDP	4
1.3.3.2 Growth Rate of GSDP	5
1.4 Packages required	5
2 Data Cleaning	7
2.1 World	7
2.1.1 GDP of G20 Countries	7
2.1.1.1 Missing Data Analysis	7
2.1.1.2 Imputation	8
2.1.1.3 Changing Data Types	8
2.2 Contribution of Sectors	9
2.2.1 GVA Amount	9
2.2.1.1 Missing Data Analysis	9
2.2.1.2 Imputation	10
2.2.1.3 Changing Data Types	10
2.2.2 Growth Rate of GVA	10
2.2.2.1 Missing Data Analysis	10
2.2.2.2 Imputation	11
2.2.2.3 Changing Data Types	12
2.3 Contribution of States	12
2.3.1 GSDP	12
2.3.1.1 Missing Data Analysis	12
2.3.1.2 Imputation	13
2.3.1.3 Changing Data Types	14



2.3.2	Growth Rate of GSDP	15
2.3.2.1	Missing Data Analysis	15
2.3.2.2	Imputation	16
2.3.2.3	Changing Data Types	17
3	Feature Engineering	18
3.1	World	18
3.1.1	G20	18
3.1.1.1	Feature Extraction	18
3.2	Contribution of Sectors	18
3.2.1	GVA	18
3.2.1.1	Feature Selection	18
3.2.1.2	Feature Extraction	20
3.2.2	Growth Rate of GVA	21
3.2.2.1	Feature Selection	21
3.3	Contribution of States	23
3.3.1	GSDP	23
3.3.1.1	Feature Selection	23
3.3.1.2	Feature Extraction	25
3.3.2	Growth Rate of GSDP	26
3.3.2.1	Feature Selection	26
4	Visualization	27
4.1	World	27
4.1.1	GDP of G20 Countries	27
4.1.1.1	Univariate Analysis	27
4.1.1.2	Bivariate Analysis	40
4.1.1.3	Multivariate analysis	40
4.2	Contribution of Sectors	42
4.2.1	GVA	42
4.2.1.1	Univariate Analysis	42
4.2.1.2	Bivariate Analysis	52
4.2.1.3	Multivariate analysis	53
4.2.2	Growth Rate of GVA	58
4.2.2.1	Univariate Analysis	58
4.2.2.2	Bivariate Analysis	66
4.2.2.3	Multivariate analysis	67
4.3	Contributions of States	68
4.3.1	GSDP	68
4.3.1.1	Univariate Analysis	68
4.3.1.2	Bivariate Analysis	91
4.3.1.3	Multivariate analysis	91
4.3.2	Growth Rate of GSDP	98
4.3.2.1	Univariate Analysis	98
4.3.2.2	Bivariate Analysis	120
4.3.2.3	Multivariate analysis	122



5 Model Fitting	126
5.1 Contributions of Sectors	126
5.1.1 GVA	126
5.1.1.1 Linear Regression	126
5.1.1.2 Polynomial Extrapolation	128
5.1.1.3 Comparison of Models	130
5.2 Contributions of States	133
5.2.1 GSDP	133
5.2.1.1 Linear Regression	133
5.2.1.2 Polynomial Extrapolation	139
5.2.1.3 Comparison of Models	144
6 Conclusion & future scope	152
6.1 Findings and Observations	152
6.1.1 World	152
6.1.2 Contribution of Sectors	152
6.1.3 Contribution of States	153
6.1.4 Predictions	153
6.2 Challenges	155
6.3 Future plan	155
6.4 Conclusion	156

List of Figures

1.1	GDP Data of G20 Countries	2
1.2	National Annual GVA	3
1.3	National Annual Growth Rate	4
1.4	State GSDP	5
1.5	State GSDP Growth Rate	5
2.1	Columns Information After Transformation	8
2.2	Columns Information After Transformation	9
2.3	Columns Information After Transformation	10
2.4	Columns Information after Transformation	10
2.5	Growth Rate of GVA - Missing Value Table	11
2.6	Growth Rate of GVA - Missing Value Matrix	11
2.7	Imputed Value Table	12
2.8	Columns Information after Transformation	12
2.9	GSDP - Missing Value Tables	13
2.10	GSDP - Missing Value Matrix	13
2.11	GSDP - Imputed Table	14
2.12	GSDP - Imputed Bar	14
2.13	Growth Rate of GSDP - After Transformation	15
2.14	Growth Rate of GSDP - Missing Value Tables	15
2.15	Growth Rate of GSDP - Missing Value Matrix	16
2.16	Growth Rate of GSDP - Imputed Tables	16
2.17	Growth Rate of GSDP - Interpolated Matrix	17
2.18	Growth Rate of GSDP - After Transformation	17
3.1	Format in billion \$	18
3.2	Transposed Dataframe of G20 Countries	18
3.3	Dataframe with Dropped Columns	19
3.4	Dataframe with Sectors	19
3.5	Dataframe with Categories	20
3.6	Data Distribution of Contribution of Sectors along with Outliers	21
3.7	Comparison of Data Distribution Before and After Normalisation	21
3.8	Dataframe with Dropped Columns	22
3.9	Dataframe with Dropped Columns	22
3.10	Dataframe with Sectors	22
3.11	Transposed Dataframe of GSDP Dataframe	23
3.12	Data of Each Zone	25



3.13 Data Distribution of CSDP of States along with Outliers	25
3.14 Comparison of Data Distribution Before and After Normalisation	26
3.15 Transposed Dataframe of Growth Rate of CSDP Dataframe	26
3.16 Adding GDP Column	26
4.1 Data Distribution of GDP of Argentina	27
4.2 Value Distribution of GDP of Argentina	27
4.3 Data Distribution of GDP of Australia	28
4.4 Value Distribution of GDP of Australia	28
4.5 Data Distribution of GDP of Brazil	28
4.6 Value Distribution of GDP of Brazil	29
4.7 Data Distribution of GDP of Canada	29
4.8 Value Distribution of GDP of Canada	29
4.9 Data Distribution of GDP of China	30
4.10 Value Distribution of GDP of China	30
4.11 Data Distribution of GDP of France	30
4.12 Value Distribution of GDP of France	31
4.13 Data Distribution of GDP of Germany	31
4.14 Value Distribution of GDP of Germany	31
4.15 Data Distribution of GDP of India	32
4.16 Value Distribution of GDP of India	32
4.17 Data Distribution of GDP of Indonesia	32
4.18 Value Distribution of GDP of Indonesia	33
4.19 Data Distribution of GDP of Italy	33
4.20 Value Distribution of GDP of Italy	33
4.21 Data Distribution of GDP of Japan	34
4.22 Value Distribution of GDP of Japan	34
4.23 Data Distribution of GDP of Mexico	34
4.24 Value Distribution of GDP of Mexico	35
4.25 Data Distribution of GDP of Russian	35
4.26 Value Distribution of GDP of Russian	35
4.27 Data Distribution of GDP of Saudi Arabia	36
4.28 Value Distribution of GDP of Saudi Arabia	36
4.29 Data Distribution of GDP of South Africa	36
4.30 Value Distribution of GDP of South Africa	37
4.31 Data Distribution of GDP of South Korea	37
4.32 Value Distribution of GDP of South Korea	37
4.33 Data Distribution of GDP of Turkey	38
4.34 Value Distribution of GDP of Turkey	38
4.35 Data Distribution of GDP of United Kingdom	38
4.36 Value Distribution of GDP of United Kingdom	39
4.37 Data Distribution of GDP of United States	39
4.38 Value Distribution of GDP of United States	39
4.39 Comparison with India's GDP	40
4.40 Growth of the GDP of G20 Countries Over the Years in Terms of Absolute Value	41
4.41 GDP of G20 Countries in 2023	41



4.42 Data Distribution of Agriculture, forestry & fishing	42
4.43 Value Distribution of Agriculture,forestry & fishing	42
4.44 Data Distribution of Mining & quarrying	43
4.45 Value Distribution of Mining & quarrying	43
4.46 Data Distribution of Manufacturing	43
4.47 Value Distribution of Manufacturing	44
4.48 Data Distribution of Electricity, gas, water supply & other utility services	44
4.49 Value Distribution of Electricity, gas, water supply & other utility services	44
4.50 Data Distribution of Construction	45
4.51 Value Distribution of Construction	45
4.52 Data Distribution of Trade, hotels, transport, communication and services related to broadcasting	45
4.53 Value Distribution of Trade, hotels, transport, communication and services related to broadcasting	46
4.54 Data Distribution of Financial , real estate & prof servs	46
4.55 Value Distribution of Financial , real estate & prof servs	46
4.56 Data Distribution of Public Administration, defence and other services	47
4.57 Value Distribution of Public Administration, defence and other services	47
4.58 Data Distribution of CVA at basic prices	47
4.59 Value Distribution of CVA at basic prices	48
4.60 Data Distribution of NVA at basic prices	48
4.61 Value Distribution of NVA at basic prices	48
4.62 Data Distribution of GNI	49
4.63 Value Distribution of GNI	49
4.64 Data Distribution of NNI	49
4.65 Value Distribution of NNI	50
4.66 Data Distribution of Per capita income(Rs.)	50
4.67 Value Distribution of Per capita income(Rs.)	50
4.68 Data Distribution of Net taxes on Products	51
4.69 Value Distribution of Net taxes on Products)	51
4.70 Data Distribution of GDP	51
4.71 Value Distribution of GDP	52
4.72 Relation of CVA, GNI, NNI and NVA With GDP	52
4.73 Growth of Contribution of Sectors Over The Years in Terms of Absolute Value	53
4.74 Contribution of Sectors Over The Years - Stacked	53
4.75 Contribution of Sectors Over the Years in Terms of %	54
4.76 Growth of Contribution of Categories Over The Years in Terms of Absolute Value	55
4.77 Contribution of Categories Over the Years - Side By Side	56
4.78 Contribution of Categories Over the Years- Stacked	56
4.79 Contribution of Categories Over the Years in Terms of %	57
4.80 Data Distribution of Growth Rate of Agriculture, forestry & fishing	58
4.81 Trend of Growth of Agriculture,forestry & fishing	58
4.82 Data Distribution of Growth Rate of Mining & quarrying	59
4.83 Trend of Growth of Mining & quarrying	59
4.84 Data Distribution of Growth Rate of Manufacturing	59
4.85 Trend of Growth of Manufacturing	60



4.86 Data Distribution of Growth Rate of Electricity, gas, water supply & other utility services	60
4.87 Trend of Growth of Electricity, gas, water supply & other utility services	60
4.88 Data Distribution of Growth Rate of Construction	61
4.89 Trend of Growth of Construction	61
4.90 Data Distribution of Growth Rate of Trade, hotels, transport, communication and services related to broadcasting	61
4.91 Trend of Growth of Trade, hotels, transport, communication and services related to broadcasting	62
4.92 Data Distribution of Growth Rate of Financial , real estate & prof servs	62
4.93 Trend of Growth of Financial , real estate & prof servs	62
4.94 Data Distribution of Growth Rate of Public Administration, defence and other services	63
4.95 Trend of Growth of Public Administration, defence and other services	63
4.96 Data Distribution of Grwoth Rate of GVA at basic prices	63
4.97 Trend of Growth of GVA at basic prices	64
4.98 Data Distribution of Growth Rate of Per capita income(Rs.)	64
4.99 Trend of Growth of Per capita income(Rs.)	64
4.100Data Distribution of Growth Rate of Net taxes on Products	65
4.101Trend of Growth of Net taxes on Products)	65
4.102Data Distribution of Growth Rate of GDP	65
4.103Trend of Growth of GDP	66
4.104Comparison of Growth Rate of Various Sectors With GDP's Growth Rate	66
4.105Growth of Sectors Over The Years	67
4.106Contribution of Sectors Over The Years in Terms of Growth Rate	67
4.107Correlation of Growth of Sectors	68
4.108Data Distribution of GSDP of Andhra Pradesh	69
4.109Value Distribution of GSDP of Andhra Pradesh	69
4.110Data Distribution of GSDP of Arunachal Pradesh	69
4.111Value Distribution of GSDP of Arunachal Pradesh	70
4.112Data Distribution of GSDP of Assam	70
4.113Value Distribution of GSDP of Assam	70
4.114Data Distribution of GSDP of Bihar	71
4.115Value Distribution of GSDP of Bihar	71
4.116Data Distribution of GSDP of Chhattisgarh	71
4.117Value Distribution of GSDP of Chhattisgarh	72
4.118Data Distribution of GSDP of Goa	72
4.119Value Distribution of GSDP of Goa	72
4.120Data Distribution of GSDP of Gujarat	73
4.121Value Distribution of GSDP of Gujarat	73
4.122Data Distribution of GSDP of Haryana	73
4.123Value Distribution of GSDP of Haryana	74
4.124Data Distribution of GSDP of Himachal Pradesh	74
4.125Value Distribution of GSDP of Himachal Pradesh	74
4.126Data Distribution of GSDP of Jharkhand	75
4.127Value Distribution of GSDP of Jharkhand	75
4.128Data Distribution of GSDP of Karnataka	75
4.129Value Distribution of GSDP of Karnataka	76



4.130Data Distribution of GSDP of Kerala	76
4.131Value Distribution of GSDP of Kerala	76
4.132Data Distribution of GSDP of Madhya Pradesh	77
4.133Value Distribution of GSDP of Madhya Pradesh	77
4.134Data Distribution of GSDP of Maharashtra	77
4.135Value Distribution of GSDP of Maharashtra	78
4.136Data Distribution of GSDP of Manipur	78
4.137Value Distribution of GSDP of Manipur	78
4.138Data Distribution of GSDP of Meghalaya	79
4.139Value Distribution of GSDP of Meghalaya	79
4.140Data Distribution of GSDP of Mizoram	79
4.141Value Distribution of GSDP of Mizoram	80
4.142Data Distribution of GSDP of Nagaland	80
4.143Value Distribution of GSDP of Nagaland	80
4.144Data Distribution of GSDP of Odisha	81
4.145Value Distribution of GSDP of Odisha	81
4.146Data Distribution of GSDP of Punjab	81
4.147Value Distribution of GSDP of Punjab	82
4.148Data Distribution of GSDP of Rajasthan	82
4.149Value Distribution of GSDP of Rajasthan	82
4.150Data Distribution of GSDP of Sikkim	83
4.151Value Distribution of GSDP of Sikkim	83
4.152Data Distribution of GSDP of Tamil Nadu	83
4.153Value Distribution of GSDP of Tamil Nadu	84
4.154Data Distribution of GSDP of Telangana	84
4.155Value Distribution of GSDP of Telangana	84
4.156Data Distribution of GSDP of Tripura	85
4.157Value Distribution of GSDP of Tripura	85
4.158Data Distribution of GSDP of Uttar Pradesh	85
4.159Value Distribution of GSDP of Uttar Pradesh	86
4.160Data Distribution of GSDP of Uttarakhand	86
4.161Value Distribution of GSDP of Uttarakhand	86
4.162Data Distribution of GSDP of West Bengal	87
4.163Value Distribution of GSDP of West Bengal	87
4.164Data Distribution of GSDP of Andaman & Nicobar Islands	87
4.165Value Distribution of GSDP of Andaman & Nicobar Islands	88
4.166Data Distribution of GSDP of Chandigarh	88
4.167Value Distribution of GSDP of Chandigarh	88
4.168Data Distribution of GSDP of Delhi	89
4.169Value Distribution of GSDP of Delhi	89
4.170Data Distribution of GSDP of Jammu & Kashmir-UT	89
4.171Value Distribution of GSDP of Jammu & Kashmir-UT	90
4.172Data Distribution of GSDP of Puducherry	90
4.173Value Distribution of GSDP of Puducherry	90
4.174Growth of GSDP of States Over The Years in Terms of Absolute Value	91
4.175Contribution of States to National GDP Over The Years	92



4.176Contribution of States Over The Years in Terms of %	93
4.177GSDP of States of India in 2023	94
4.178Growth of Contribution of Zones of India Over The Years in Terms of Absolute Value	95
4.179Contribution of Zones Of India Over the Years - Side By Side	95
4.180Contribution of Zones Of India Over the Years- Stacked	96
4.181Contribution of Zones Of India Over the Years in Terms of %	97
4.182Data Distribution of Growth Rate of GSDP of Andhra Pradesh	98
4.183Trend of Growth Rate of GSDP of Andhra Pradesh	98
4.184Data Distribution of Growth Rate of GSDP of Arunachal Pradesh	99
4.185Trend of Growth Rate of GSDP of Arunachal Pradesh	99
4.186Data Distribution of Growth Rate of GSDP of Assam	99
4.187Trend of Growth Rate of GSDP of Assam	100
4.188Data Distribution of Growth Rate of GSDP of Bihar	100
4.189Trend of Growth Rate of GSDP of Bihar	100
4.190Data Distribution of Growth Rate of GSDP of Chhattisgarh	101
4.191Trend of Growth Rate of Chhattisgarh	101
4.192Data Distribution of Growth Rate of GSDP of Goa	101
4.193Trend of Growth Rate of GSDP of Goa	102
4.194Data Distribution of Growth Rate of GSDP of Gujarat	102
4.195Trend of Growth Rate of GSDP of Gujarat	102
4.196Data Distribution of Growth Rate of GSDP of Haryana	103
4.197Trend of Growth Rate of GSDP of Haryana	103
4.198Data Distribution of Growth Rate of GSDP of Himachal Pradesh	103
4.199Trend of Growth Rate of GSDP of Himachal Pradesh	104
4.200Data Distribution of Growth Rate of GSDP of Jharkhand	104
4.201Trend of Growth Rate of GSDP of Jharkhand	104
4.202Data Distribution of Growth Rate of GSDP of Karnataka	105
4.203Trend of Growth Rate of GSDP of Karnataka	105
4.204Data Distribution of Growth Rate of GSDP of Kerala	105
4.205Trend of Growth Rate of GSDP of Kerala	106
4.206Data Distribution of Growth Rate of GSDP of Madhya Pradesh	106
4.207Trend of Growth Rate of GSDP of Madhya Pradesh	106
4.208Data Distribution of Growth Rate of GSDP of Maharashtra	107
4.209Trend of Growth Rate of GSDP of Maharashtra	107
4.210Data Distribution of Growth Rate of GSDP of Manipur	107
4.211Trend of Growth Rate of GSDP of Manipur	108
4.212Data Distribution of Growth Rate of GSDP of Meghalaya	108
4.213Trend of Growth Rate of GSDP of Meghalaya	108
4.214Data Distribution of Growth Rate of GSDP of Mizoram	109
4.215Trend of Growth Rate of GSDP of Mizoram	109
4.216Data Distribution of Growth Rate of GSDP of Nagaland	109
4.217Trend of Growth Rate of GSDP of Nagaland	110
4.218Data Distribution of Growth Rate of GSDP of Odisha	110
4.219Trend of Growth Rate of GSDP of Odisha	110
4.220Data Distribution of Growth Rate of GSDP of Punjab	111
4.221Trend of Growth Rate of GSDP of Punjab	111



4.222Data Distribution of Growth Rate of GSDP of Rajasthan	111
4.223Trend of Growth Rate of GSDP of Rajasthan	112
4.224Data Distribution of Growth Rate of GSDP of Sikkim	112
4.225Trend of Growth Rate of GSDP of Sikkim	112
4.226Data Distribution of Growth Rate of GSDP of Tamil Nadu	113
4.227Trend of Growth Rate of GSDP of Tamil Nadu	113
4.228Data Distribution of Growth Rate of GSDP of Telangana	113
4.229Trend of Growth Rate of GSDP of Telangana	114
4.230Data Distribution of Growth Rate of GSDP of Tripura	114
4.231Trend of Growth Rate of GSDP of Tripura	114
4.232Data Distribution of Growth Rate of GSDP of Uttar Pradesh	115
4.233Trend of Growth Rate of GSDP of Uttar Pradesh	115
4.234Data Distribution of Growth Rate of GSDP of Uttarakhand	115
4.235Trend of Growth Rate of GSDP of Uttarakhand	116
4.236Data Distribution of Growth Rate of GSDP of West Bengal	116
4.237Trend of Growth Rate of GSDP of West Bengal	116
4.238Data Distribution of Growth Rate of GSDP of Andaman & Nicobar Islands	117
4.239Trend of Growth Rate of GSDP of Andaman & Nicobar Islands	117
4.240Data Distribution of Growth Rate of GSDP of Chandigarh	117
4.241Trend of Growth Rate of GSDP of Chandigarh	118
4.242Data Distribution of Growth Rate of GSDP of Delhi	118
4.243Trend of Growth Rate of GSDP of Delhi	118
4.244Data Distribution of Growth Rate of GSDP of Jammu & Kashmir-UT*	119
4.245Trend of Growth Rate of GSDP of Jammu & Kashmir-UT*	119
4.246Data Distribution of Growth Rate of GSDP of Puducherry	119
4.247Trend of Growth Rate of GSDP of Puducherry	120
4.248Comparison of Growth Rate of GSDP of Various States of India With GDP's Growth Rate	121
4.249Growth of GSDP of Staes of India Over The Years	122
4.250GSDP Growth Rate of States of India in 2023	123
4.251Growth Rate of GSDP of States of India Over the Decade	124
4.252Correlation of Growth Rate of States	125
 5.1 Predicted Contribution of Sectors till 2028	126
5.2 Residual of LR model for Contribution of Sectors in Terms of Absolute Value	127
5.3 Residual of LR model for Contribution of Sectors in Terms of Percent	127
5.4 RMSE of LR model for Contribution of Sectors in Terms of Percent	128
5.5 R ² of LR model for Contribution of Sectors in Terms of Percent	128
5.6 Predicted Contribution of Sectors till 2028	129
5.7 Residual of PE model for Contribution of Sectors in Terms of Absolute Value	129
5.8 Residual of PE model for Contribution of Sectors in Terms of Percent	130
5.9 RMSE of PE model for Contribution of Sectors in Terms of Percent	130
5.10 R ² of PE model for Contribution of Sectors in Terms of Percent	130
5.11 Comparison of Predicted Contribution of Sectors till 2028	131
5.12 Comparison of Residuals for Contribution of Sectors in Terms of Absolute Value	131
5.13 Comparison of RMSE for Contribution of Sectors in Terms of Absolute Value	132
5.14 Comparison of R ² for Contribution of Sectors in Terms of Absolute Value	133



5.15 Predicted Contribution of States till 2029	134
5.16 Predicted Contribution of States till 2029	135
5.17 Predicted Contribution of States till 2029	136
5.18 Residual of LR model for Contribution of States in Terms of Absolute Value	137
5.19 Residual of LR model for Contribution of States in Terms of Percent	137
5.20 RMSE of LR model for Contribution of States in Terms of Percent	138
5.21 R ² of LR model for Contribution of States in Terms of Percent	138
5.22 Predicted Contribution of States till 2029	140
5.23 Predicted Contribution of States till 2029	141
5.24 Predicted Contribution of States till 2029	142
5.25 Residual of PE model for Contribution of States in Terms of Absolute Value	143
5.26 Residual of PE model for Contribution of States in Terms of Percent	143
5.27 RMSE of PE model for Contribution of States in Terms of Percent	144
5.28 R ² of PE model for Contribution of States in Terms of Percent	144
5.29 Comparison of Predicted Contribution of States till 2029	145
5.30 Comparison of Predicted Contribution of States till 2029	146
5.31 Comparison of Predicted Contribution of States till 2029	147
5.32 Comparison of Residuals for Contribution of States in Terms of Absolute Value	148
5.33 Comparison of Residuals for Contribution of States in Terms of Absolute Value	149
5.34 Comparison of Residuals for Contribution of States in Terms of Absolute Value	150
5.35 Comparison of RMSE for Contribution of States in Terms of Absolute Value	151
5.36 Comparison of R ² for Contribution of States in Terms of Absolute Value	151
6.1 GDP Growth of G20 Countries and it's Trend Over the Decade	154
6.2 GVA Growth of various Sectors and it's Trend Over the Decade	154
6.3 GSDP Growth and Growth Rate of States of India and it's Trend Over the Decade	155

List of Tables

Abstract

The economy of a country is a complex interplay of resources allocation, production, and consumption, reflecting its overall health and well-being. Central to understanding this economic landscape is the Gross Domestic Product (GDP), which measures the market value of all final goods and services produced within a nation over a specific time frame. This project focuses on analyzing annual and quarterly GDP data, aiming to forecast future trends and gain insights into the economic performance at both national and state levels. By exploring the relationship between GDP and Gross Value Added (GVA), the project seeks to provide a comprehensive understanding of the factors driving economic growth and development. The findings of this analysis are crucial for policymakers, businesses, and investors in making informed decisions to foster sustainable economic prosperity.

Chapter 1. Introduction

1.1 Project idea

The economy of a country depicts the relationship between people, businesses and government as they allot limited resources to fulfill human needs. It signifies economic health, growth and well being of a country.

One of the most important measures of the economy is Gross Domestic Product. According to [1] Gross domestic product (GDP) is a monetary measure of the market value of all the final goods and services produced and rendered in a specific time period by a country. The GDP helps us get an idea of how the resources are being allocated and what output they are producing.

The GDP is made up various factors, one of the major one being Gross Value Added (GVA). Gross value added is the measure of the value of goods and services produced in an area, industry or sector of an economy [2]. The relation between GDP and GVA in [2] is given as

$$GVA = GDP + \text{subsidies on products} - \text{taxes on products}$$

The project's objective is to analyse the annual and quarterly GDP data and forecast GDP data for a few upcoming years. Additionally, we also aim to analyze state-wise data to get better insights. Analyzing a country's GDP is significant as it gives us a sense of the direction in which the country is heading.

1.2 Data Collection

Data collection from authoritative sources like the Ministry of Statistics and Programme Implementation by the Government of India[3] and the World Bank[4] ensures reliability and credibility. These institutions employ rigorous methodologies to gather and analyze data, providing valuable insights for policymaking, research, and decision-making processes. By accessing data from such reputable sources, one can enhance the accuracy and robustness of their analyses and contribute to evidence-based decision-making on a global scale.

1.3 Dataset Description

Throughout the course of this project we would be using datasets, consisting of data related to GDP. All data has been taken at constant prices with the base year as 2011-12.



1.3.1 World

The datasets under this section contain data related to the world GDP. The datasets are

1.3.1.1 GDP of G20 Countries

DATASET NAME: GDP-World 2022

This dataset contained the GDP values of all the countries in the world from 1960 to 2022. We trimmed it down to 19 countries, including India, which are part of G20, the 20th member being the European Union. We chose to do so because India's role at the summit is increasing, and these factors have affected our economy. To keep consistency across our project, we selected data from the year 2010 to 2022. The value given is in USD.

Below is a snapshot of the data and information about columns.

	Unnamed: 0	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
0	Argentina	4.240000e+11	5.300000e+11	5.460000e+11	5.520000e+11	5.260000e+11	5.950000e+11	5.580000e+11	6.440000e+11	5.250000e+11	4.480000e+11	3.860000e+11	4.880000e+11	6.310000e+11
1	Australia	1.150000e+12	1.400000e+12	1.550000e+12	1.580000e+12	1.470000e+12	1.350000e+12	1.210000e+12	1.330000e+12	1.430000e+12	1.390000e+12	1.330000e+12	1.560000e+12	1.690000e+12
2	Brazil	2.210000e+12	2.620000e+12	2.470000e+12	2.460000e+12	1.800000e+12	1.800000e+12	2.060000e+12	1.920000e+12	1.870000e+12	1.480000e+12	1.650000e+12	1.920000e+12	1.920000e+12
3	Canada	1.620000e+12	1.790000e+12	1.830000e+12	1.850000e+12	1.810000e+12	1.560000e+12	1.530000e+12	1.650000e+12	1.730000e+12	1.740000e+12	1.660000e+12	2.010000e+12	2.160000e+12
4	China	6.090000e+12	7.550000e+12	8.530000e+12	9.570000e+12	1.050000e+13	1.110000e+13	1.120000e+13	1.230000e+13	1.390000e+13	1.430000e+13	1.470000e+13	1.780000e+13	1.800000e+13
5	France	2.650000e+12	2.870000e+12	2.680000e+12	2.810000e+12	2.460000e+12	2.470000e+12	2.600000e+12	2.790000e+12	2.730000e+12	2.650000e+12	2.960000e+12	2.780000e+12	2.780000e+12
6	Germany	3.400000e+12	3.750000e+12	3.530000e+12	3.730000e+12	3.890000e+12	3.360000e+12	3.470000e+12	3.690000e+12	3.970000e+12	3.890000e+12	3.890000e+12	4.280000e+12	4.080000e+12
7	India	1.680000e+12	1.820000e+12	1.830000e+12	1.860000e+12	2.040000e+12	2.100000e+12	2.290000e+12	2.650000e+12	2.700000e+12	2.840000e+12	2.670000e+12	3.150000e+12	3.420000e+12
8	Indonesia	7.550000e+11	8.930000e+11	9.180000e+11	9.130000e+11	8.910000e+11	8.610000e+11	9.320000e+11	1.020000e+12	1.040000e+12	1.120000e+12	1.060000e+12	1.190000e+12	1.320000e+12
9	Italy	2.140000e+12	2.290000e+12	2.090000e+12	2.140000e+12	2.160000e+12	1.840000e+12	1.880000e+12	1.960000e+12	2.090000e+12	2.010000e+12	1.900000e+12	2.160000e+12	2.050000e+12

Figure 1.1: GDP Data of G20 Countries

1.3.2 Contribution of Sectors

The datasets under this section contain data related to the GVA and the growth rate of various sectors that comprise the GDP. The values are taken by keeping

1.3.2.1 GVA

DATASET NAME: ANNUAL ESTIMATES OF GDP AT CONSTANT PRICES, 2011-12 SERIES - Amount of GVA at Basic Prices from Items

This dataset contains the GDP values of the various sectors in terms of Crore Rs. The various sectors are

- Agriculture, forestry & fishing
- Mining & quarrying
- Manufacturing
- Electricity, gas, water supply & other utility services
- Construction
- Trade, hotels, transport, communication and services related to broadcasting
- Financial , real estate & prof servs



- Public Administration, defence and other services

Collectively the GVA of these sectors add up to give the GVA of the nation. Additionally, we also have some parameters like

- NVA at basic prices (Net Value Added)
- GNI (Gross National Income)
- NNI (Net Value Income)
- Per capita income(Rs.)
- Net taxes on Products

Finally, the values of the GVA column and Net Taxes column add up to the GDP value. The data is from the financial year 2011-12 to the financial year 2023-24.

Below is a snapshot of the data and information about columns.

Year	Agriculture, forestry & fishing	Mining & quarrying	Manufacturing	Electricity, gas, water supply & other utility services	Construction	Trade, hotels, restaurants & service related to broadcasting	Financial communication	real estate & prof servs	Administration, defence and other services	GVA at basic prices	NVA at basic prices	GNI	NNI	Per capita income(Rs.)	Net taxes on Products	GDP
0 2011-12	15,01,947	2,61,035	14,09,985	1,86,668	7,77,335	14,13,116	15,30,877	10,25,982	81,06,946	71,89,771	86,59,505	77,42,320	63462	6,29,383	87,36,329	
1 2012-13	15,24,288	2,62,609	14,86,873	1,91,635	7,80,050	15,51,143	16,80,031	10,69,646	85,46,275	75,35,614	91,04,662	80,94,001	65538	6,66,741	92,19,017	
2 2013-14	16,09,198	2,63,107	15,60,709	1,99,601	8,00,771	16,52,062	18,67,407	11,10,794	90,63,649	79,63,039	96,79,027	85,78,417	68572	7,37,721	96,01,370	
3 2014-15	16,05,715	2,88,685	16,83,938	2,14,047	8,35,229	18,07,689	20,73,714	12,03,115	97,12,133	85,83,489	1,04,02,987	92,24,343	72805	8,15,541	1,05,27,674	
4 2015-16	16,16,146	3,17,974	19,03,850	2,24,158	8,65,335	19,92,825	22,94,787	12,76,797	1,04,91,870	92,20,980	1,12,34,571	99,63,681	77659	8,77,623	1,13,69,493	
5 2016-17	17,26,004	3,49,248	20,54,764	2,46,496	9,16,445	21,46,379	24,92,967	13,95,882	1,13,28,285	99,46,758	1,21,63,619	1,07,82,092	83003	9,79,909	1,23,08,193	
6 2017-18	18,40,023	3,29,612	22,09,428	2,72,650	9,64,306	23,68,419	25,37,190	15,12,542	1,20,24,171	1,05,44,250	1,29,98,695	1,15,08,774	87586	11,10,412	1,31,44,582	
7 2018-19	18,78,598	3,26,815	23,28,992	2,94,147	10,26,789	25,38,757	27,14,222	16,25,477	1,27,33,798	1,11,19,543	1,38,40,474	1,22,26,019	92133	12,89,116	1,39,92,914	
8 2019-20	19,94,326	3,17,134	22,59,706	3,00,798	10,43,429	26,90,060	28,98,243	17,32,404	1,32,36,100	1,15,64,922	1,43,92,900	1,26,61,722	94420	12,98,540	1,45,34,641	
9 2020-21	20,74,212	2,91,075	23,29,160	2,88,213	9,95,371	21,53,888	29,54,130	16,01,295	1,26,67,345	1,08,54,586	1,34,93,976	1,16,61,217	86034	10,07,525	1,36,94,869	
10 2021-22	21,70,106	3,09,276	25,61,033	3,17,966	11,93,532	24,80,380	31,22,847	17,21,699	1,38,76,840	1,19,21,052	1,48,27,920	1,28,72,132	94054	11,45,006	1,50,21,846	
11 2022-23	22,72,250	3,15,256	25,04,663	3,47,973	13,06,256	27,77,723	34,05,474	18,75,304	1,48,04,901	1,27,20,270	1,58,31,133	1,37,46,502	99404	12,66,528	1,60,71,429	
12 2023-24	22,87,339	3,40,821	27,17,235	3,74,125	14,45,603	29,57,058	36,84,959	20,20,579	1,58,27,708	1,35,99,059	1,70,34,386	1,48,05,796	1,06,134	14,62,573	1,72,90,281	

Figure 1.2: National Annual GVA

1.3.2.2 Growth Rate of GVA

DATASET NAME: ANNUAL ESTIMATES OF GDP AT CONSTANT PRICES, 2011-12 SERIES - Growth Rate of GVA at Basic Prices from Items

Similar to the above dataset, this dataset contains the growth rate of GVA in various sectors. The columns of this dataset are the same as those mentioned in the GVA dataset under the Contribution of Sectors section. In the columns of GVA and GDP too, the growth rate of those parameters is mentioned. The growth rate is calculated by keeping the base year as 2011-12, and the data ranges from the financial year 2011-12 to the financial year 2023-24.

Below is a snapshot of the data and information about columns.



Figure 1.3: National Annual Growth Rate

1.3.3 Contribution of States

1.3.3.1 GSDP

DATASET NAME: GROSS STATE DOMESTIC PRODUCT AT CURRENT PRICES; BASE YEAR 2011-12

GSDP stands for Gross State Domestic Product. It is the GDP of a state. This dataset contains data about the GDP of various states and UTs. The states and UTs are

- Andhra Pradesh
 - Arunachal Pradesh
 - Assam
 - Bihar
 - Chhattisgarh
 - Goa
 - Gujarat
 - Haryana
 - Himachal Pradesh
 - Jharkhand
 - Karnataka
 - Kerala
 - Madhya Pradesh
 - Maharashtra
 - Manipur
 - Meghalaya
 - Mizoram
 - Nagaland
 - Odisha
 - Punjab
 - Rajasthan
 - Sikkim
 - Tamil Nadu
 - Telangana
 - Tripura
 - Uttar Pradesh
 - Uttarakhand
 - West Bengal
 - Andaman & Nicobar Islands
 - Chandigarh
 - Delhi
 - Jammu & Kashmir-UT*
 - Puducherry

The data is in Crore Rs and is taken by keeping 2011-12 as the base year. The data ranges from the financial year 2011-12 to 2023-24.

Below is a snapshot of the data and information about columns.



Year	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chhattisgarh	Goa	Gujarat	Haryana	Himachal Pradesh	Jharkhand	J&K	Telangana	Tripura	Uttar Pradesh	Uttarakhand	West Bengal	Andaman & Nicobar Islands	Chandigarh	Delhi	Jammu & Kashmir-UT*	Puducherry
2011-12	379402	11063.0	149175	247144.0	158074	42367.0	615606.0	297539	72720	150918.0	...	359434	19208.0	724050	115828	520485	3978.0	18768.0	343798	78256	166818.0
2012-13	380629	11299.0	147342	256851.0	165977	35850.0	682550.0	320912	77284	163250.0	...	370113	20873.0	758205	123710	542191	4156.0	20285.0	366428	80767	17310.0
2013-14	407115	12339.0	154525	269450.0	182580	31568.0	734284.0	347507	82847	165816.0	...	389957	23819.0	802070	134182	558353	4488.0	22105.0	392908	85102	19170.0
2014-15	444564	14383.0	165212	279482.0	185813	40116.0	811428.0	370535	89060	186334.0	...	416332	26965.0	834432	141278	574364	4742.0	22870.0	428355	82327	18207.0
2015-16	498606	14246.0	191109	294488.0	190584	46091.0	894465.0	413405	96274	174881.0	...	464542	26787.0	908241	132699	609545	5092.0	24922.0	475623	97001	19060.0
2016-17	540212	14893.0	202081	318757.0	213705	51249.0	981342.0	456709	102055	193174.0	...	567446	30538.0	1011500	167703	658416	5752.0	26917.0	511765	100203	20474.0
2017-18	594737	15572.0	219919	344028.0	220136	52653.0	1086570.0	482036	109406	210887.0	...	557410	33093.0	1056399	180956	694981	6464.0	28480.0	542015	106424	22318.0
2018-19	626614	16668.0	231040	381383.0	244579	50063.0	1183020.0	532996	116414	229274.0	...	608401	36754.0	1097353	186083	738920	6867.0	29866.0	56327	115062	26209.0
2019-20	649810	19140.0	240707	398329.0	251549	54812.0	1265277.0	545124	121227	231755.0	...	640968	38063.0	1141630	189740	761794	7266.0	31679.0	586168	119191	26374.0
2020-21	659678	18435.0	247819	368733.0	251192	52534.0	1241118.0	496087	115958	219483.0	...	602435	36404.0	1092787	166786	704045	6715.0	28596.0	533634	112400	24137.0
2021-22	722727	19707.0	278835	399930.0	280409	55342.0	1366922.0	544317	124770	243348.0	...	669229	39487.0	1200093	184274	787439	7172.0	30587.0	580396	119174	27131.0
2022-23	764685	21085.0	300799	442473.0	302119	59857.0	1475629.0	587198	133372	259800.0	...	715157	42997.0	1318027	198341	839805	7172.0	33879.0	625981	129673	28849.0
2023-24	820894	21085.0	318559	442472.0	321945	59857.0	1475629.0	634027	142800	259800.0	...	767936	42997.0	1422358	213978	904088	7172.0	33879.0	672247	139171	28849.0

13 rows × 33 columns

Figure 1.4: State GSDP

1.3.3.2 Growth Rate of GSDP

DATASET NAME: Growth Rate of GROSS STATE DOMESTIC PRODUCT AT CURRENT PRICES; BASE YEAR 2011-12

This dataset includes the growth rate of GSDP of the states and UTs mentioned in the description of the GSDP dataset under the Contribution of States section. The data is taken by keeping 2011-12 as the base year. The data ranges from the financial year 2011-12 to 2023-24.

Below is a snapshot of the data and information about columns.

Year	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chhattisgarh	Goa	Gujarat	Haryana	Himachal Pradesh	Jharkhand	J&K	Telangana	Tripura	Uttar Pradesh	Uttarakhand	West Bengal	Andaman & Nicobar Islands	Chandigarh	Delhi	Jammu & Kashmir-UT*	Puducherry
0	2012-13	0.32	2.14	2.91	3.93	5.00	-15.38	10.89	7.86	6.41	...	2.97	8.67	4.72	7.27	4.17	4.48	8.08	6.64	3.21	2.93
1	2013-14	6.96	9.21	4.88	4.98	10.00	-11.94	7.56	8.29	7.06	...	5.36	9.32	5.79	8.47	3.01	7.99	8.97	7.17	5.37	10.74
2	2014-15	9.20	16.56	6.92	3.65	1.77	27.08	10.51	6.63	7.50	...	6.76	18.17	4.03	5.29	2.83	5.64	3.46	9.02	-3.21	-5.03
3	2015-16	12.16	-0.99	15.67	6.08	2.57	14.89	10.23	11.57	8.10	...	11.58	-0.66	8.85	8.08	6.13	7.39	9.02	11.03	17.76	4.69
4	2016-17	8.34	4.58	5.74	7.52	12.13	11.19	9.71	10.48	7.04	...	9.34	14.00	11.37	9.83	7.20	12.96	7.96	7.60	3.30	7.42
5	2017-18	10.09	4.56	8.83	7.91	3.01	2.74	10.72	5.55	6.16	...	9.74	8.37	4.44	7.90	6.36	12.39	5.80	5.91	6.41	9.00
6	2018-19	5.36	7.04	5.06	10.86	11.10	0.78	8.88	10.57	6.41	...	9.15	11.06	3.88	2.83	6.32	6.23	4.87	4.30	7.91	17.44
7	2019-20	3.70	14.83	4.18	4.44	2.85	3.30	6.95	2.28	4.13	...	5.35	3.56	4.03	1.97	3.10	5.81	6.07	3.69	-0.99	0.63
8	2020-21	1.52	-3.69	2.95	-7.43	-0.14	-4.16	-1.91	-9.00	-4.35	...	-6.01	-4.36	-4.28	-12.10	-7.58	-7.58	-9.73	-8.96	-1.33	-8.48
9	2021-22	9.56	6.90	12.52	8.46	11.71	5.35	10.06	9.72	7.60	...	11.09	8.47	9.82	10.49	11.85	6.81	6.96	8.76	6.03	12.41
10	2022-23	5.81	6.99	7.88	10.64	7.67	8.18	8.03	7.88	6.89	...	6.86	8.89	9.83	7.63	6.65	NaN	10.76	7.85	8.73	6.33
11	2023-24	7.35	NaN	5.90	NaN	6.56	NaN	NaN	7.97	7.07	...	7.38	NaN	7.99	7.58	7.65	NaN	NaN	7.39	7.41	NaN

Figure 1.5: State GSDP Growth Rate

1.4 Packages required

We have used various packages during our project to help us analyse and visualise the data efficiently. The package are

**1. NumPy (np)**

NumPy is a fundamental package for scientific computing with Python. We have used it for numerical computations, array manipulation, and linear algebra operations.

2. Pandas (pd)

Pandas is a powerful data manipulation and analysis library. We have used it for data cleaning, transformation, exploration, and analysis tasks.

3. Seaborn (sns)

Seaborn is a statistical data visualization library based on Matplotlib. We have used it for creating visually appealing plots for exploring relationships in data.

4. Matplotlib.pyplot (plt)

Matplotlib is a comprehensive plotting library for Python. We have used for creating various types of plots and visualizations, including line plots, scatter plots, histograms, etc.

5. Missingno (msno) Missingno is a Python library for visualizing missing data in datasets.

It provides informative visualizations such as bar charts, heatmaps, and dendograms to identify missing values in datasets. It helped us to know the completeness of data.

6. Scikit-learn (sklearn) Scikit-learn is a machine-learning library for Python. We have used for tasks such as regression, and model selection.

Chapter 2. Data Cleaning

Data Cleaning is the process of analyzing missing values and outliers in a dataset and handling them according to the needs of the objective. [Missing value Description]

From the datasets mentioned above, we encountered missing values in some of them, while some of them didn't have any missing values. Given below is the missing data analysis for datasets in which we have encountered missing values.

For outliers, we can't replace them with some other values as they will not accurately represent reality.

2.1 World

2.1.1 GDP of G20 Countries

2.1.1.1 Missing Data Analysis

Here we have encountered no missing values



	Missing Values	% of Total Values
Unnamed: 0	0	0.0
2010	0	0.0
2011	0	0.0
2012	0	0.0
2013	0	0.0
2014	0	0.0
2015	0	0.0
2016	0	0.0
2017	0	0.0
2018	0	0.0
2019	0	0.0
2020	0	0.0
2021	0	0.0
2022	0	0.0

Figure 2.1: Columns Information After Transformation

2.1.1.2 Imputation

As no missing values have been encountered, no imputation is required.

2.1.1.3 Changing Data Types

Here all the columns transformed into same data type columns.



```
<class 'pandas.core.frame.DataFrame'>
Index: 13 entries, 2010 to 2022
Data columns (total 19 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Argentina       13 non-null      object  
 1   Australia       13 non-null      float64 
 2   Brazil          13 non-null      float64 
 3   Canada          13 non-null      float64 
 4   China           13 non-null      float64 
 5   France          13 non-null      float64 
 6   Germany         13 non-null      float64 
 7   India           13 non-null      float64 
 8   Indonesia        13 non-null      float64 
 9   Italy            13 non-null      float64 
 10  Japan           13 non-null      float64 
 11  Mexico          13 non-null      float64 
 12  Russian          13 non-null      float64 
 13  Saudi Arabia    13 non-null      float64 
 14  South Africa    13 non-null      float64 
 15  South Korea     13 non-null      float64 
 16  Turkey          13 non-null      float64 
 17  United Kingdom  13 non-null      float64 
 18  United States   13 non-null      float64 
dtypes: float64(18), object(1)
memory usage: 2.0+ KB
```

Figure 2.2: Columns Information After Transformation

2.2 Contribution of Sectors

2.2.1 GVA Amount

2.2.1.1 Missing Data Analysis

Here we have encountered no missing values



	Missing Values	% of Total Values
Year	0	0.0
Agriculture,forestry & fishing	0	0.0
Mining & quarrying	0	0.0
Manufacturing	0	0.0
Electricity, gas, water supply & other utility services	0	0.0
Construction	0	0.0
Trade, hotels, transport, communication and services related to broadcasting	0	0.0
Financial , real estate & prof servs	0	0.0
Public Administration, defence and other services	0	0.0
GVA at basic prices	0	0.0
NVA at basic prices	0	0.0
GNI	0	0.0
NNI	0	0.0
Per capita income(Rs.)	0	0.0
Net taxes on Products	0	0.0
GDP	0	0.0

Figure 2.3: Columns Information After Transformation

2.2.1.2 Imputation

As no missing values have been encountered, no imputation is required.

2.2.1.3 Changing Data Types

All the columns are in desired datatypes.

Agriculture,forestry & fishing	int64
Mining & quarrying	int64
Manufacturing	int64
Electricity, gas, water supply & other utility services	int64
Construction	int64
Trade, hotels, transport, communication and services related to broadcasting	int64
Financial , real estate & prof servs	int64
Public Administration, defence and other services	int64
GVA at basic prices	int64
Per capita income(Rs.)	int64
Net taxes on Products	int64
GDP	int64
dtype: object	

Figure 2.4: Columns Information after Transformation

2.2.2 Growth Rate of GVA

2.2.2.1 Missing Data Analysis



	Missing Values	% of Total Values
Agriculture,forestry & fishing	1	7.692308
Mining & quarrying	1	7.692308
Manufacturing	1	7.692308
Electricity, gas, water supply & other utility services	1	7.692308
Construction	1	7.692308
Trade, hotels, transport, communication and services related to broadcasting	1	7.692308
Financial , real estate & prof servs	1	7.692308
Public Administration, defence and other services	1	7.692308
GVA at basic prices	1	7.692308
Per capita income(Rs.)	1	7.692308
Net taxes on Products	1	7.692308
GDP	1	7.692308

Figure 2.5: Growth Rate of GVA - Missing Value Table

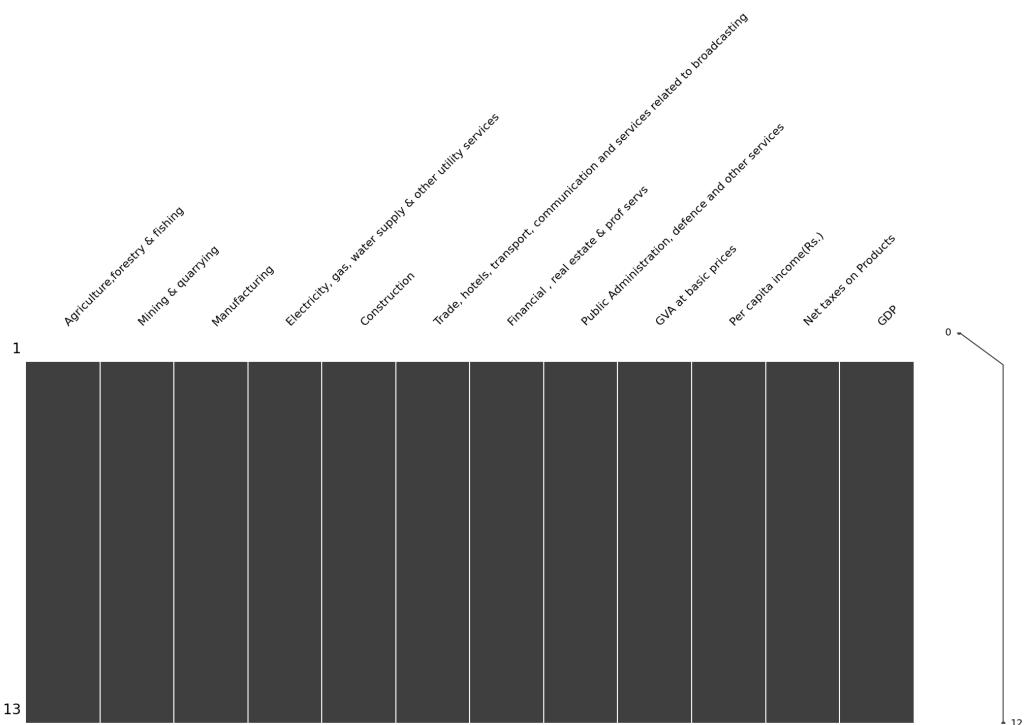


Figure 2.6: Growth Rate of GVA - Missing Value Matrix

Here, we can observe that we have one missing value in each column, all belonging to the same year. As data is absent completely for a year, we plan to drop that row from the analysis.

2.2.2.2 Imputation

As observed in the missing value analysis all values for the year 2011-12 are missing. As 2011-12 is the first row of dataset and to determine its value we would require data from the previous years which we don't have. So we dropped the entire row.



	Missing Values	% of Total Values
Agriculture,forestry & fishing	0	0.0
Mining & quarrying	0	0.0
Manufacturing	0	0.0
Electricity, gas, water supply & other utility services	0	0.0
Construction	0	0.0
Trade, hotels, transport, communication and services related to broadcasting	0	0.0
Financial , real estate & prof servs	0	0.0
Public Administration, defence and other services	0	0.0
GVA at basic prices	0	0.0
Per capita income(Rs.)	0	0.0
Net taxes on Products	0	0.0
GDP	0	0.0

Figure 2.7: Imputed Value Table

2.2.2.3 Changing Data Types

Here all the columns transformed into same data type columns.

```
Agriculture,forestry & fishing          float64
Mining & quarrying                     float64
Manufacturing                          float64
Electricity, gas, water supply & other utility services float64
Construction                           float64
Trade, hotels, transport, communication and services related to broadcasting float64
Financial , real estate & prof servs    float64
Public Administration, defence and other services   float64
GVA at basic prices                    float64
Per capita income(Rs.)                 float64
Net taxes on Products                  float64
GDP                                    float64
dtype: object
```

Figure 2.8: Columns Information after Transformation

2.3 Contribution of States

2.3.1 GSDP

2.3.1.1 Missing Data Analysis



	Missing Values	% of Total Values
Andhra Pradesh	0	0.000000
Arunachal Pradesh	1	7.692308
Assam	0	0.000000
Bihar	1	7.692308
Chhattisgarh	0	0.000000
Goa	1	7.692308
Gujarat	1	7.692308
Haryana	0	0.000000
Himachal Pradesh	0	0.000000
Jharkhand	1	7.692308
Karnataka	0	0.000000
Kerala	1	7.692308
Madhya Pradesh	0	0.000000
Maharashtra	2	15.384615
Manipur	2	15.384615
Meghalaya	0	0.000000

(a) GSDP - Missing Value Table 1

Mizoram	2	15.384615
Nagaland	1	7.692308
Odisha	0	0.000000
Punjab	0	0.000000
Rajasthan	0	0.000000
Sikkim	1	7.692308
Tamil Nadu	0	0.000000
Telangana	0	0.000000
Tripura	1	7.692308
Uttar Pradesh	0	0.000000
Uttarakhand	0	0.000000
West Bengal	0	0.000000
Andaman & Nicobar Islands	2	15.384615
Chandigarh	1	7.692308
Delhi	0	0.000000
Jammu & Kashmir-UT*	0	0.000000
Puducherry	1	7.692308

(b) GSDP - Missing Value Table 2

Figure 2.9: GSDP - Missing Value Tables

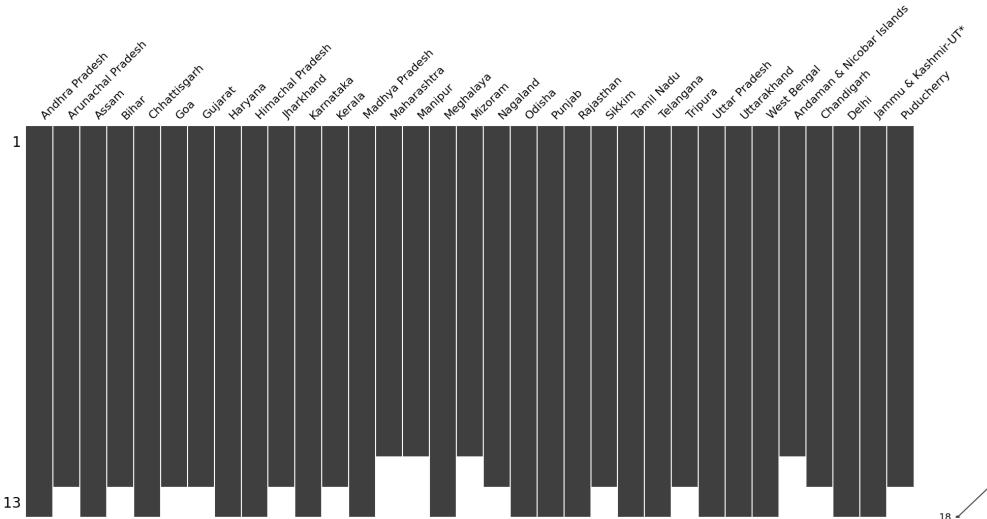


Figure 2.10: GSDP - Missing Value Matrix

We can observe here that we have some NULL values which are missing completely at random.

2.3.1.2 Imputation

To handle missing values in the GSDP for a particular state in a particular year, we adopted a simple approach. We calculated the average growth rate of the GSDP for that state over the last three available years (with respect to the missing year). Then, we added this average growth rate to the GSDP value of the most recent year with available data (with respect to the missing year).



Andhra Pradesh	0	Mizoram	0
Arunachal Pradesh	0	Nagaland	0
Assam	0	Odisha	0
Bihar	0	Punjab	0
Chhattisgarh	0	Rajasthan	0
Goa	0	Sikkim	0
Gujarat	0	Tamil Nadu	0
Haryana	0	Telangana	0
Himachal Pradesh	0	Tripura	0
Jharkhand	0	Uttar Pradesh	0
Karnataka	0	Uttarakhand	0
Kerala	0	West Bengal	0
Madhya Pradesh	0	Andaman & Nicobar Islands	0
Maharashtra	0	Chandigarh	0
Manipur	0	Delhi	0
Meghalaya	0	Jammu & Kashmir-UT*	0
		Puducherry	0

(a) GSDP - Imputed Table1

(b) GSDP - Imputed Table2

Figure 2.11: GSDP - Imputed Table

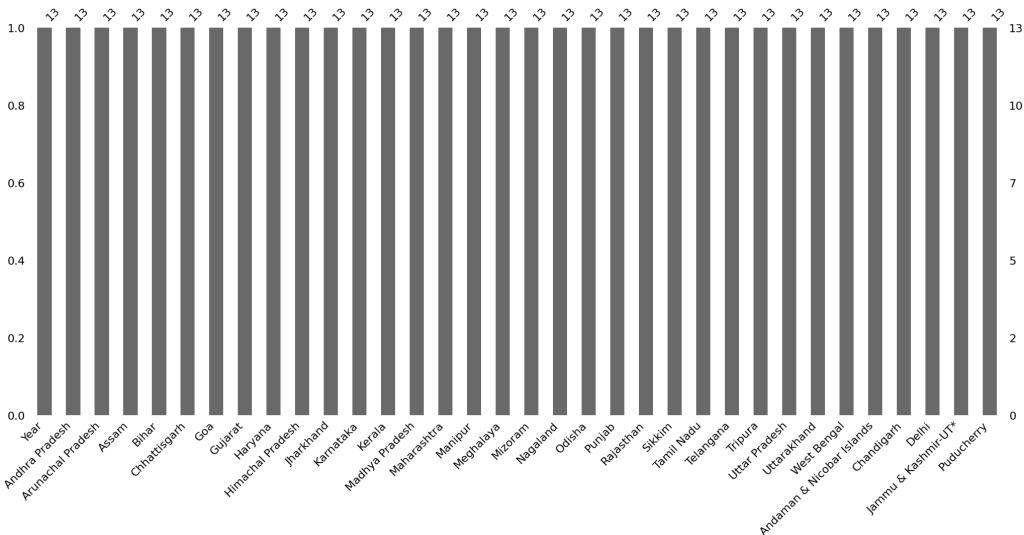


Figure 2.12: GSDP - Imputed Bar

2.3.1.3 Changing Data Types

Here all the columns transformed into same data type columns.



Andhra Pradesh	int64	Nagaland	int64
Arunachal Pradesh	int64	Odisha	int64
Assam	int64	Punjab	int64
Bihar	int64	Rajasthan	int64
Chhattisgarh	int64	Sikkim	int64
Goa	int64	Tamil Nadu	int64
Gujarat	int64	Telangana	int64
Haryana	int64	Tripura	int64
Himachal Pradesh	int64	Uttar Pradesh	int64
Jharkhand	int64	Uttarakhand	int64
Karnataka	int64	West Bengal	int64
Kerala	int64	Andaman & Nicobar Islands	int64
Madhya Pradesh	int64	Chandigarh	int64
Maharashtra	int64	Delhi	int64
Manipur	int64	Jammu & Kashmir-UT*	int64
Meghalaya	int64	Puducherry	int64
Mizoram	int64		
			dtype: object

Figure 2.13: Growth Rate of GSDP - After Transformation

2.3.2 Growth Rate of GSDP

2.3.2.1 Missing Data Analysis

	Missing Values	% of Total Values
Andhra Pradesh	0	0.000000
Arunachal Pradesh	1	8.333333
Assam	0	0.000000
Bihar	1	8.333333
Chhattisgarh	0	0.000000
Goa	1	8.333333
Gujarat	1	8.333333
Haryana	0	0.000000
Himachal Pradesh	0	0.000000
Jharkhand	1	8.333333
Karnataka	0	0.000000
Kerala	1	8.333333
Madhya Pradesh	0	0.000000
Maharashtra	2	16.666667
Manipur	2	16.666667
Meghalaya	0	0.000000

(a) Growth Rate of GSDP - Missing Value Table 1

Mizoram	2	16.666667
Nagaland	1	8.333333
Odisha	0	0.000000
Punjab	0	0.000000
Rajasthan	0	0.000000
Sikkim	1	8.333333
Tamil Nadu	0	0.000000
Telangana	0	0.000000
Tripura	1	8.333333
Uttar Pradesh	0	0.000000
Uttarakhand	0	0.000000
West Bengal	0	0.000000
Andaman & Nicobar Islands	2	16.666667
Chandigarh	1	8.333333
Delhi	0	0.000000
Jammu & Kashmir-UT*	0	0.000000
Puducherry	1	8.333333

(b) Growth Rate of GSDP - Missing Value Table 2

Figure 2.14: Growth Rate of GSDP - Missing Value Tables

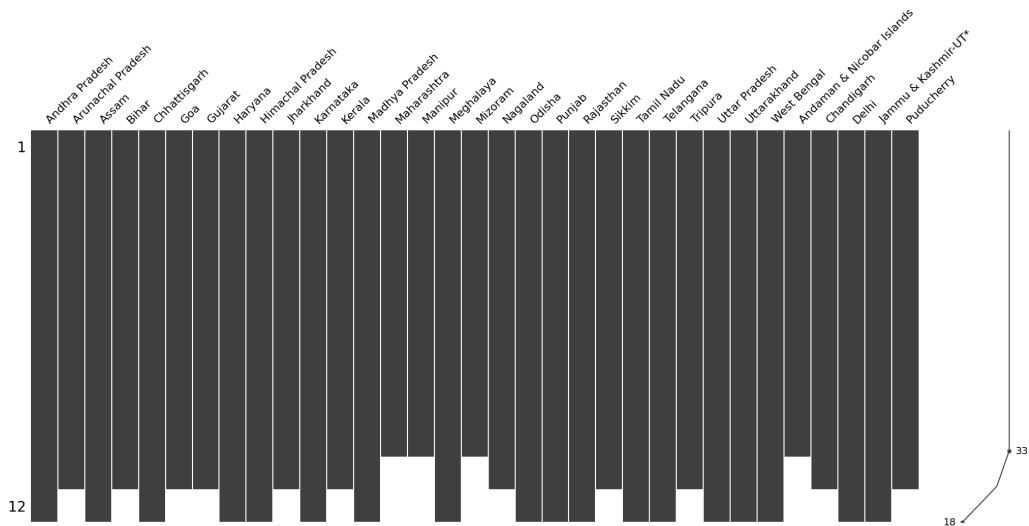


Figure 2.15: Growth Rate of GSDP - Missing Value Matrix

Similar to the missing value analysis of GSDP, here too we encounter some missing values which are missing completely at random.

2.3.2.2 Imputation

Here too we have used a similar approach as described in the imputation for the GSDP dataset of taking average growth rate of last three available years and then adding it to the recent most available year.

	Missing Values	% of Total Values
Andhra Pradesh	0	0.0
Arunachal Pradesh	0	0.0
Assam	0	0.0
Bihar	0	0.0
Chhattisgarh	0	0.0
Goa	0	0.0
Gujarat	0	0.0
Haryana	0	0.0
Himachal Pradesh	0	0.0
Jharkhand	0	0.0
Karnataka	0	0.0
Kerala	0	0.0
Madhya Pradesh	0	0.0
Maharashtra	0	0.0
Manipur	0	0.0
Meghalaya	0	0.0

(a) Growth Rate of GSDP - Imputed Table 1

Mizoram	0	0.0
Nagaland	0	0.0
Odisha	0	0.0
Punjab	0	0.0
Rajasthan	0	0.0
Sikkim	0	0.0
Tamil Nadu	0	0.0
Telangana	0	0.0
Tripura	0	0.0
Uttar Pradesh	0	0.0
Uttarakhand	0	0.0
West Bengal	0	0.0
Andaman & Nicobar Islands	0	0.0
Chandigarh	0	0.0
Delhi	0	0.0
Jammu & Kashmir-UT*	0	0.0
Puducherry	0	0.0

(b) Growth Rate of GSDP - Imputed Table 2

Figure 2.16: Growth Rate of GSDP - Imputed Tables

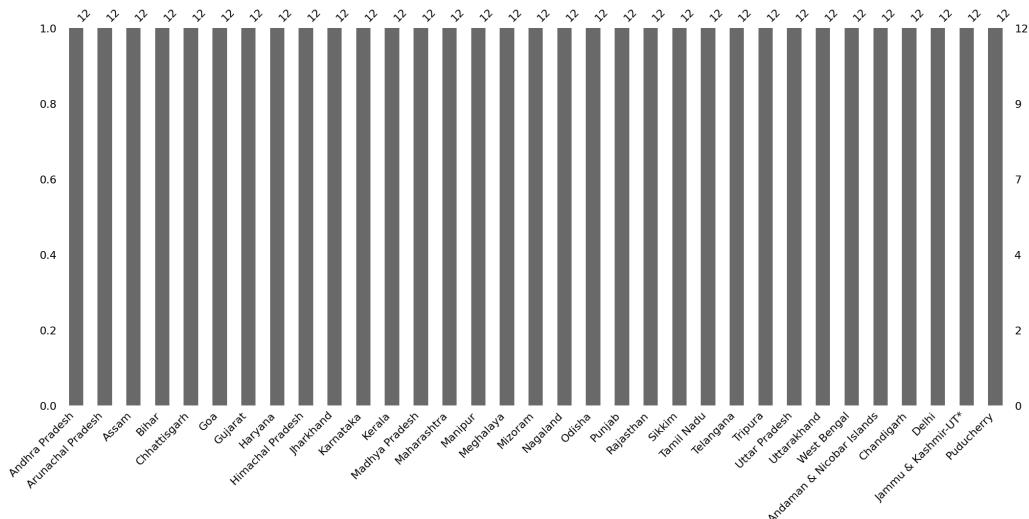


Figure 2.17: Growth Rate of GSDP - Interpolated Matrix

2.3.2.3 Changing Data Types

Here all the columns transformed into same data type columns.

Andhra Pradesh	float64	Nagaland	float64
Arunachal Pradesh	float64	Odisha	float64
Assam	float64	Punjab	float64
Bihar	float64	Rajasthan	float64
Chhattisgarh	float64	Sikkim	float64
Goa	float64	Tamil Nadu	float64
Gujarat	float64	Telangana	float64
Haryana	float64	Tripura	float64
Himachal Pradesh	float64	Uttar Pradesh	float64
Jharkhand	float64	Uttarakhand	float64
Karnataka	float64	West Bengal	float64
Kerala	float64	Andaman & Nicobar Islands	float64
Madhya Pradesh	float64	Chandigarh	float64
Maharashtra	float64	Delhi	float64
Manipur	float64	Jammu & Kashmir-UT*	float64
Meghalaya	float64	Puducherry	float64
Mizoram	float64		

Figure 2.18: Growth Rate of GSDP - After Transformation

Chapter 3. Feature Engineering

3.1 World

3.1.1 G20

3.1.1.1 Feature Extraction

3.1.1.1.1 Converting in format of billion \$

We have divided the value of GDP by 1e9 to convert them in form of billion dollars.

	Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
0	Argentina	4.236270e-07	5.301580e-07	5.459820e-07	5.520250e-07	5.263200e-07	5.947490e-07	5.575320e-07	6.436280e-07	5.248200e-07	4.477550e-07	3.857410e-07	4.879030e-07	6.311330e-07
1	Australia	1.148890e-06	1.398700e-06	1.547650e-06	1.577300e-06	1.468600e-06	1.351770e-06	1.207580e-06	1.326880e-06	1.429730e-06	1.394670e-06	1.330380e-06	1.559030e-06	1.692060e-06
2	Brazil	2.208840e-06	2.616160e-06	2.465230e-06	2.472820e-06	2.456040e-06	1.802210e-06	1.795690e-06	2.063510e-06	1.916930e-06	1.873290e-06	1.476110e-06	1.649620e-06	1.920100e-06
3	Canada	1.617340e-06	1.793330e-06	1.828370e-06	1.846600e-06	1.805750e-06	1.556510e-06	1.527990e-06	1.649270e-06	1.725330e-06	1.743730e-06	1.655680e-06	2.007470e-06	2.161480e-06
4	China	6.087190e-06	7.551550e-06	8.532190e-06	9.570470e-06	1.047560e-05	1.106160e-05	1.123330e-05	1.231050e-05	1.389490e-05	1.428000e-05	1.468770e-05	1.782050e-05	1.796320e-05

Figure 3.1: Format in billion \$

3.1.1.1.2 Transposing Dataframe

We have stored a transposed copy of the dataframe into a new dataframe which will help us in creating multivariate plots.

Year	Argentina	Australia	Brazil	Canada	China	France	Germany	India	Indonesia	Italy	Japan	Mexico	Russian	Saudi Arabia	South Africa	South Korea	Turkey	United Kingdom	United States
2010	0.0	0.000001	0.000002	0.000002	0.000003	0.000003	0.000002	0.000001	0.000002	0.000006	0.000001	0.000002	0.000001	0.0	0.000001	0.000001	0.000002	0.000015	
2011	0.000001	0.000001	0.000002	0.000002	0.000003	0.000003	0.000002	0.000002	0.000002	0.000006	0.000001	0.000002	0.000001	0.0	0.000001	0.000001	0.000003	0.000016	
2012	0.000001	0.000002	0.000002	0.000002	0.000003	0.000004	0.000003	0.000001	0.000004	0.000006	0.000001	0.000002	0.000001	0.0	0.000001	0.000001	0.000003	0.000016	
2013	0.000001	0.000002	0.000002	0.000002	0.000003	0.000004	0.000002	0.000001	0.000005	0.000006	0.000001	0.000002	0.000001	0.0	0.000001	0.000001	0.000003	0.000017	

Figure 3.2: Transposed Dataframe of G20 Countries

3.2 Contribution of Sectors

3.2.1 GVA

3.2.1.1 Feature Selection

3.2.1.1.1 Dropping Unnecessary Columns

Here, the columns of NVA, GNI and NNI are not much of use so we dropped them.



Year	Agriculture,forestry & fishing	Mining & quarrying	Manufacturing	Electricity, gas, water supply & other utility services	Construction	Trade, hotels, transport, communication and services related to broadcasting	Financial estate & prof servs	Public Administration, defence and other services	GVA at basic prices	Per capita income(Rs.)	Net taxes on products	GDP
2011-12	1501947	261035	1409986	188668	777335	1413116	1530877	1025982	8106946	63462	629383	8736329
2012-13	1524288	262609	1489873	191835	780050	1551143	1680031	1069646	8546275	65538	666741	9213017
2013-14	1609198	263107	1560709	199601	800771	1652062	1867407	1110794	905349	68572	737721	9801370
2014-15	1605715	288685	1683938	214047	835229	1807689	2073714	1203115	9712133	72805	815441	10527674
2015-16	1616146	317974	1903850	224158	865335	1992825	2294787	1276797	10491670	77659	977632	11309493
2016-17	1726004	349248	2054764	246496	916445	2146379	2492967	1395982	11328285	83003	979099	12308193
2017-18	1840023	329612	2209428	272850	964306	2368419	2537190	1512542	12034171	87586	1110412	13144582
2018-19	1878598	326815	2328992	294147	1026789	2538757	2714222	1625477	12733798	92133	1259116	13992914
2019-20	1994326	317134	2259706	300798	1043429	2690000	2898243	1732404	13236100	94420	1288540	14534641
2020-21	2074212	291075	2329160	288213	99571	2153888	2954130	1601295	12687345	86034	1007525	1364869
2021-22	2177016	309276	2561033	317966	1193532	2480380	3122847	1721699	13876840	94054	1145006	15021846
2022-23	2272250	315256	2504663	347973	1306256	2777723	3405474	1875304	14804901	99404	1266528	16071429
2023-24	2287329	340821	2717235	374125	1445603	2957058	3684059	2020579	15827708	106134	1462573	17290281

Figure 3.3: Dataframe with Dropped Columns

3.2.1.1.2 Sector Selection

For the purpose of univariate visualization, multivariate visualization and model fitting, we created a new dataframe consisting of the 8 sectors with the year as the index.

Year	Agriculture,forestry & fishing	Mining & quarrying	Manufacturing	Electricity, gas, water supply & other utility services	Construction	Trade, hotels, transport, communication and services related to broadcasting	Financial estate & prof servs	Public Administration, defence and other services	GDP
2011-12	1501947	261035	1409986	188668	777335	1413116	1530877	1025982	8736329
2012-13	1524288	262609	1489873	191835	780050	1551143	1680031	1069646	9213017
2013-14	1609198	263107	1560709	199601	800771	1652062	1867407	1110794	9801370
2014-15	1605715	288685	1683938	214047	835229	1807689	2073714	1203115	10527674
2015-16	1616146	317974	1903850	224158	865335	1992825	2294787	1276797	11309493
2016-17	1726004	349248	2054764	246496	916445	2146379	2492967	1395982	12308193
2017-18	1840023	329612	2209428	272850	964306	2368419	2537190	1512542	13144582
2018-19	1878598	326815	2328992	294147	1026789	2538757	2714222	1625477	13992914
2019-20	1994326	317134	2259706	300798	1043429	2690000	2898243	1732404	14534641
2020-21	2074212	291075	2329160	288213	99571	2153888	2954130	1601295	1364869
2021-22	2177016	309276	2561033	317966	1193532	2480380	3122847	1721699	15021846
2022-23	2272250	315256	2504663	347973	1306256	2777723	3405474	1875304	16071429
2023-24	2287329	340821	2717235	374125	1445603	2957058	3684059	2020579	17290281

Figure 3.4: Dataframe with Sectors

The GDP is calculated by combining the contribution of three major categories - Primary, Secondary, and Tertiary, based on how they interact with the natural resources. These categories are further made up of sectors which we have discussed above. Following is the distribution of sectors we have followed to make up these categories

1. Primary

- Agriculture,forestry & fishing
- Mining & quarrying

2. Secondary

- Manufacturing
- Electricity, gas, water supply & other utility services



- Construction

3. Tertiary

- Trade, hotels, transport, communication and services related to broadcasting
- Financial , real estate & prof servs
- Public Administration, defence and other services

3.2.1.1.3 Category Creation

As explained above, we have divided the sectors into categories. We have stored the data of the categories in a new dataframe.

Year	Primary	Secondary	Tertiary
2011-12	1762982	2373989	3969975
2012-13	1786897	2458558	4300820
2013-14	1872305	2561081	4630263
2014-15	1894400	2733214	5084518
2015-16	1934120	2993343	5564409
2016-17	2075252	3217705	6035328
2017-18	2169635	3446384	6418151
2018-19	2205413	3649928	6878456
2019-20	2311460	3603933	7320707
2020-21	2365287	3612744	6709313
2021-22	2479382	4072531	7324926
2022-23	2587506	4158892	8058501
2023-24	2628150	4536963	8662596

Figure 3.5: Dataframe with Categories

3.2.1.2 Feature Extraction

3.2.1.2.1 Handling Outliers

The given below plot shows the data distribution of each sector along with outliers. To ensure better model fitting we have handles outliers and have replaced outliers beyond $Q1 - 1.5 * IQR$ with the lower bound and those beyond $Q3 + 1.5 * IQR$ with the upper bound.

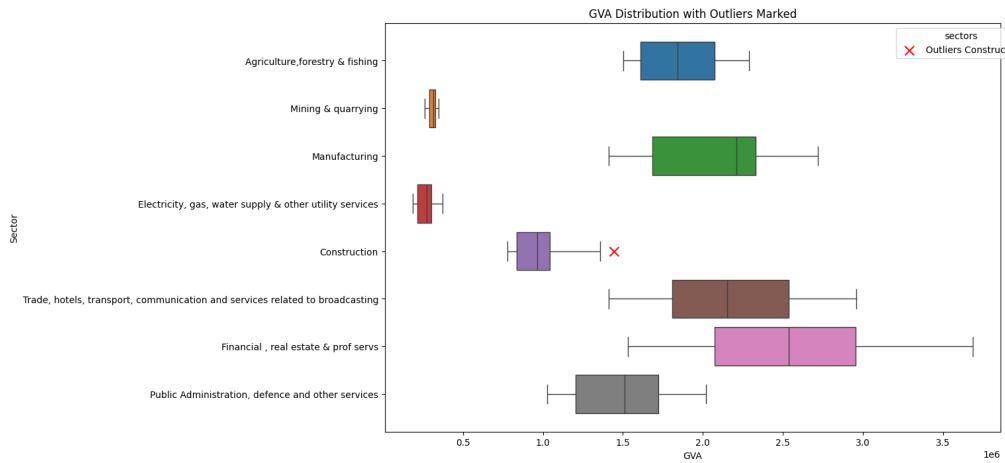


Figure 3.6: Data Distribution of Contribution of Sectors along with Outliers

3.2.1.2.2 Normalising Data

Also to ensure that sectors with higher amounts don't overshadow those with lower amounts, we have performed normalisation on the data.

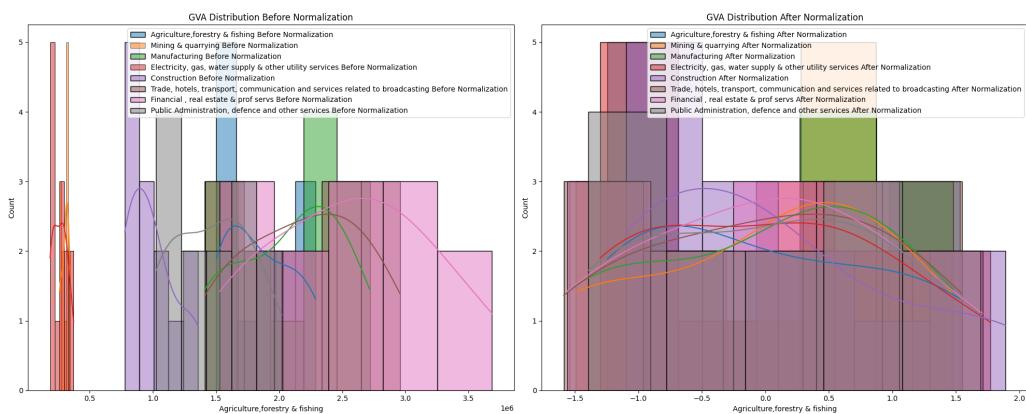


Figure 3.7: Comparison of Data Distribution Before and After Normalisation

3.2.2 Growth Rate of GVA

3.2.2.1 Feature Selection

3.2.2.1.1 Dropping Unnecessary Columns

Here too, the columns of NVA, GNI and NNI are not much of use so we dropped them.



Year	Agriculture,forestry & fishing	Mining & quarrying	Manufacturing	Electricity, gas, water supply & other utility services	Construction	Trade, hotels, transport, communication and services related to broadcasting	Financial , real estate & prof servs	Public Administration, defence and other services	GVA at basic prices	Per capita income(₹.)	Net taxes on products	GDP
2012-13	1.5	0.6	5.5	2.7	0.3	9.8	9.7	4.3	5.4	3.3	5.9	5.5
2013-14	5.6	0.2	5.0	4.2	2.7	6.5	11.2	3.8	6.1	4.6	10.6	6.4
2014-15	-0.2	9.7	7.9	7.2	4.3	9.4	11.0	8.3	7.2	6.2	10.5	7.4
2015-16	0.6	10.1	13.1	4.7	3.6	10.2	10.7	6.1	8.0	6.7	7.6	8.0
2016-17	6.8	9.8	7.9	10.0	5.9	7.7	8.6	9.3	8.0	6.9	11.7	8.3
2017-18	6.6	-5.6	7.5	10.6	5.2	10.3	1.8	8.3	6.2	5.5	13.3	6.8
2018-19	2.1	-0.8	5.4	7.9	6.5	7.2	7.0	7.5	5.8	5.2	13.4	6.5
2019-20	6.2	-3.0	-3.0	2.3	1.6	6.0	6.8	6.6	3.9	2.5	3.1	3.9
2020-21	4.0	-8.2	3.1	-4.2	-4.6	-19.9	1.9	-7.6	-4.1	-8.9	-22.4	-5.8
2021-22	4.6	6.3	10.0	10.3	19.9	15.2	5.7	7.5	9.4	9.3	13.6	9.7
2022-23	4.7	1.9	-2.2	9.4	9.4	12.0	9.1	8.9	6.7	5.7	10.6	7.0
2023-24	0.7	8.1	8.5	7.5	10.7	6.5	8.2	7.7	6.9	6.8	15.5	7.6

Figure 3.8: Dataframe with Dropped Columns

3.2.2.1.2 Dropping Unnecessary Columns

Here too, the columns of NVA, GNI and NNI are not much of use so we dropped them.

Year	Agriculture,forestry & fishing	Mining & quarrying	Manufacturing	Electricity, gas, water supply & other utility services	Construction	Trade, hotels, transport, communication and services related to broadcasting	Financial , real estate & prof servs	Public Administration, defence and other services	GVA at basic prices	Per capita income(₹.)	Net taxes on products	GDP
2012-13	1.5	0.6	5.5	2.7	0.3	9.8	9.7	4.3	5.4	3.3	5.9	5.5
2013-14	5.6	0.2	5.0	4.2	2.7	6.5	11.2	3.8	6.1	4.6	10.6	6.4
2014-15	-0.2	9.7	7.9	7.2	4.3	9.4	11.0	8.3	7.2	6.2	10.5	7.4
2015-16	0.6	10.1	13.1	4.7	3.6	10.2	10.7	6.1	8.0	6.7	7.6	8.0
2016-17	6.8	9.8	7.9	10.0	5.9	7.7	8.6	9.3	8.0	6.9	11.7	8.3
2017-18	6.6	-5.6	7.5	10.6	5.2	10.3	1.8	8.3	6.2	5.5	13.3	6.8
2018-19	2.1	-0.8	5.4	7.9	6.5	7.2	7.0	7.5	5.8	5.2	13.4	6.5
2019-20	6.2	-3.0	-3.0	2.3	1.6	6.0	6.8	6.6	3.9	2.5	3.1	3.9
2020-21	4.0	-8.2	3.1	-4.2	-4.6	-19.9	1.9	-7.6	-4.1	-8.9	-22.4	-5.8
2021-22	4.6	6.3	10.0	10.3	19.9	15.2	5.7	7.5	9.4	9.3	13.6	9.7
2022-23	4.7	1.9	-2.2	9.4	9.4	12.0	9.1	8.9	6.7	5.7	10.6	7.0
2023-24	0.7	8.1	8.5	7.5	10.7	6.5	8.2	7.7	6.9	6.8	15.5	7.6

Figure 3.9: Dataframe with Dropped Columns

3.2.2.1.3 Sector Selection

For the purpose of univariate visualization, multivariate visualization, we created a new dataframe consisting of the 8 sectors with the year as the index.

Year	Agriculture,forestry & fishing	Mining & quarrying	Manufacturing	Electricity, gas, water supply & other utility services	Construction	Trade, hotels, transport, communication and services related to broadcasting	Financial , real estate & prof servs	Public Administration, defence and other services
2012-13	1.5	0.6	5.5	2.7	0.3	9.8	9.7	4.3
2013-14	5.6	0.2	5.0	4.2	2.7	6.5	11.2	3.8
2014-15	-0.2	9.7	7.9	7.2	4.3	9.4	11.0	8.3
2015-16	0.6	10.1	13.1	4.7	3.6	10.2	10.7	6.1
2016-17	6.8	9.8	7.9	10.0	5.9	7.7	8.6	9.3
2017-18	6.6	-5.6	7.5	10.6	5.2	10.3	1.8	8.3
2018-19	2.1	-0.8	5.4	7.9	6.5	7.2	7.0	7.5
2019-20	6.2	-3.0	-3.0	2.3	1.6	6.0	6.8	6.6
2020-21	4.0	-8.2	3.1	-4.2	-4.6	-19.9	1.9	-7.6
2021-22	4.6	6.3	10.0	10.3	19.9	15.2	5.7	7.5
2022-23	4.7	1.9	-2.2	9.4	9.4	12.0	9.1	8.9
2023-24	0.7	8.1	8.5	7.5	10.7	6.5	8.2	7.7

Figure 3.10: Dataframe with Sectors



3.3 Contribution of States

3.3.1 GSDP

3.3.1.1 Feature Selection

3.3.1.1.1 Transposing Dataframe

We have stored a transposed copy of the dataframe into a new dataframe which will help us in creating multivariate plots.

Year	2011-12	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22	2022-23	2023-24
Andhra Pradesh	379402	380629	407115	444564	498606	540212	594737	626614	649810	659678	722727	764685	820894
Arunachal Pradesh	11063	11299	12339	14383	14240	14893	15572	16668	19140	18435	19707	21085	21803
Assam	143175	147342	154525	165212	191109	202081	219919	231040	240707	247819	278835	300799	318559
Bihar	247144	256851	269650	279482	296488	318797	344028	381383	398329	368733	399930	442473	459682
Chhattisgarh	158074	165977	182580	185813	190584	213705	220136	244579	251549	251192	280609	302119	321945

Figure 3.11: Transposed Dataframe of GSDP Dataframe

3.3.1.1.2 Creating Zones

For the sake of better visualisation we have divided the state into zones based on the part of India they are located in. The division is as follows

- North:
 - Haryana
 - Himachal Pradesh
 - Punjab
 - Uttar Pradesh
 - Uttarakhand
- Central:
 - Chhattisgarh
 - Madhya Pradesh
- East:
 - Bihar
 - Jharkhand
 - Odisha
 - West Bengal
- North-East:



- Arunachal Pradesh
- Assam
- Manipur
- Meghalaya
- Mizoram
- Nagaland
- Sikkim
- Tripura

- West:

- Goa
- Gujarat
- Maharashtra
- Rajasthan

- South:

- Andhra Pradesh
- Karnataka
- Kerala
- Tamil Nadu
- Telangana

- UT:

- Andaman & Nicobar Islands
- Chandigarh
- Delhi
- Jammu & Kashmir-UT*
- Puducherry



	South	North-East	East	Central	West	North	UT	grid
Year								bar
2011-12	2460380	236880	1149534	473636	2373179	1476265	461618	
2012-13	2573292	244928	1205655	517660	2531006	1561034	489146	
2013-14	2756295	259469	1259893	547714	2703697	1666056	523773	
2014-15	2923196	280676	1311045	569757	2916218	1747430	556546	
2015-16	3213242	310552	1373143	609320	3158180	1900671	621708	

Figure 3.12: Data of Each Zone

3.3.1.2 Feature Extraction

3.3.1.2.1 Handling Outliers

The given below plot shows the data distribution of GSDP of each state along with outliers. To ensure better model fitting we have handles outliers and have replaced outliers beyond $Q1 - 1.5 * IQR$ with the lower bound and those beyond $Q3 + 1.5 * IQR$ with the upper bound.

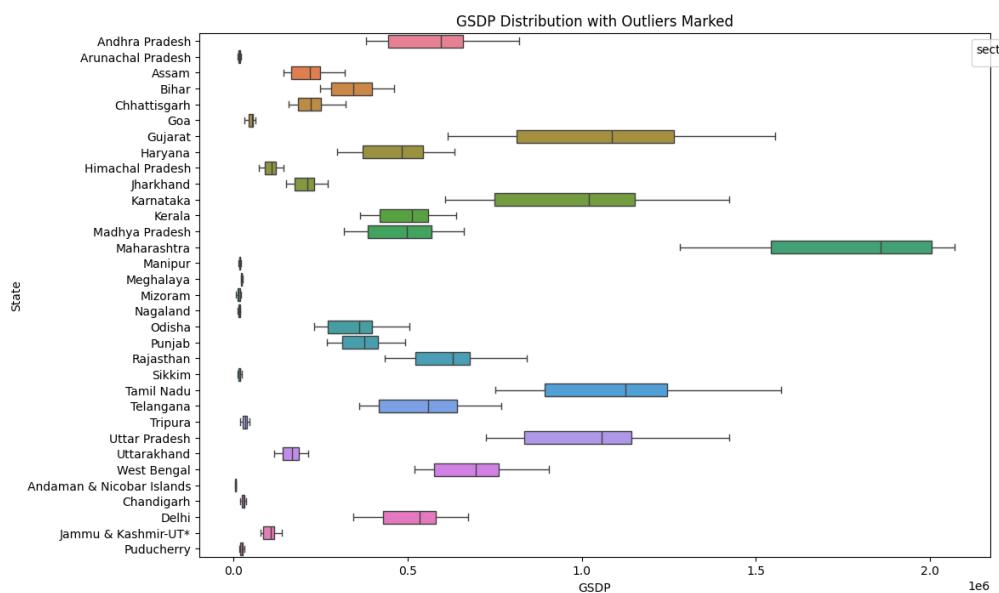


Figure 3.13: Data Distribution of GSDP of States along with Outliers

3.3.1.2.2 Normalising Data

Also to ensure that states with higher GSDP don't overshadow those with lower GSDP, we have performed normalisation on the data.

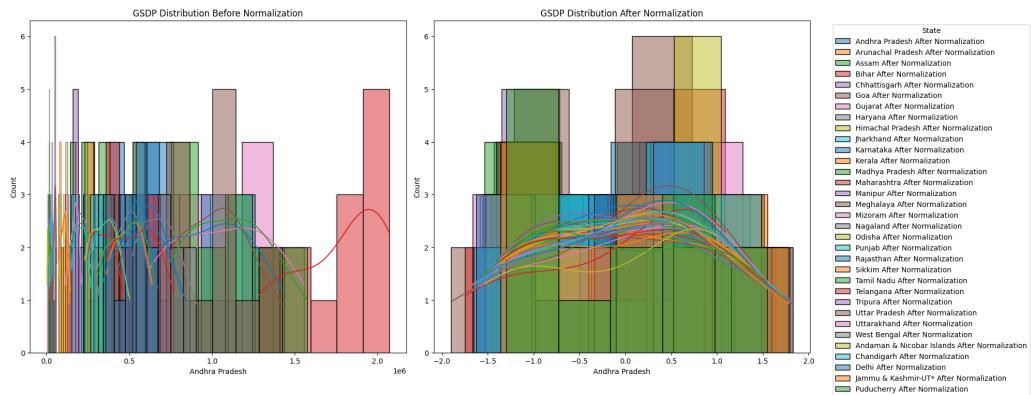


Figure 3.14: Comparison of Data Distribution Before and After Normalisation

3.3.2 Growth Rate of GSDP

3.3.2.1 Feature Selection

3.3.2.1.1 Transposing Dataframe

We have stored a transposed copy of the dataframe into a new dataframe which will help us in creating multivariate plots.

Year	2012-13	2013-14	2014-15	2015-16	2016-17	2017-18	2018-19	2019-20	2020-21	2021-22	2022-23	2023-24
Andhra Pradesh	0.32	6.96	9.20	12.16	8.34	10.09	5.36	3.70	1.52	9.56	5.81	7.35
Arunachal Pradesh	2.14	9.21	16.56	-0.99	4.58	4.56	7.04	14.83	-3.69	6.90	6.99	3.40
Assam	2.91	4.88	6.92	15.67	5.74	8.83	5.06	4.18	2.95	12.52	7.88	5.90
Bihar	3.93	4.98	3.65	6.08	7.52	7.91	10.86	4.44	-7.43	8.46	10.64	3.89
Chhattisgarh	5.00	10.00	1.77	2.57	12.13	3.01	11.10	2.85	-0.14	11.71	7.67	6.56

Next steps: [View recommended plots](#)

Figure 3.15: Transposed Dataframe of Growth Rate of GSDP Dataframe

3.3.2.1.2 Adding GDP Column to Datfaramne

To the existing dataframe of the growth rate we have added the column of growth rate of GDP. This dataframe will help us in bivariate visualisation.

Year	Andhra Pradesh	Arunachal Pradesh	Assam	Bihar	Chhattisgarh	Goa	Gujarat	Haryana	Jharkhand	Jharkhand	Tripura	Uttar Pradesh	Uttarakhand	West Bengal	Andaman & Nicobar Islands	Chandigarh	Delhi	Jammu & Kashmir-UT*	Puducherry	GDP	
2012-13	0.32	2.14	2.91	3.93	5.00	-15.380000	10.890000	7.86	6.41	8.17	...	8.670000	4.72	7.27	4.17	4.480000	8.000000	6.64	3.21	2.93	5.5
2013-14	6.96	9.21	4.88	4.98	10.00	-11.940000	7.560000	8.29	7.06	1.57	...	9.320000	5.79	8.47	3.01	7.990000	8.970000	7.17	5.37	10.74	6.4
2014-15	9.20	16.56	6.92	3.65	1.77	27.080000	10.510000	6.63	7.50	12.49	...	18.170000	4.03	5.29	2.83	5.640000	3.460000	9.02	-3.21	-5.03	7.4
2015-16	12.16	-0.99	15.67	6.08	2.57	14.800000	10.230000	11.57	8.10	-6.25	...	-0.660000	8.85	8.08	6.13	7.390000	9.020000	11.03	17.76	4.69	8.0
2016-17	8.34	4.58	5.74	7.52	12.11	11.190000	9.710000	10.48	7.04	10.46	...	14.000000	11.37	9.83	7.20	12.960000	7.960000	7.60	3.30	7.42	8.3
2017-18	10.09	4.56	8.83	7.91	3.01	2.740000	10.720000	5.55	6.16	9.01	...	0.370000	4.44	7.90	6.36	12.390000	5.800000	5.91	6.41	9.00	6.8
2018-19	5.36	7.04	5.06	10.86	11.10	0.780000	0.880000	10.57	6.41	8.87	...	11.000000	3.88	2.83	6.32	6.230000	4.870000	4.30	7.91	17.44	6.5
2019-20	3.70	14.83	4.18	4.44	2.85	3.300000	6.950000	2.28	4.13	1.08	...	3.560000	4.03	1.97	3.10	5.810000	8.070000	3.69	-0.99	0.63	3.9
2020-21	1.52	-3.69	2.95	-7.43	-0.14	4.160000	-1.910000	9.00	-4.35	-5.30	...	-4.360000	-4.28	-12.10	-7.58	-7.580000	-8.730000	-6.96	-1.33	-6.48	-5.8
2021-22	9.56	9.90	12.52	8.46	11.71	5.360000	10.060000	9.72	7.80	10.87	...	8.470000	9.82	10.49	11.85	6.810000	6.960000	8.76	8.03	12.41	9.7
2022-23	5.81	6.99	7.88	10.64	7.67	8.160000	8.030000	7.88	6.89	6.76	...	8.890000	9.83	7.63	6.65	1.680000	10.760000	7.85	8.73	6.33	7.0
2023-24	7.35	3.40	5.90	3.89	6.56	3.123333	5.393333	7.97	7.07	4.11	...	4.333333	7.99	7.58	7.65	0.303333	2.663333	7.39	7.41	3.42	7.6

Figure 3.16: Adding GDP Column

Chapter 4. Visualization

4.1 World

4.1.1 GDP of G20 Countries

4.1.1.1 Univariate Analysis

Here we will analyse the GDP of the G20 countries individually.

4.1.1.1.1 Argentina

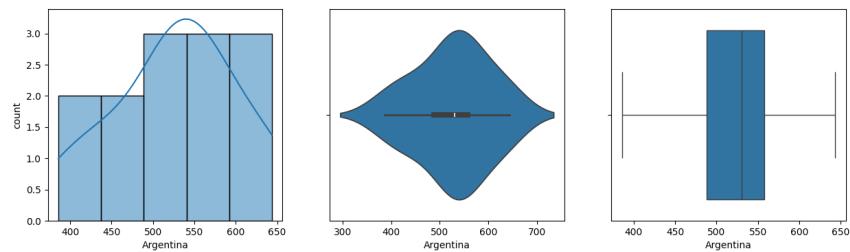


Figure 4.1: Data Distribution of GDP of Argentina

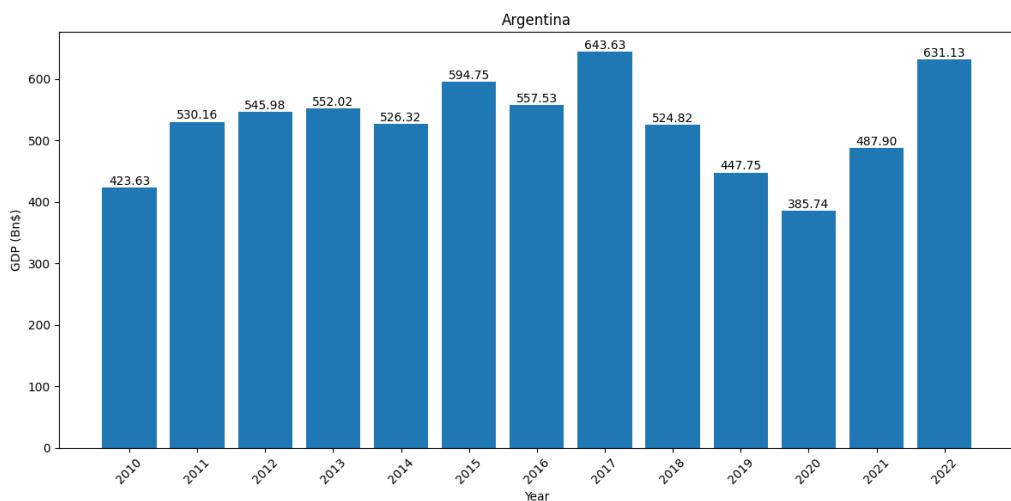


Figure 4.2: Value Distribution of GDP of Argentina



4.1.1.1.2 Australia

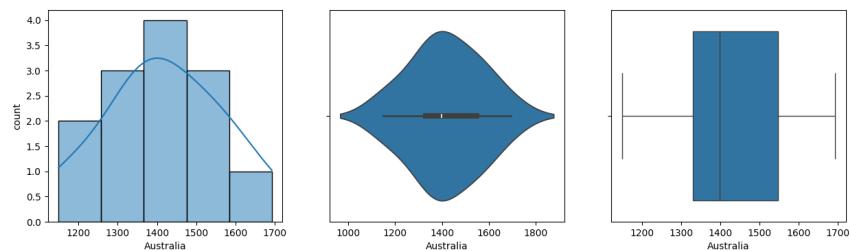


Figure 4.3: Data Distribution of GDP of Australia

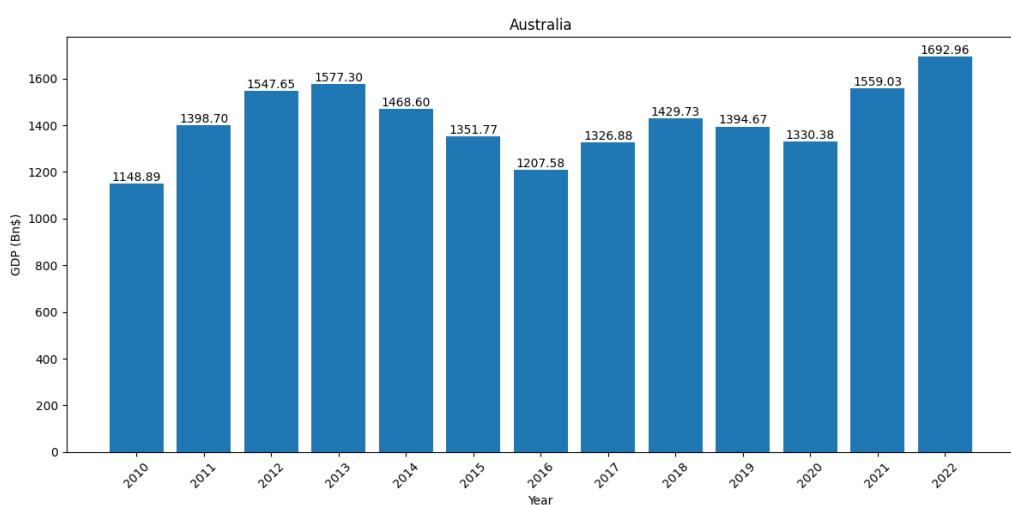


Figure 4.4: Value Distribution of GDP of Australia

4.1.1.1.3 Brazil

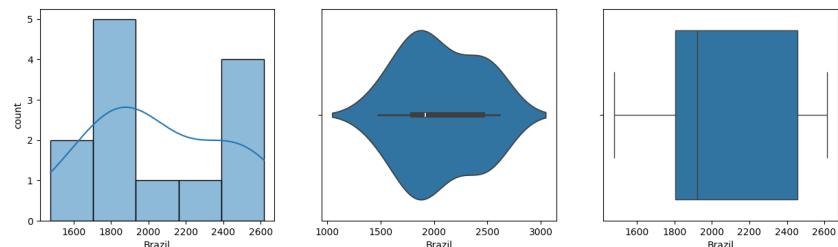


Figure 4.5: Data Distribution of GDP of Brazil

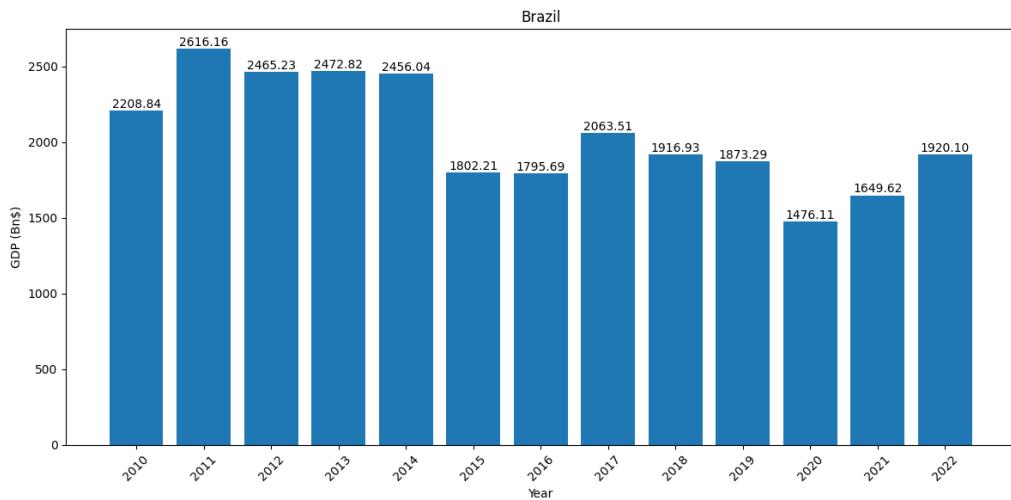


Figure 4.6: Value Distribution of GDP of Brazil

4.1.1.1.4 Canada

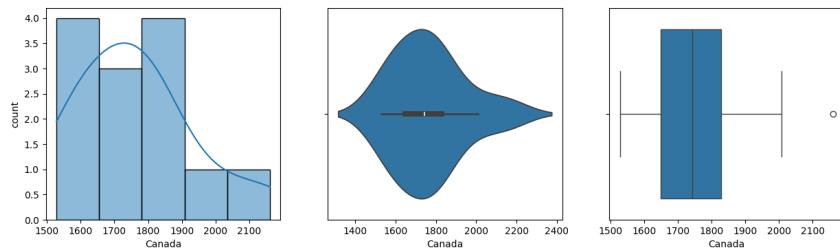


Figure 4.7: Data Distribution of GDP of Canada

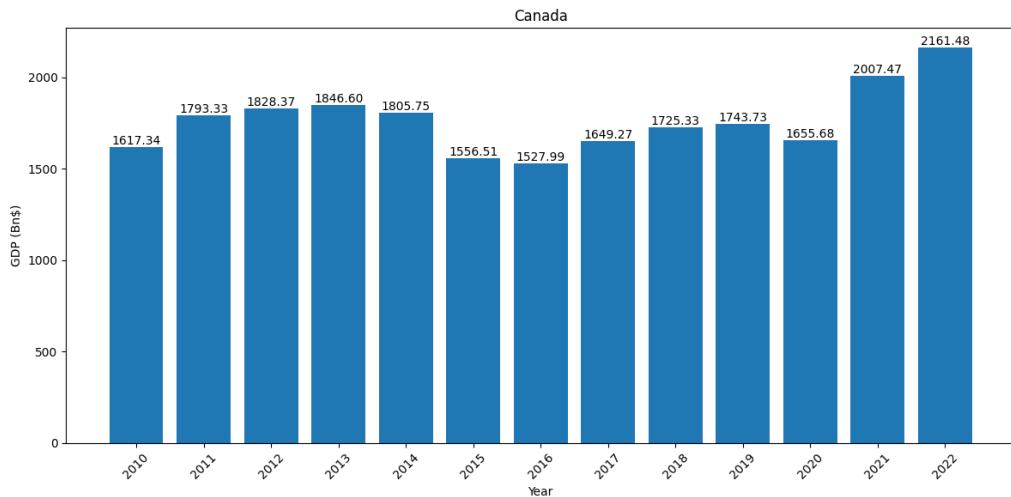


Figure 4.8: Value Distribution of GDP of Canada

4.1.1.1.5 China

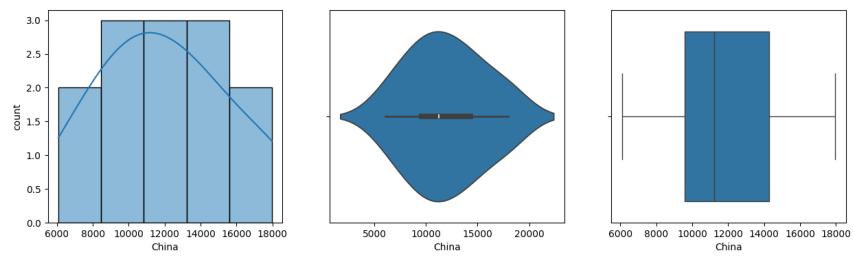


Figure 4.9: Data Distribution of GDP of China

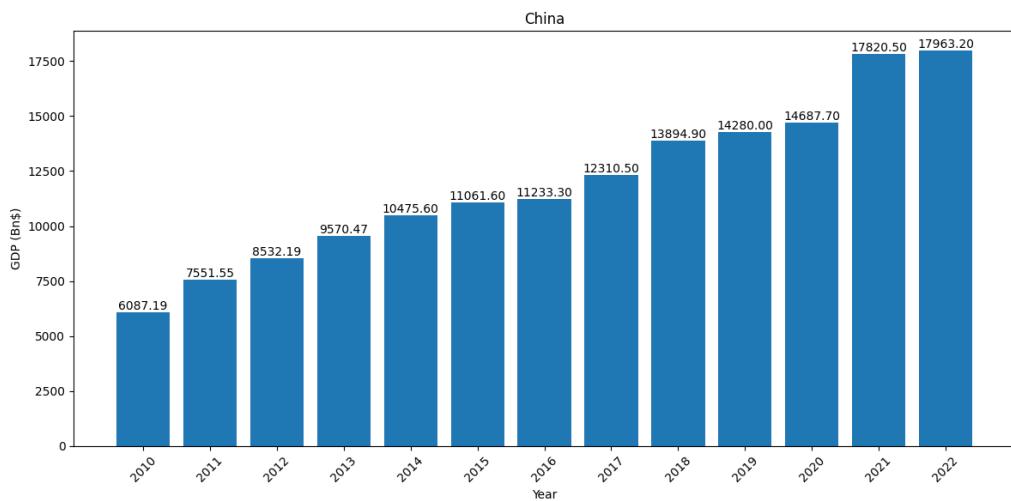


Figure 4.10: Value Distribution of GDP of China

4.1.1.1.6 France

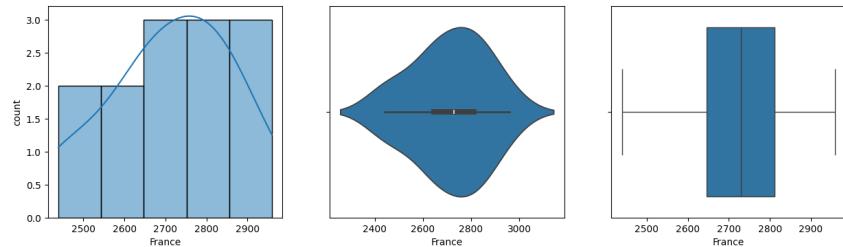


Figure 4.11: Data Distribution of GDP of France

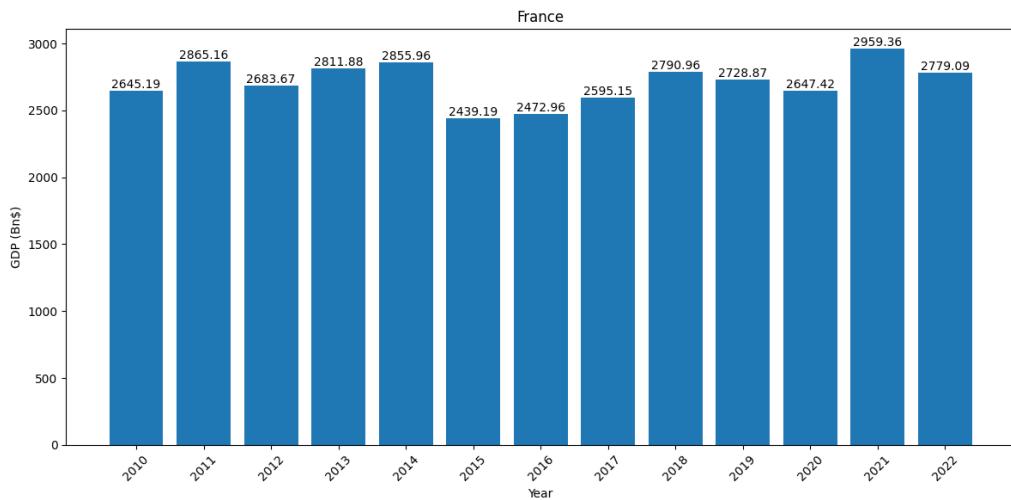


Figure 4.12: Value Distribution of GDP of France

4.1.1.1.7 Germany

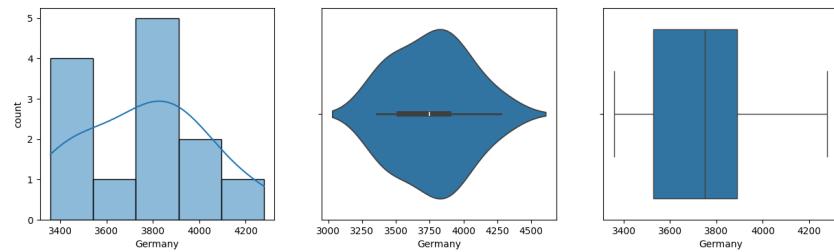


Figure 4.13: Data Distribution of GDP of Germany

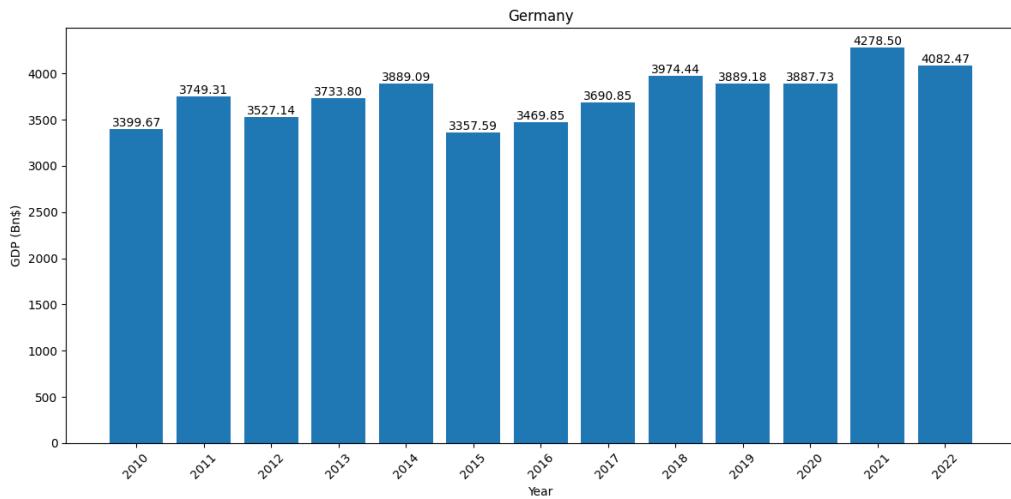


Figure 4.14: Value Distribution of GDP of Germany

4.1.1.1.8 India

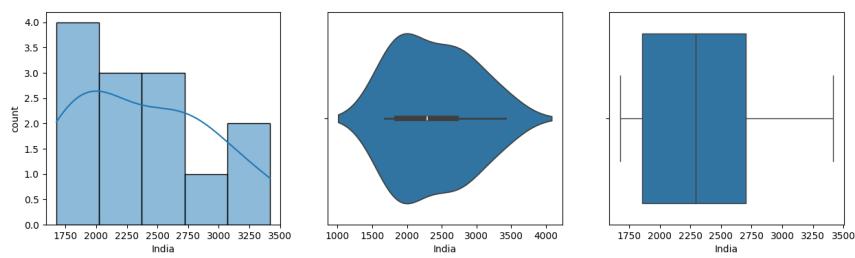


Figure 4.15: Data Distribution of GDP of India

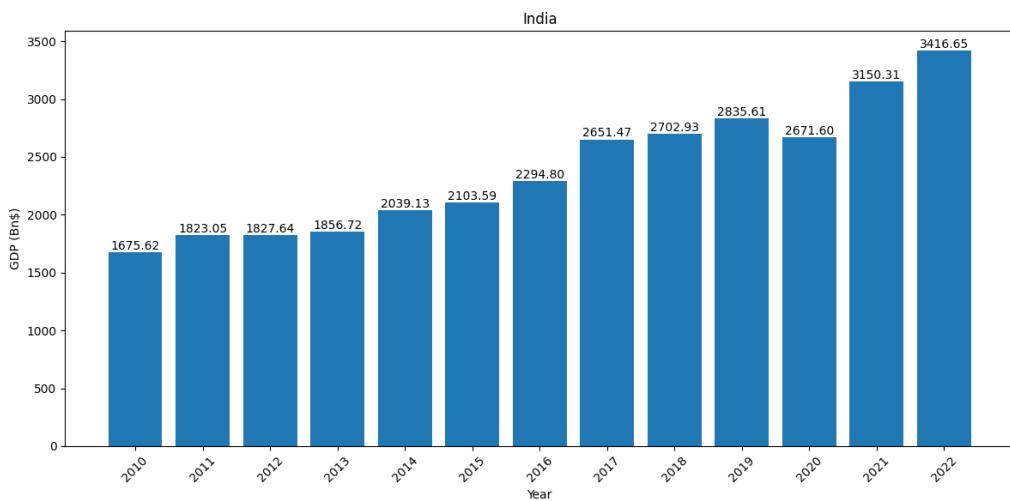


Figure 4.16: Value Distribution of GDP of India

4.1.1.1.9 Indonesia

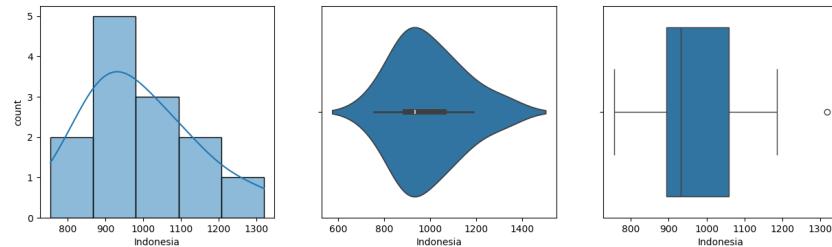


Figure 4.17: Data Distribution of GDP of Indonesia

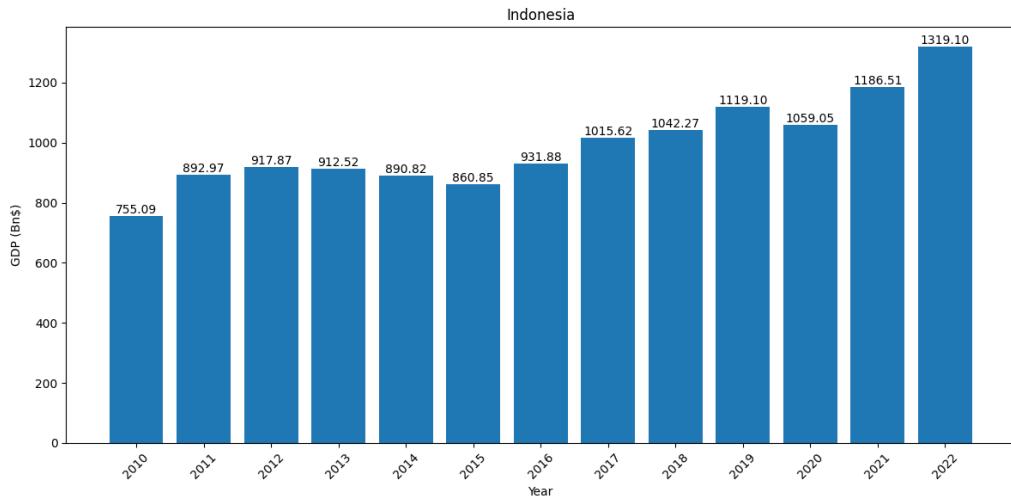


Figure 4.18: Value Distribution of GDP of Indonesia

4.1.1.1.10 Italy

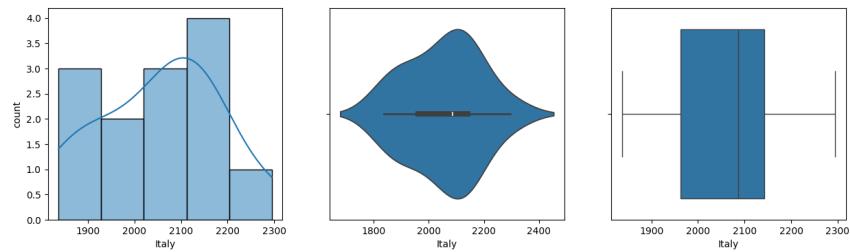


Figure 4.19: Data Distribution of GDP of Italy

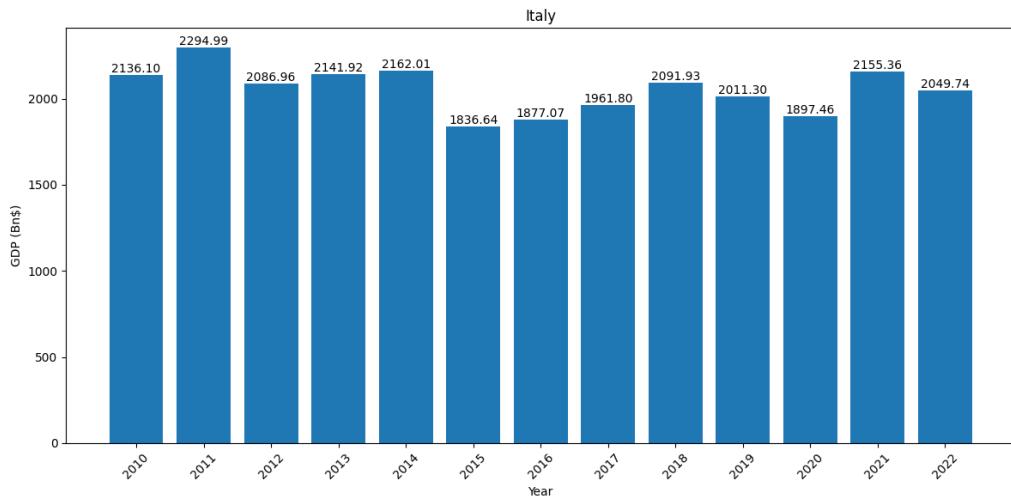


Figure 4.20: Value Distribution of GDP of Italy

4.1.1.1.11 Japan

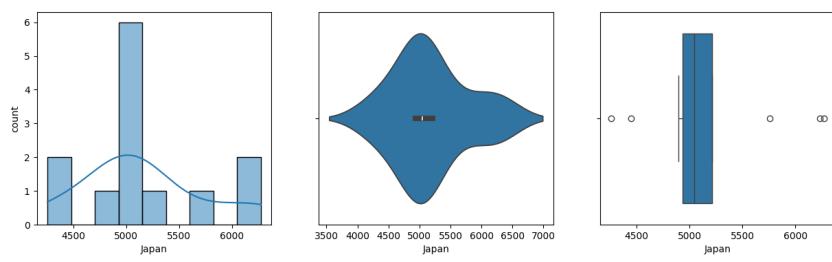


Figure 4.21: Data Distribution of GDP of Japan

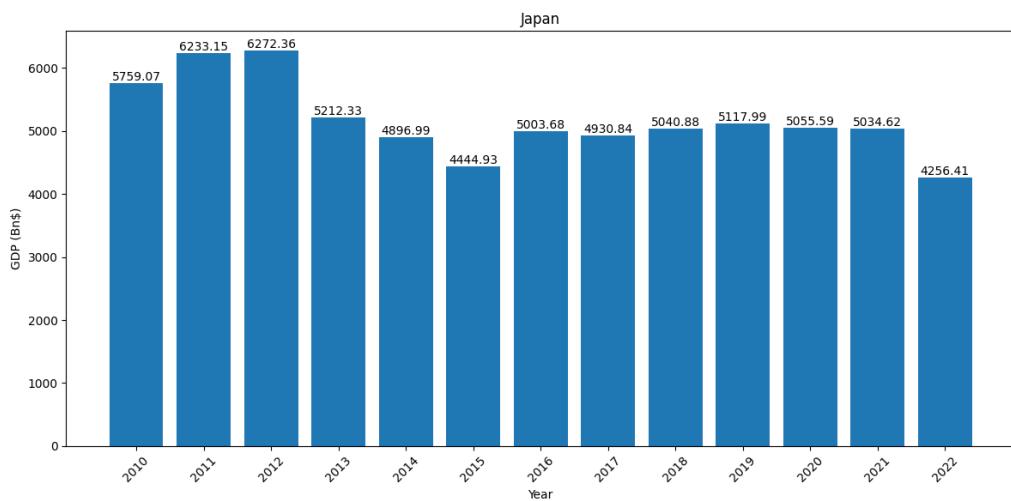


Figure 4.22: Value Distribution of GDP of Japan

4.1.1.12 Mexico

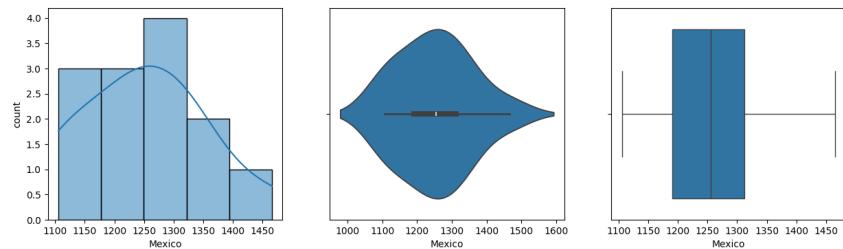


Figure 4.23: Data Distribution of GDP of Mexico

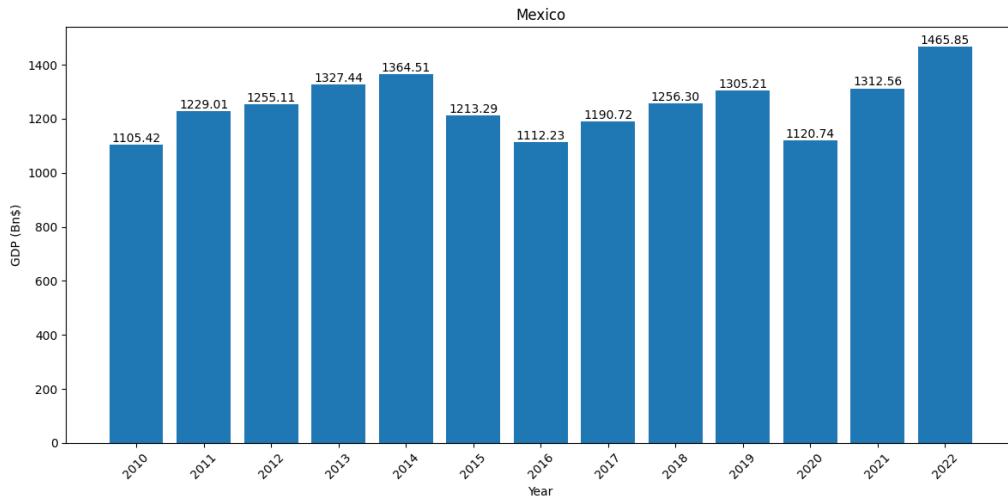


Figure 4.24: Value Distribution of GDP of Mexico

4.1.1.13 Russian

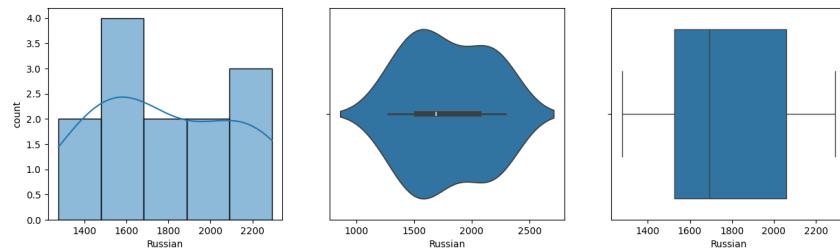


Figure 4.25: Data Distribution of GDP of Russian

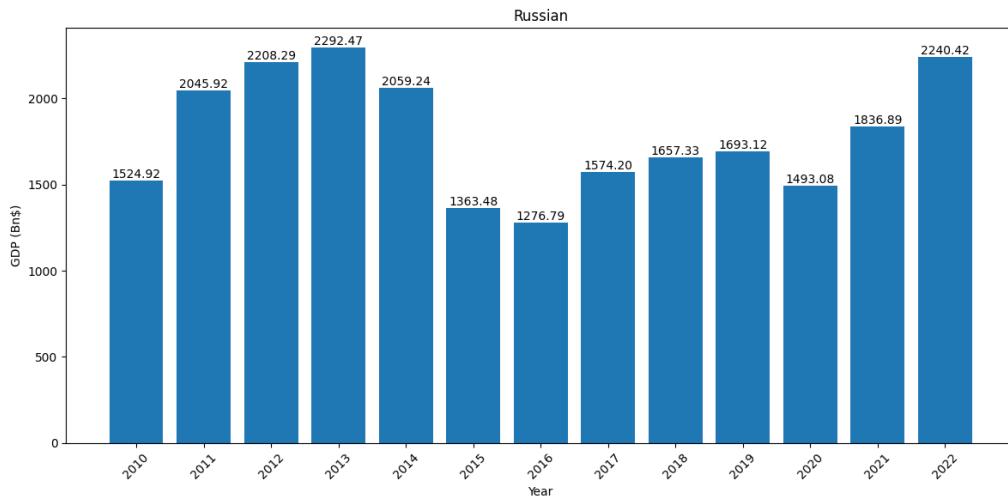


Figure 4.26: Value Distribution of GDP of Russian

4.1.1.14 Saudi Arabia

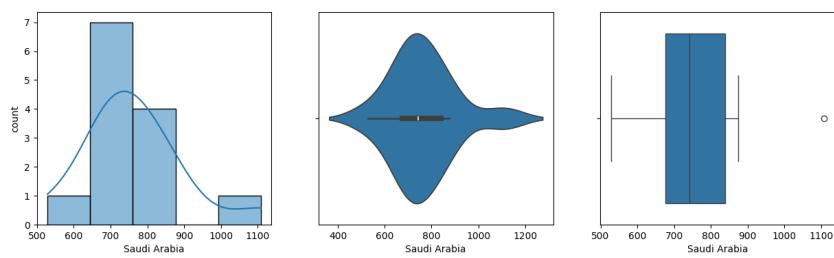


Figure 4.27: Data Distribution of GDP of Saudi Arabia

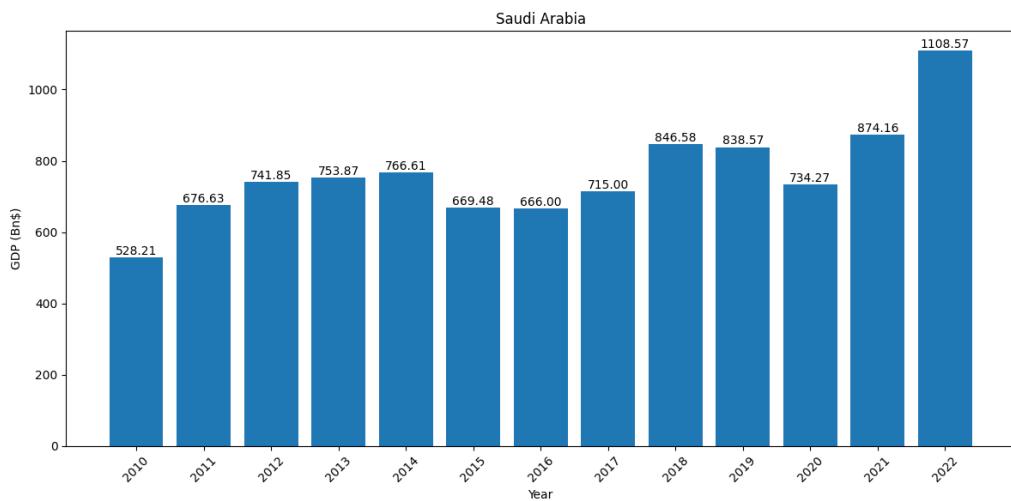


Figure 4.28: Value Distribution of GDP of Saudi Arabia

4.1.1.15 South Africa

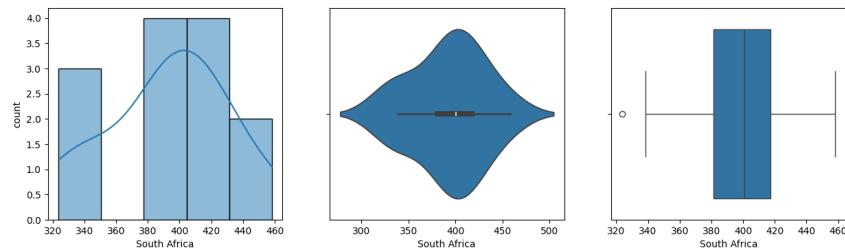


Figure 4.29: Data Distribution of GDP of South Africa

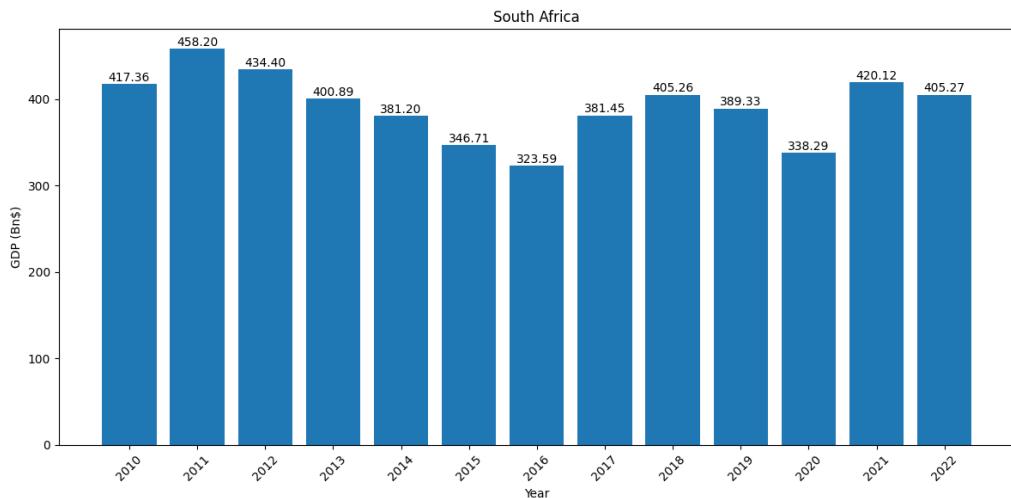


Figure 4.30: Value Distribution of GDP of South Africa

4.1.1.16 South Korea

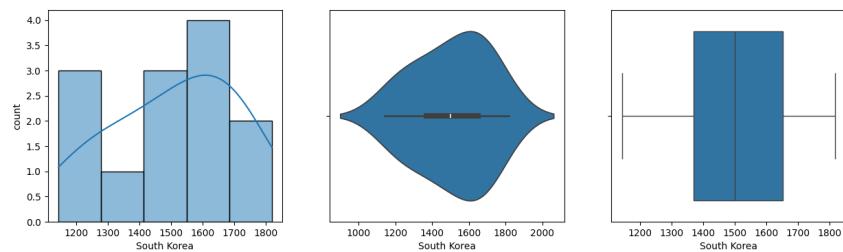


Figure 4.31: Data Distribution of GDP of South Korea

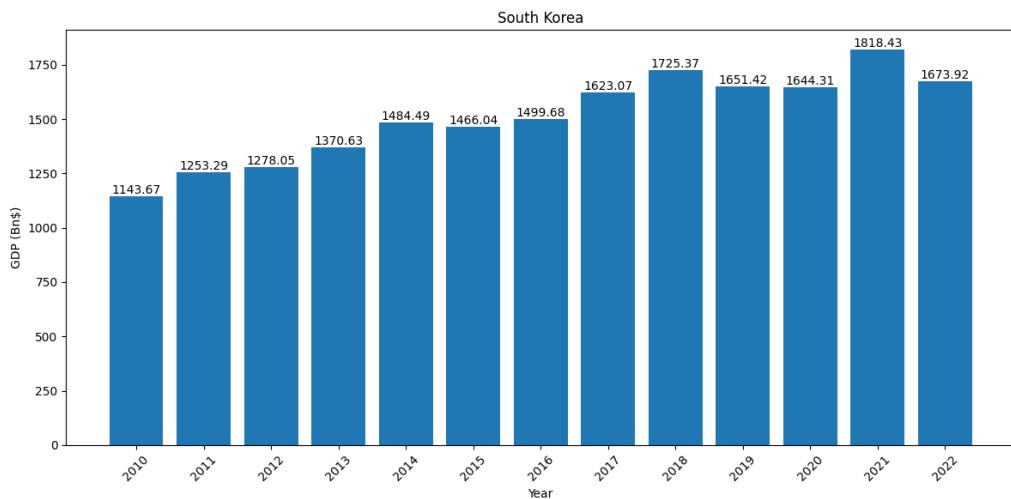


Figure 4.32: Value Distribution of GDP of South Korea

4.1.1.17 Turkey

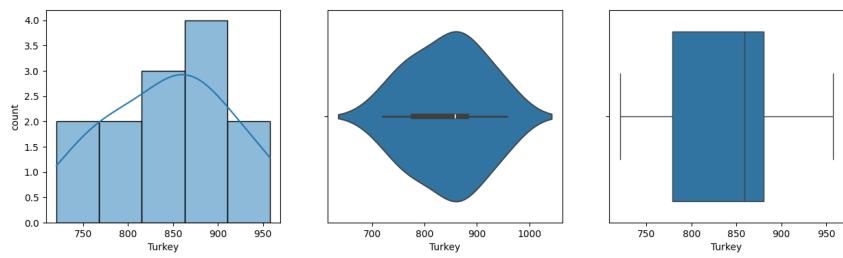


Figure 4.33: Data Distribution of GDP of Turkey

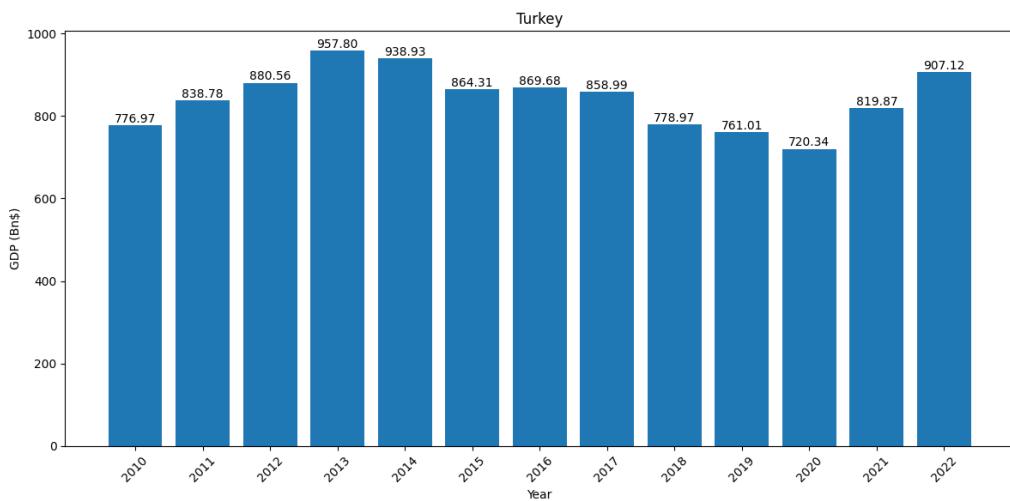


Figure 4.34: Value Distribution of GDP of Turkey

4.1.1.1.18 United Kingdom

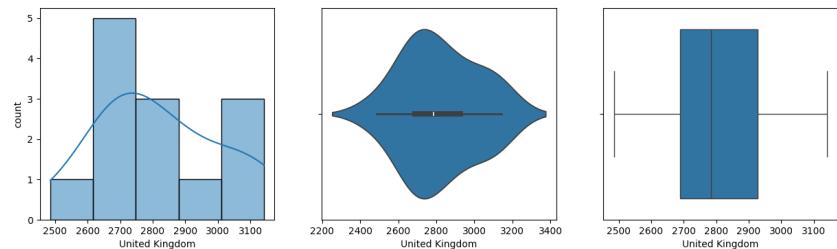


Figure 4.35: Data Distribution of GDP of United Kingdom

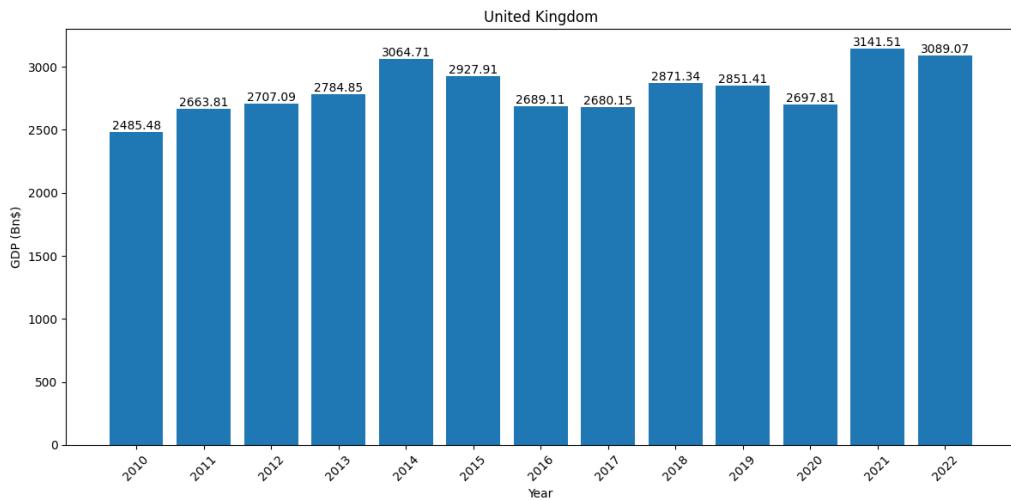


Figure 4.36: Value Distribution of GDP of United Kingdom

4.1.1.1.19 United States

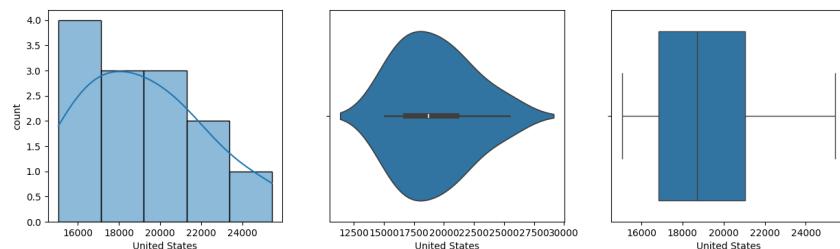


Figure 4.37: Data Distribution of GDP of United States

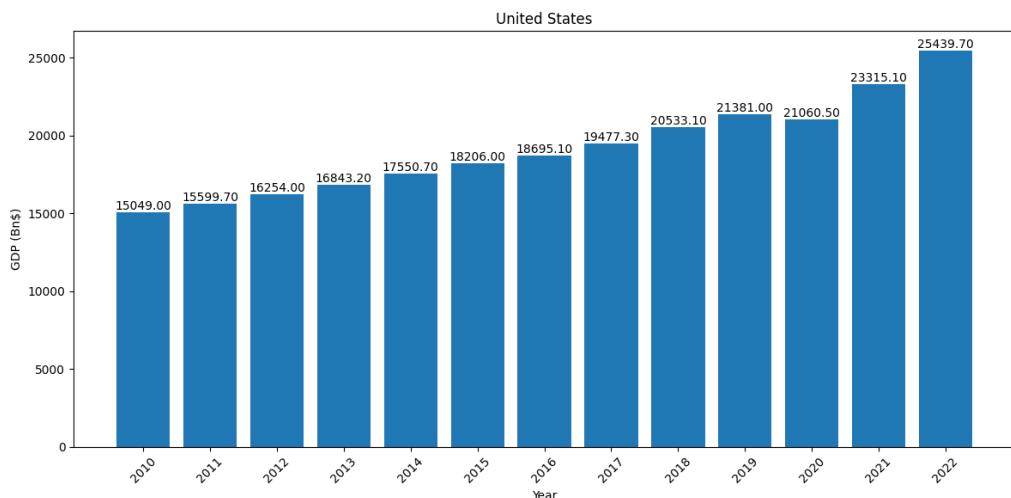


Figure 4.38: Value Distribution of GDP of United States



4.1.1.2 Bivariate Analysis

Here we will analyse the GDP between two countries.

4.1.1.2.1 Comparison with India's GDP

Given below is the comparison of the GDP we of G20 countries with the GDP of India over several years.

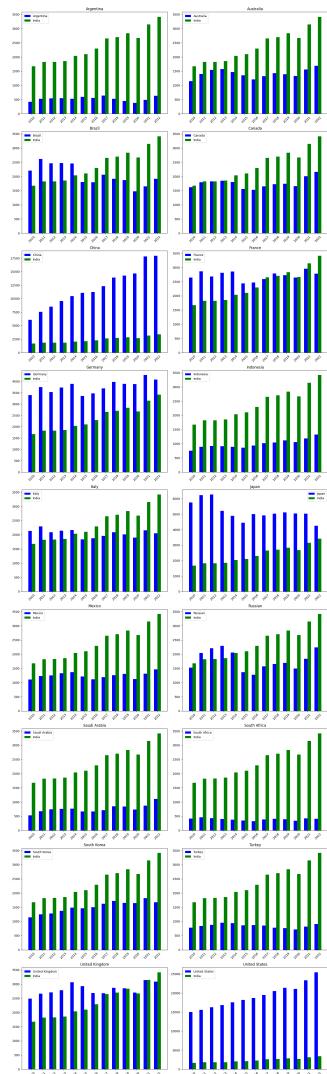


Figure 4.39: Comparison with India's GDP

4.1.1.3 Multivariate analysis

Here we will compare the GDP of various G20 countries together.

4.1.1.3.1 Trend of GDP of G20 Countries

Given below is the growth of the GDP of G20 countries over the years in terms of absolute value.

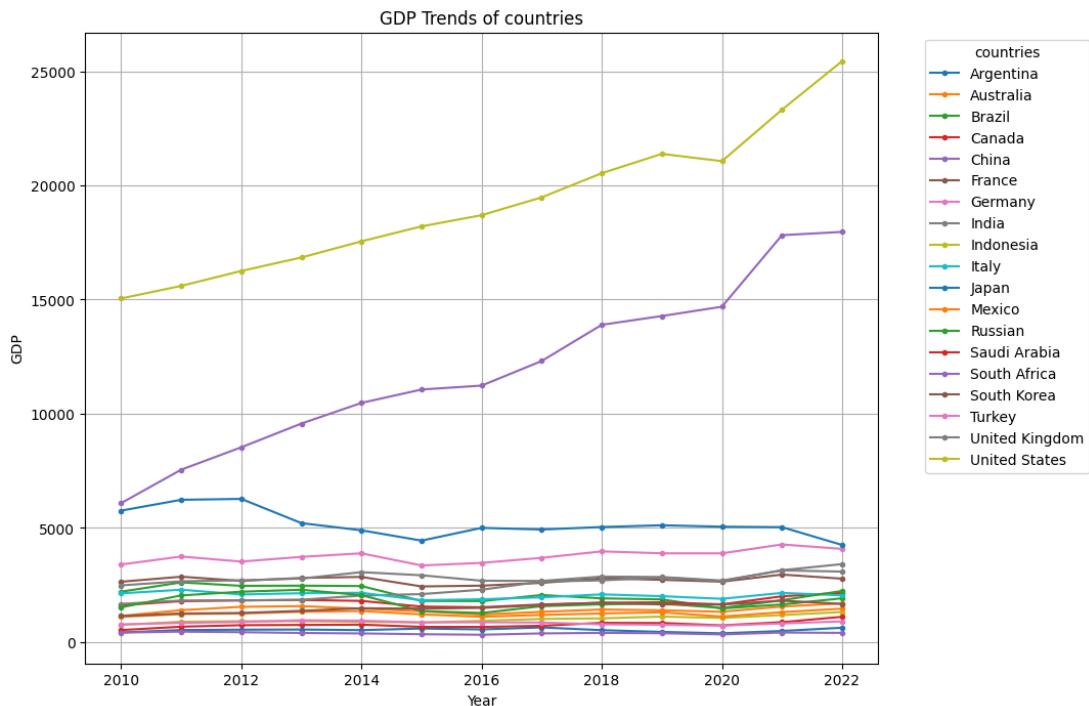


Figure 4.40: Growth of the GDP of G20 Countries Over the Years in Terms of Absolute Value

4.1.1.3.2 GDP of G20 Countries in 2023

Given below is a sorted representation of the 2023's GDP of the G20 countries.

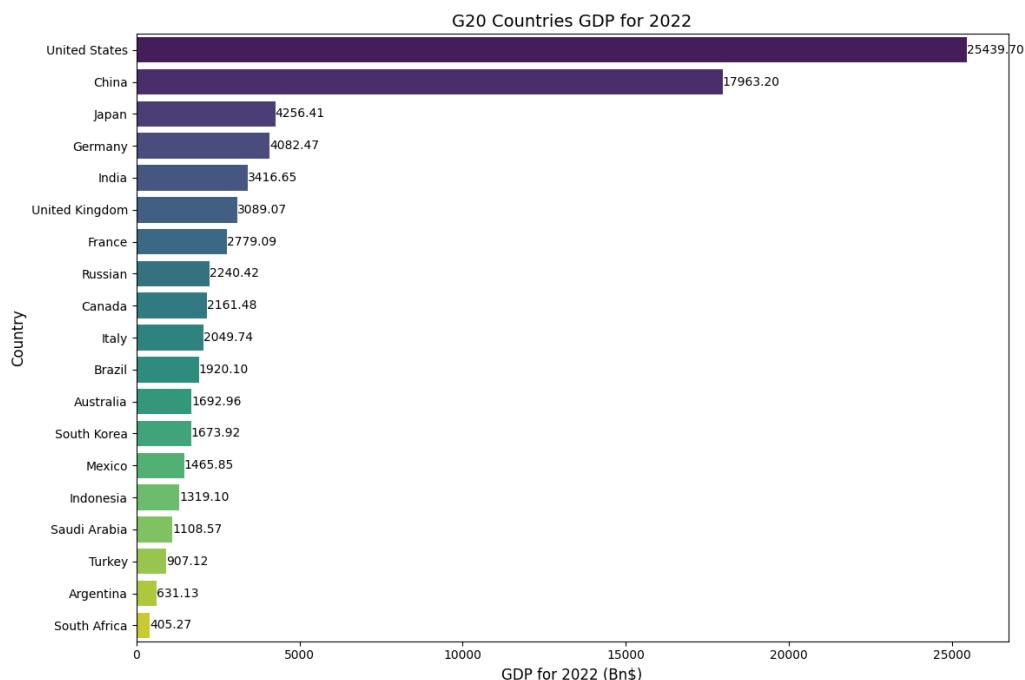


Figure 4.41: GDP of G20 Countries in 2023



4.2 Contribution of Sectors

4.2.1 GVA

4.2.1.1 Univariate Analysis

Here we will analyse the contribution of all the sectors individually.

4.2.1.1.1 Agriculture, forestry & fishing

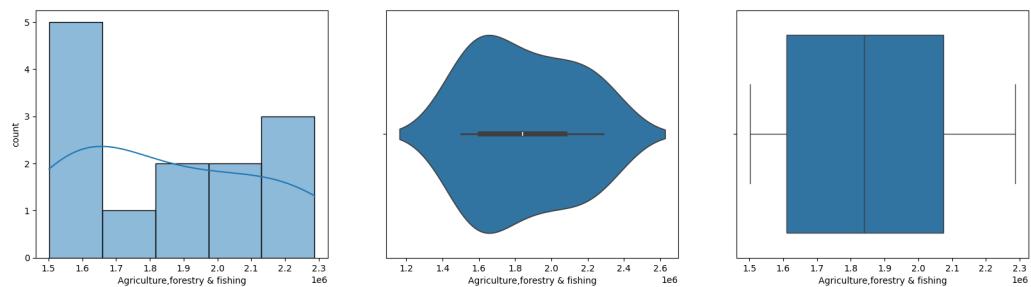


Figure 4.42: Data Distribution of Agriculture, forestry & fishing

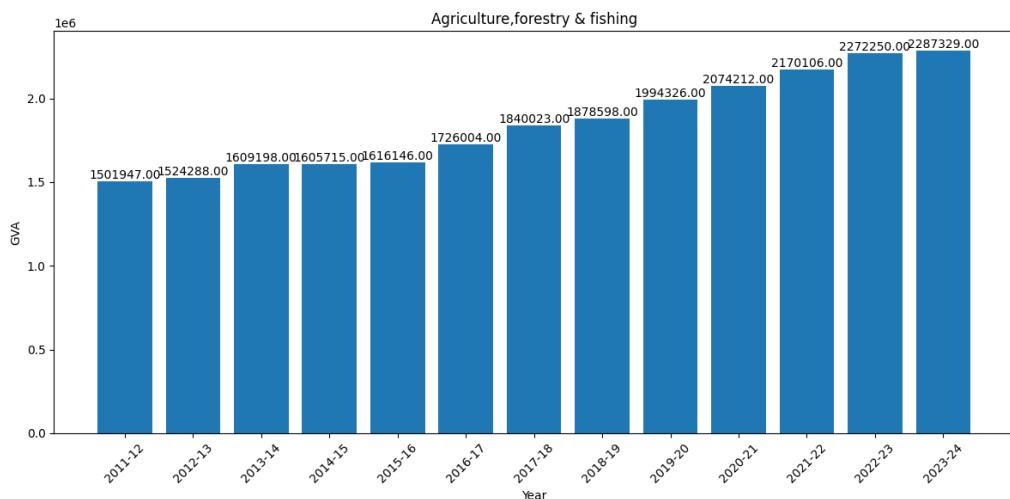


Figure 4.43: Value Distribution of Agriculture,forestry & fishing

4.2.1.1.2 Mining & quarrying

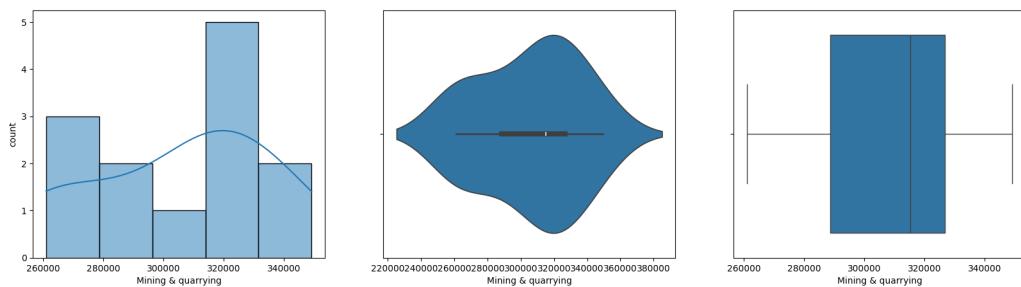


Figure 4.44: Data Distribution of Mining & quarrying

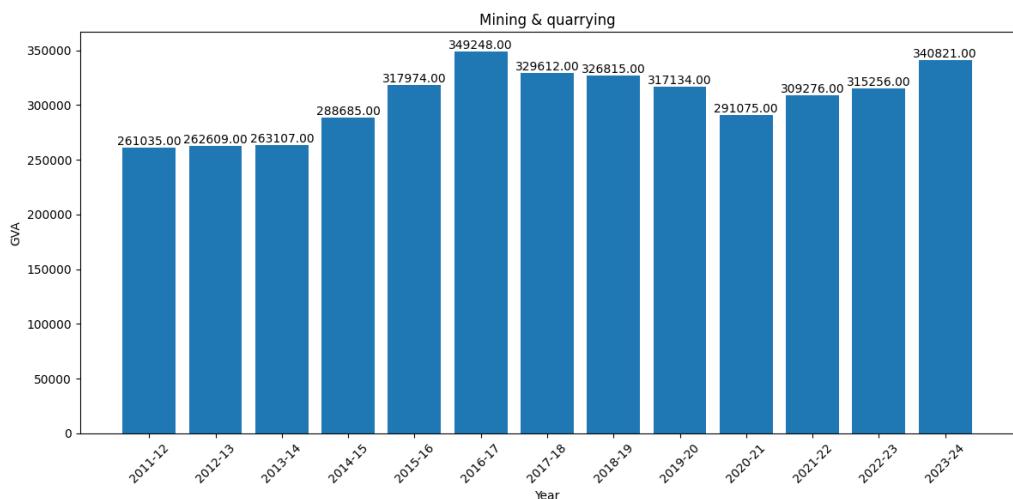


Figure 4.45: Value Distribution of Mining & quarrying

4.2.1.1.3 Manufacturing

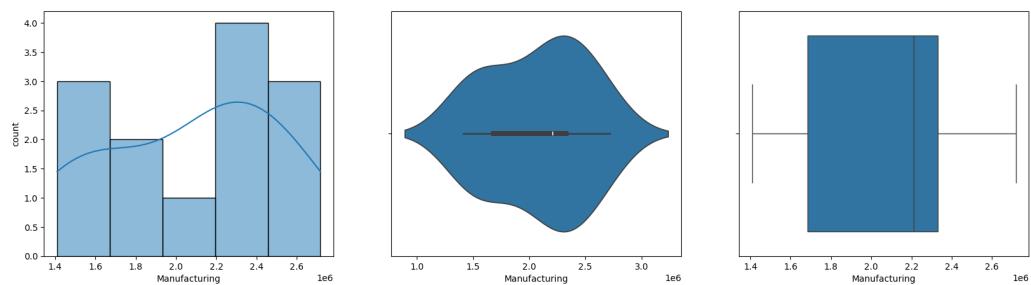


Figure 4.46: Data Distribution of Manufacturing

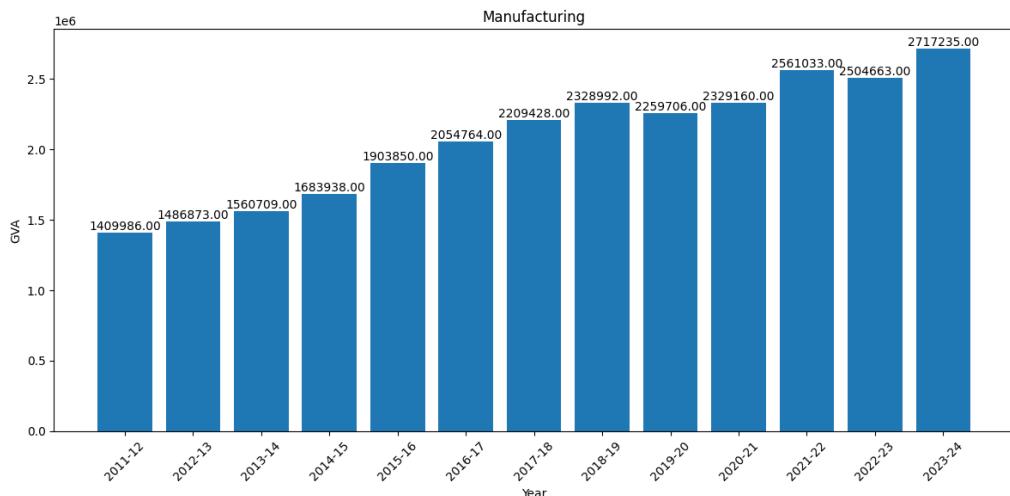


Figure 4.47: Value Distribution of Manufacturing

4.2.1.1.4 Electricity, gas, water supply & other utility services

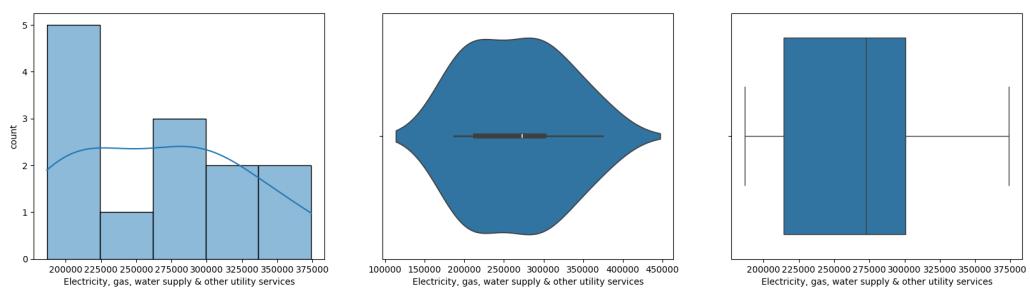


Figure 4.48: Data Distribution of Electricity, gas, water supply & other utility services

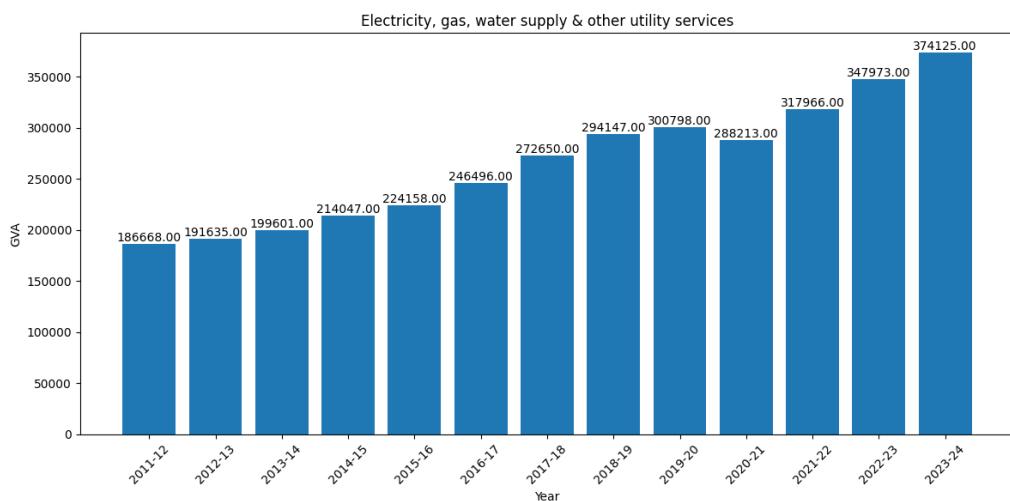


Figure 4.49: Value Distribution of Electricity, gas, water supply & other utility services

4.2.1.1.5 Construction

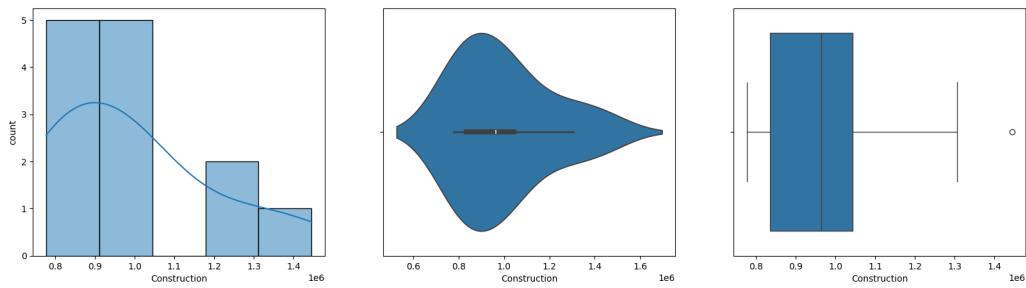


Figure 4.50: Data Distribution of Construction

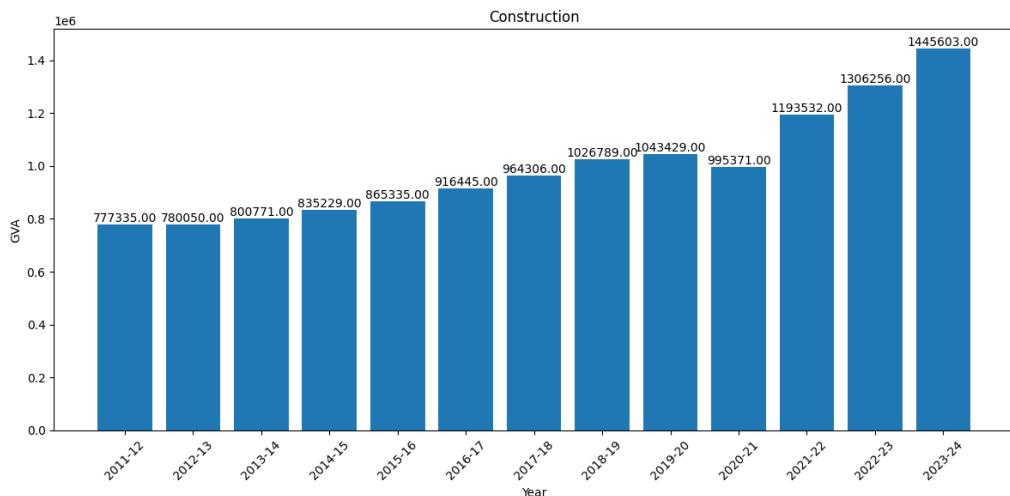


Figure 4.51: Value Distribution of Construction

4.2.1.1.6 Trade, hotels, transport, communication and services related to broadcasting

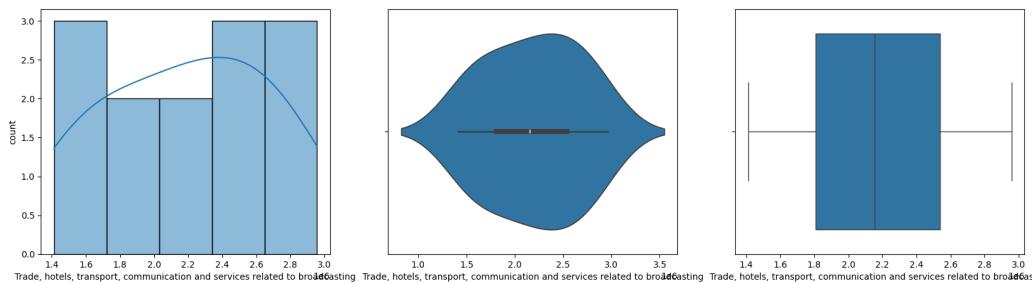


Figure 4.52: Data Distribution of Trade, hotels, transport, communication and services related to broadcasting

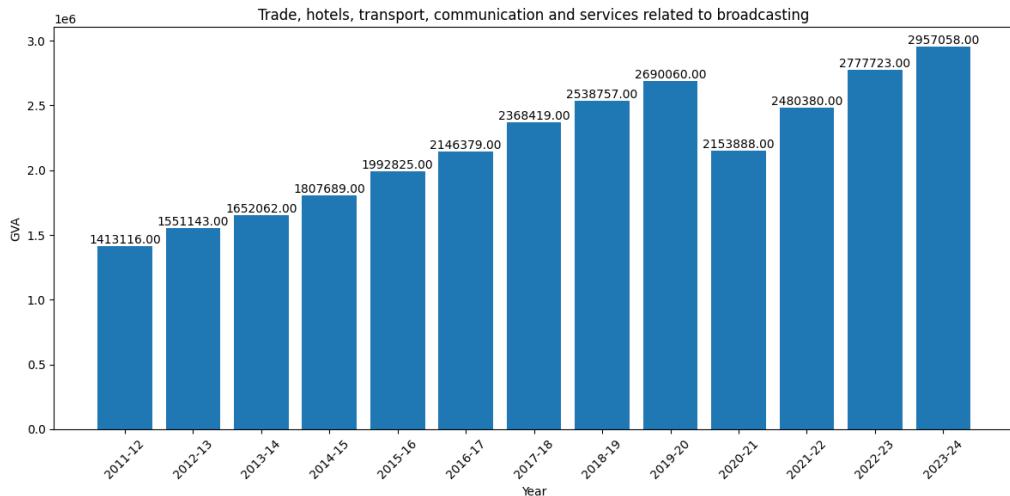


Figure 4.53: Value Distribution of Trade, hotels, transport, communication and services related to broadcasting

4.2.1.1.7 Financial , real estate & prof servs

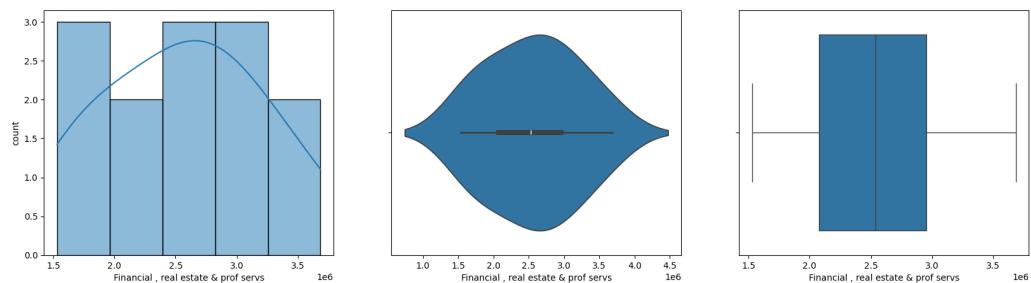


Figure 4.54: Data Distribution of Financial , real estate & prof servs

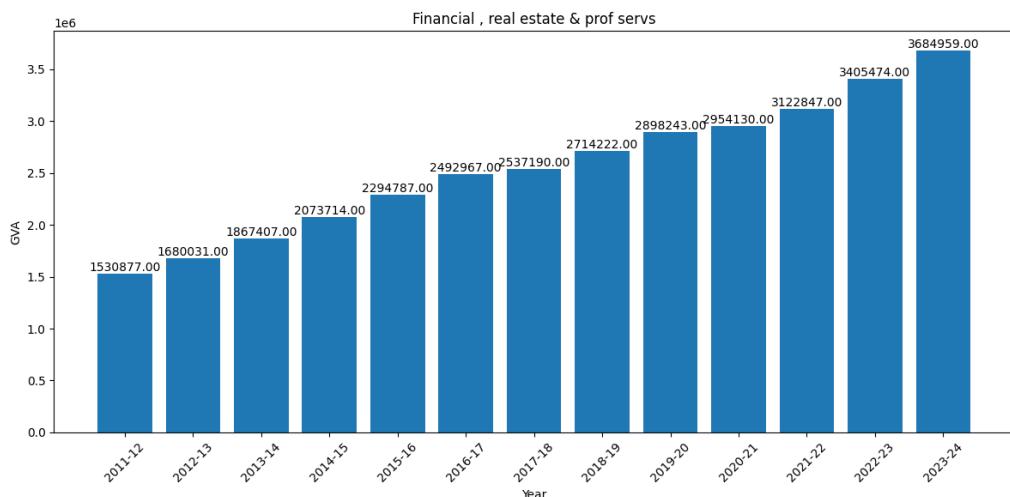


Figure 4.55: Value Distribution of Financial , real estate & prof servs

4.2.1.1.8 Public Administration, defence and other services

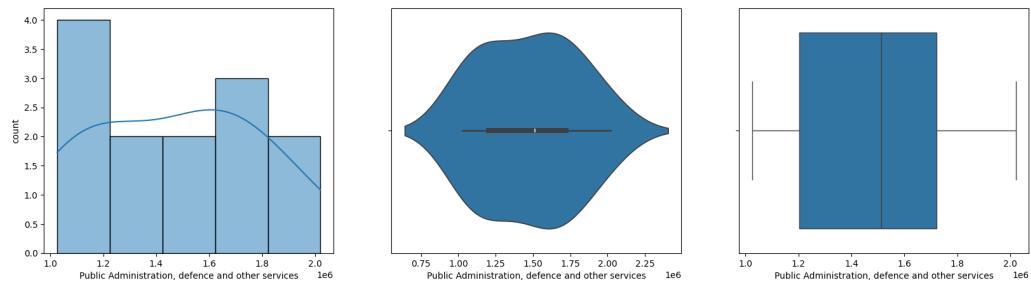


Figure 4.56: Data Distribution of Public Administration, defence and other services

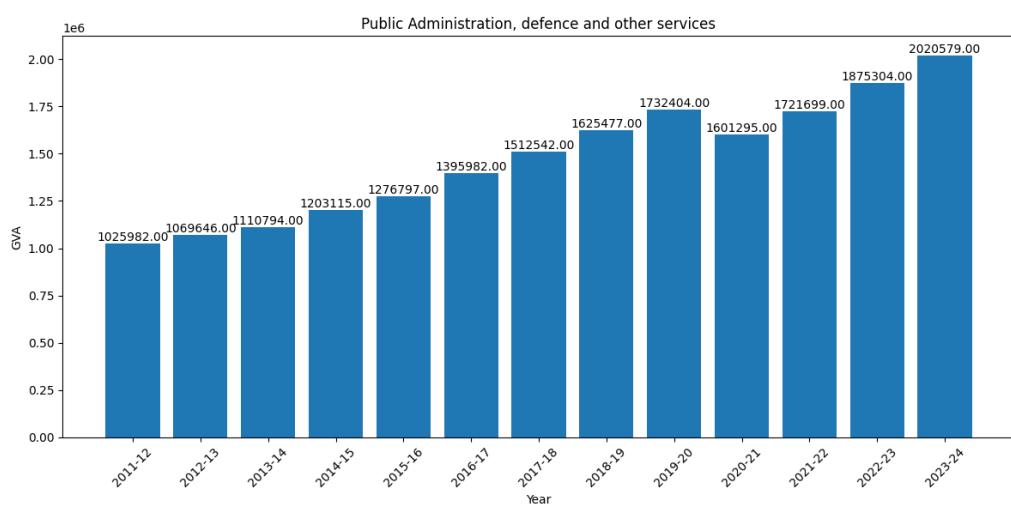


Figure 4.57: Value Distribution of Public Administration, defence and other services

4.2.1.1.9 GVA at basic prices

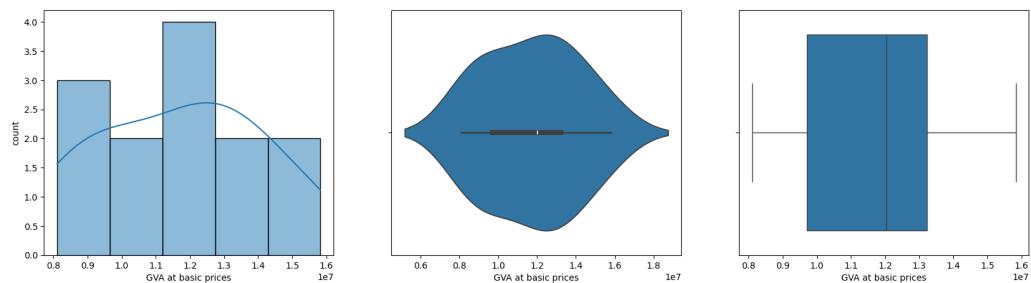


Figure 4.58: Data Distribution of GVA at basic prices

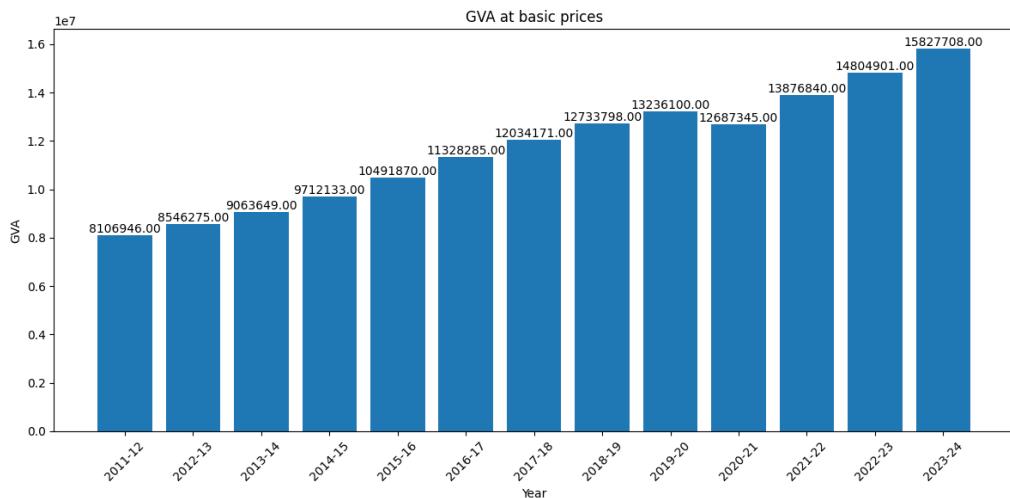


Figure 4.59: Value Distribution of GVA at basic prices

4.2.1.1.10 NVA at basic prices

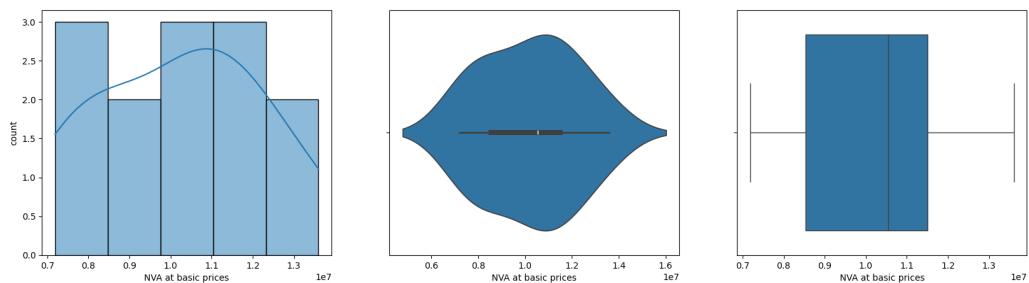


Figure 4.60: Data Distribution of NVA at basic prices

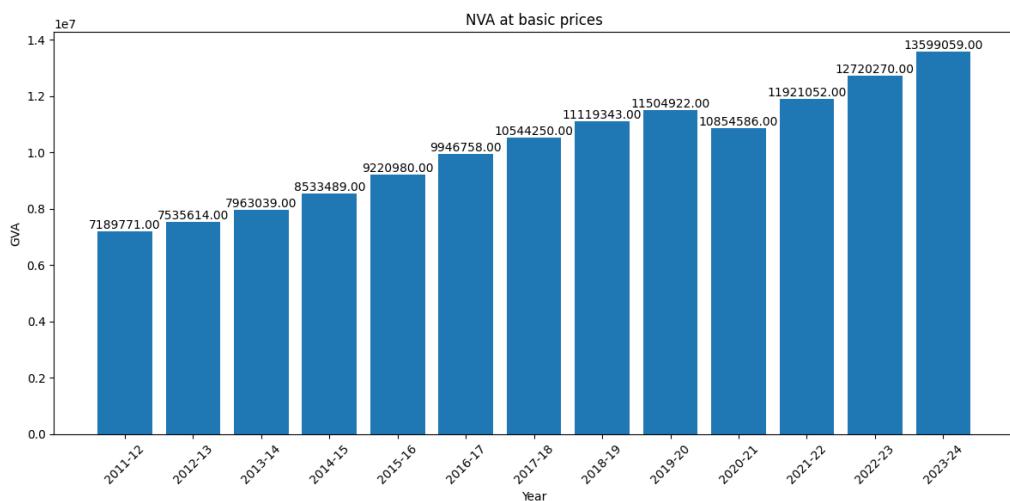


Figure 4.61: Value Distribution of NVA at basic prices

4.2.1.1.11 GNI

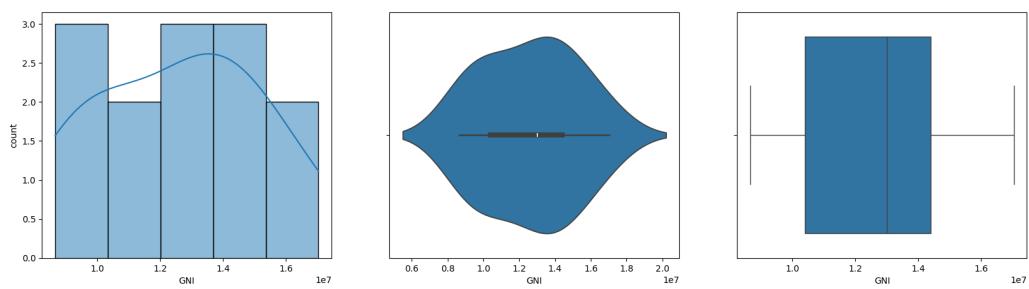


Figure 4.62: Data Distribution of GNI

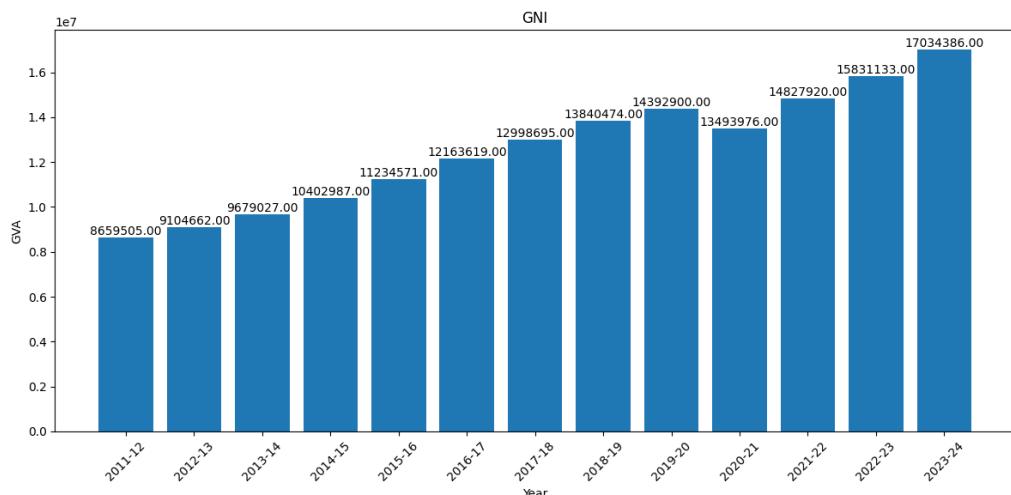


Figure 4.63: Value Distribution of GNI

4.2.1.1.12 NNI

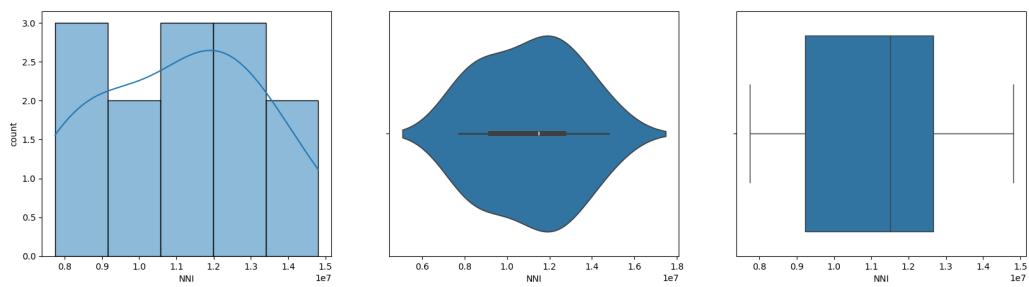


Figure 4.64: Data Distribution of NNI

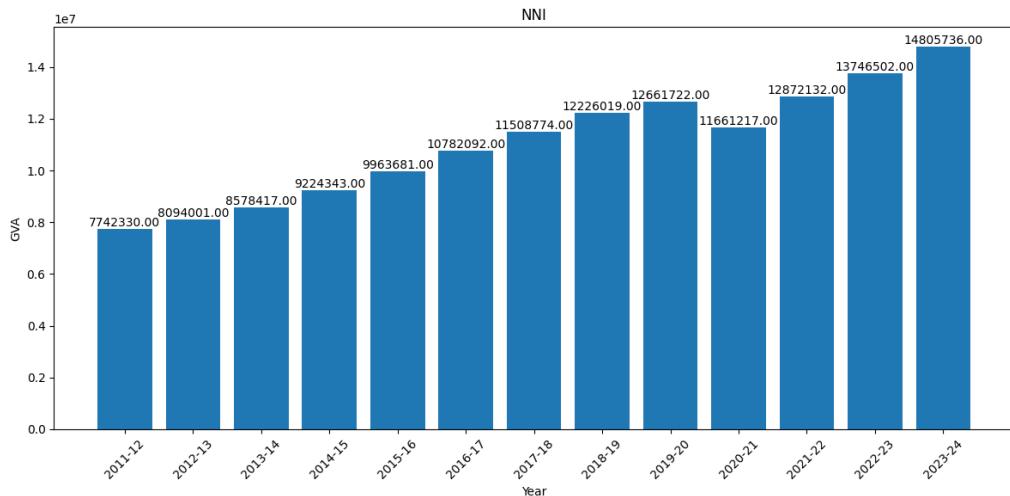


Figure 4.65: Value Distribution of NNI

4.2.1.1.13 Per capita income(Rs.)

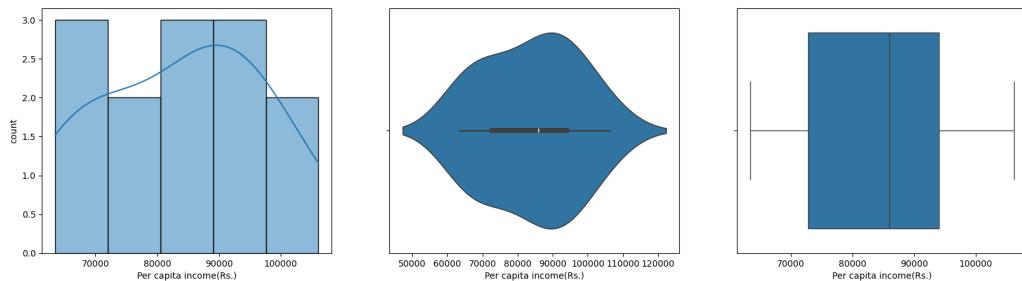


Figure 4.66: Data Distribution of Per capita income(Rs.)

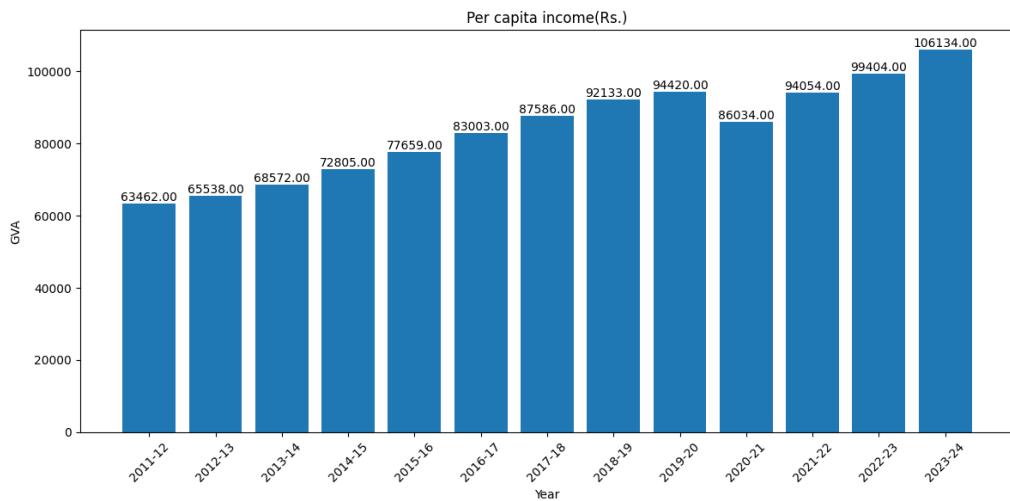


Figure 4.67: Value Distribution of Per capita income(Rs.)

4.2.1.1.14 Net taxes on Products

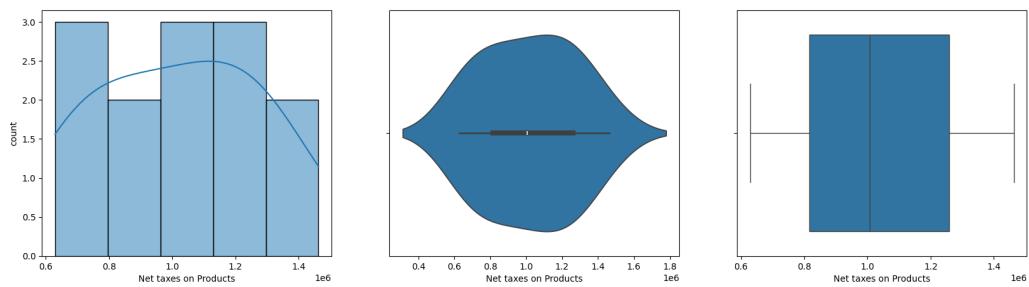


Figure 4.68: Data Distribution of Net taxes on Products

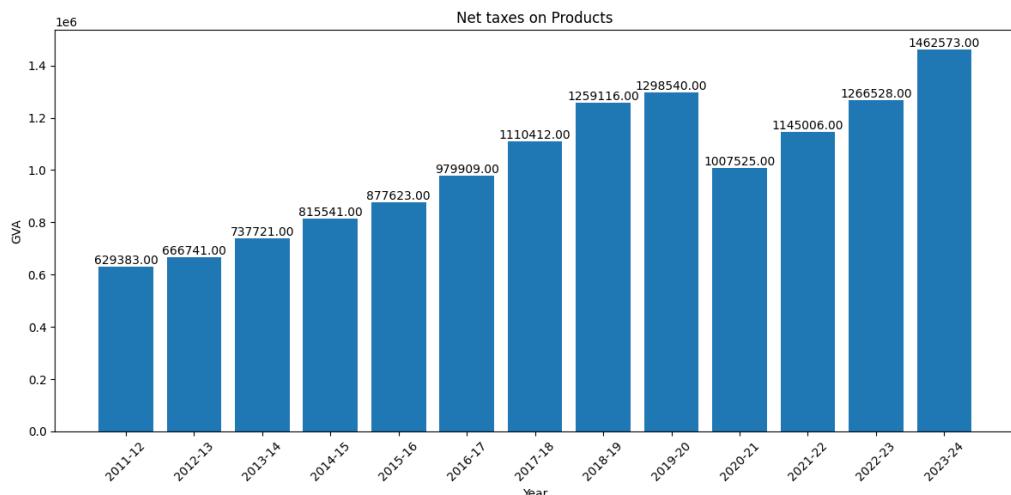


Figure 4.69: Value Distribution of Net taxes on Products)

4.2.1.1.15 GDP

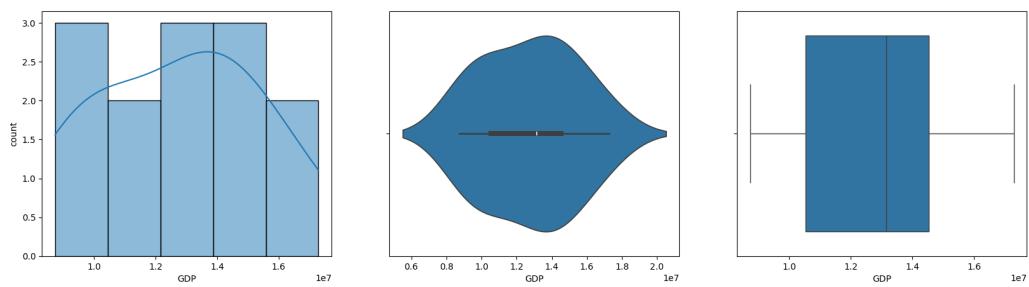


Figure 4.70: Data Distribution of GDP

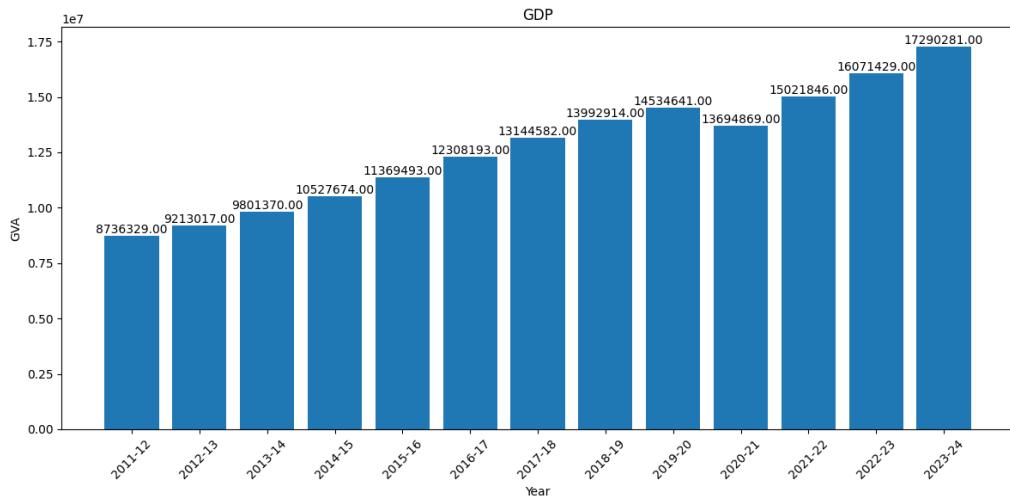


Figure 4.71: Value Distribution of GDP

4.2.1.2 Bivariate Analysis

Here we will analyse the relationship between two sectors

4.2.1.2.1 Relation with GDP

Given below is the relation between GVA, NNI, GNI, and NVA with GDP individually and how they have grown over the years with respect to GDP.

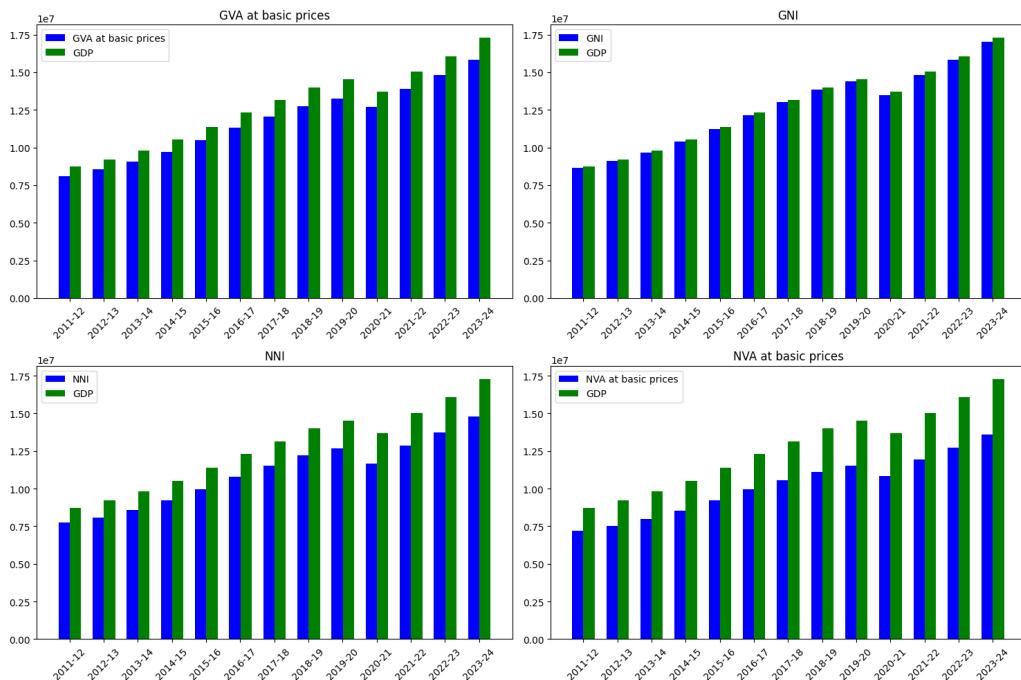


Figure 4.72: Relation of GVA, GNI, NNI and NVA With GDP



4.2.1.3 Multivariate analysis

Here we will analyse multiple sectors together.

4.2.1.3.1 Trend of Contribution of Sectors

Given below is a graph representing the growth of the contribution of each sector towards the national GDP in terms of absolute value.

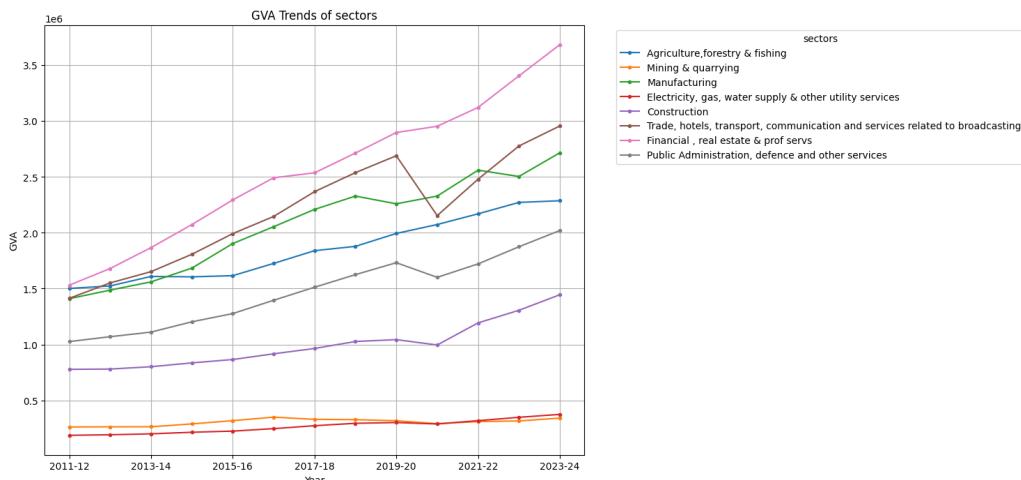


Figure 4.73: Growth of Contribution of Sectors Over The Years in Terms of Absolute Value

4.2.1.3.2 Contribution of Sectors

The plots given below represent the contribution of sectors to the national GDP.

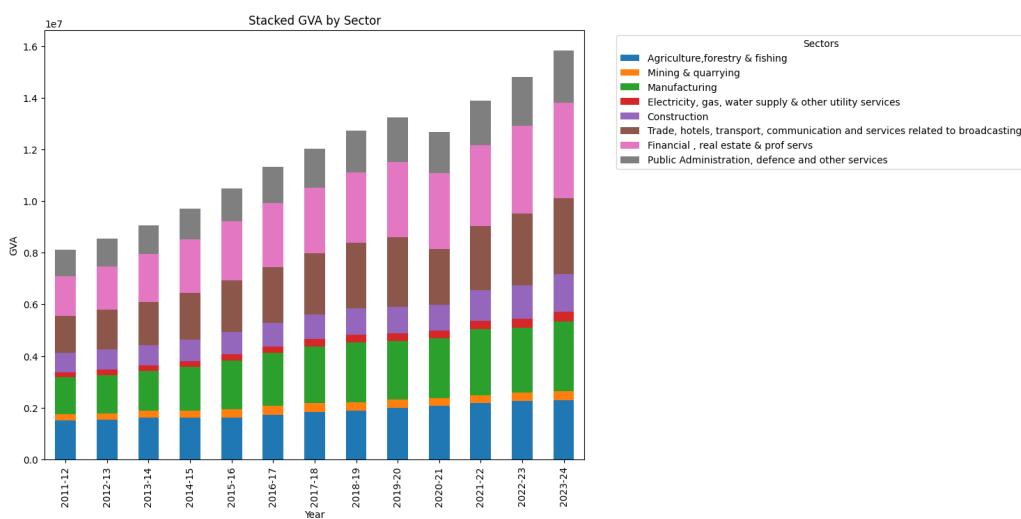


Figure 4.74: Contribution of Sectors Over The Years - Stacked

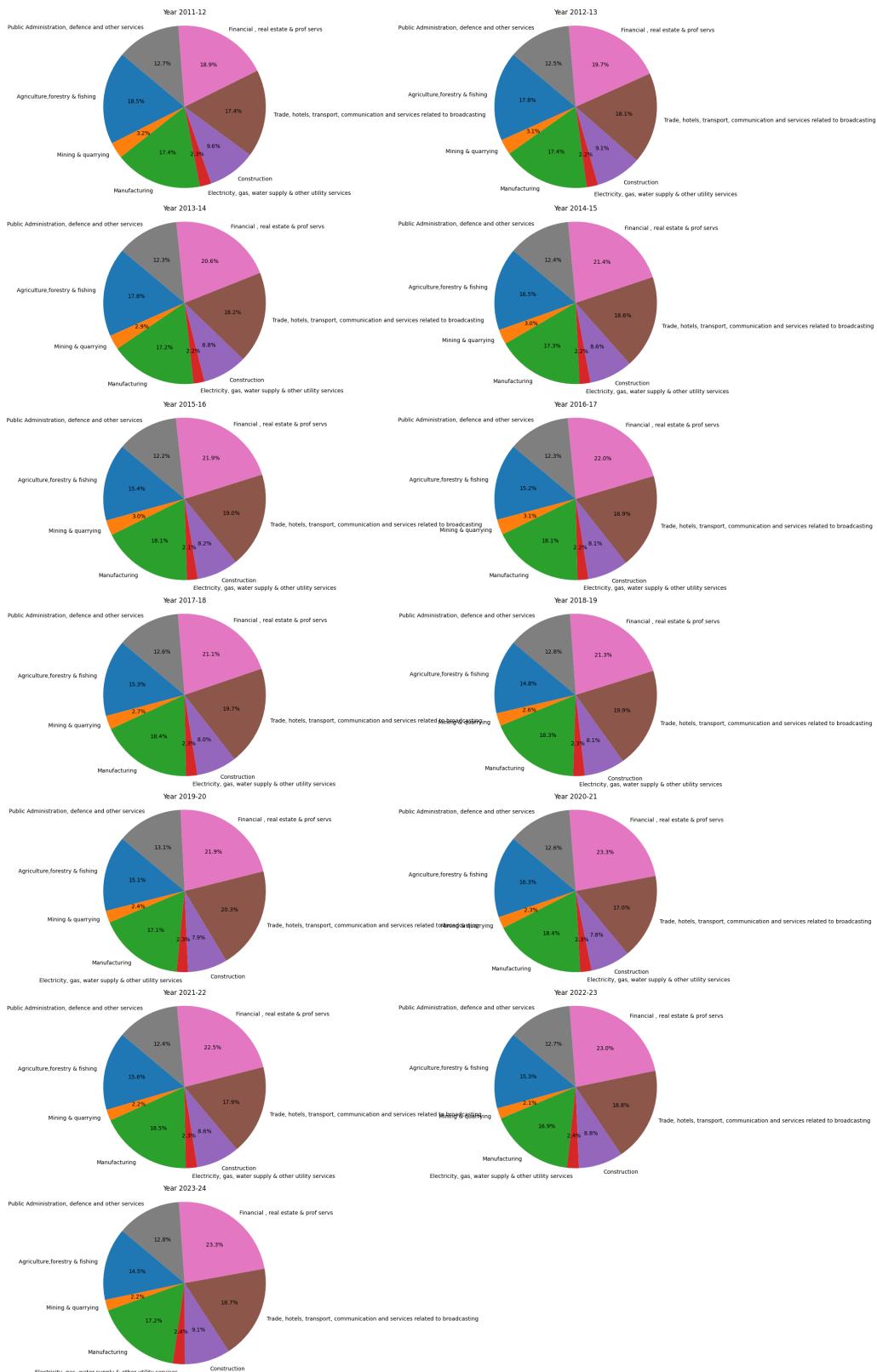


Figure 4.75: Contribution of Sectors Over the Years in Terms of %

Here, we will analyze the contributions of these categories and how they have performed over the

years.

4.2.1.3.3 Trend of Contribution of Categories

The plot here represents the growth of the contribution of categories over the years in terms of absolute value.

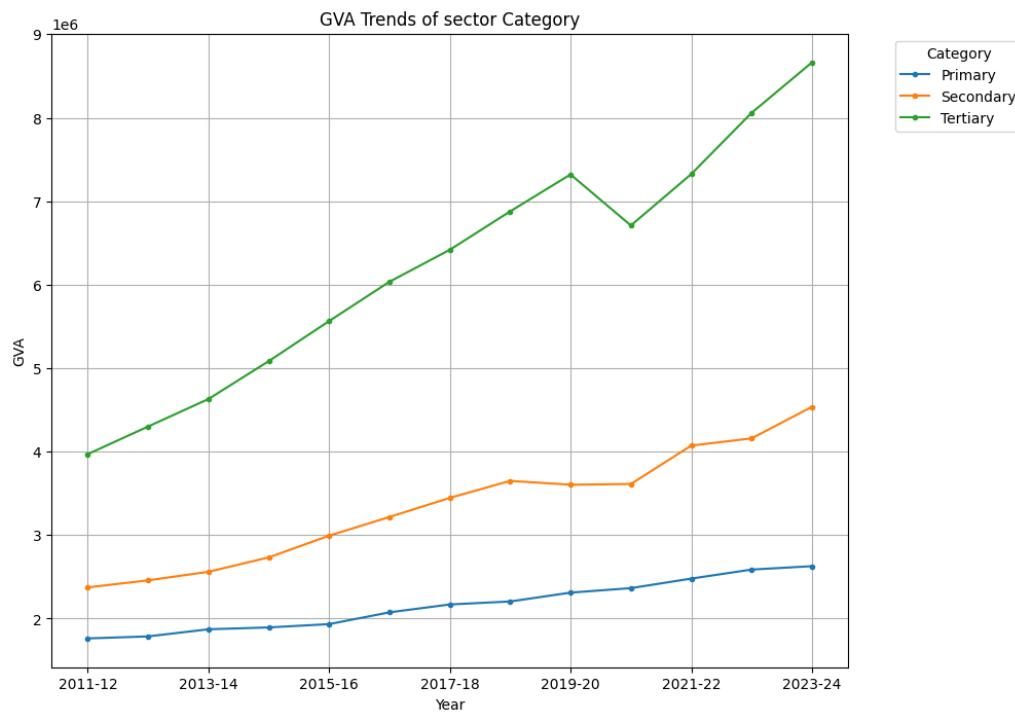


Figure 4.76: Growth of Contribution of Categories Over The Years in Terms of Absolute Value

4.2.1.3.4 Contribution of Categories

The plots below represent the categories' contribution to the national GDP.

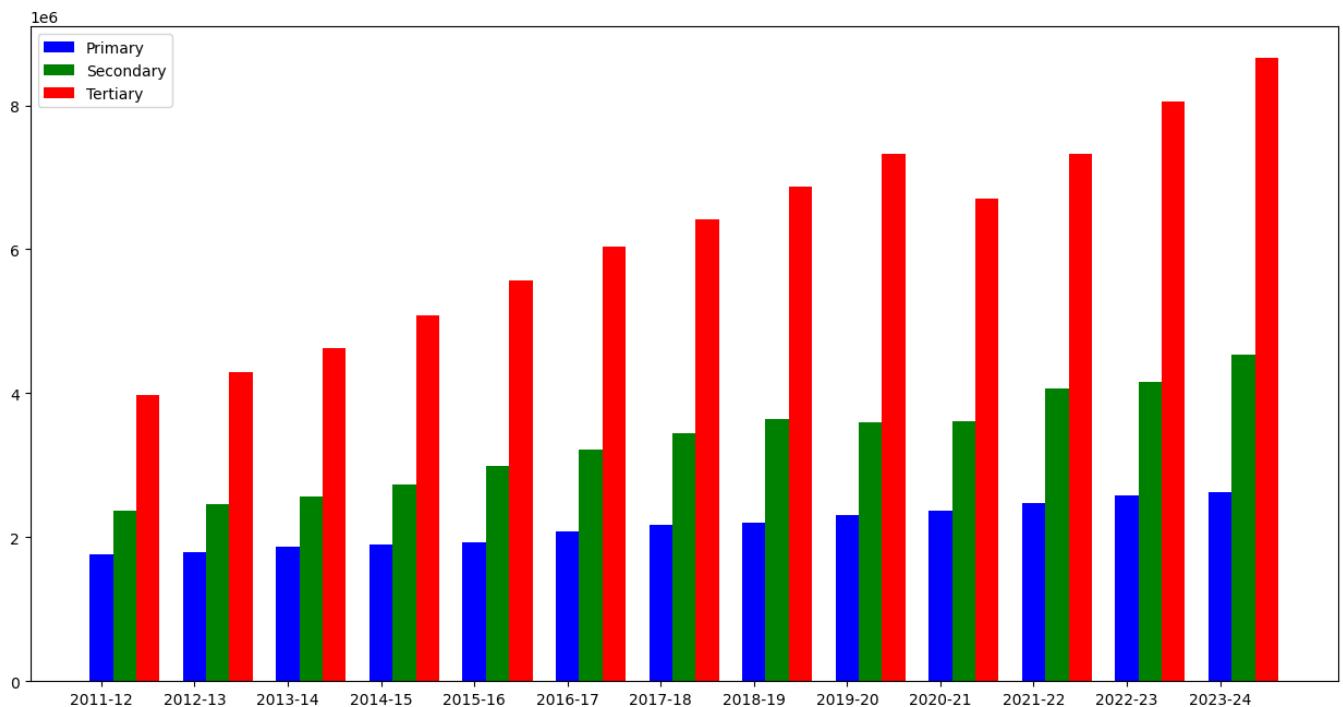


Figure 4.77: Contribution of Categories Over the Years - Side By Side

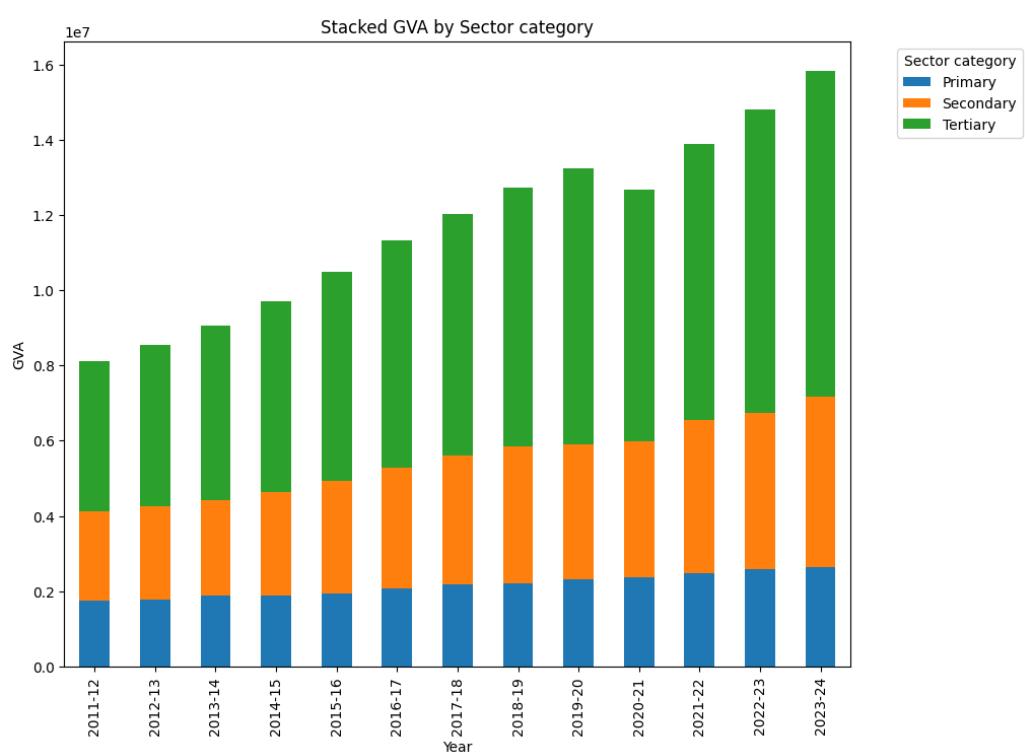


Figure 4.78: Contribution of Categories Over the Years- Stacked

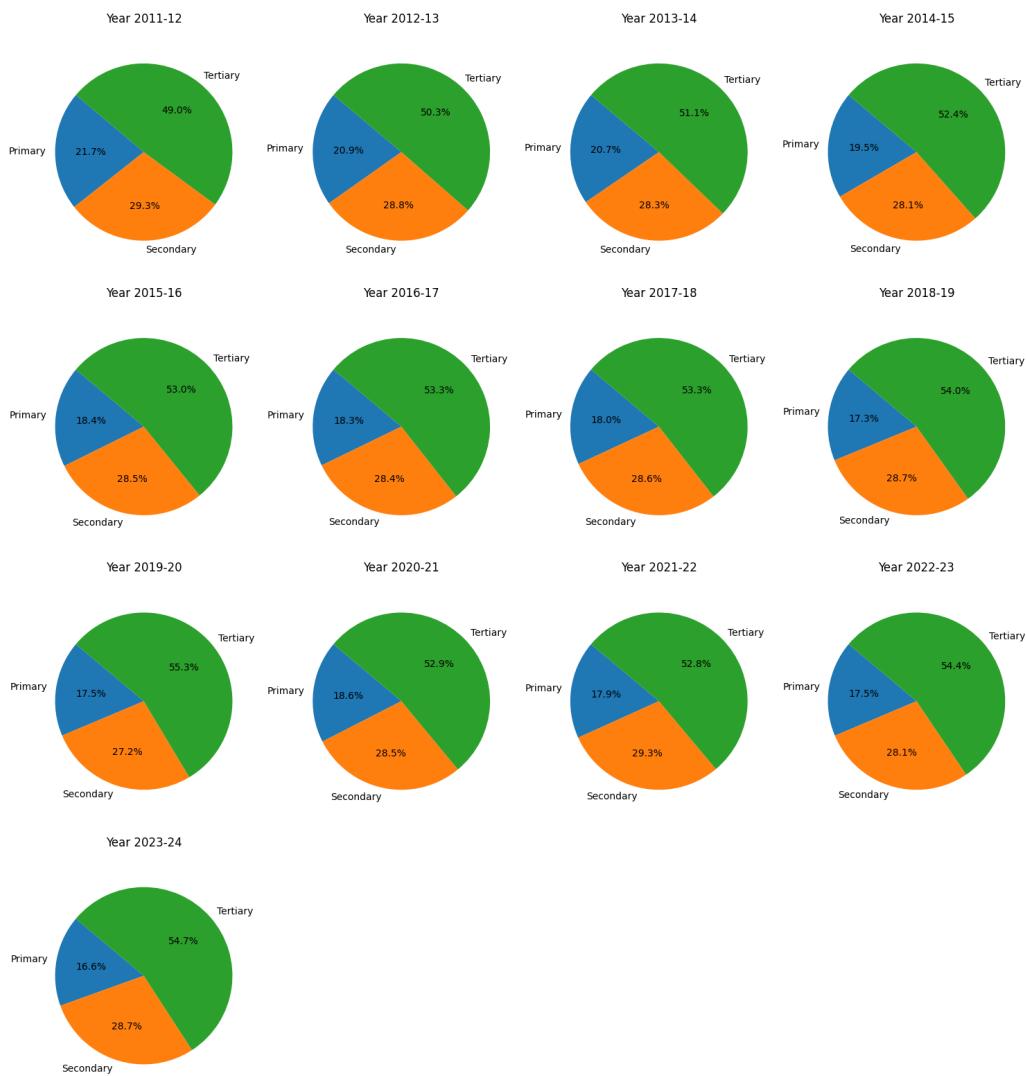


Figure 4.79: Contribution of Categories Over the Years in Terms of %



4.2.2 Growth Rate of GVA

4.2.2.1 Univariate Analysis

Now we will analyze the growth rate of all the sectors individually.

4.2.2.1.1 Agriculture, forestry & fishing

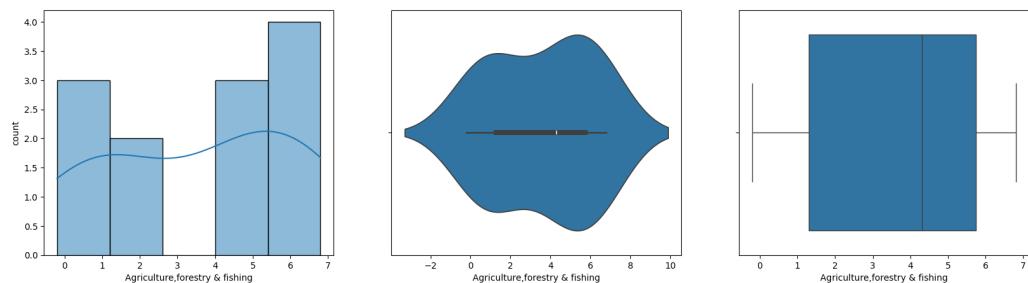


Figure 4.80: Data Distribution of Growth Rate of Agriculture, forestry & fishing

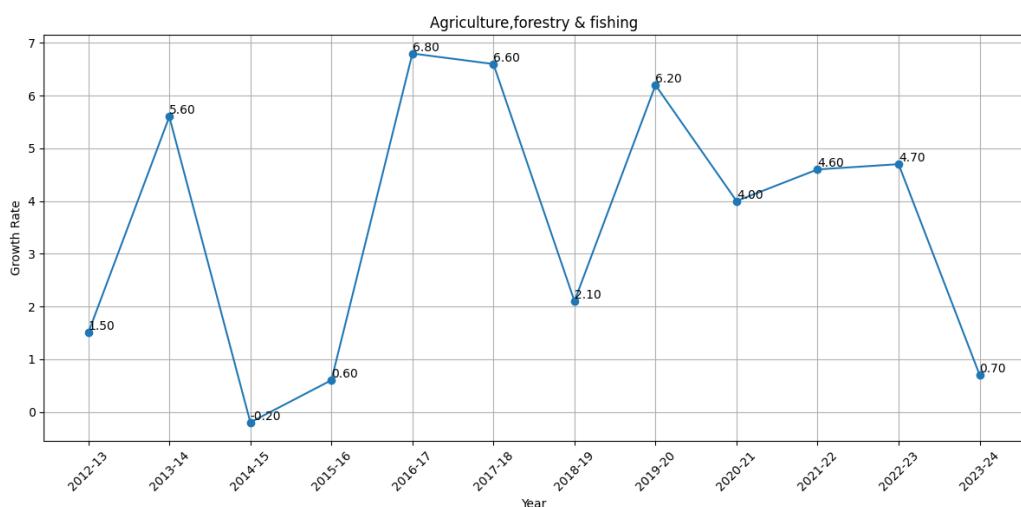


Figure 4.81: Trend of Growth of Agriculture,forestry & fishing

4.2.2.1.2 Mining & quarrying

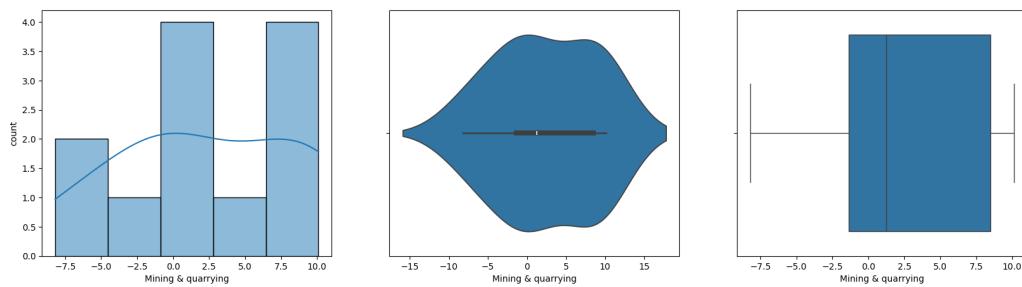


Figure 4.82: Data Distribution of Growth Rate of Mining & quarrying

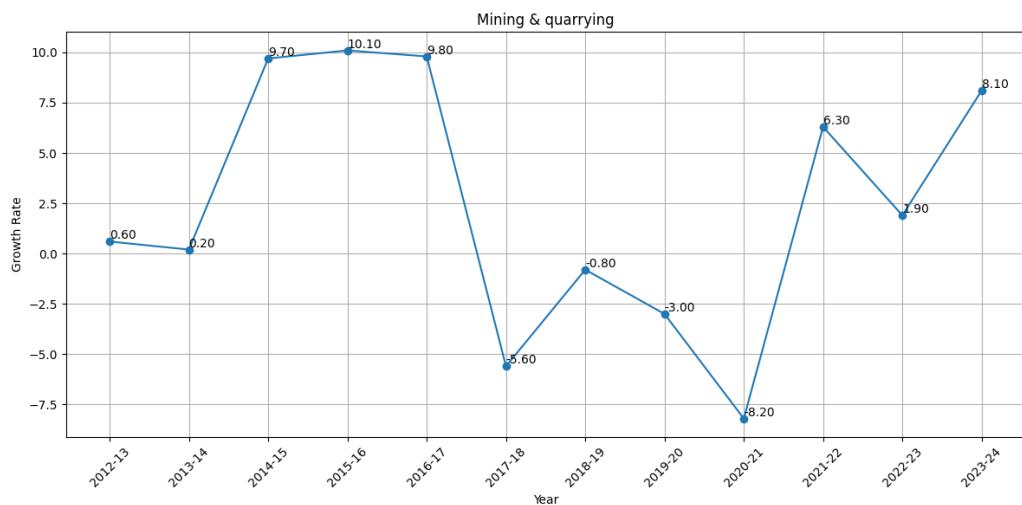


Figure 4.83: Trend of Growth of Mining & quarrying

4.2.2.1.3 Manufacturing

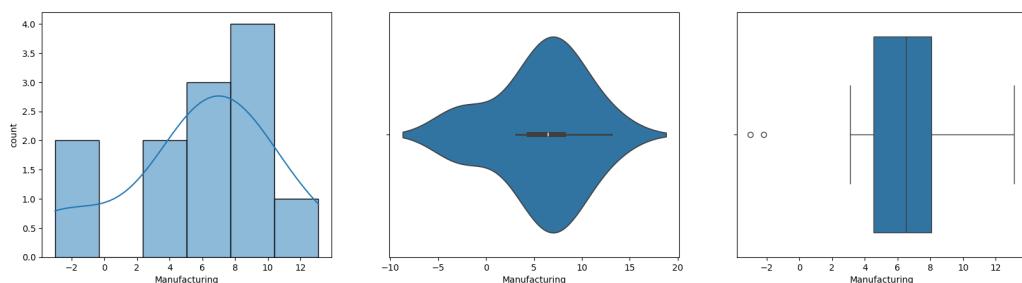


Figure 4.84: Data Distribution of Growth Rate of Manufacturing

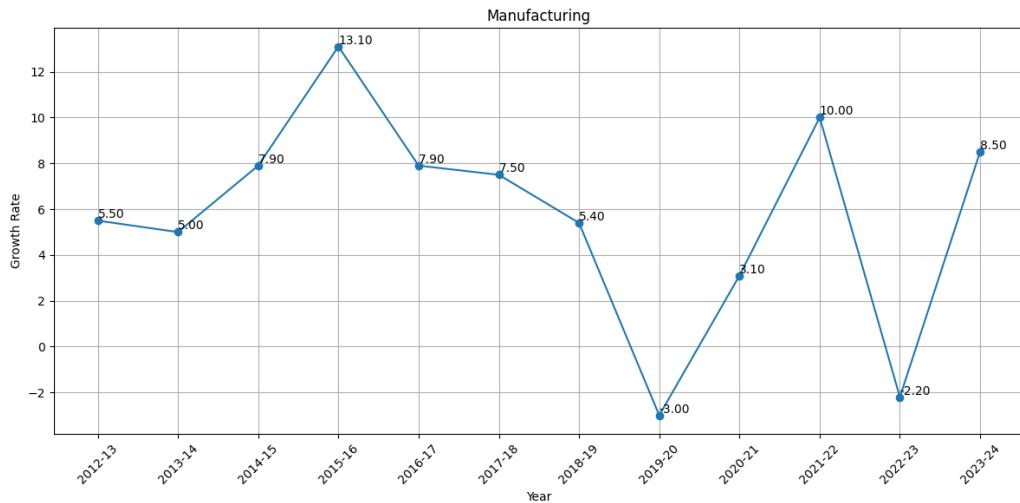


Figure 4.85: Trend of Growth of Manufacturing

4.2.2.1.4 Electricity, gas, water supply & other utility services

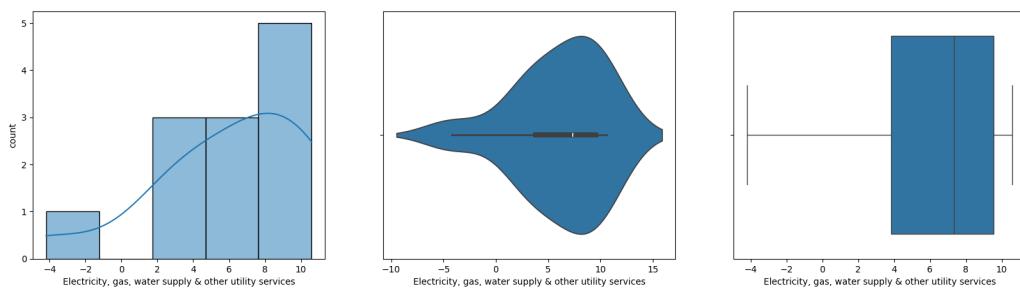


Figure 4.86: Data Distribution of Growth Rate of Electricity, gas, water supply & other utility services

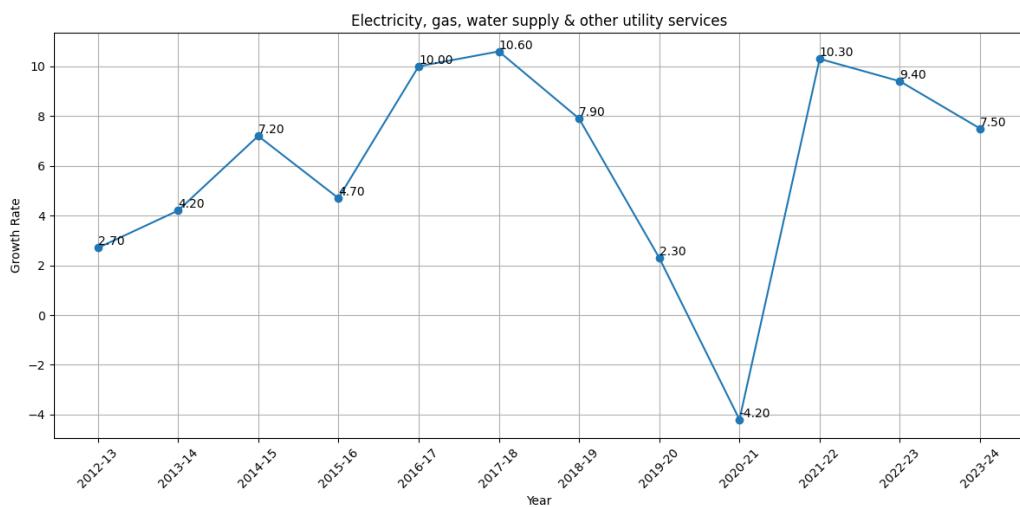


Figure 4.87: Trend of Growth of Electricity, gas, water supply & other utility services

4.2.2.1.5 Construction

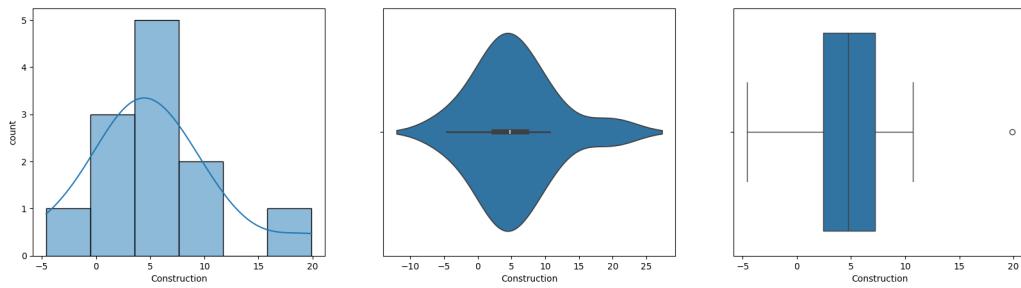


Figure 4.88: Data Distribution of Growth Rate of Construction

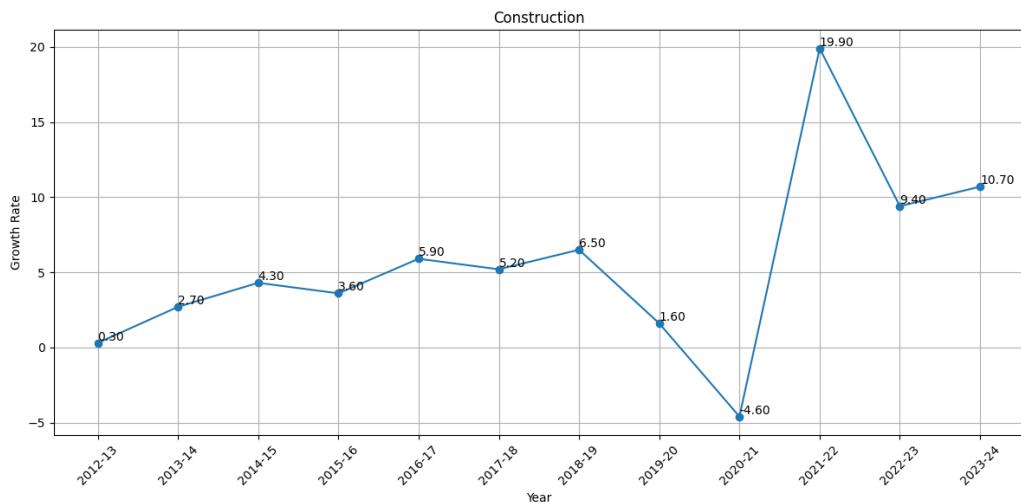


Figure 4.89: Trend of Growth of Construction

4.2.2.1.6 Trade, hotels, transport, communication and services related to broadcasting

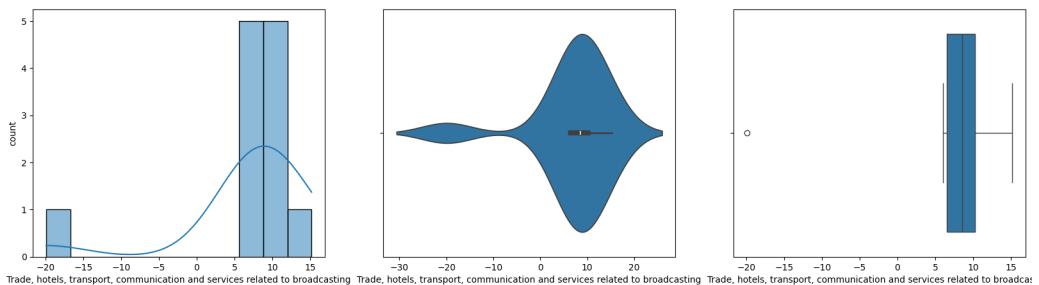


Figure 4.90: Data Distribution of Growth Rate of Trade, hotels, transport, communication and services related to broadcasting

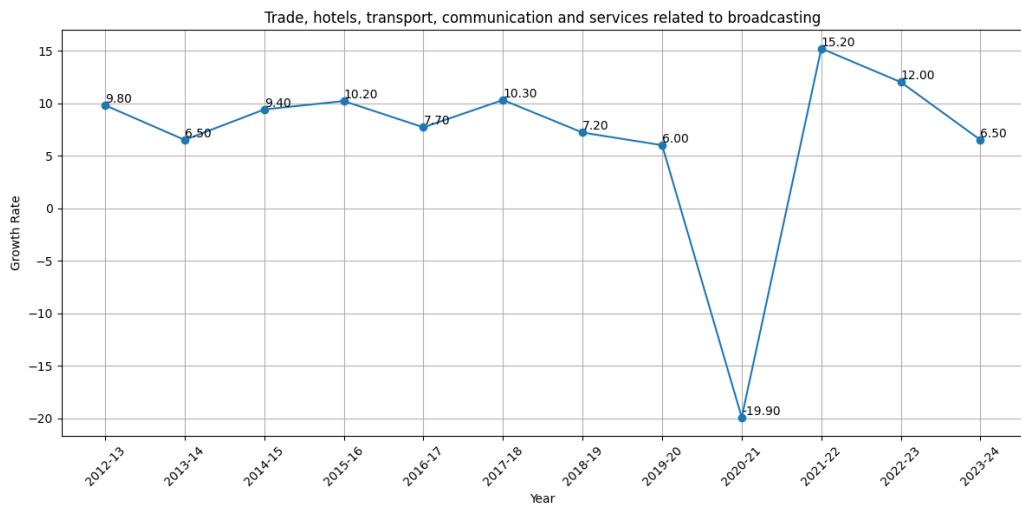


Figure 4.91: Trend of Growth of Trade, hotels, transport, communication and services related to broadcasting

4.2.2.1.7 Financial , real estate & prof servs

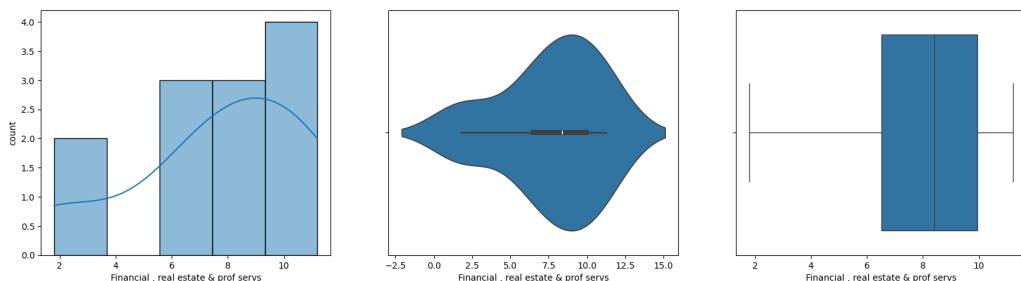


Figure 4.92: Data Distribution of Growth Rate of Financial , real estate & prof servs

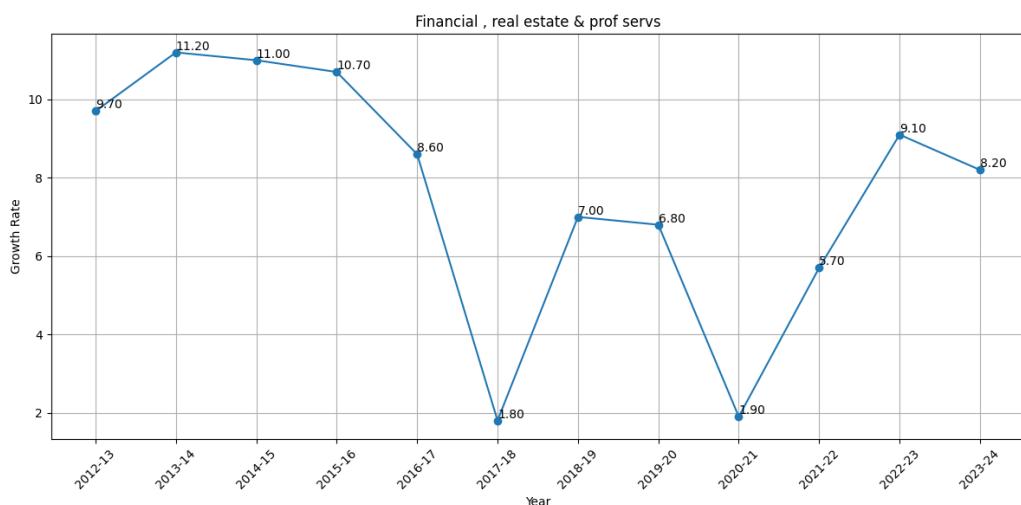


Figure 4.93: Trend of Growth of Financial , real estate & prof servs



4.2.2.1.8 Public Administration, defence and other services

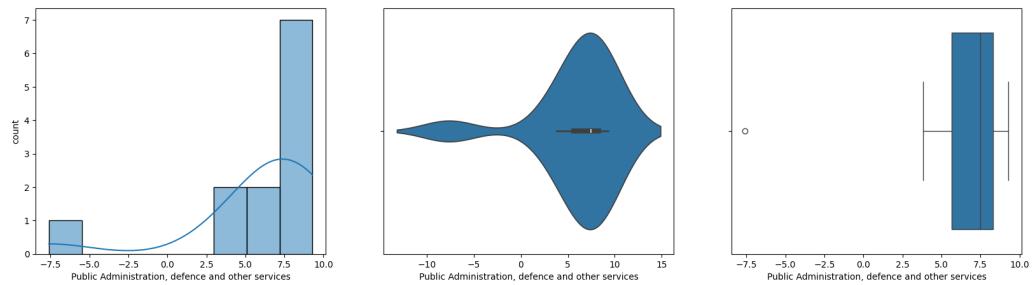


Figure 4.94: Data Distribution of Growth Rate of Public Administration, defence and other services

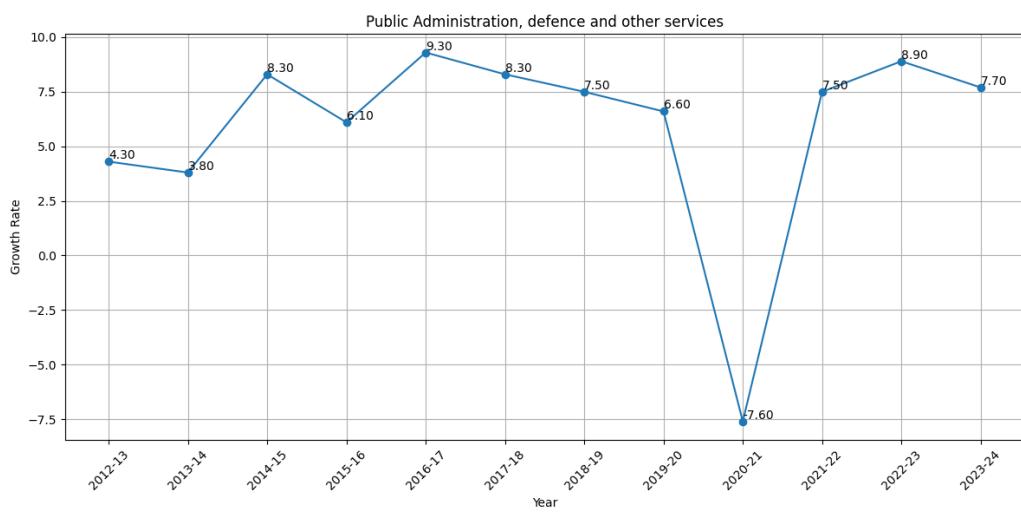


Figure 4.95: Trend of Growth of Public Administration, defence and other services

4.2.2.1.9 GVA at basic prices

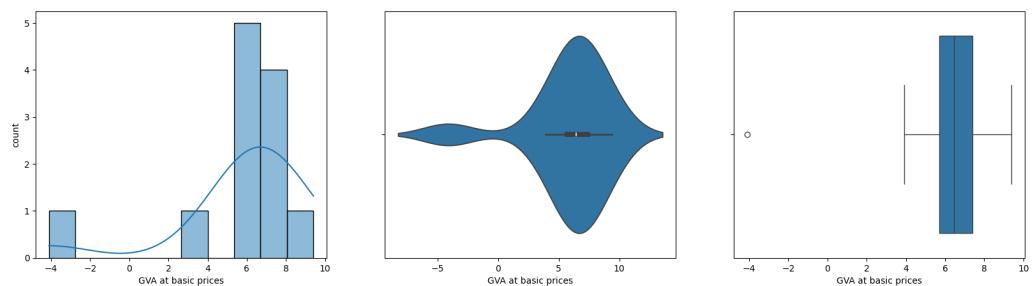


Figure 4.96: Data Distribution of Grwoth Rate of GVA at basic prices



Figure 4.97: Trend of Growth of GVA at basic prices

4.2.2.1.10 Per capita income(Rs.)

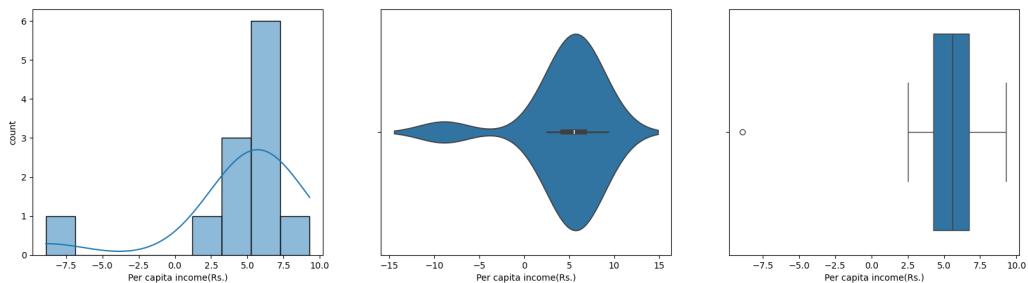


Figure 4.98: Data Distribution of Growth Rate of Per capita income(Rs.)

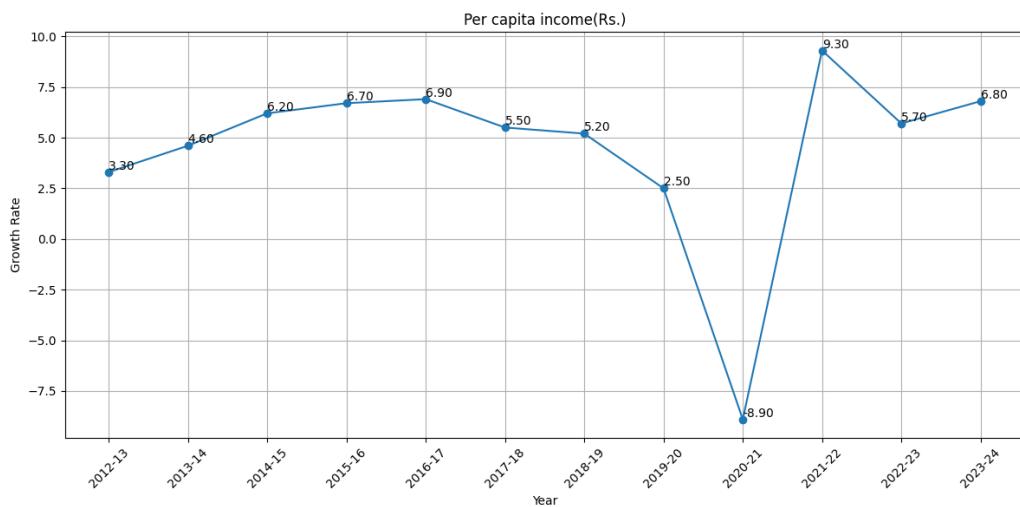


Figure 4.99: Trend of Growth of Per capita income(Rs.)

4.2.2.1.11 Net taxes on Products

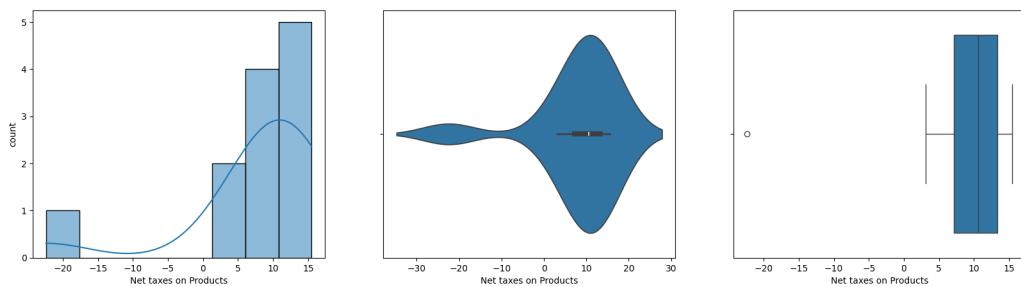


Figure 4.100: Data Distribution of Growth Rate of Net taxes on Products

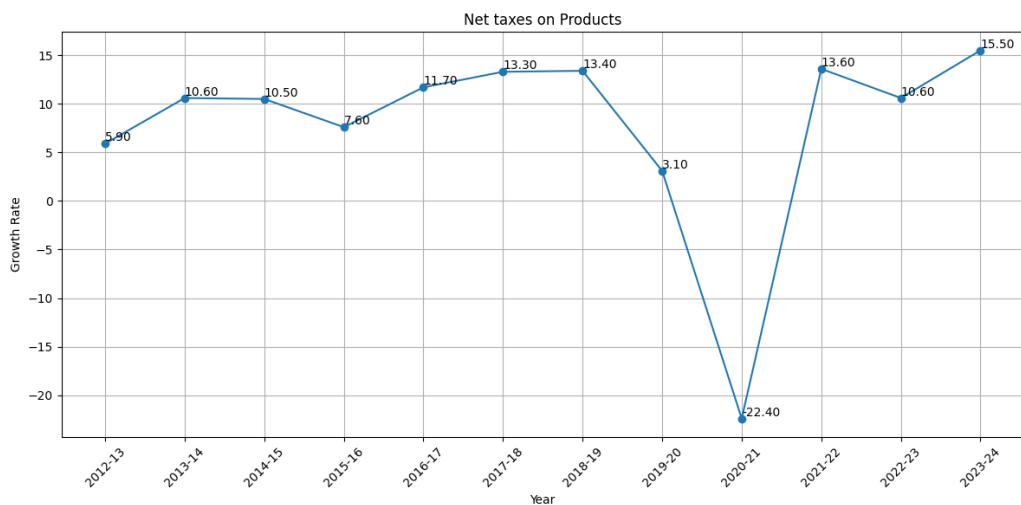


Figure 4.101: Trend of Growth of Net taxes on Products)

4.2.2.1.12 GDP

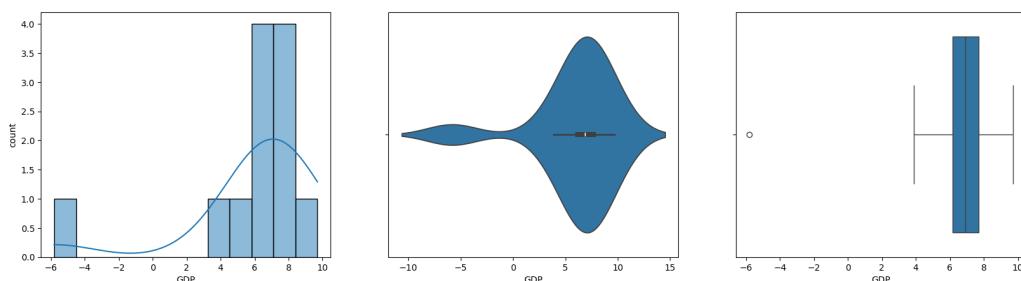


Figure 4.102: Data Distribution of Growth Rate of GDP

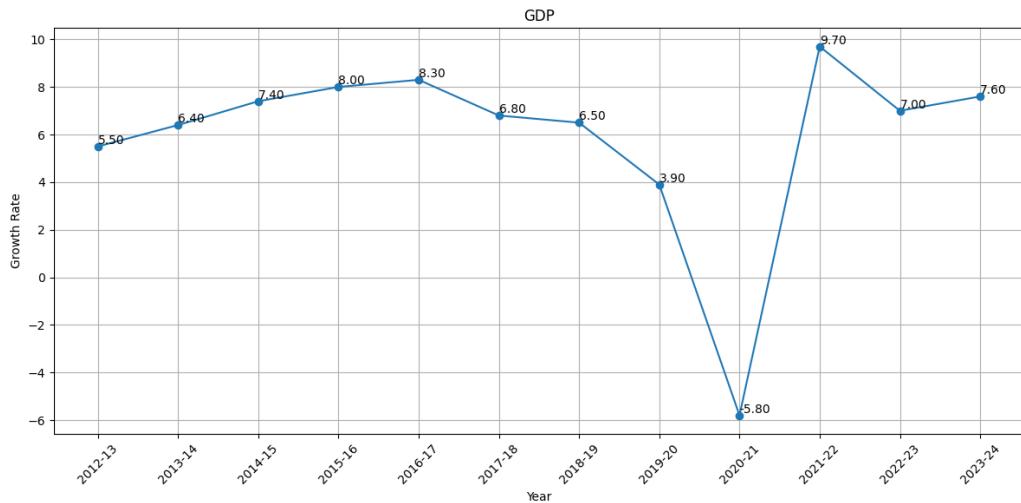


Figure 4.103: Trend of Growth of GDP

4.2.2.2 Bivariate Analysis

Here, we will analyze the relation between the growth rate of the two sectors.

4.2.2.2.1 Comparison with GDP

The comparison of the growth rate of each sector with respect to GDP growth is given below.

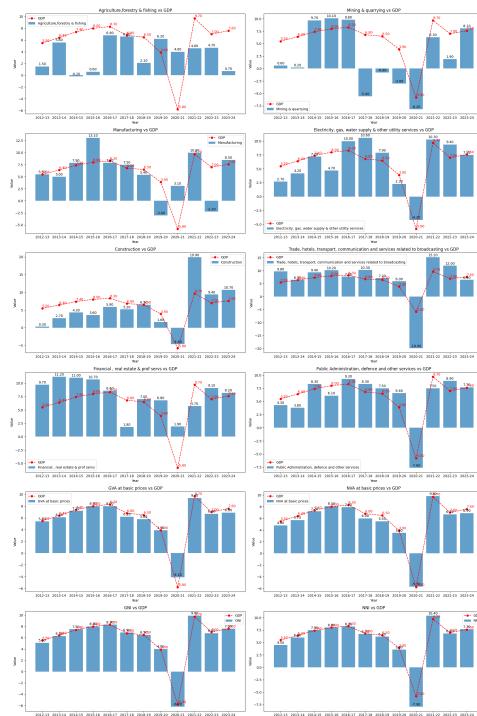


Figure 4.104: Comparison of Growth Rate of Various Sectors With GDP's Growth Rate



4.2.2.3 Multivariate analysis

Here, we will analyze the growth rate of multiple sectors together.

4.2.2.3.1 Trend of Growth of Sectors

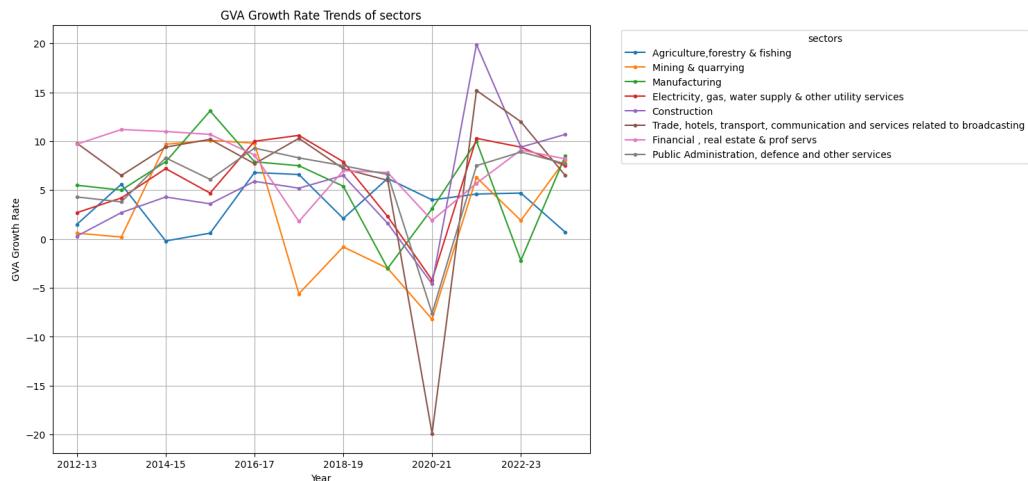


Figure 4.105: Growth of Sectors Over The Years

4.2.2.3.2 Contribution of Sectors in terms of Growth Rate

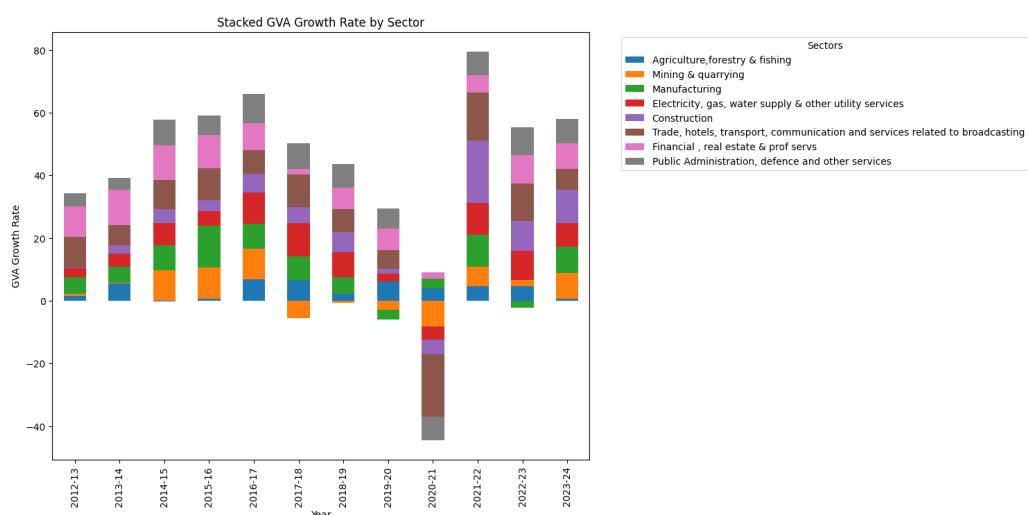


Figure 4.106: Contribution of Sectors Over The Years in Terms of Growth Rate

4.2.2.3.3 Correlation of Growth of Sectors

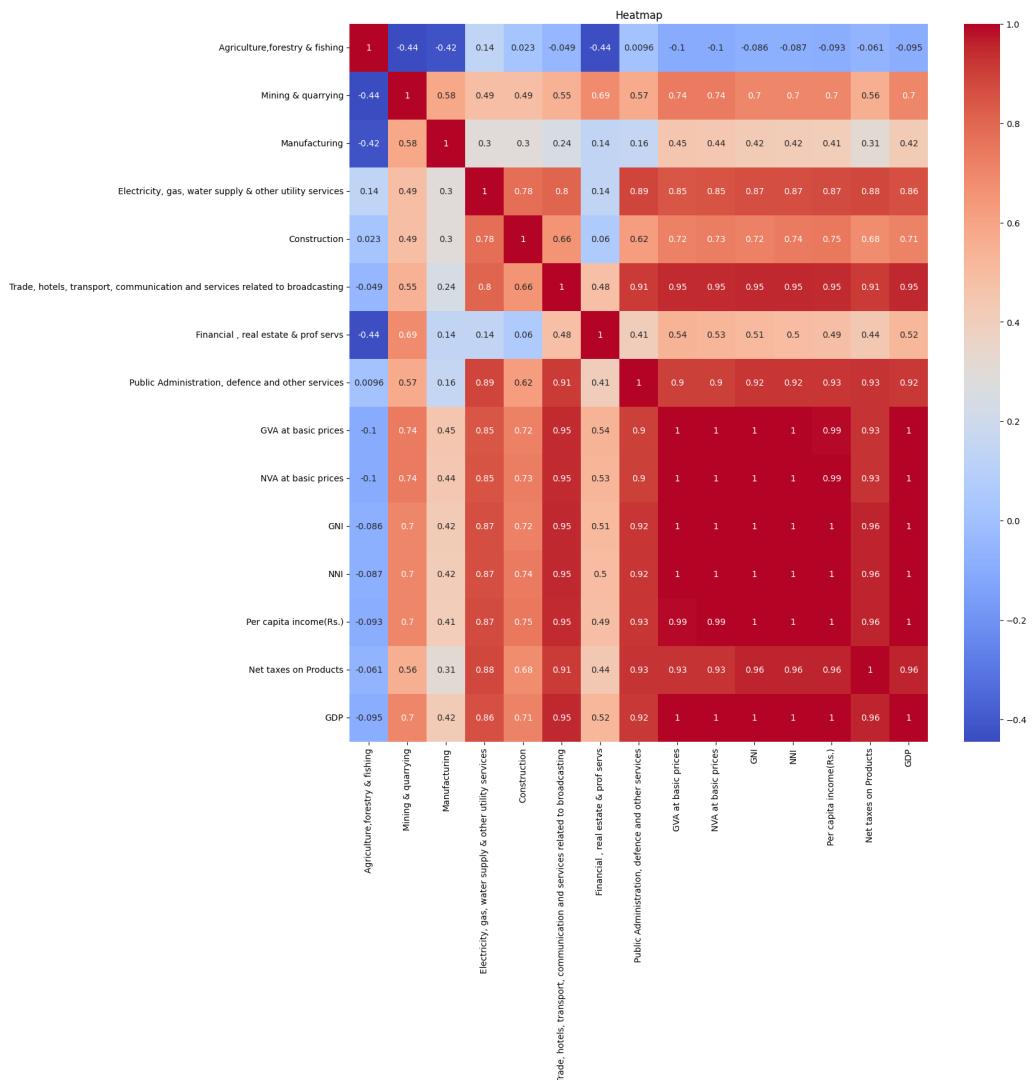


Figure 4.107: Correlation of Growth of Sectors

4.3 Contributions of States

4.3.1 GSDP

4.3.1.1 Univariate Analysis

4.3.1.1.1 Andhra Pradesh

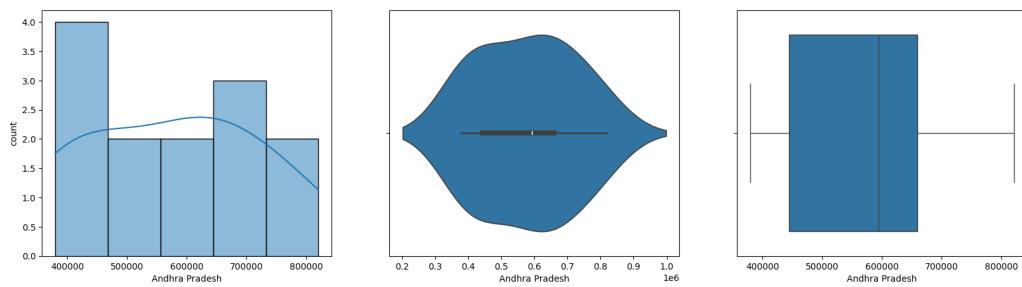


Figure 4.108: Data Distribution of GSDP of Andhra Pradesh

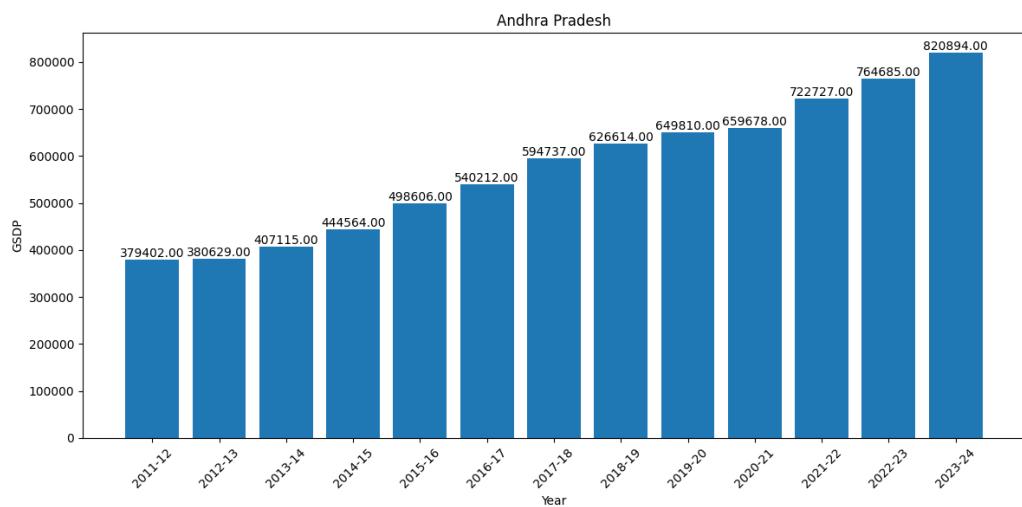


Figure 4.109: Value Distribution of GSDP of Andhra Pradesh

4.3.1.1.2 Arunachal Pradesh

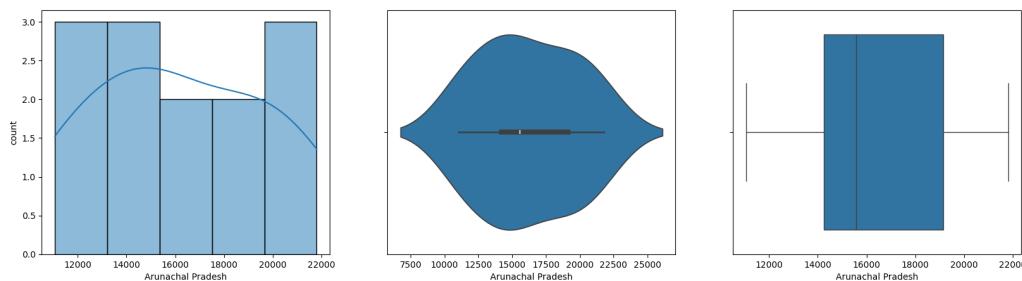


Figure 4.110: Data Distribution of GSDP of Arunachal Pradesh

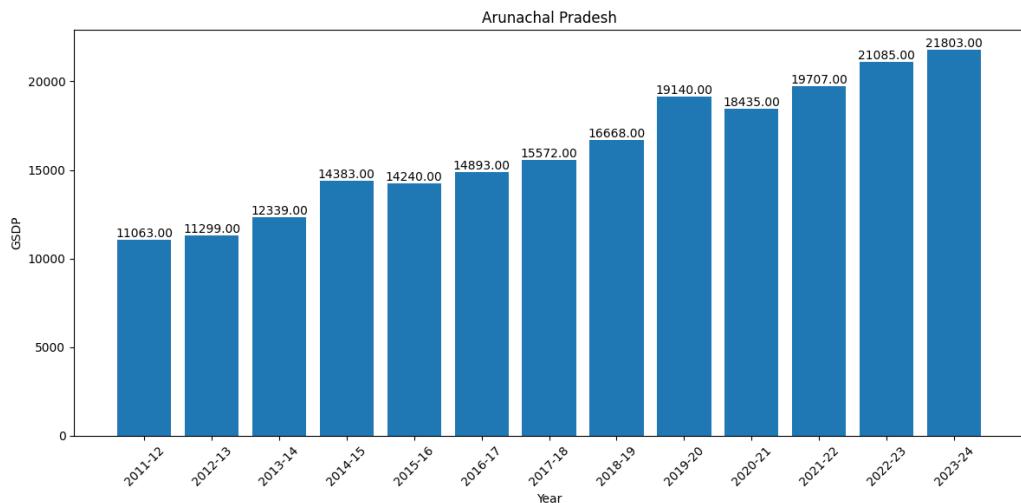


Figure 4.111: Value Distribution of GSDP of Arunachal Pradesh

4.3.1.1.3 Assam

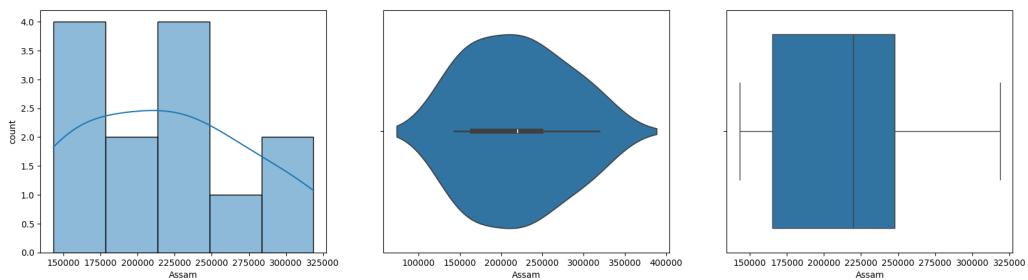


Figure 4.112: Data Distribution of GSDP of Assam

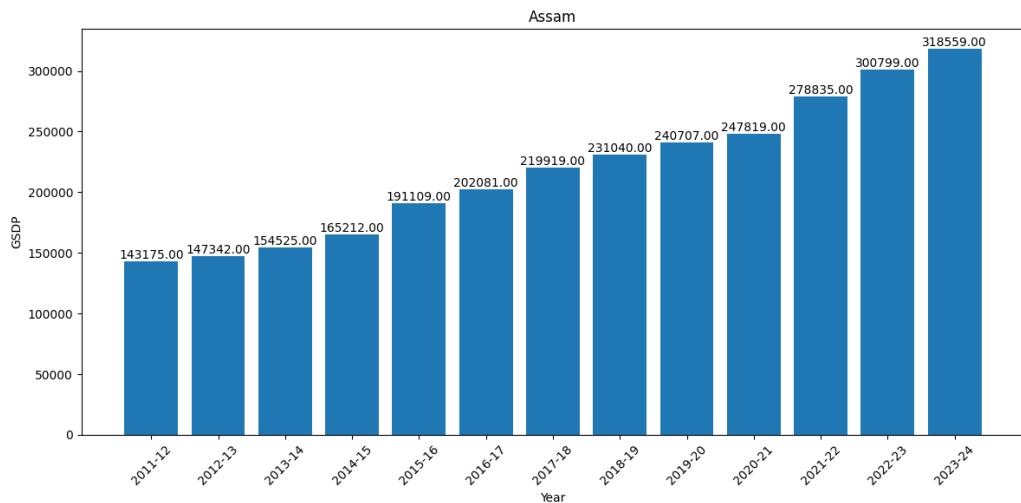


Figure 4.113: Value Distribution of GSDP of Assam

4.3.1.1.4 Bihar

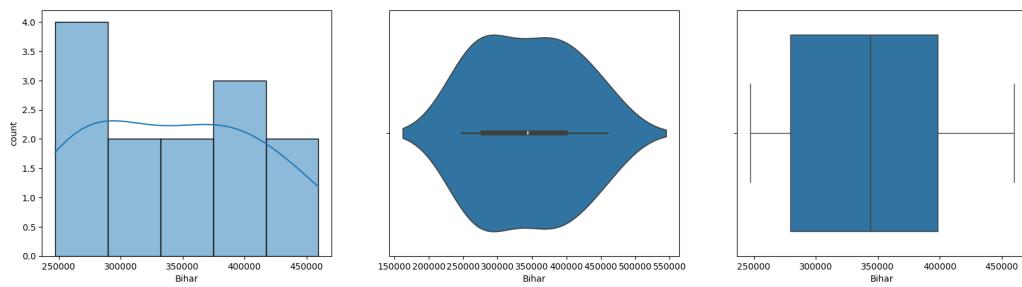


Figure 4.114: Data Distribution of GSDP of Bihar

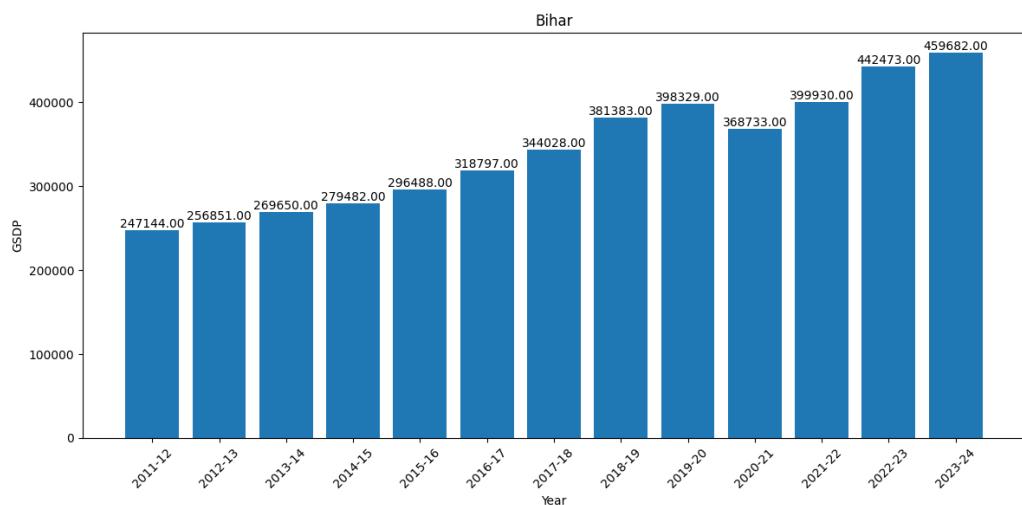


Figure 4.115: Value Distribution of GSDP of Bihar

4.3.1.1.5 Chhattisgarh

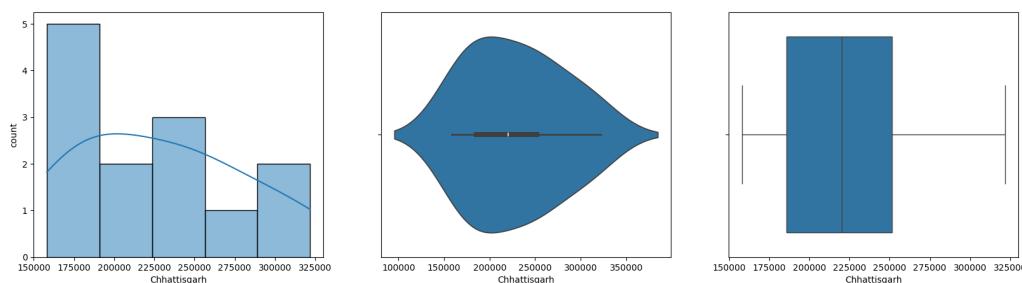


Figure 4.116: Data Distribution of GSDP of Chhattisgarh

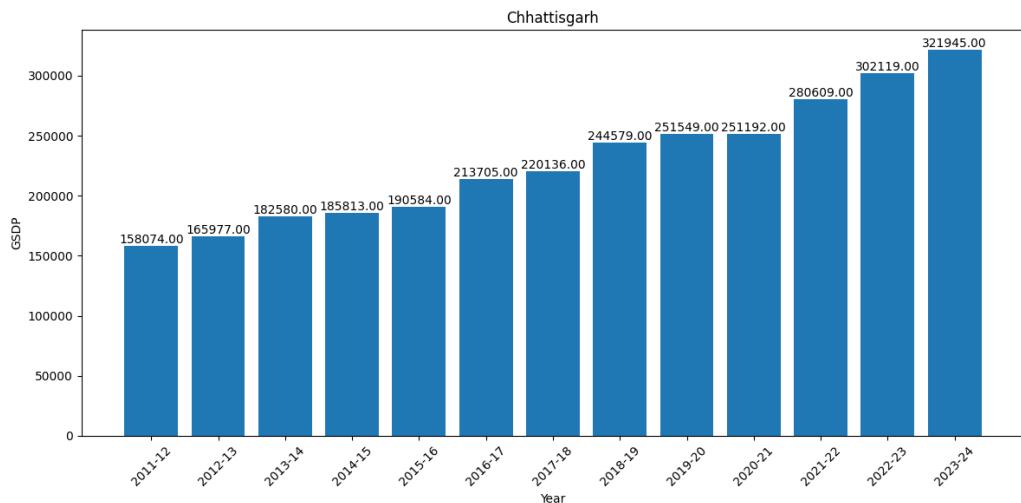


Figure 4.117: Value Distribution of GSDP of Chhattisgarh

4.3.1.1.6 Goa

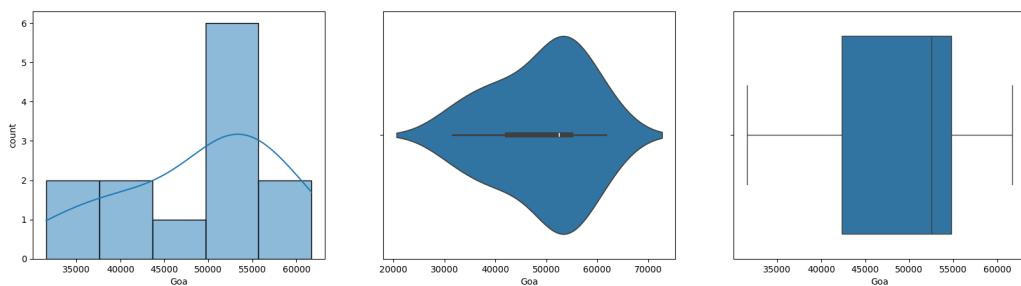


Figure 4.118: Data Distribution of GSDP of Goa

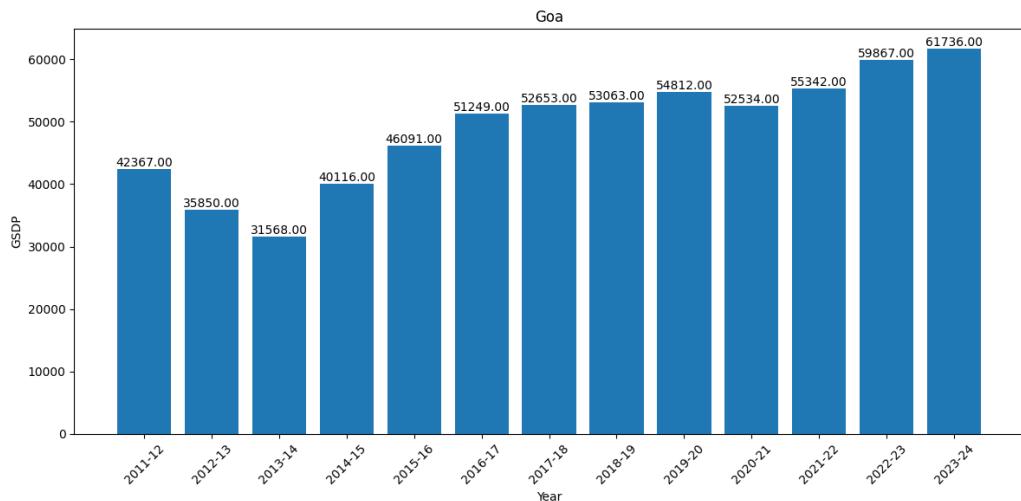


Figure 4.119: Value Distribution of GSDP of Goa

4.3.1.1.7 Gujarat

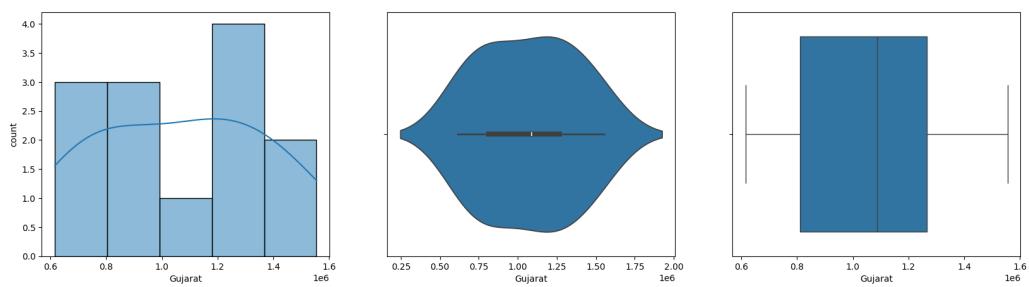


Figure 4.120: Data Distribution of GSDP of Gujarat

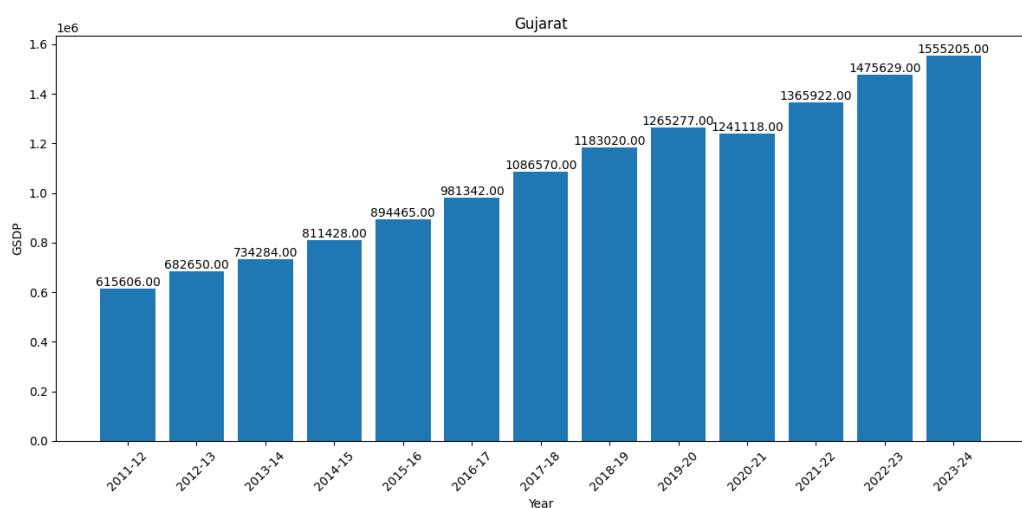


Figure 4.121: Value Distribution of GSDP of Gujarat

4.3.1.1.8 Haryana

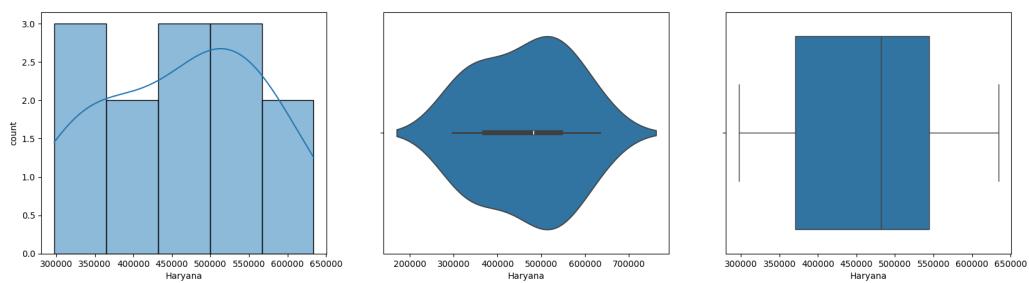


Figure 4.122: Data Distribution of GSDP of Haryana

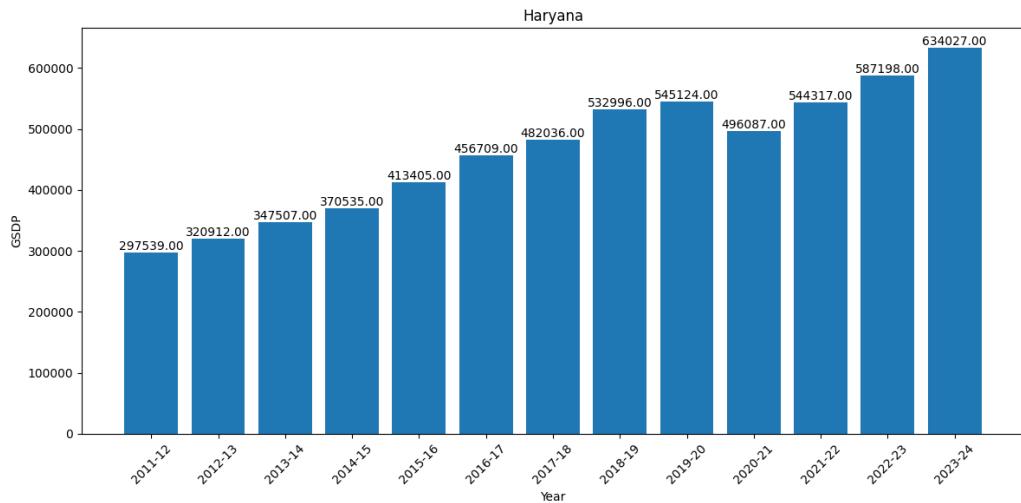


Figure 4.123: Value Distribution of GSDP of Haryana

4.3.1.1.9 Himachal Pradesh

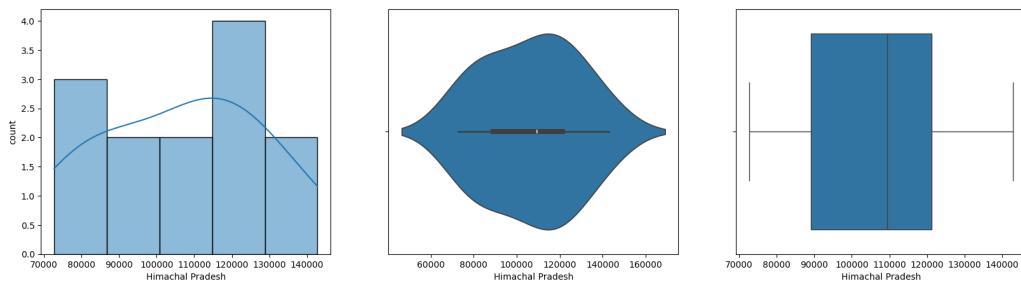


Figure 4.124: Data Distribution of GSDP of Himachal Pradesh

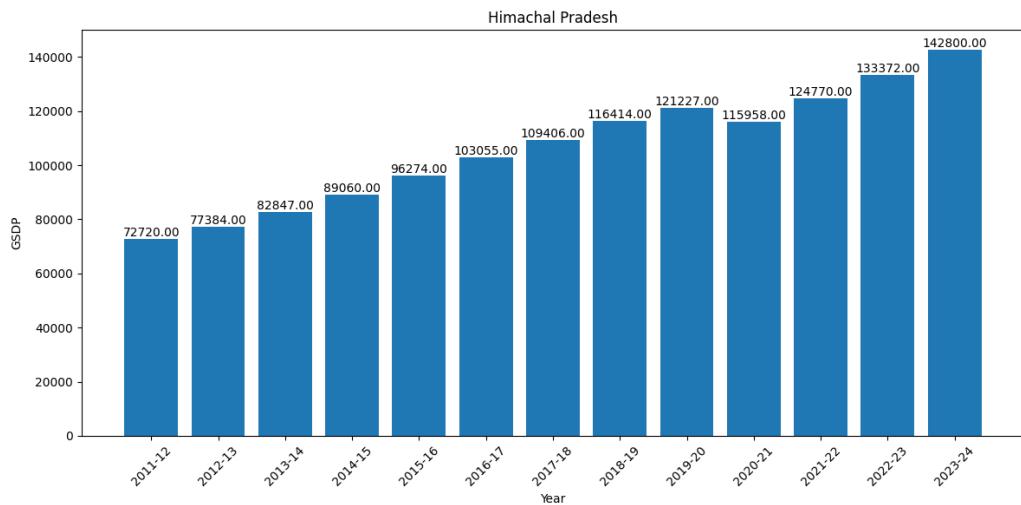


Figure 4.125: Value Distribution of GSDP of Himachal Pradesh

4.3.1.1.10 Jharkhand

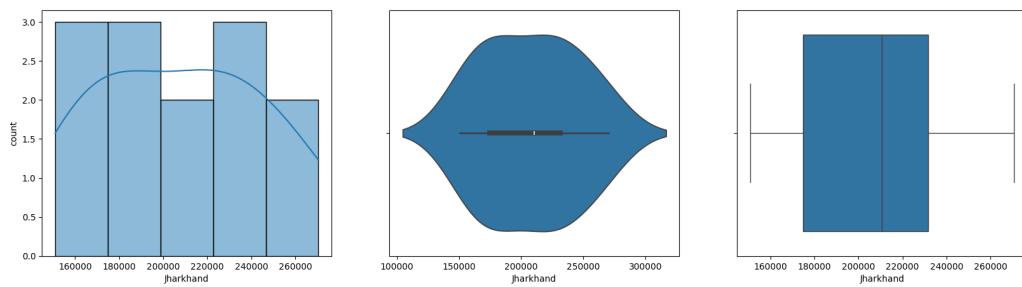


Figure 4.126: Data Distribution of GSDP of Jharkhand

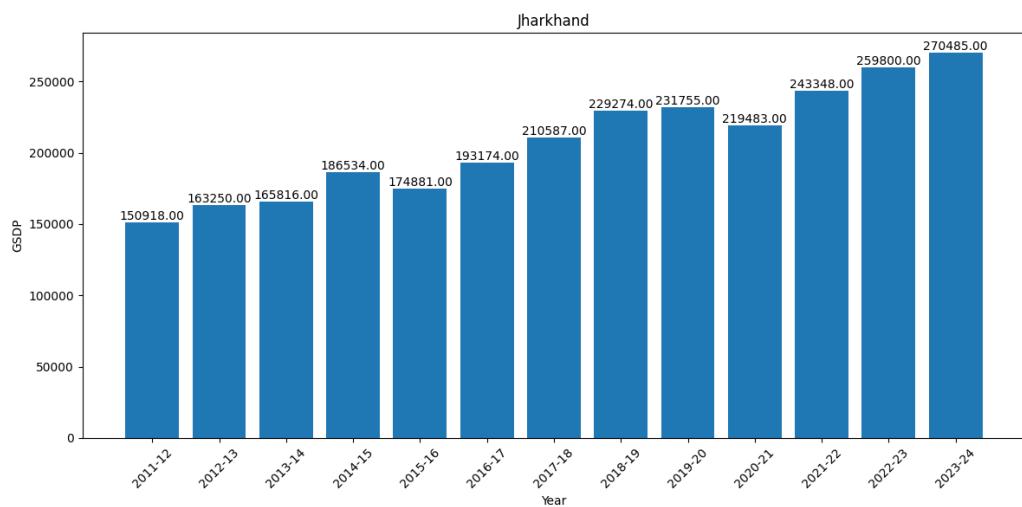


Figure 4.127: Value Distribution of GSDP of Jharkhand

4.3.1.11 Karnataka

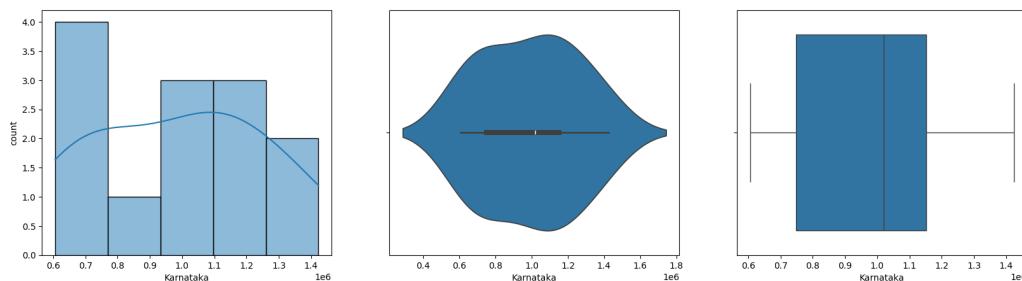


Figure 4.128: Data Distribution of GSDP of Karnataka

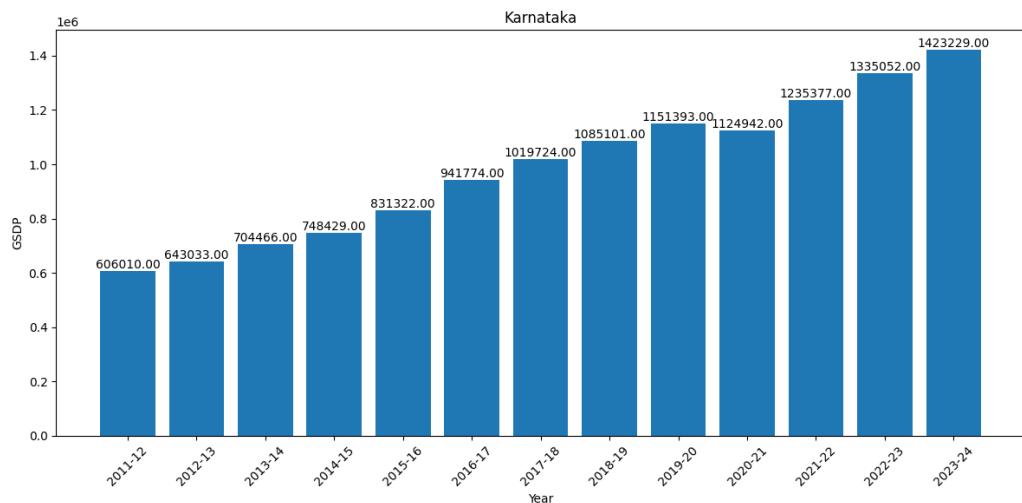


Figure 4.129: Value Distribution of GSDP of Karnataka

4.3.1.1.12 Kerala

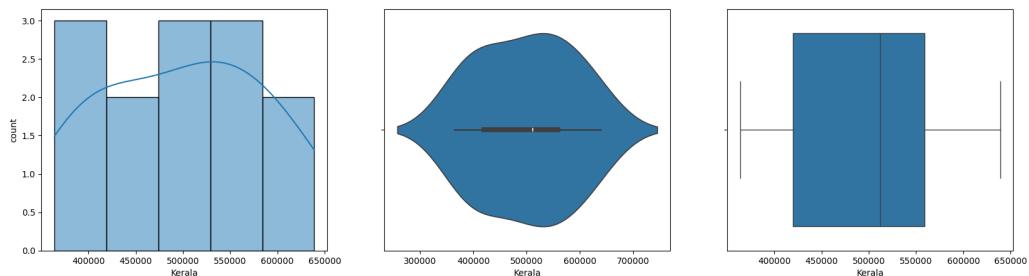


Figure 4.130: Data Distribution of GSDP of Kerala

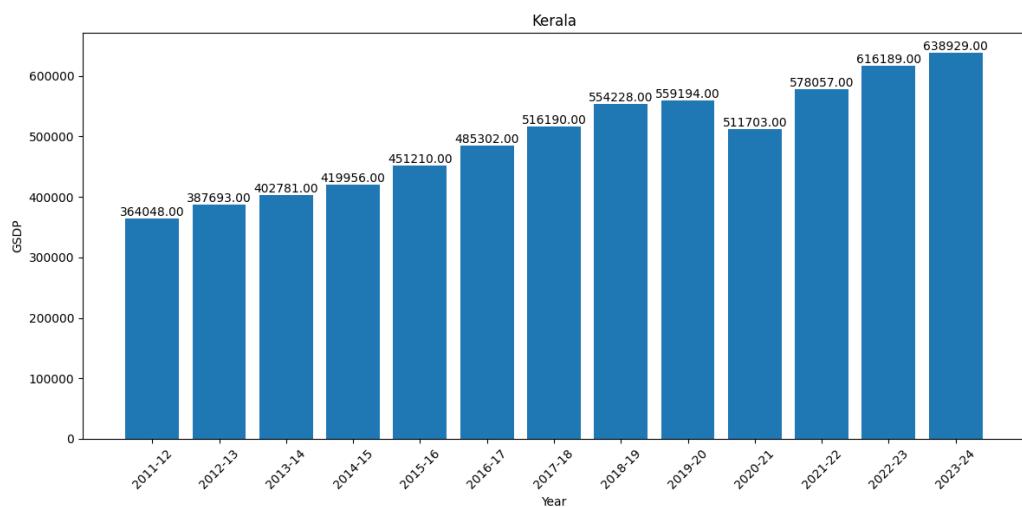


Figure 4.131: Value Distribution of GSDP of Kerala

4.3.1.1.13 Madhya Pradesh

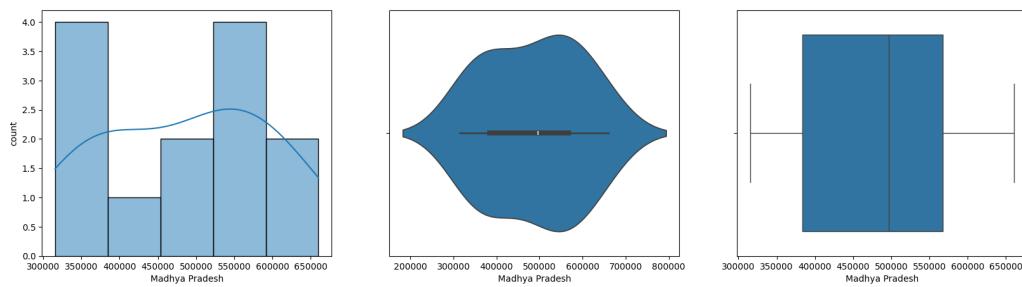


Figure 4.132: Data Distribution of GSDP of Madhya Pradesh

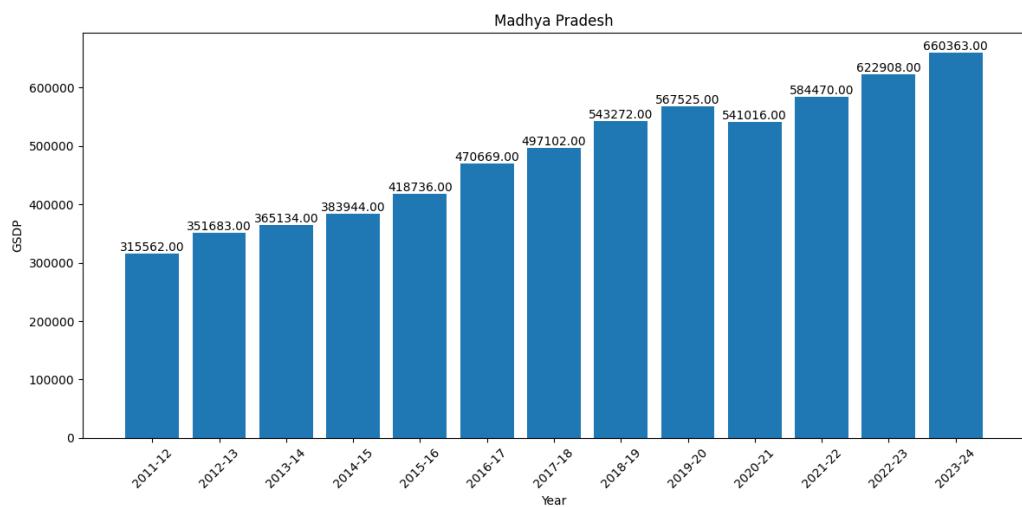


Figure 4.133: Value Distribution of GSDP of Madhya Pradesh

4.3.1.1.14 Maharashtra

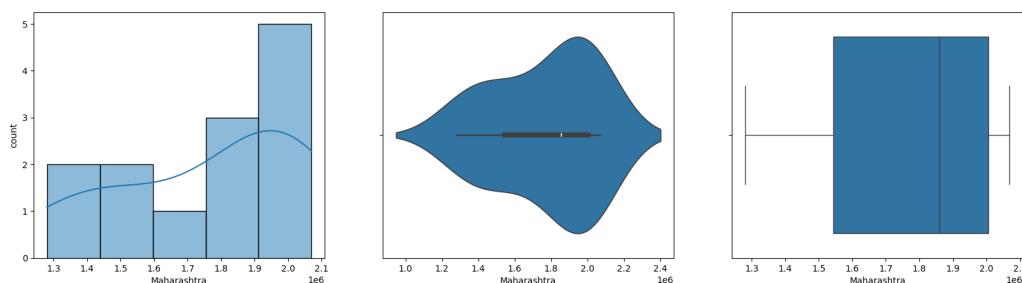


Figure 4.134: Data Distribution of GSDP of Maharashtra

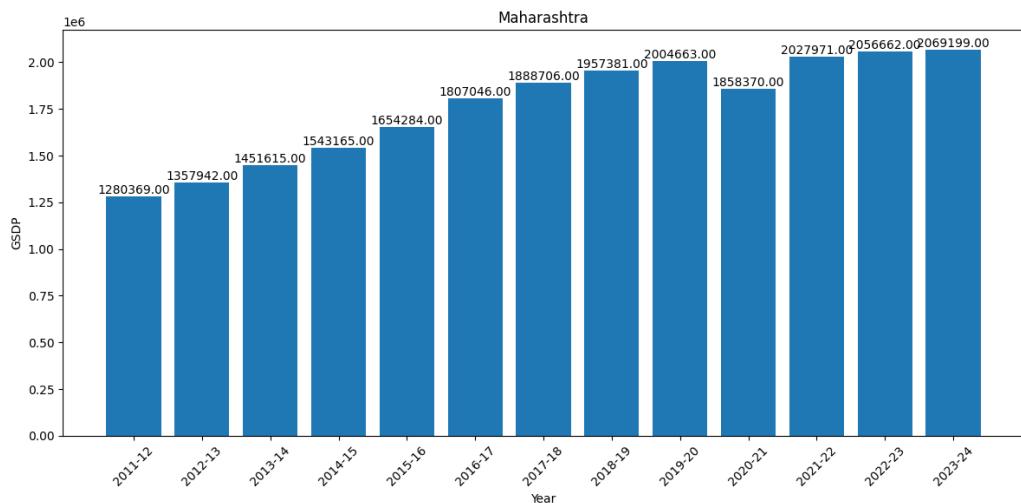


Figure 4.135: Value Distribution of GSDP of Maharashtra

4.3.1.1.15 Manipur

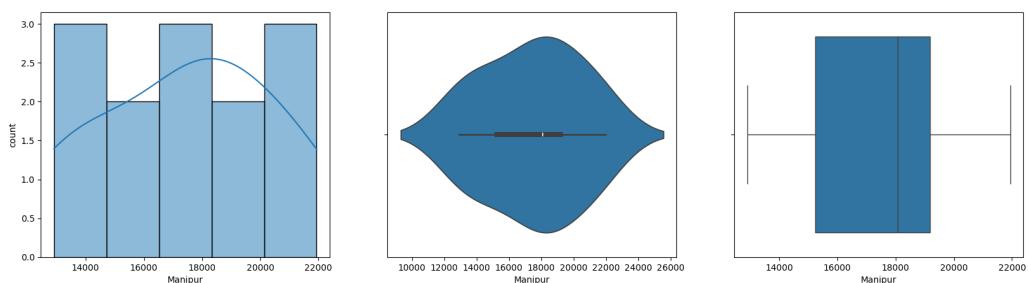


Figure 4.136: Data Distribution of GSDP of Manipur

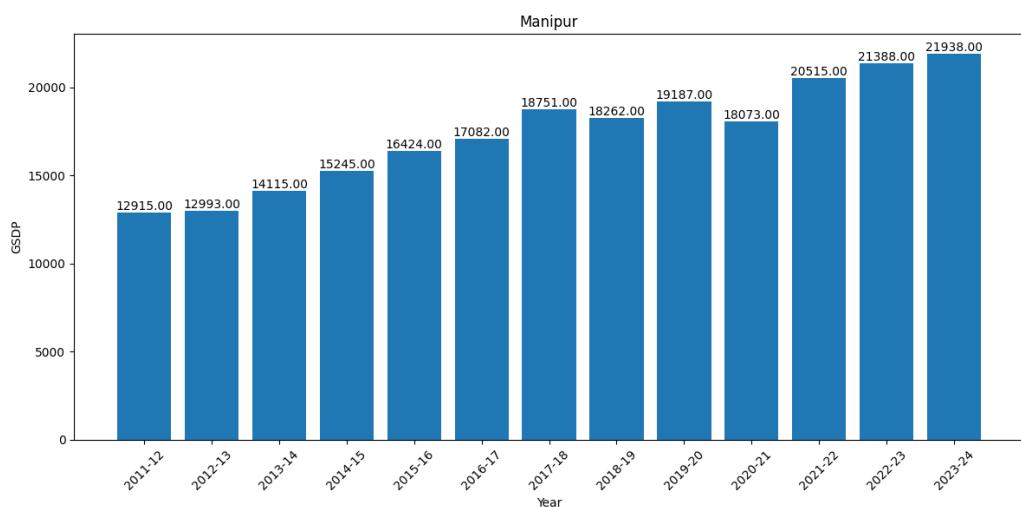


Figure 4.137: Value Distribution of GSDP of Manipur

4.3.1.1.16 Meghalaya

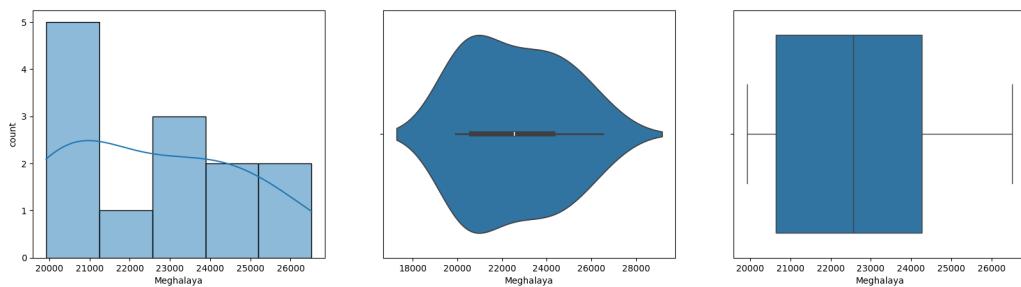


Figure 4.138: Data Distribution of GSDP of Meghalaya

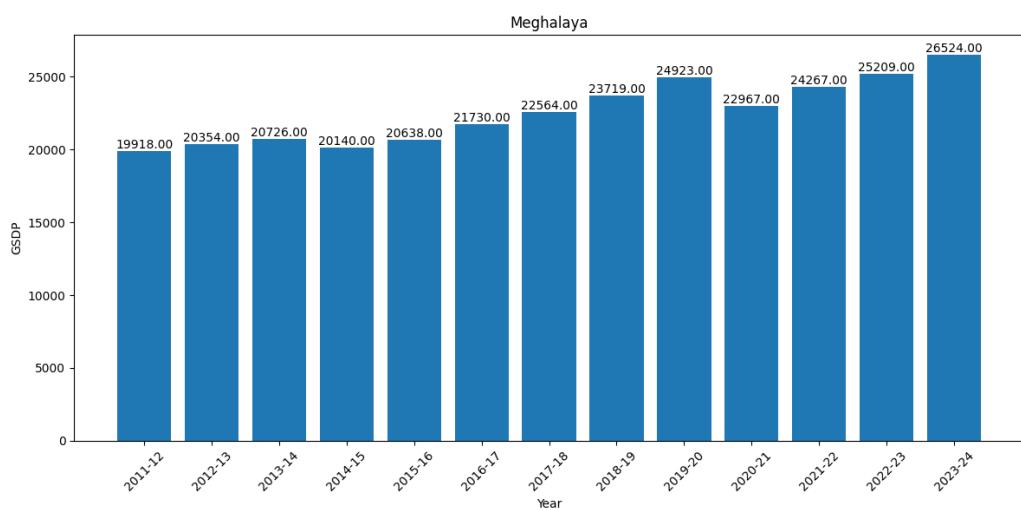


Figure 4.139: Value Distribution of GSDP of Meghalaya

4.3.1.1.17 Mizoram

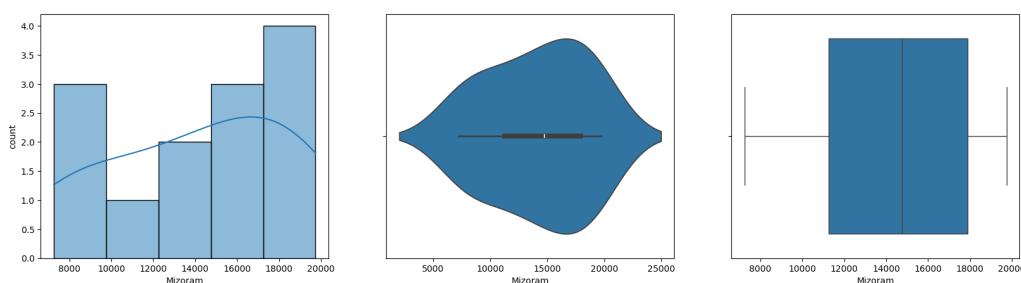


Figure 4.140: Data Distribution of GSDP of Mizoram

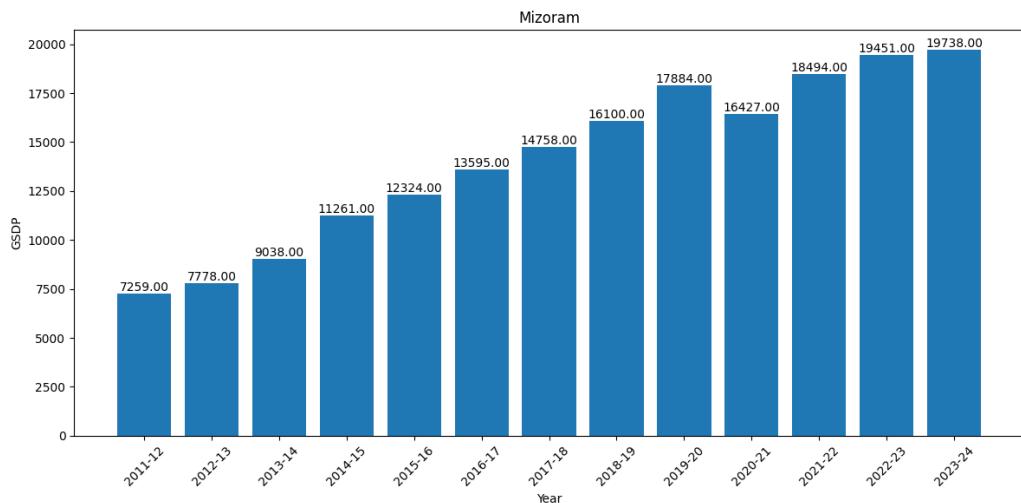


Figure 4.141: Value Distribution of GSDP of Mizoram

4.3.1.1.18 Nagaland

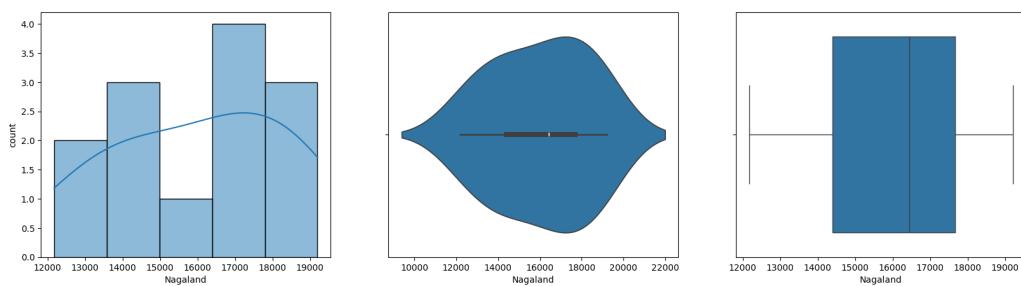


Figure 4.142: Data Distribution of GSDP of Nagaland

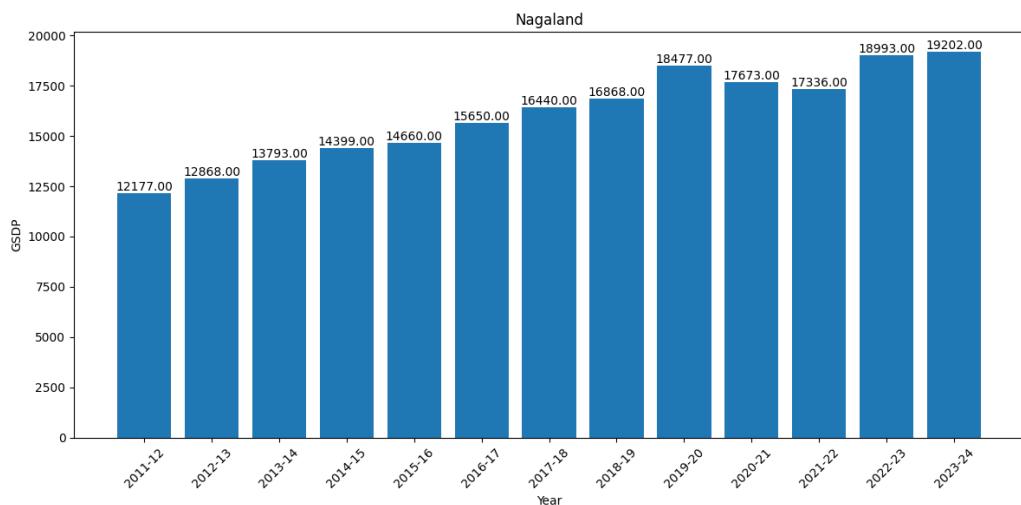


Figure 4.143: Value Distribution of GSDP of Nagaland

4.3.1.1.19 Odisha

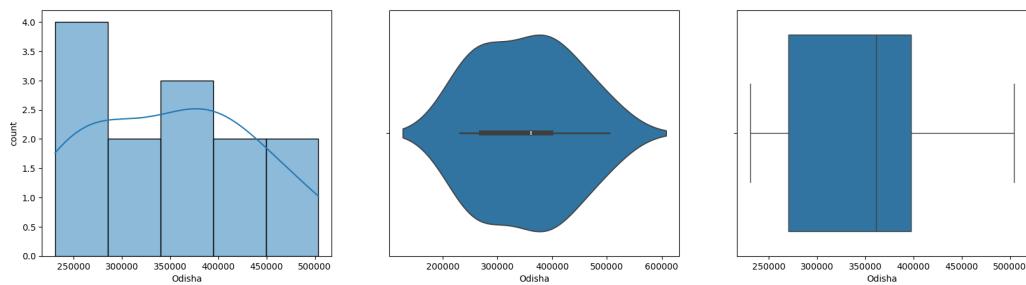


Figure 4.144: Data Distribution of GSDP of Odisha

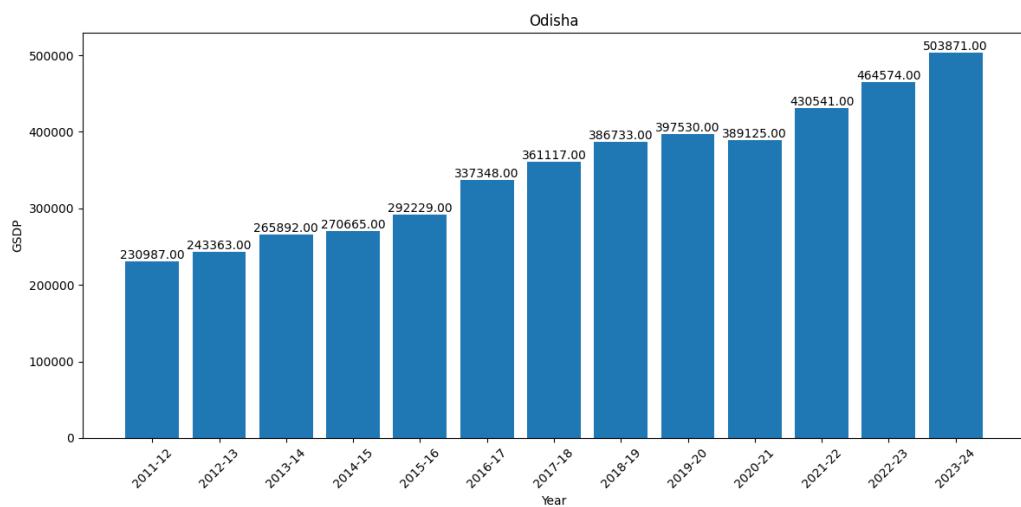


Figure 4.145: Value Distribution of GSDP of Odisha

4.3.1.1.20 Punjab

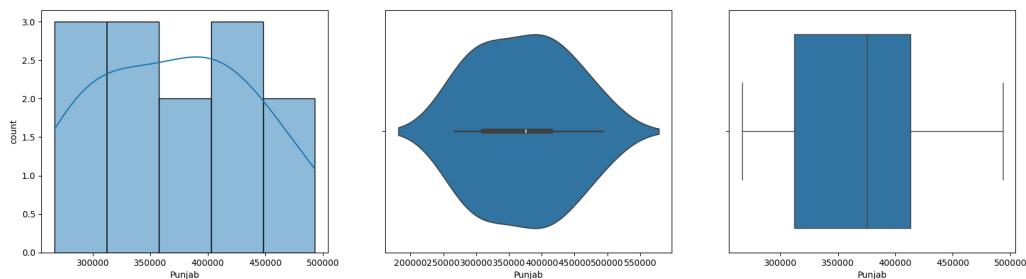


Figure 4.146: Data Distribution of GSDP of Punjab

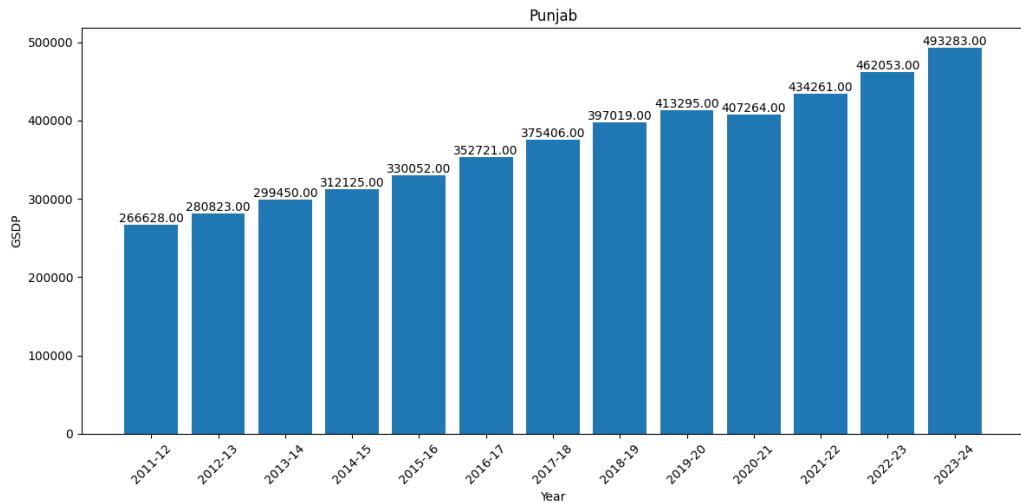


Figure 4.147: Value Distribution of GSDP of Punjab

4.3.1.1.21 Rajasthan

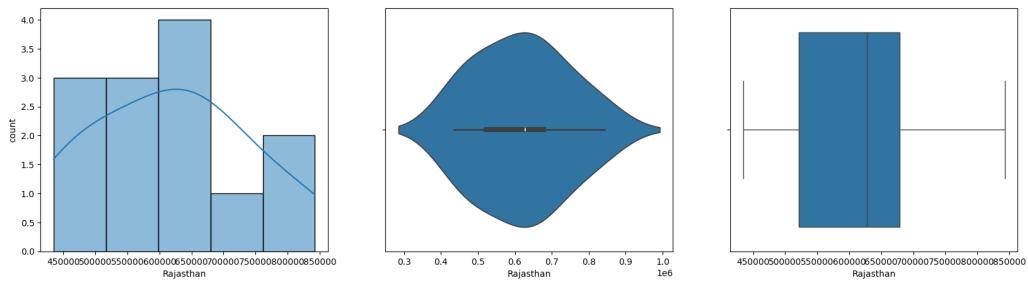


Figure 4.148: Data Distribution of GSDP of Rajasthan

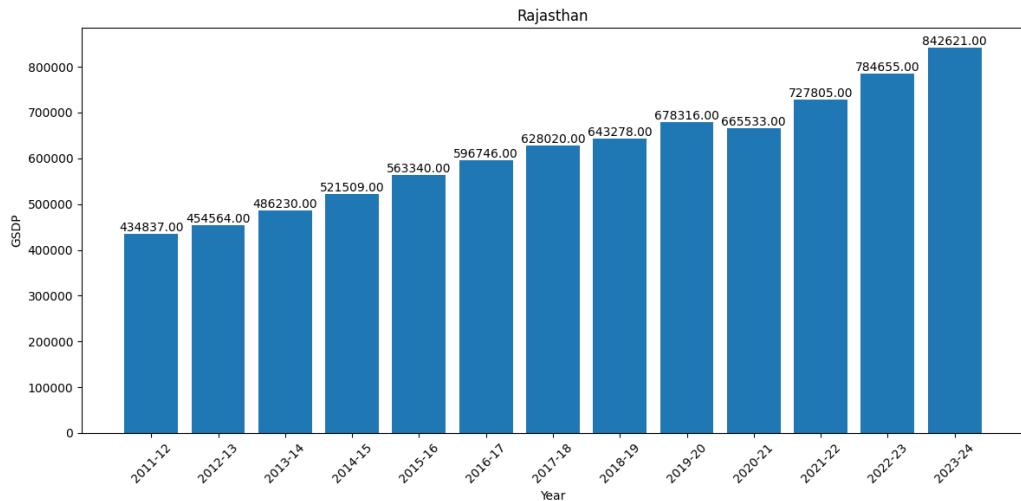


Figure 4.149: Value Distribution of GSDP of Rajasthan

4.3.1.1.22 Sikkim

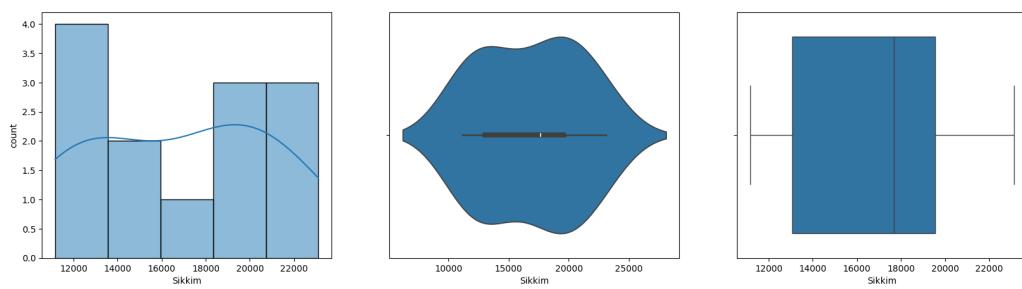


Figure 4.150: Data Distribution of GSDP of Sikkim

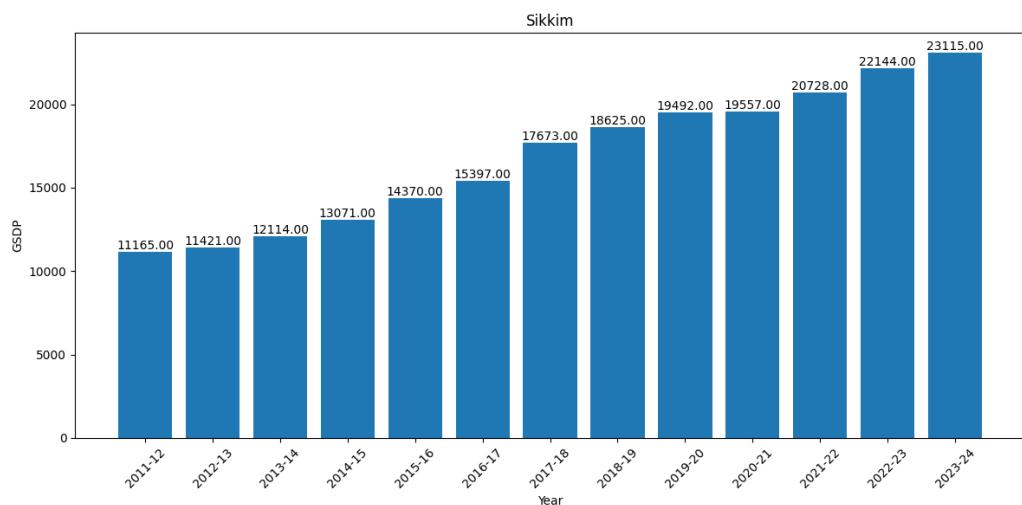


Figure 4.151: Value Distribution of GSDP of Sikkim

4.3.1.1.23 Tamil Nadu

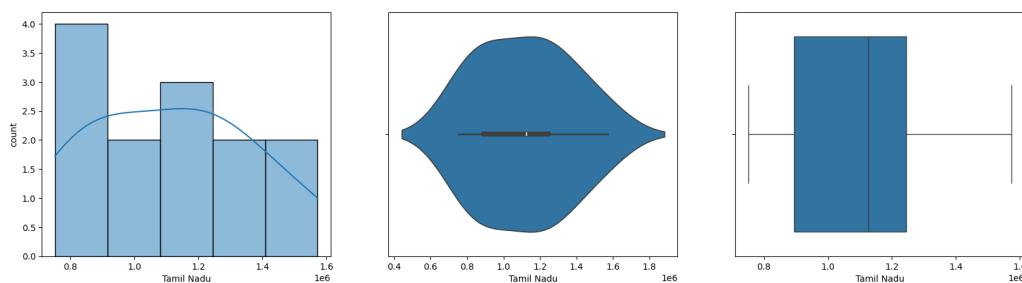


Figure 4.152: Data Distribution of GSDP of Tamil Nadu

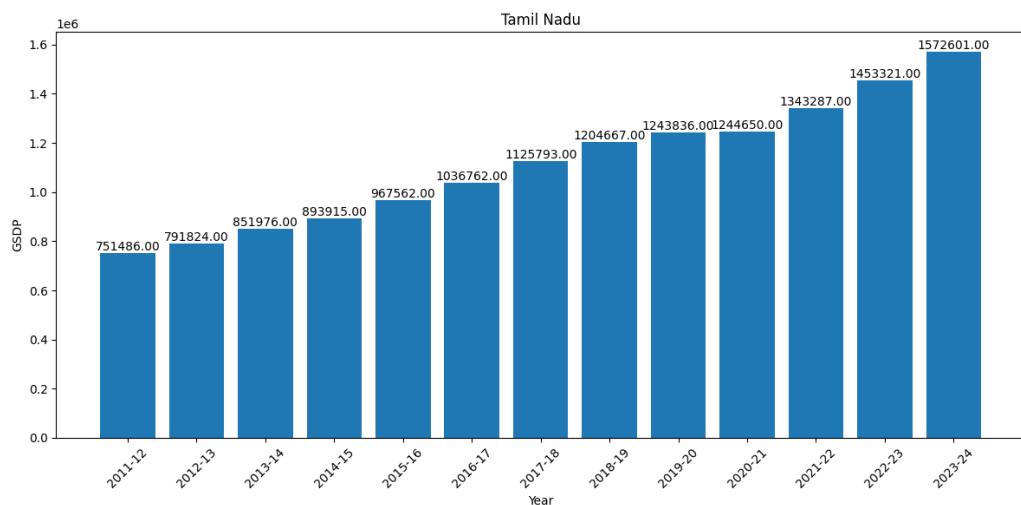


Figure 4.153: Value Distribution of GSDP of Tamil Nadu

4.3.1.1.24 Telangana

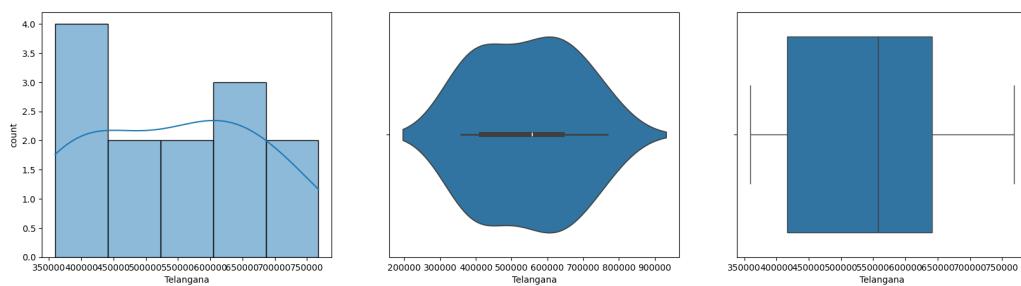


Figure 4.154: Data Distribution of GSDP of Telangana

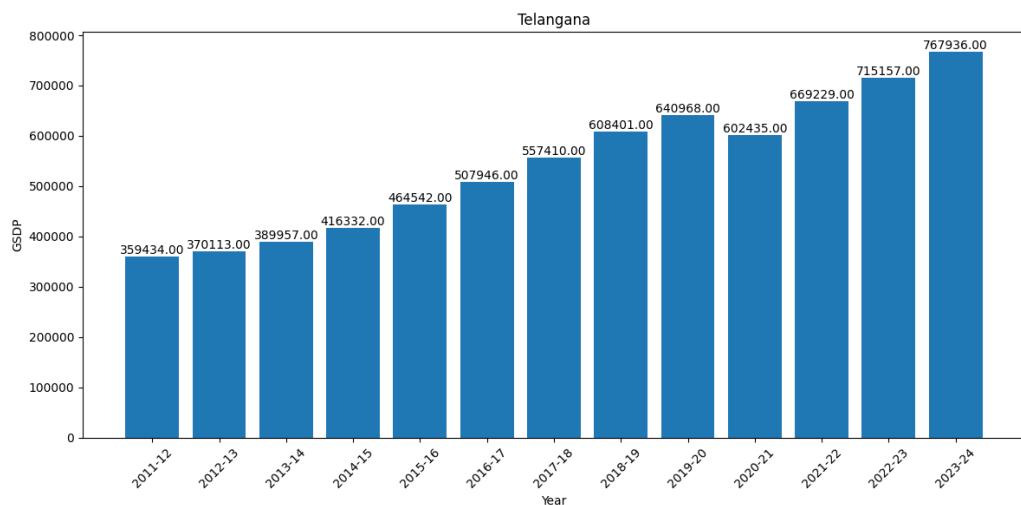


Figure 4.155: Value Distribution of GSDP of Telangana

4.3.1.1.25 Tripura

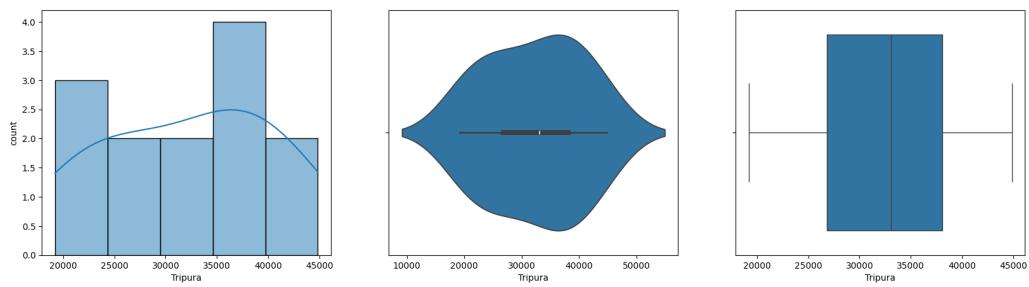


Figure 4.156: Data Distribution of GSDP of Tripura

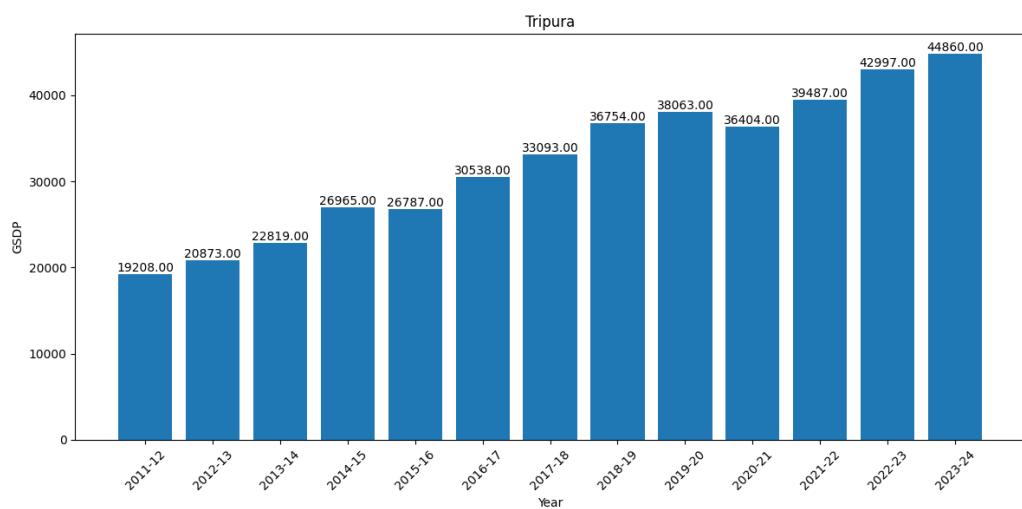


Figure 4.157: Value Distribution of GSDP of Tripura

4.3.1.1.26 Uttar Pradesh

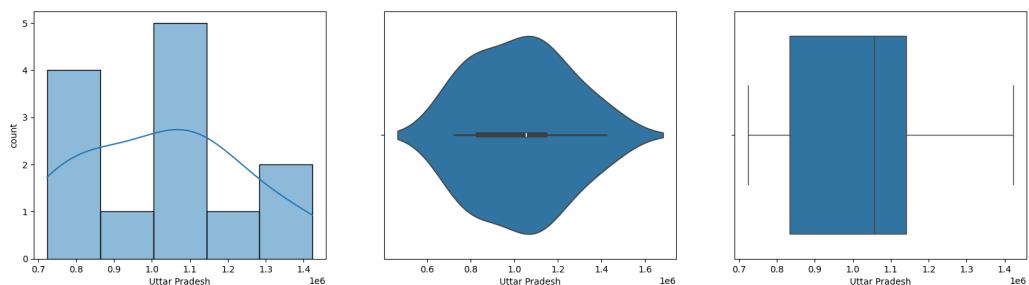


Figure 4.158: Data Distribution of GSDP of Uttar Pradesh

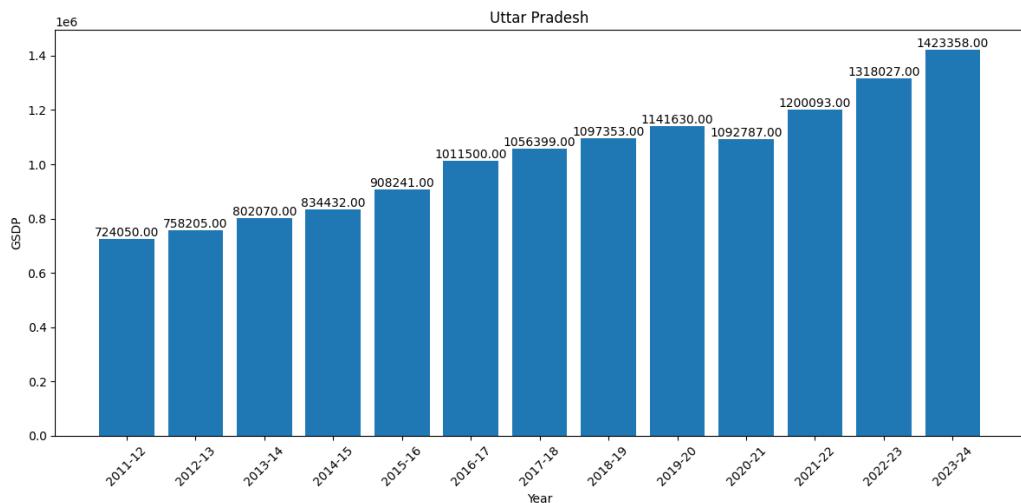


Figure 4.159: Value Distribution of GSDP of Uttar Pradesh

4.3.1.1.27 Uttarakhand

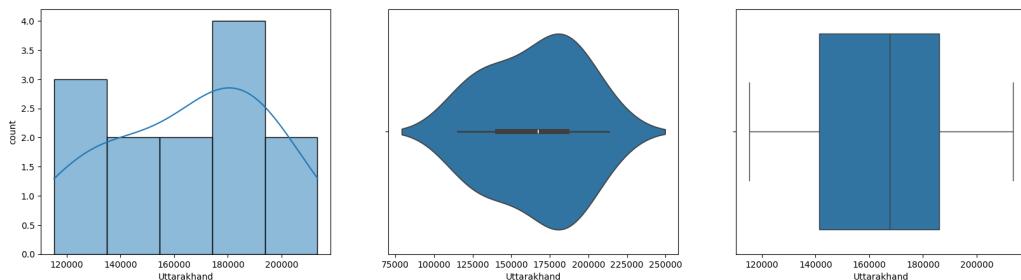


Figure 4.160: Data Distribution of GSDP of Uttarakhand

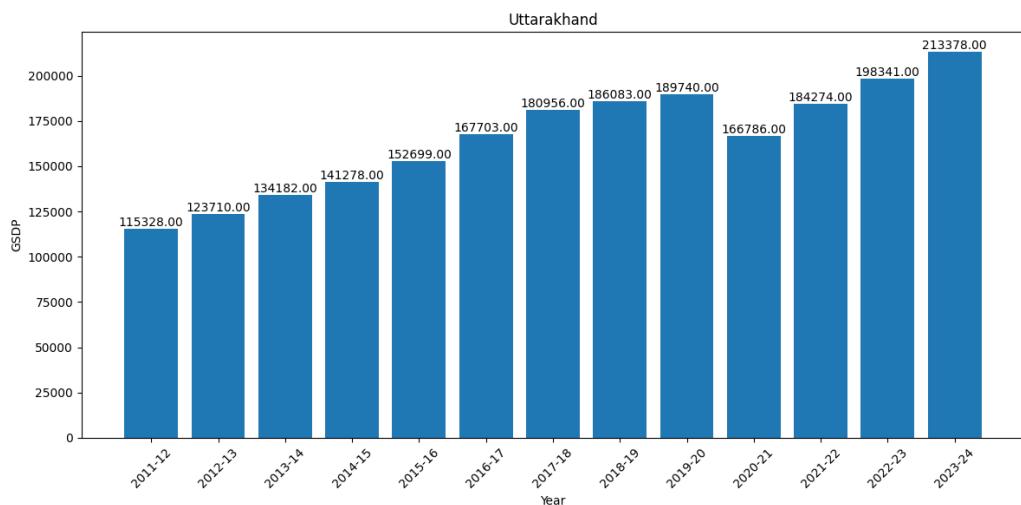


Figure 4.161: Value Distribution of GSDP of Uttarakhand

4.3.1.1.28 West Bengal

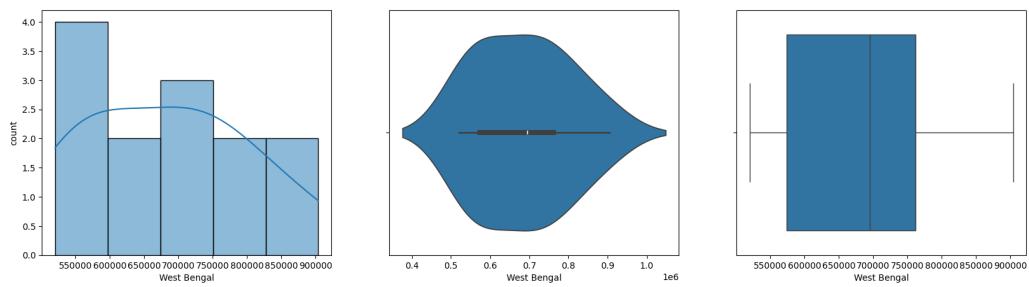


Figure 4.162: Data Distribution of GSDP of West Bengal

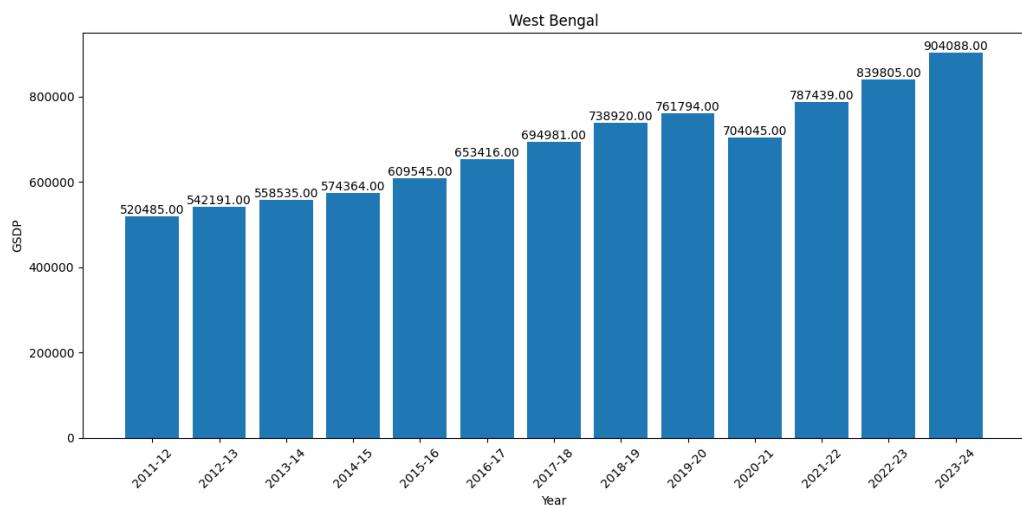


Figure 4.163: Value Distribution of GSDP of West Bengal

4.3.1.1.29 Andaman & Nicobar Islands

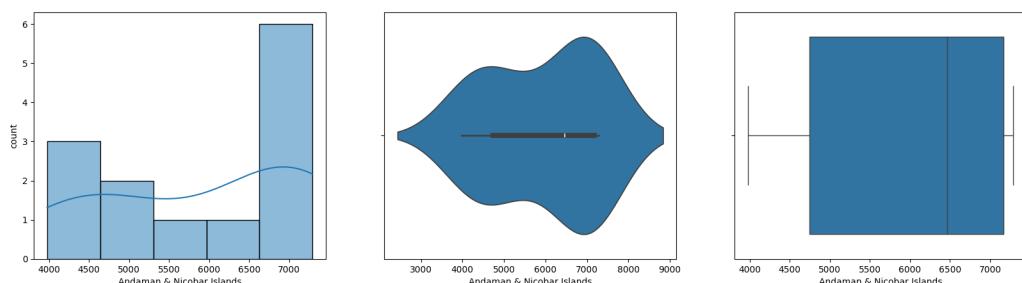


Figure 4.164: Data Distribution of GSDP of Andaman & Nicobar Islands

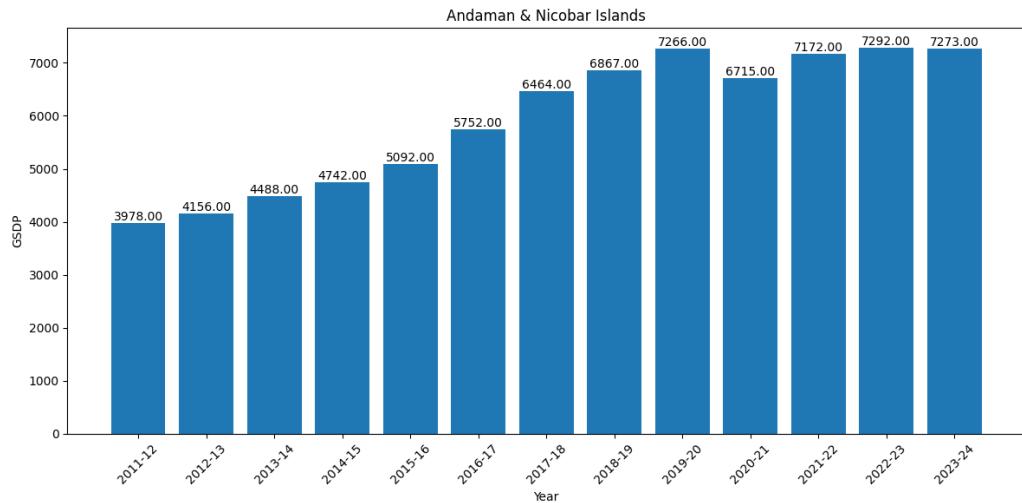


Figure 4.165: Value Distribution of GSDP of Andaman & Nicobar Islands

4.3.1.1.30 Chandigarh

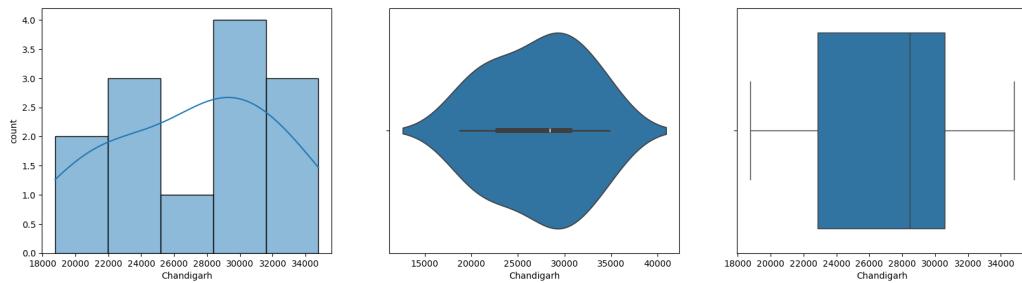


Figure 4.166: Data Distribution of GSDP of Chandigarh

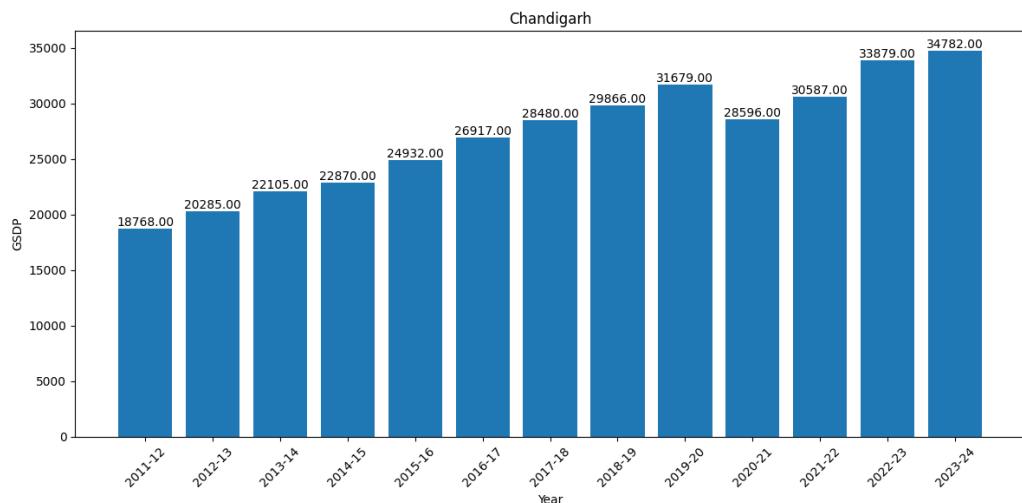


Figure 4.167: Value Distribution of GSDP of Chandigarh

paragraphDelhi

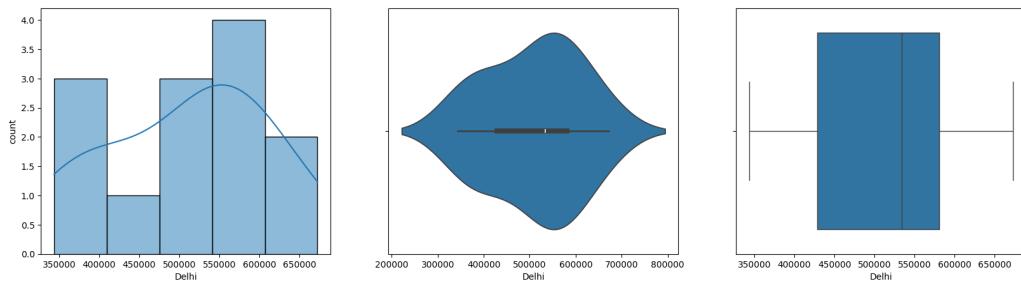


Figure 4.168: Data Distribution of GSDP of Delhi

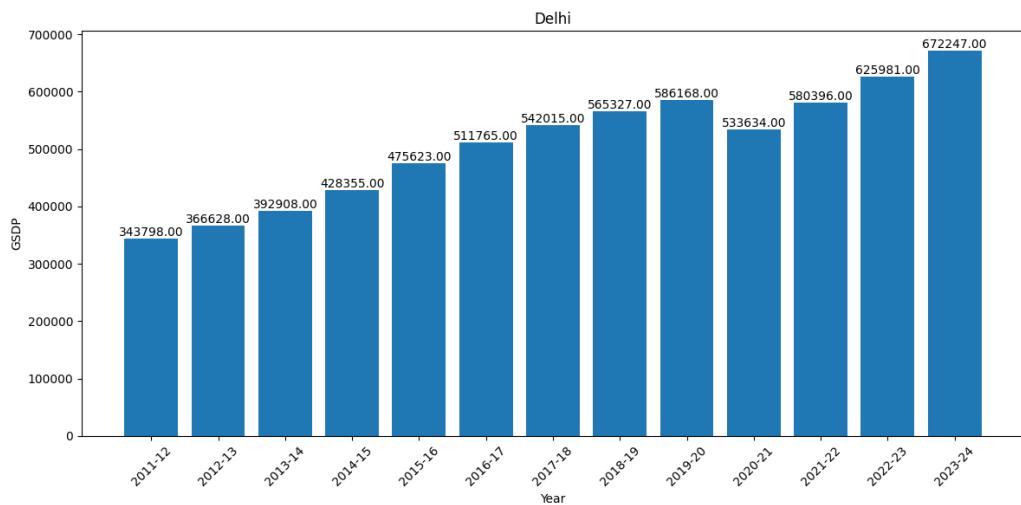


Figure 4.169: Value Distribution of GSDP of Delhi

4.3.1.1.31 Jammu & Kashmir-UT

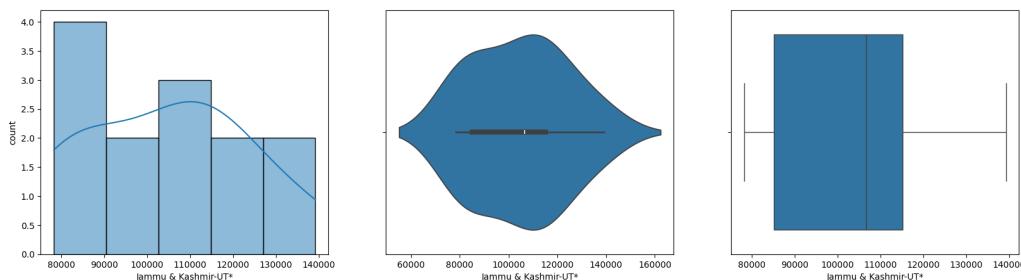


Figure 4.170: Data Distribution of CSDP of Jammu & Kashmir-UT

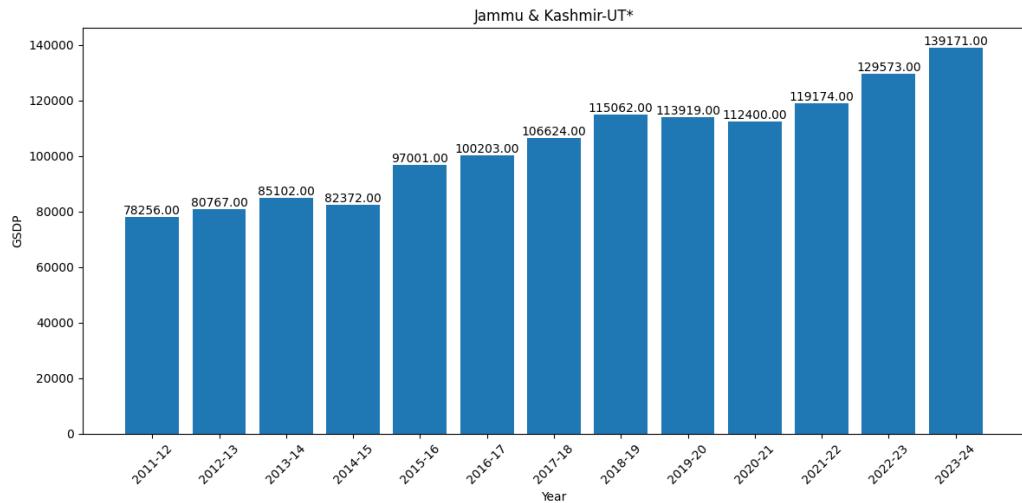


Figure 4.171: Value Distribution of GSDP of Jammu & Kashmir-UT

4.3.1.1.32 Puducherry

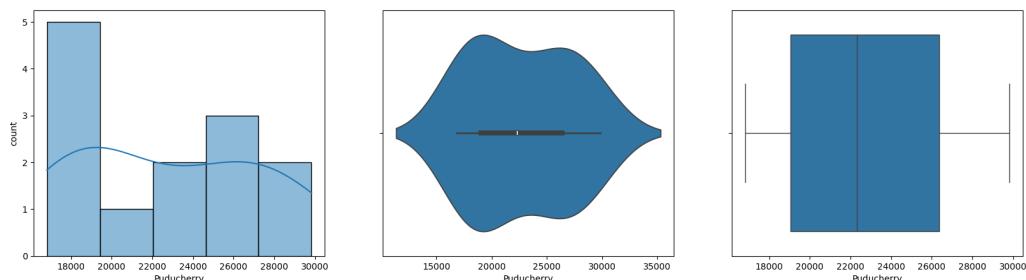


Figure 4.172: Data Distribution of GSDP of Puducherry

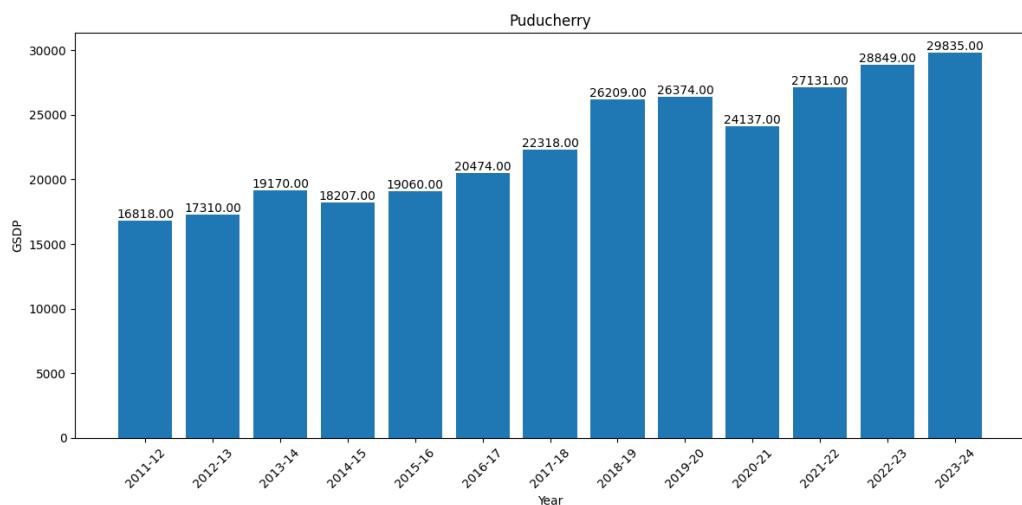


Figure 4.173: Value Distribution of GSDP of Puducherry



4.3.1.2 Bivariate Analysis

Here, we will analyze the relationship between the GSDP of two states.

4.3.1.3 Multivariate analysis

Here, we will analyze the relationship between the GSDP of multiple states.

4.3.1.3.1 Trend of GSDP

The plot given below shows the absolute growth of GSDP of various states over the years.

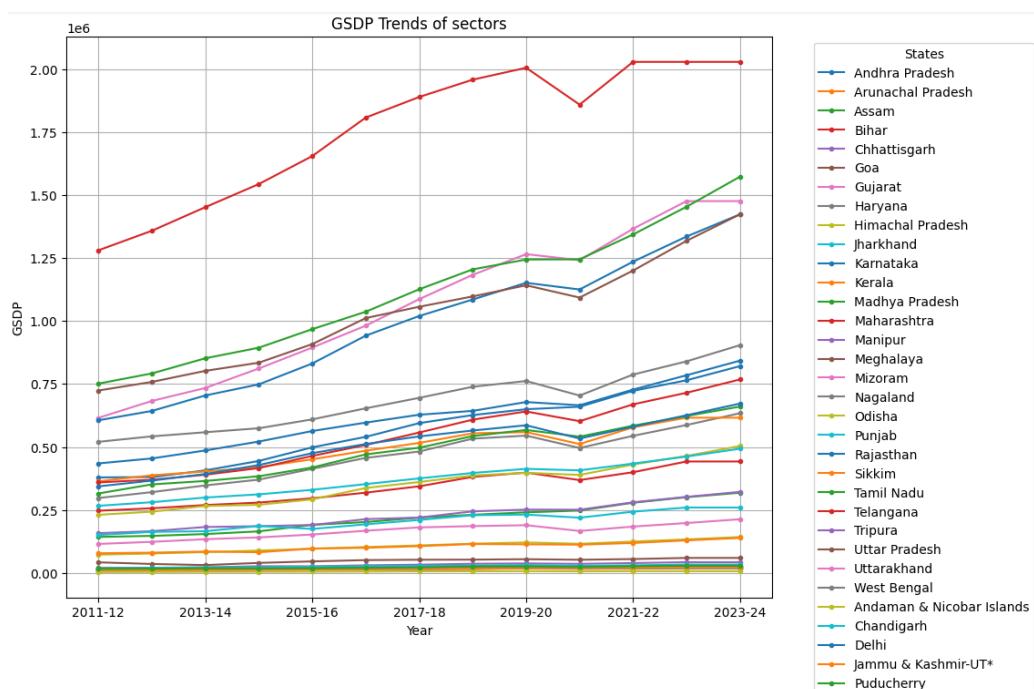


Figure 4.174: Growth of GSDP of States Over The Years in Terms of Absolute Value

4.3.1.3.2 Contribution of States

Given below are plots depicting the contribution of states towards the national GDP.

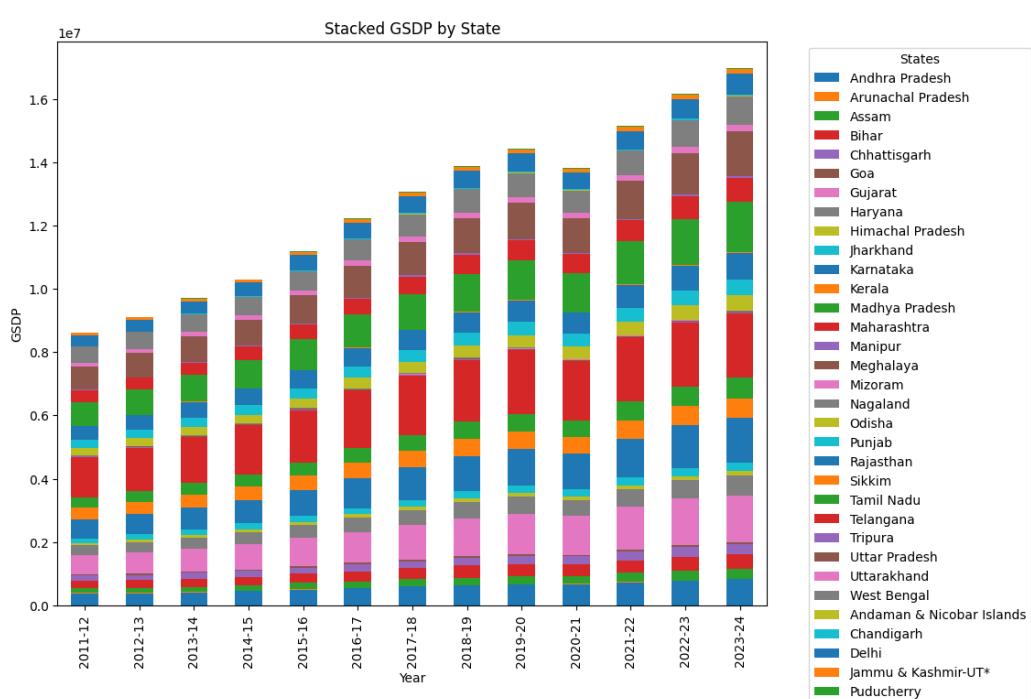


Figure 4.175: Contribution of States to National CDP Over The Years

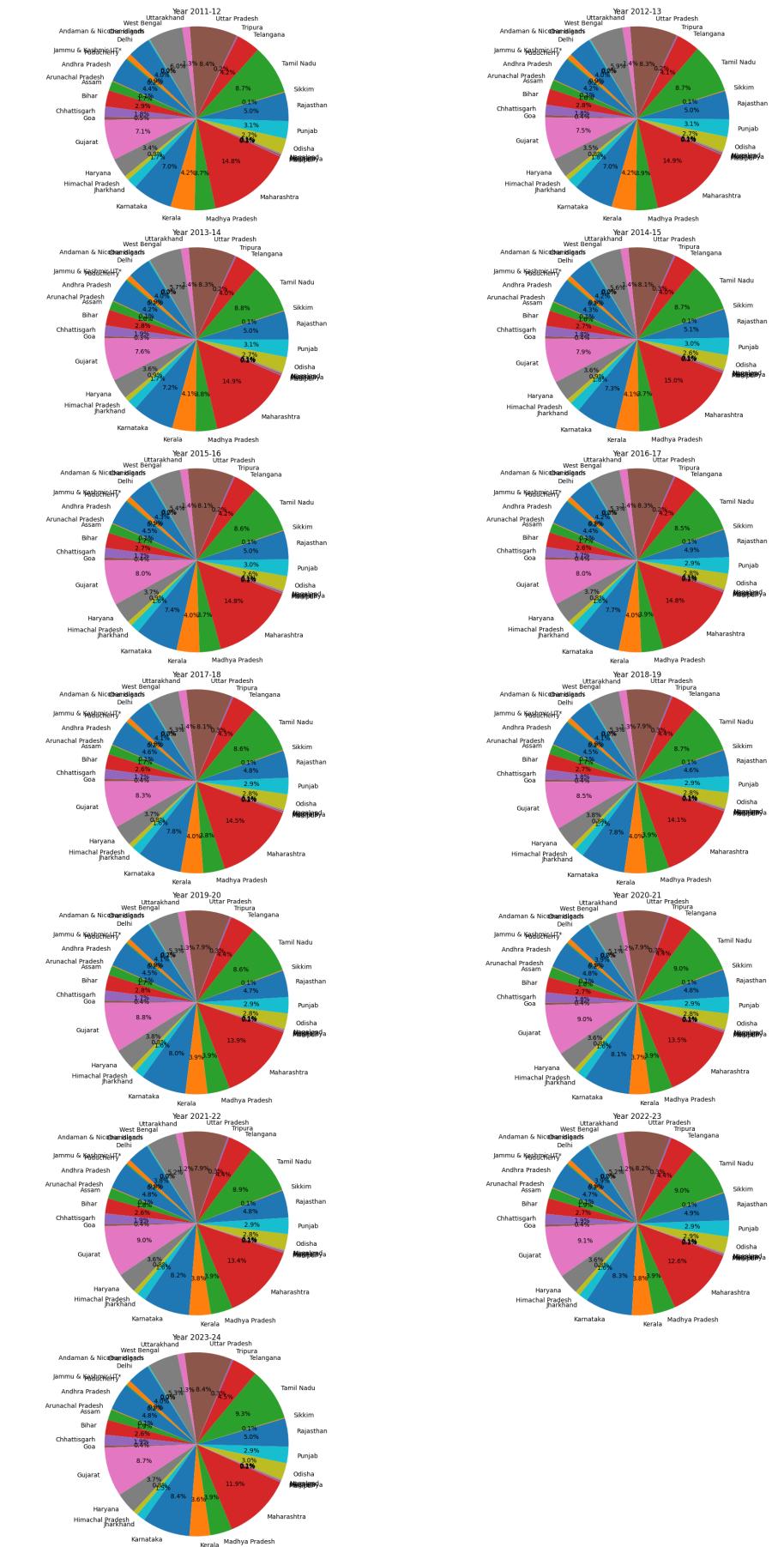


Figure 4.176: Contribution of States Over The Years in Terms of %



4.3.1.3.3 GSDP in 2023

Given below is a sorted representation of the 2023's GSDP of States of India.

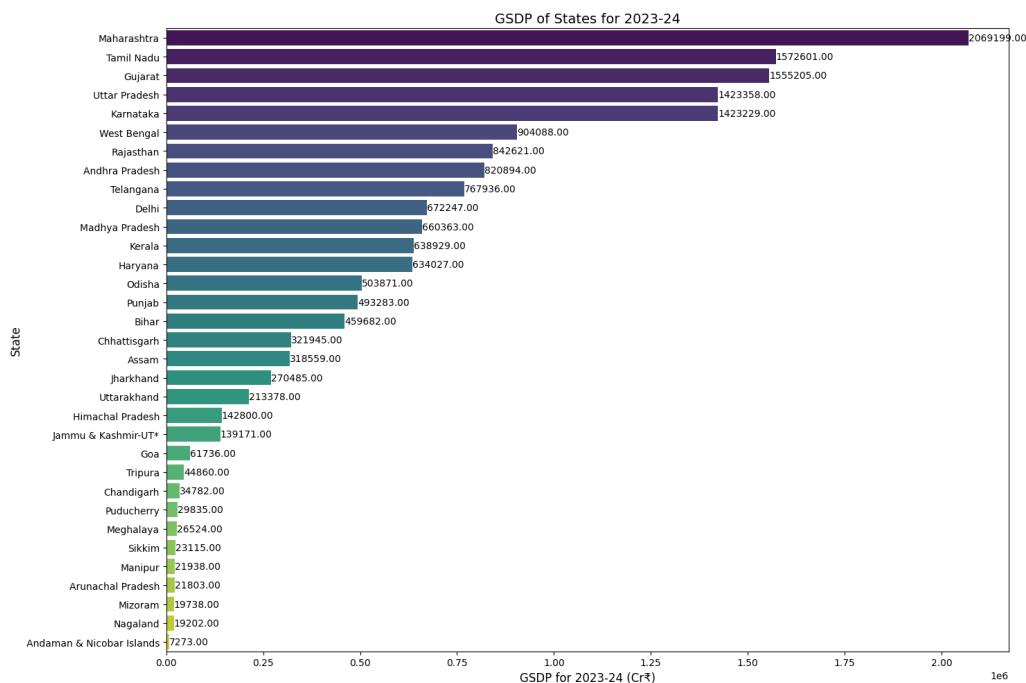


Figure 4.177: GSDP of States of India in 2023

Here, we will analyze the contributions of these Zones of India and how they have performed over the years.

4.3.1.3.4 Trend of Contribution of Zones of India

The plot here represents the growth of the contribution of Zones of India over the years in terms of absolute value.

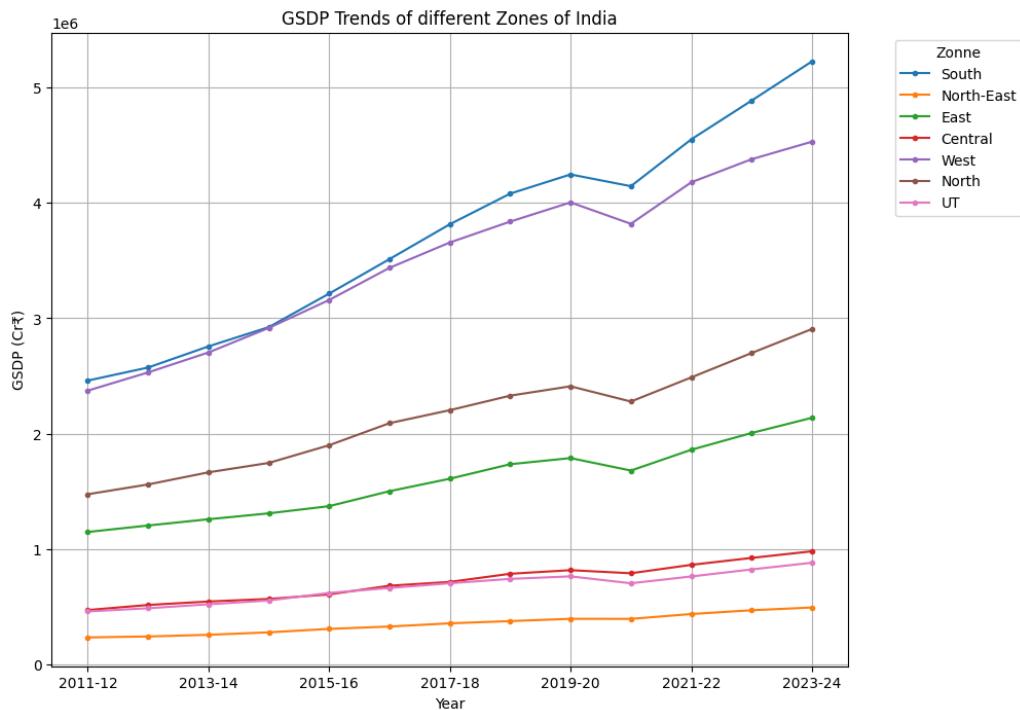


Figure 4.178: Growth of Contribution of Zones of India Over The Years in Terms of Absolute Value

4.3.1.3.5 Contribution of Zones Of India

The plots below represent contribution of Zones Of India to the national GDP.

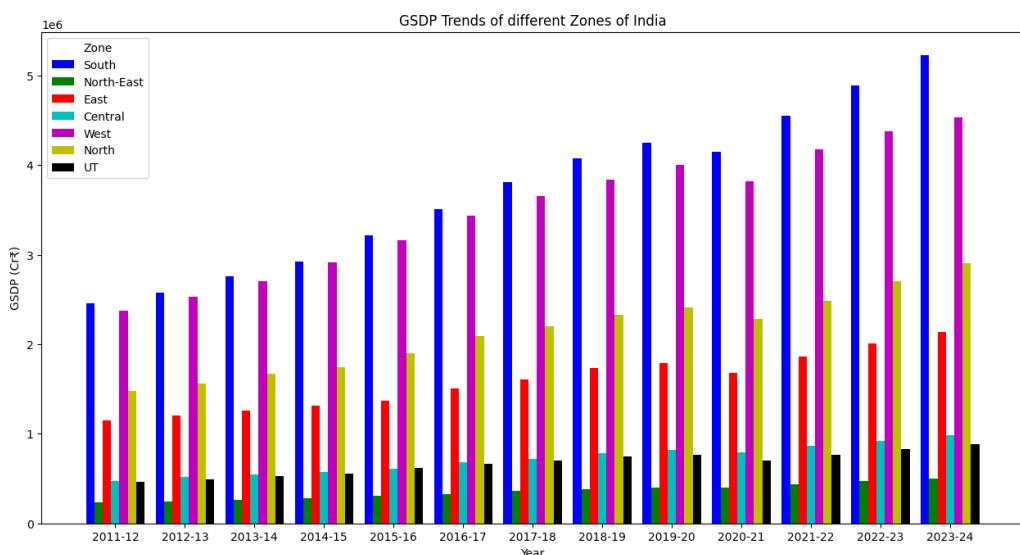


Figure 4.179: Contribution of Zones Of India Over the Years - Side By Side

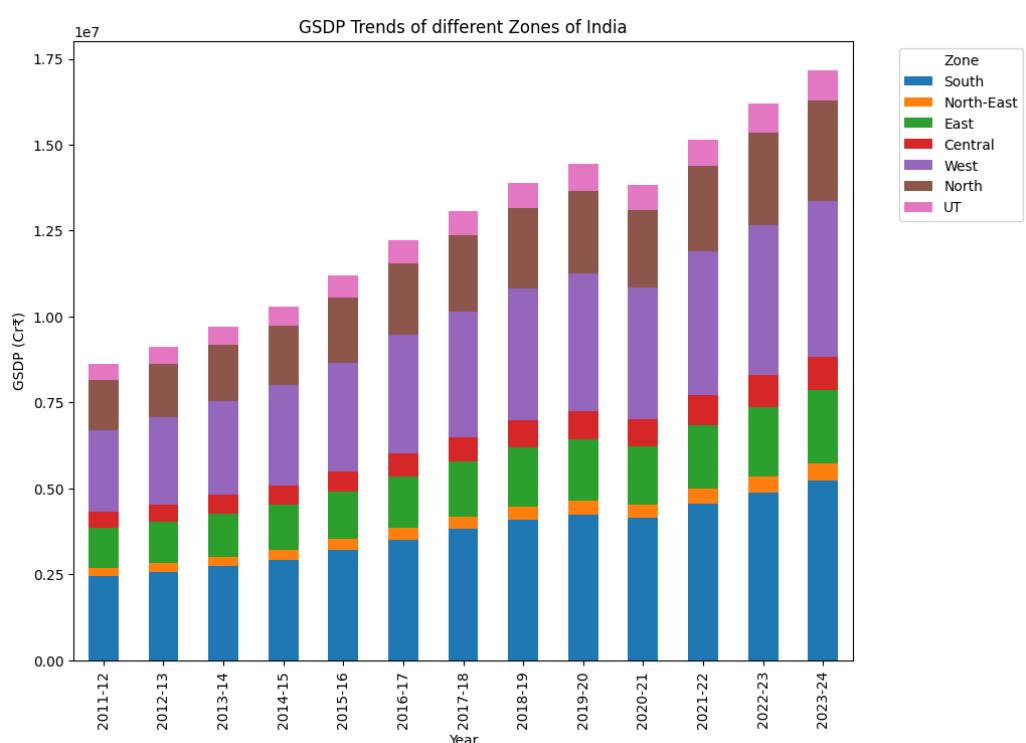


Figure 4.180: Contribution of Zones Of India Over the Years- Stacked

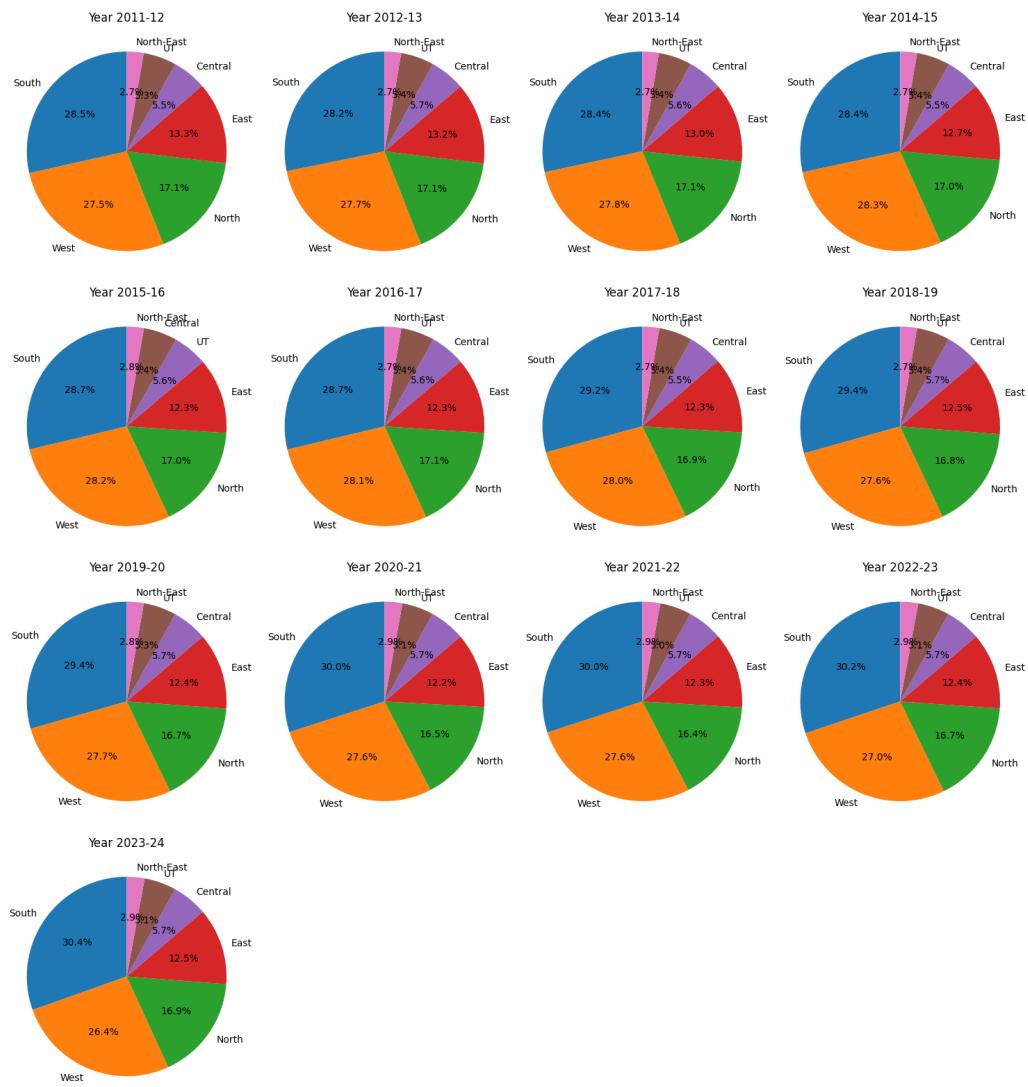


Figure 4.181: Contribution of Zones Of India Over the Years in Terms of %



4.3.2 Growth Rate of GSDP

4.3.2.1 Univariate Analysis

Here, we will analyze the Growth Rate of GSDP of each state individually.

4.3.2.1.1 Andhra Pradesh

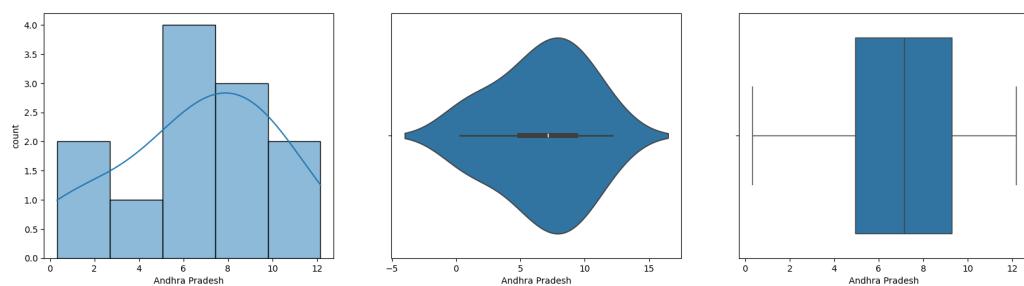


Figure 4.182: Data Distribution of Growth Rate of GSDP of Andhra Pradesh

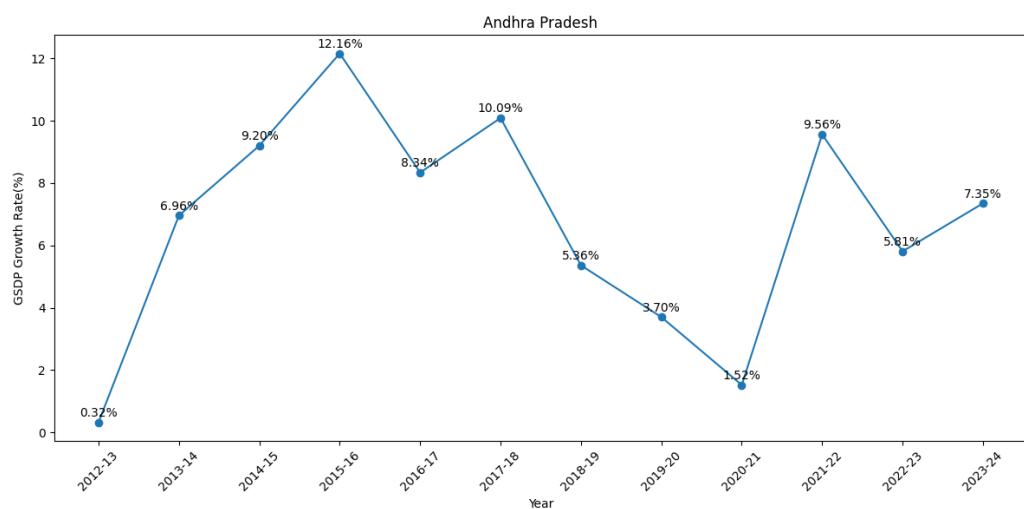


Figure 4.183: Trend of Growth Rate of GSDP of Andhra Pradesh

4.3.2.1.2 Arunachal Pradesh

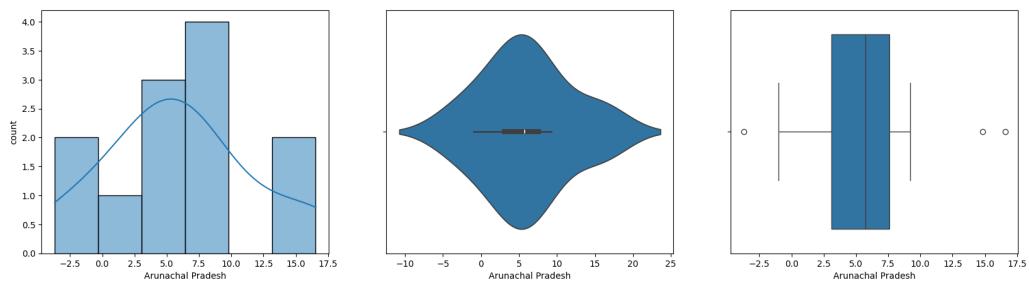


Figure 4.184: Data Distribution of Growth Rate of GSDP of Arunachal Pradesh

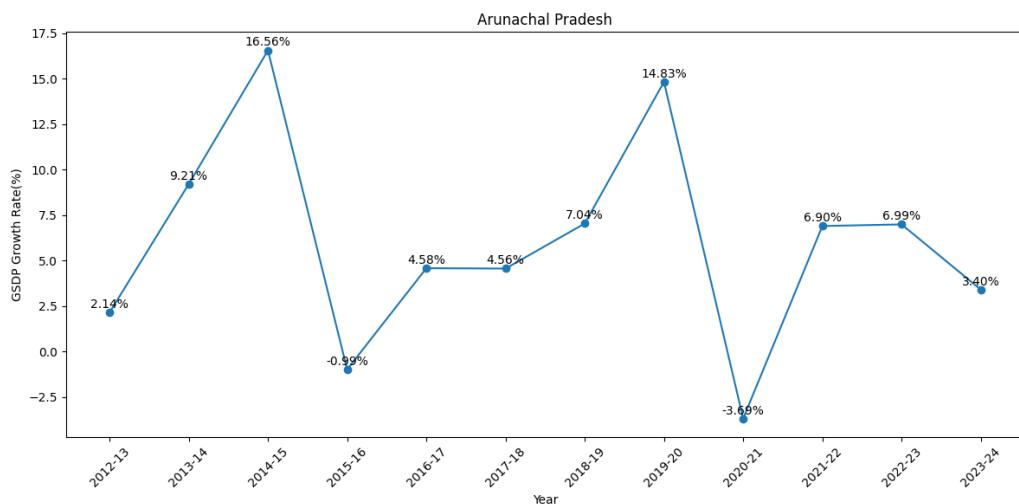


Figure 4.185: Trend of Growth Rate of GSDP of Arunachal Pradesh

4.3.2.1.3 Assam

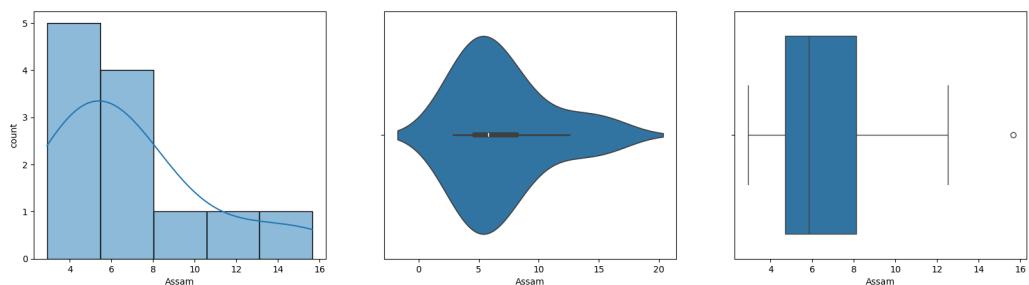


Figure 4.186: Data Distribution of Growth Rate of GSDP of Assam

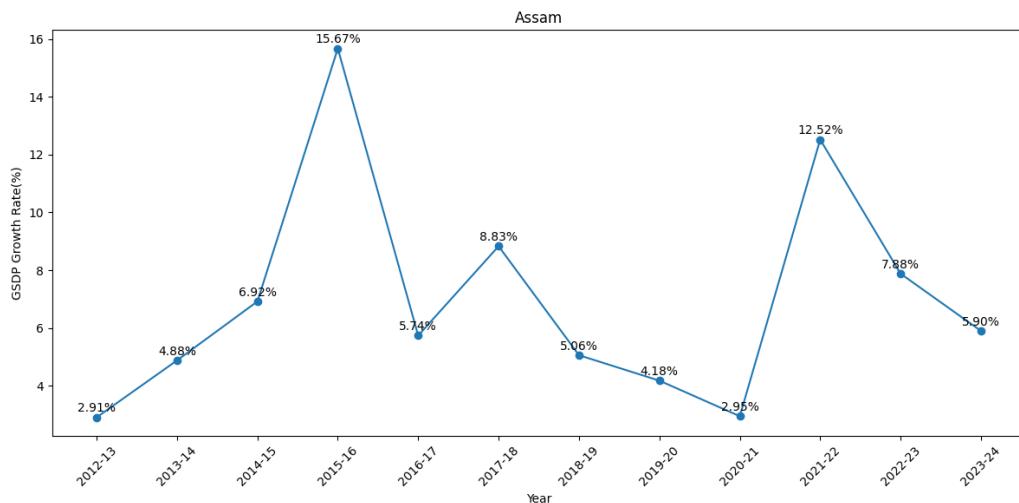


Figure 4.187: Trend of Growth Rate of GSDP of Assam

4.3.2.1.4 Bihar

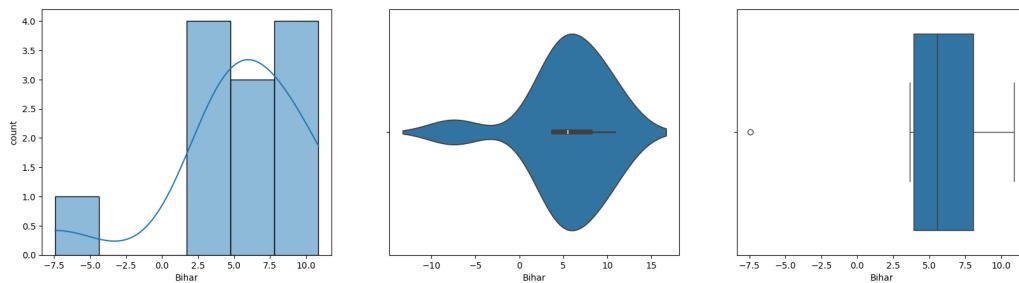


Figure 4.188: Data Distribution of Growth Rate of GSDP of Bihar

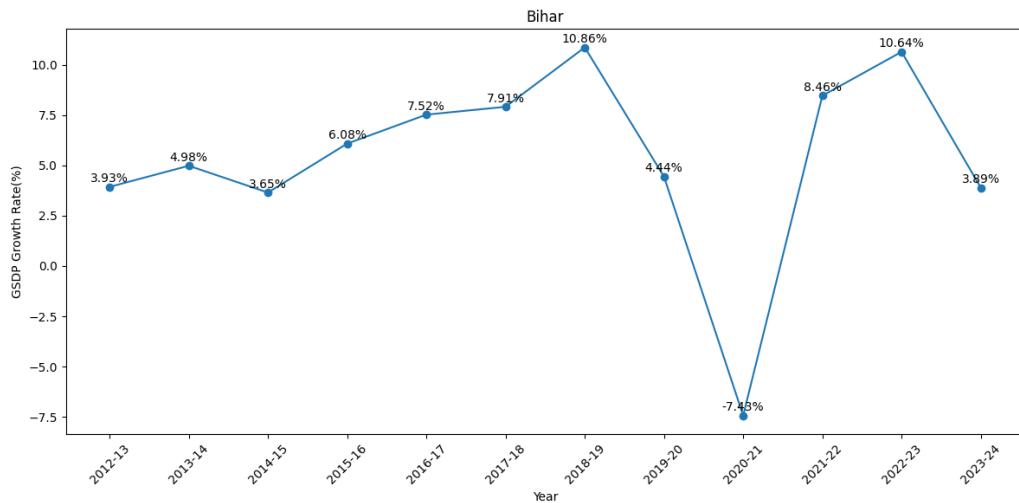


Figure 4.189: Trend of Growth Rate of GSDP of Bihar

4.3.2.1.5 Chhattisgarh

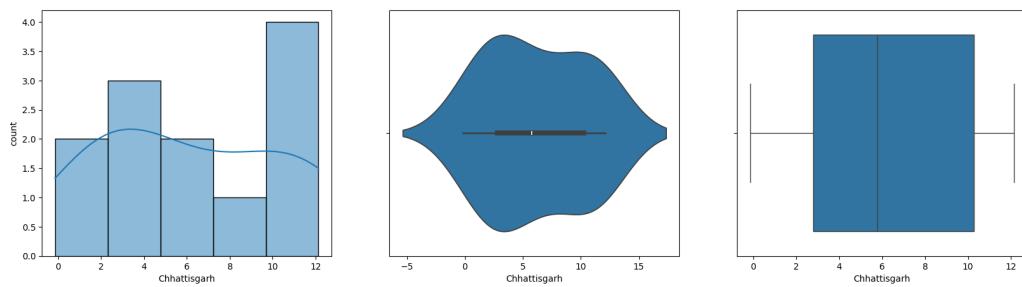


Figure 4.190: Data Distribution of Growth Rate of GSDP of Chhattisgarh

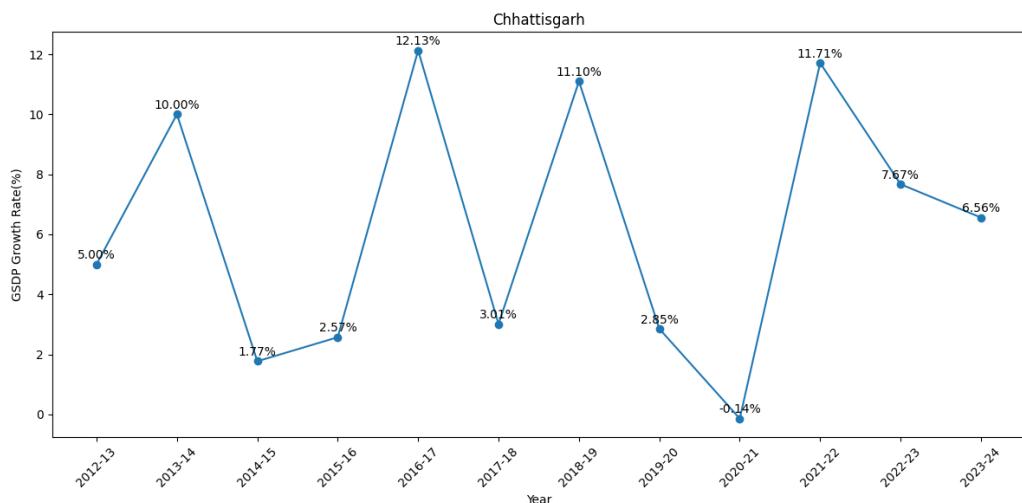


Figure 4.191: Trend of Growth Rate of GSDP of Chhattisgarh

4.3.2.1.6 Goa

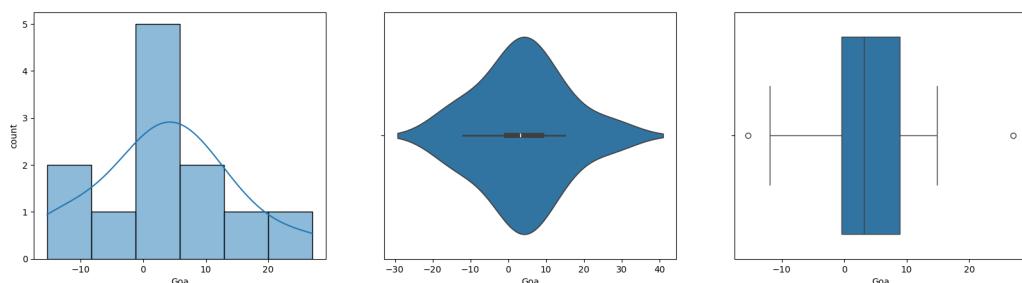


Figure 4.192: Data Distribution of Growth Rate of GSDP of Goa

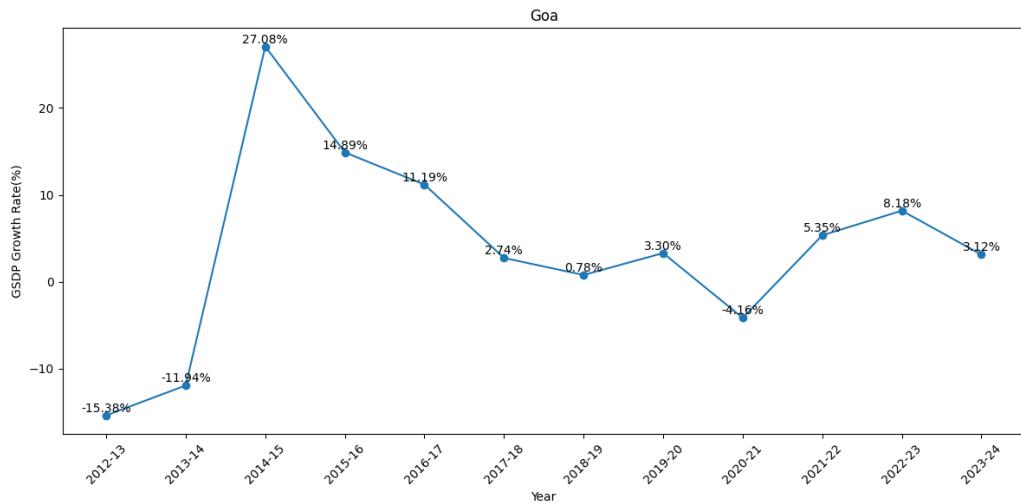


Figure 4.193: Trend of Growth Rate of GSDP of Goa

4.3.2.1.7 Gujarat

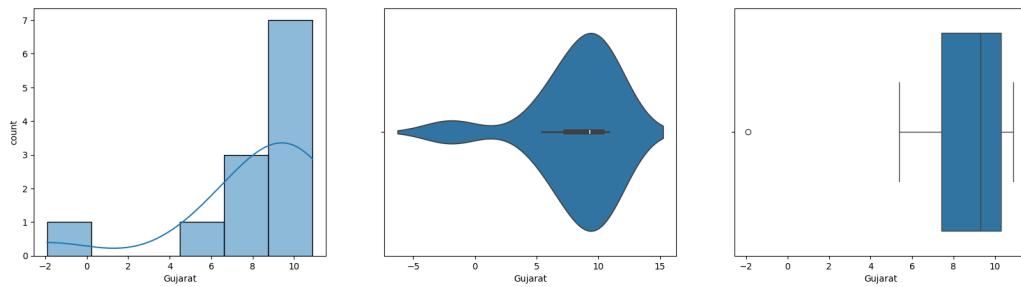


Figure 4.194: Data Distribution of Growth Rate of GSDP of Gujarat

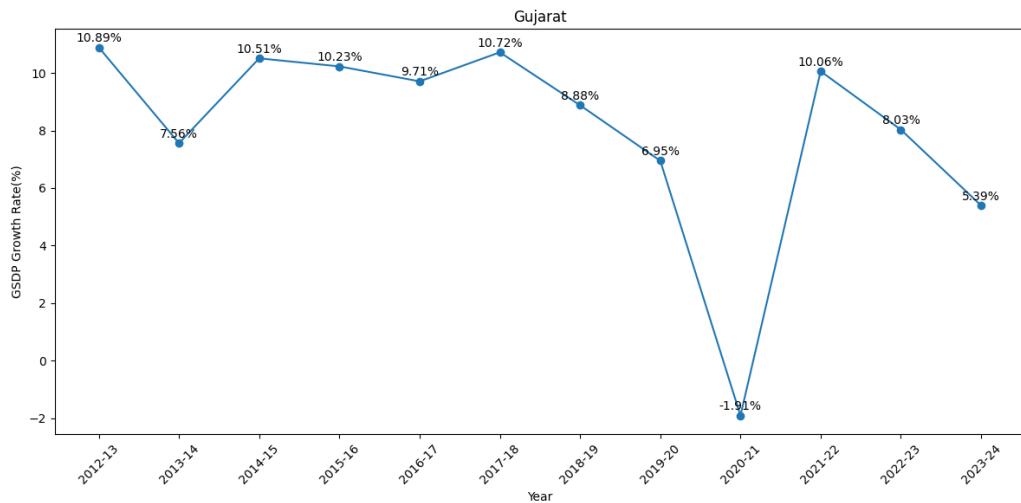


Figure 4.195: Trend of Growth Rate of GSDP of Gujarat

4.3.2.1.8 Haryana

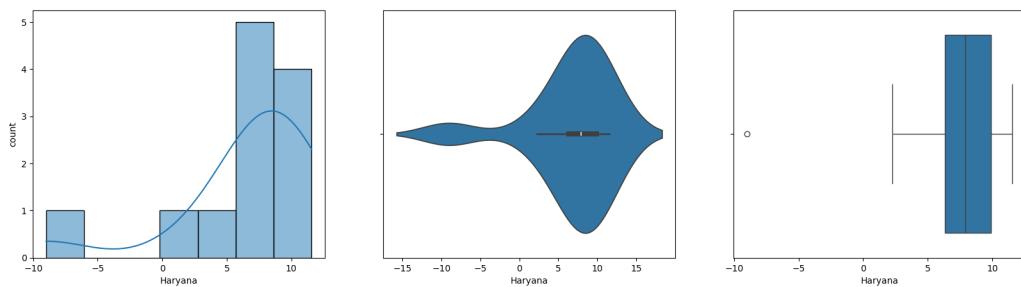


Figure 4.196: Data Distribution of Growth Rate of GSDP of Haryana

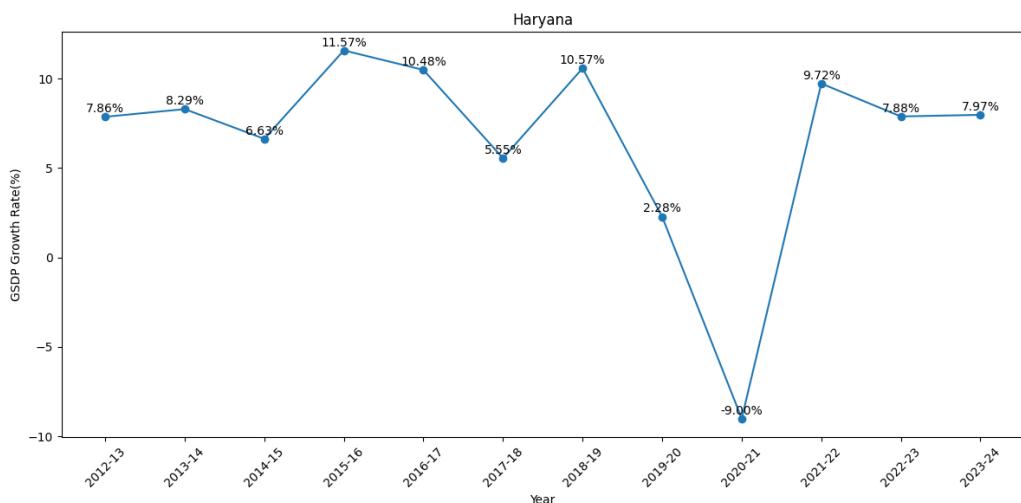


Figure 4.197: Trend of Growth Rate of GSDP of Haryana

4.3.2.1.9 Himachal Pradesh

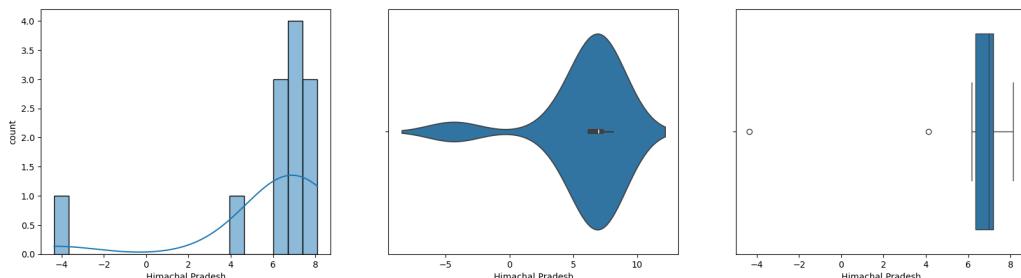


Figure 4.198: Data Distribution of Growth Rate of GSDP of Himachal Pradesh

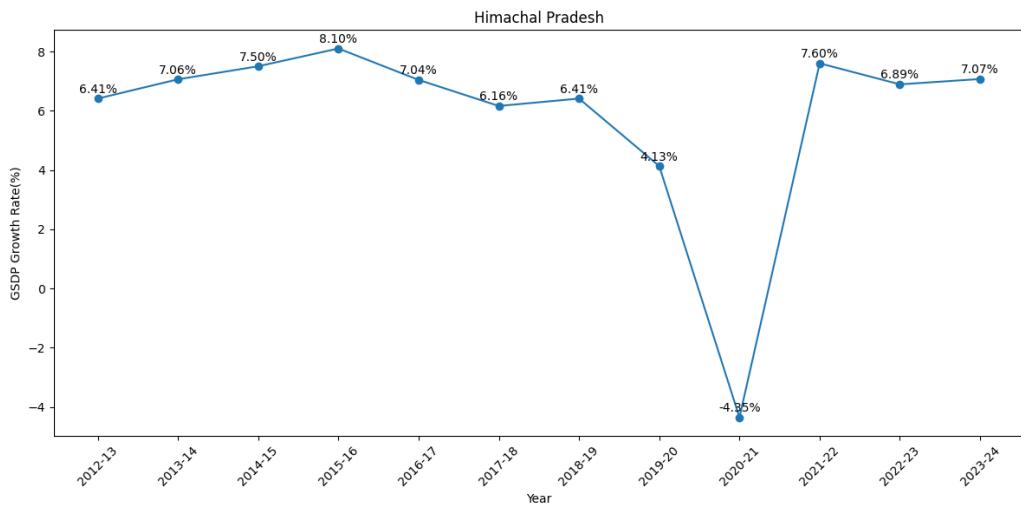


Figure 4.199: Trend of Growth Rate of CSDP of Himachal Pradesh

4.3.2.1.10 Jharkhand

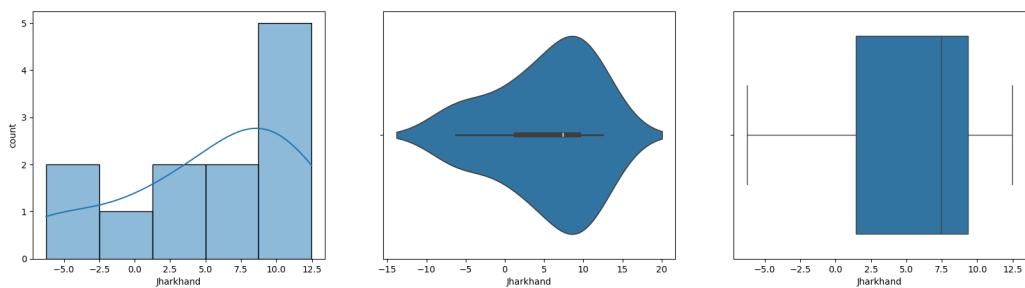


Figure 4.200: Data Distribution of Growth Rate of CSDP of Jharkhand

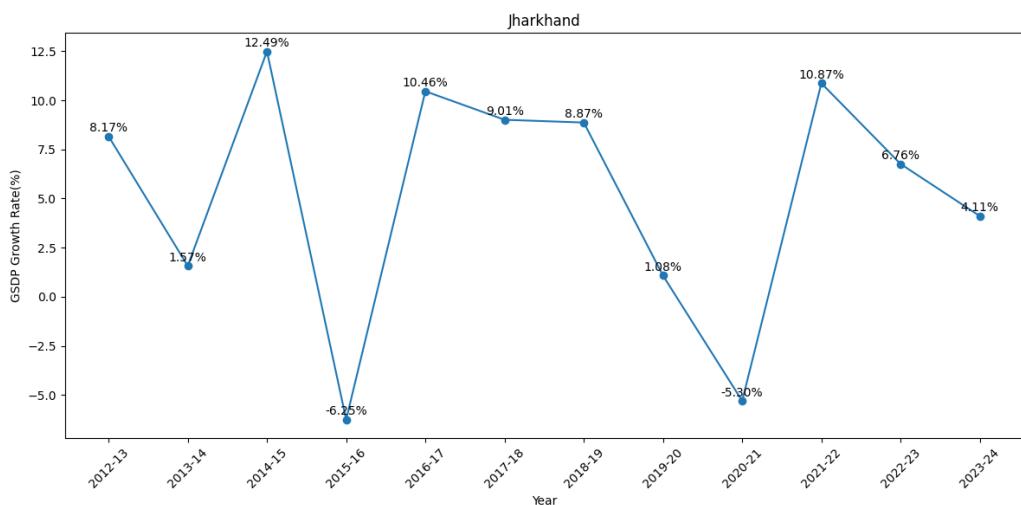


Figure 4.201: Trend of Growth Rate of CSDP of Jharkhand

4.3.2.1.11 Karnataka

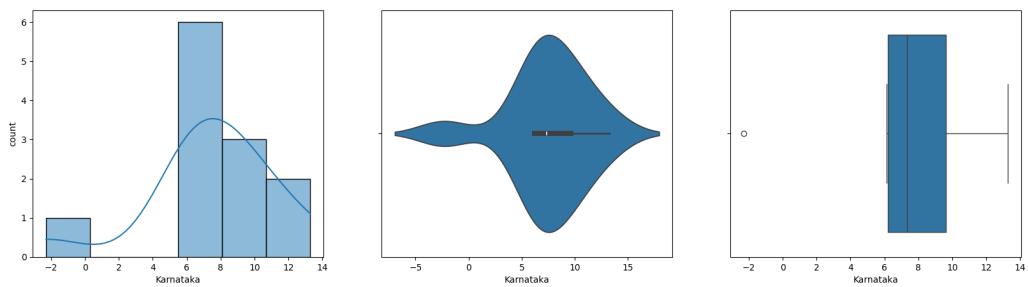


Figure 4.202: Data Distribution of Growth Rate of GSDP of Karnataka

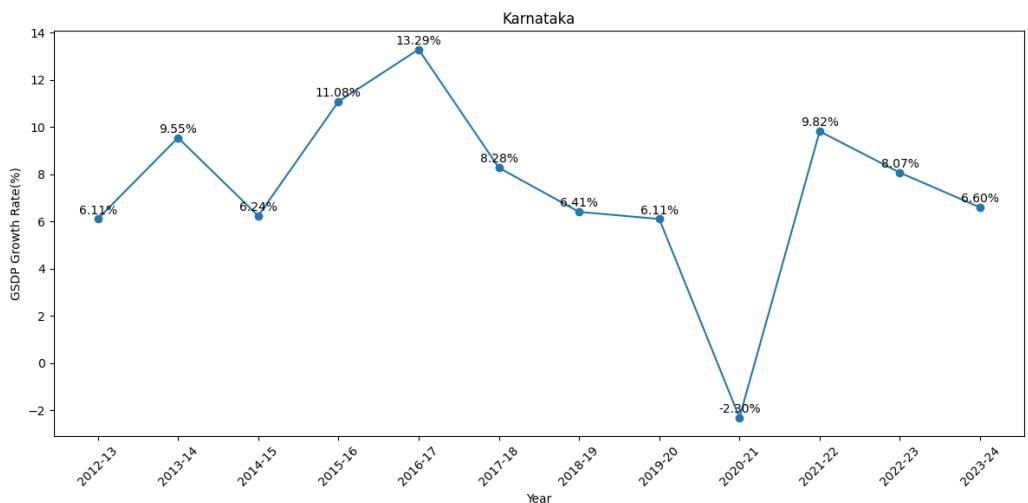


Figure 4.203: Trend of Growth Rate of GSDP of Karnataka

4.3.2.1.12 Kerala

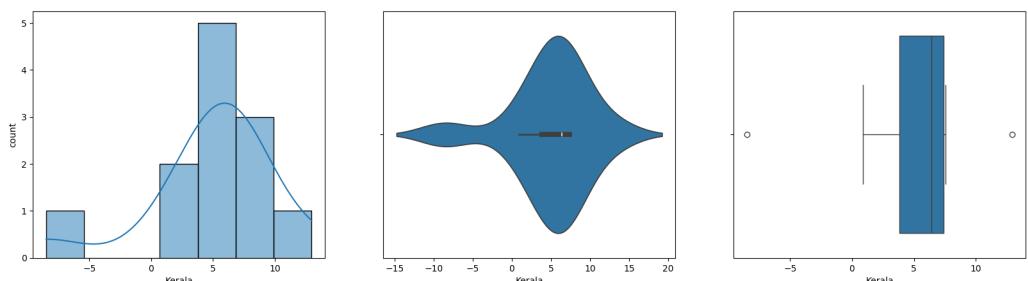


Figure 4.204: Data Distribution of Growth Rate of GSDP of Kerala

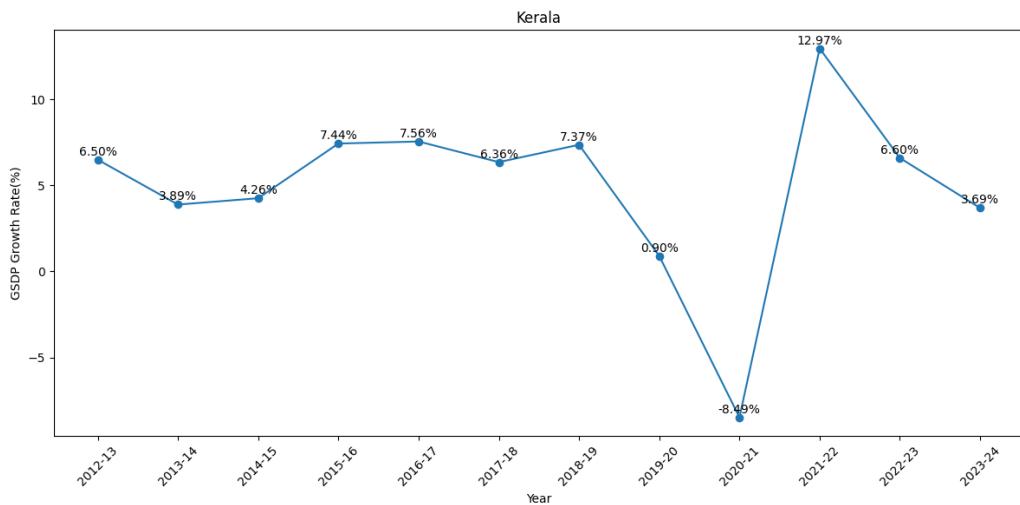


Figure 4.205: Trend of Growth Rate of GSDP of Kerala

4.3.2.1.13 Madhya Pradesh

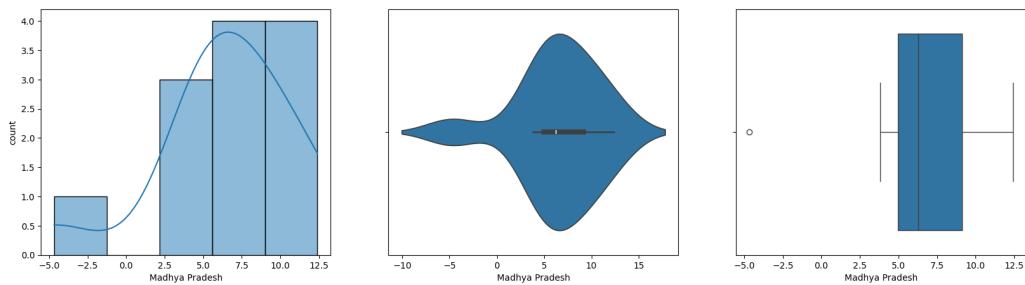


Figure 4.206: Data Distribution of Growth Rate of GSDP of Madhya Pradesh

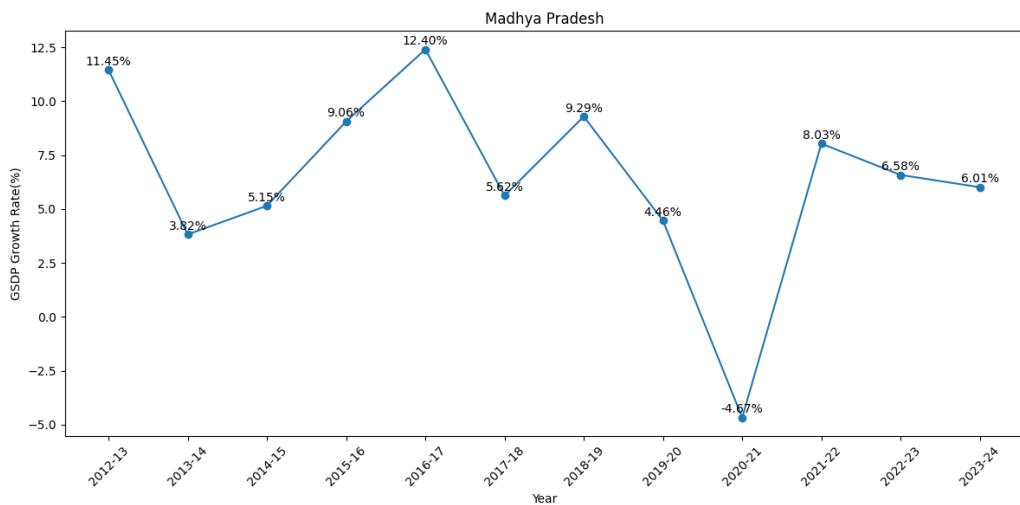


Figure 4.207: Trend of Growth Rate of GSDP of Madhya Pradesh

4.3.2.1.14 Maharashtra

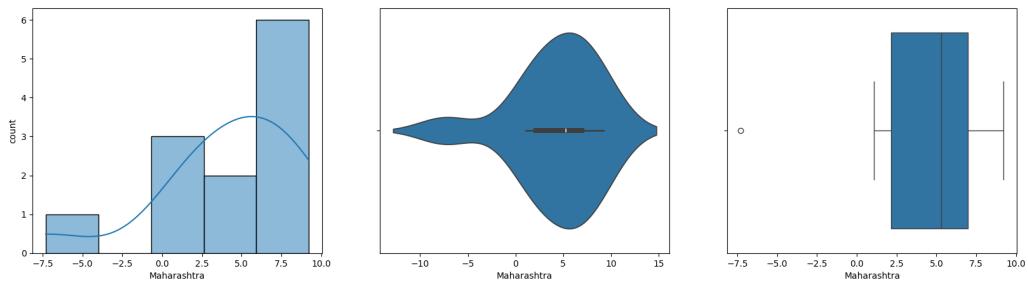


Figure 4.208: Data Distribution of Growth Rate of GSDP of Maharashtra

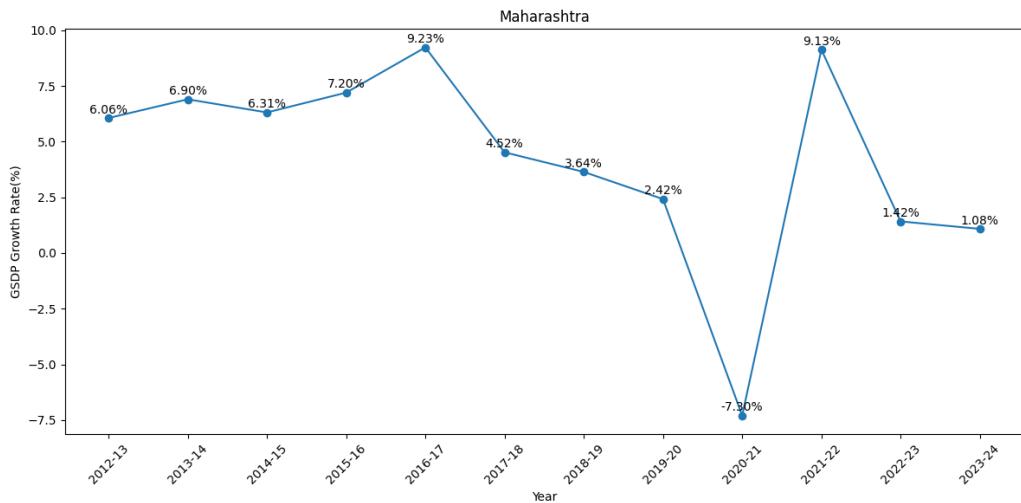


Figure 4.209: Trend of Growth Rate of GSDP of Maharashtra

4.3.2.1.15 Manipur

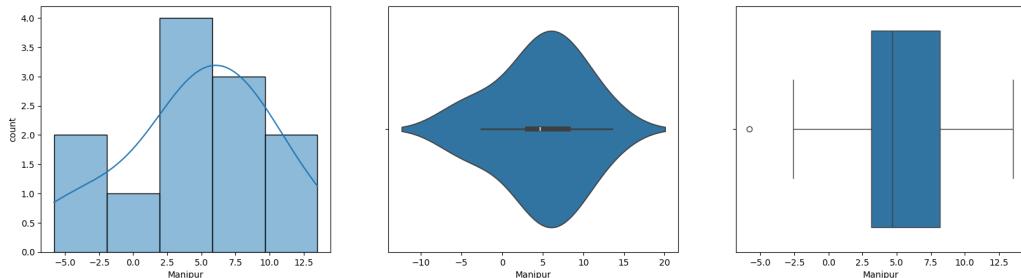


Figure 4.210: Data Distribution of Growth Rate of GSDP of Manipur

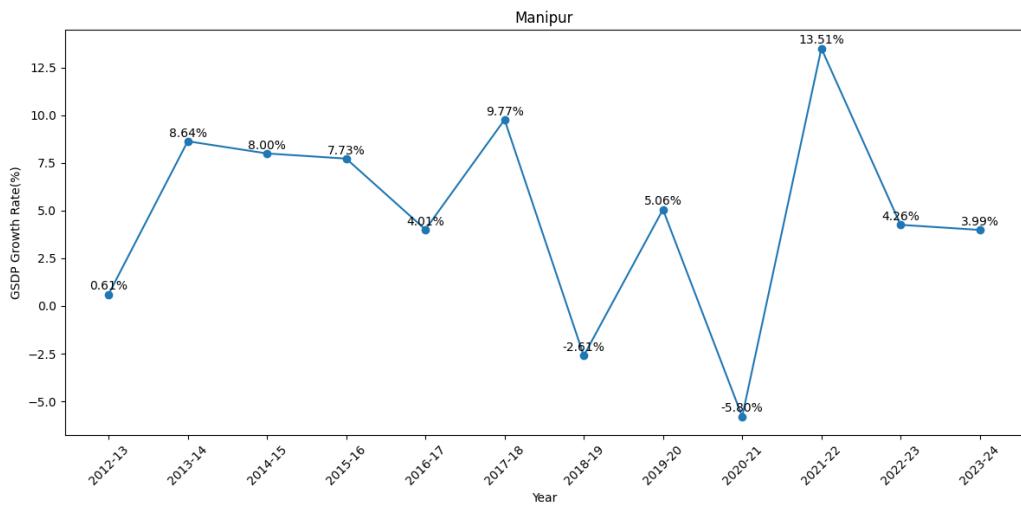


Figure 4.211: Trend of Growth Rate of GSDP of Manipur

4.3.2.1.16 Meghalaya

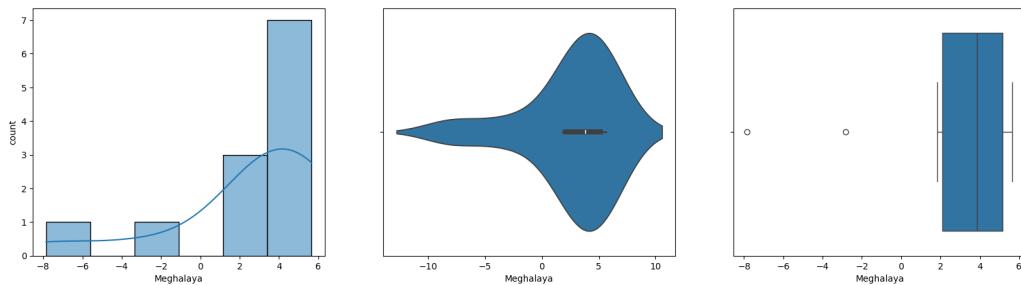


Figure 4.212: Data Distribution of Growth Rate of GSDP of Meghalaya

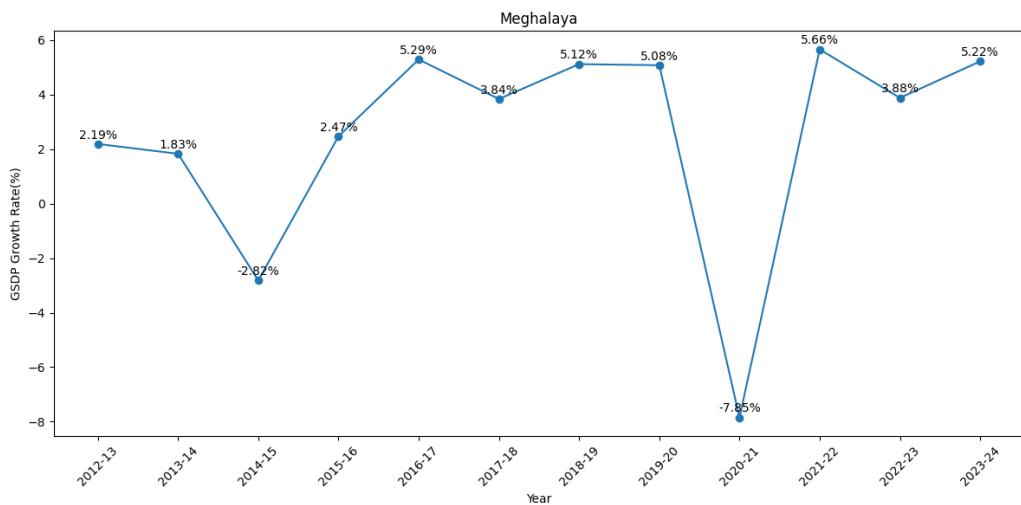


Figure 4.213: Trend of Growth Rate of GSDP of Meghalaya

4.3.2.1.17 Mizoram

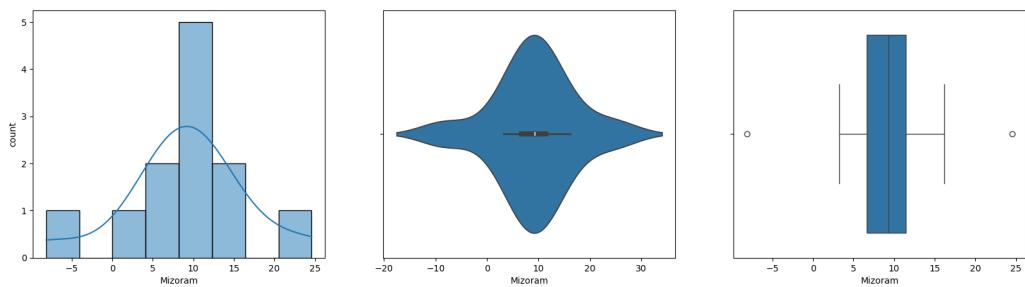


Figure 4.214: Data Distribution of Growth Rate of GSDP of Mizoram

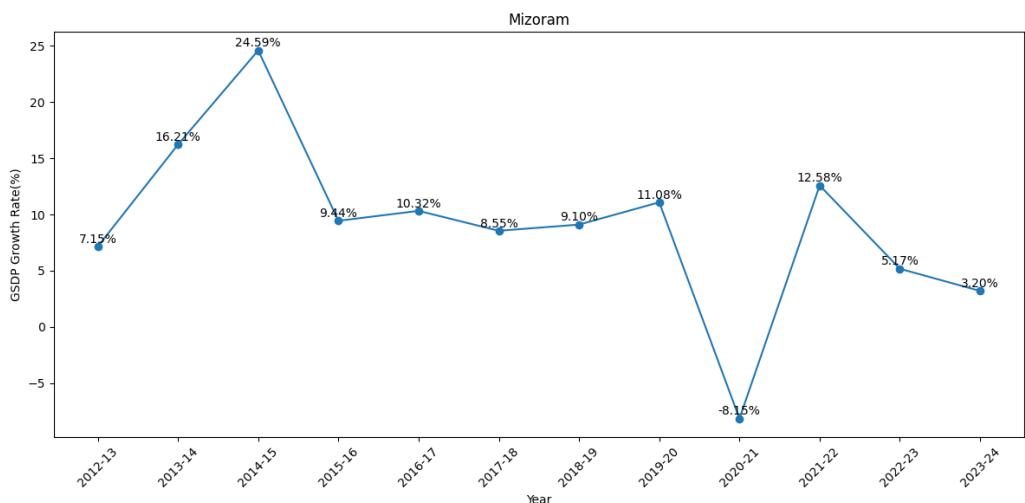


Figure 4.215: Trend of Growth Rate of GSDP of Mizoram

4.3.2.1.18 Nagaland

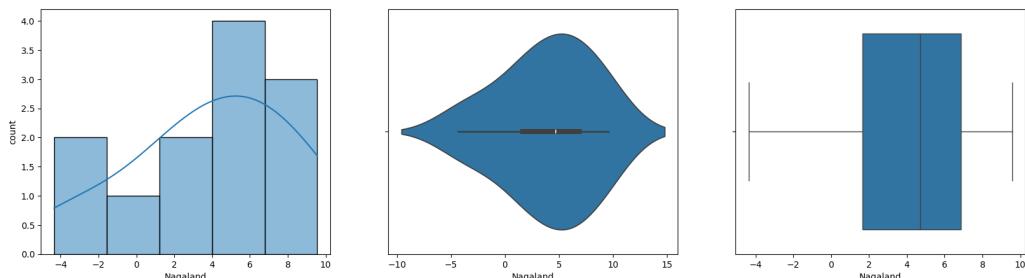


Figure 4.216: Data Distribution of Growth Rate of GSDP of Nagaland

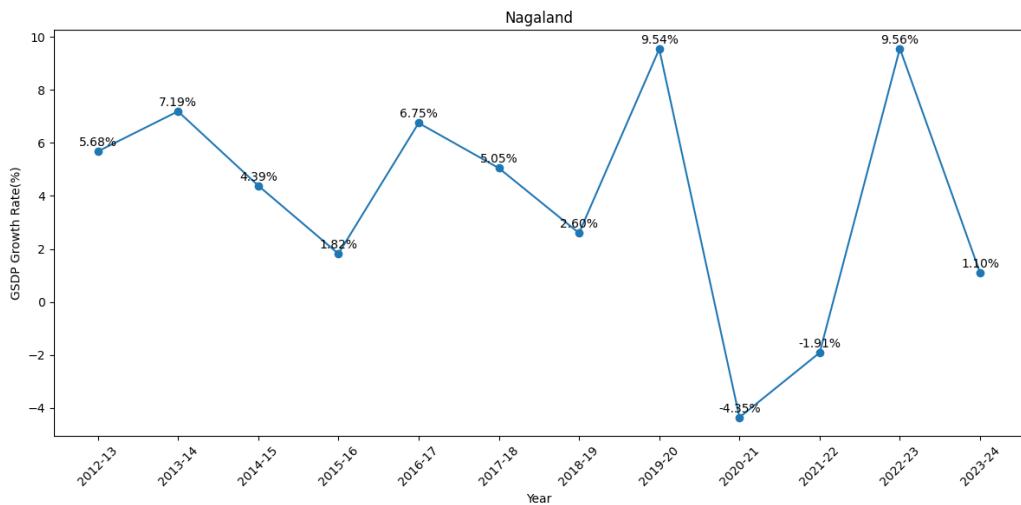


Figure 4.217: Trend of Growth Rate of GSDP of Nagaland

4.3.2.1.19 Odisha

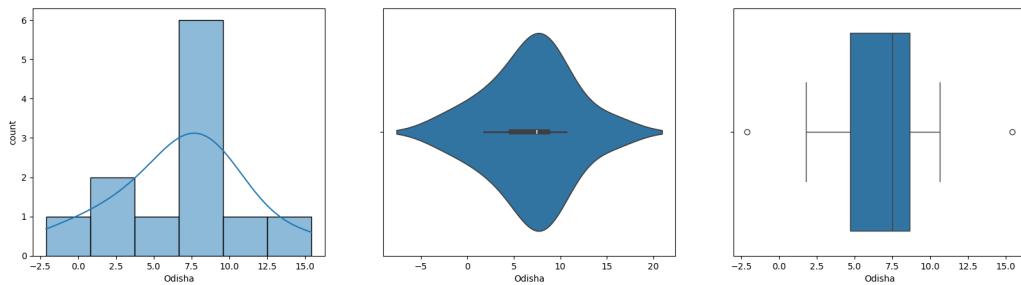


Figure 4.218: Data Distribution of Growth Rate of GSDP of Odisha

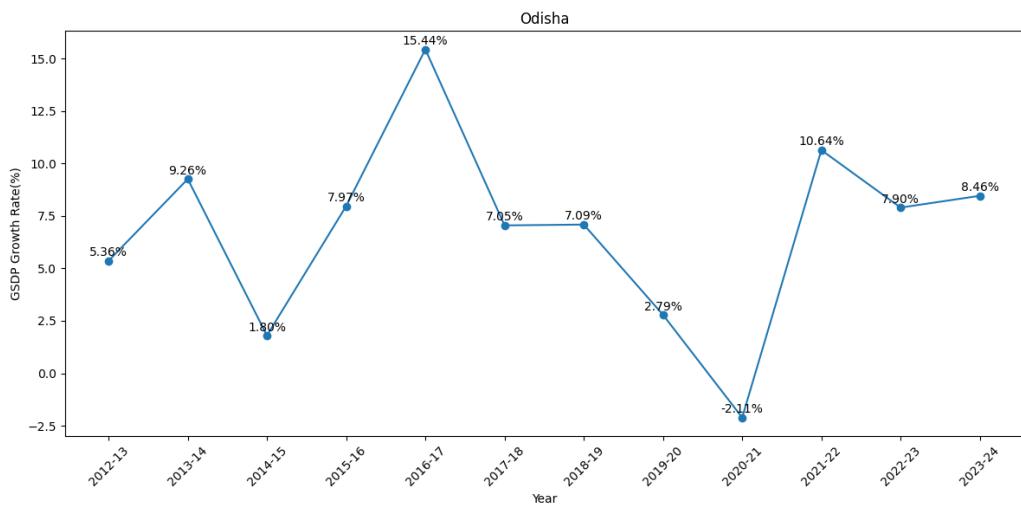


Figure 4.219: Trend of Growth Rate of GSDP of Odisha

4.3.2.1.20 Punjab

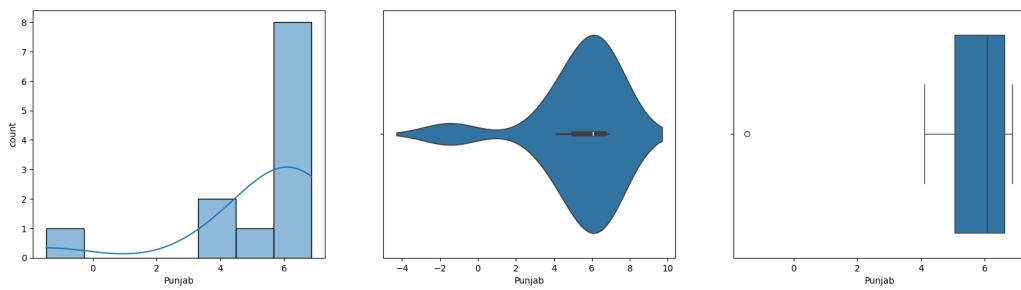


Figure 4.220: Data Distribution of Growth Rate of GSDP of Punjab

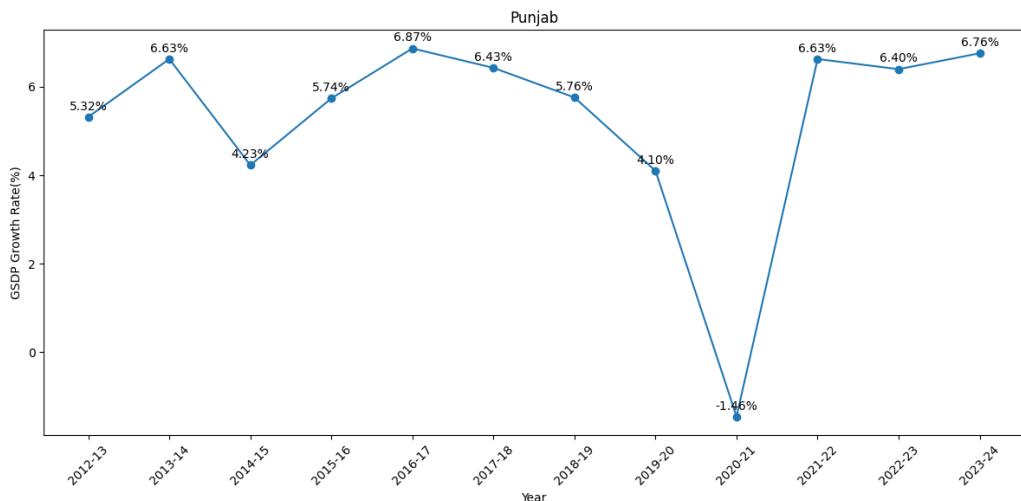


Figure 4.221: Trend of Growth Rate of GSDP of Punjab

4.3.2.1.21 Rajasthan

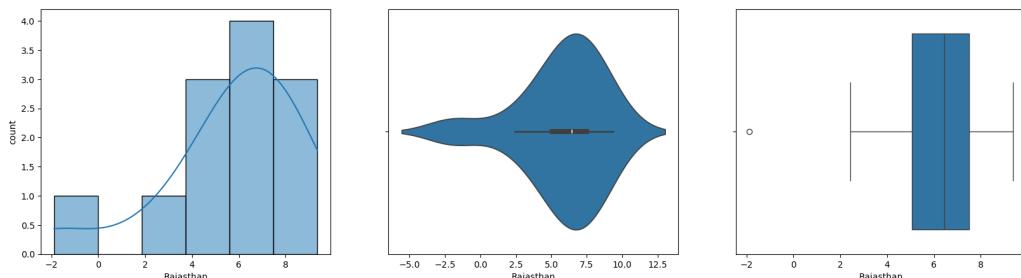


Figure 4.222: Data Distribution of Growth Rate of GSDP of Rajasthan

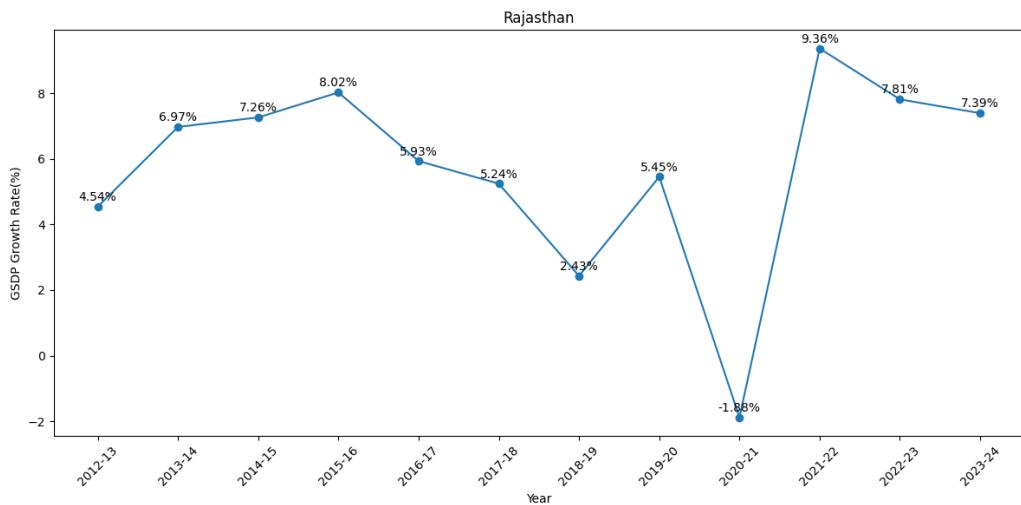


Figure 4.223: Trend of Growth Rate of GSDP of Rajasthan

4.3.2.1.22 Sikkim

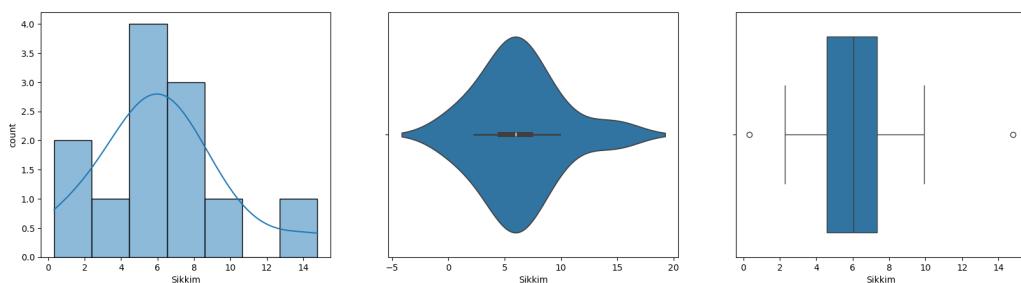


Figure 4.224: Data Distribution of Growth Rate of GSDP of Sikkim

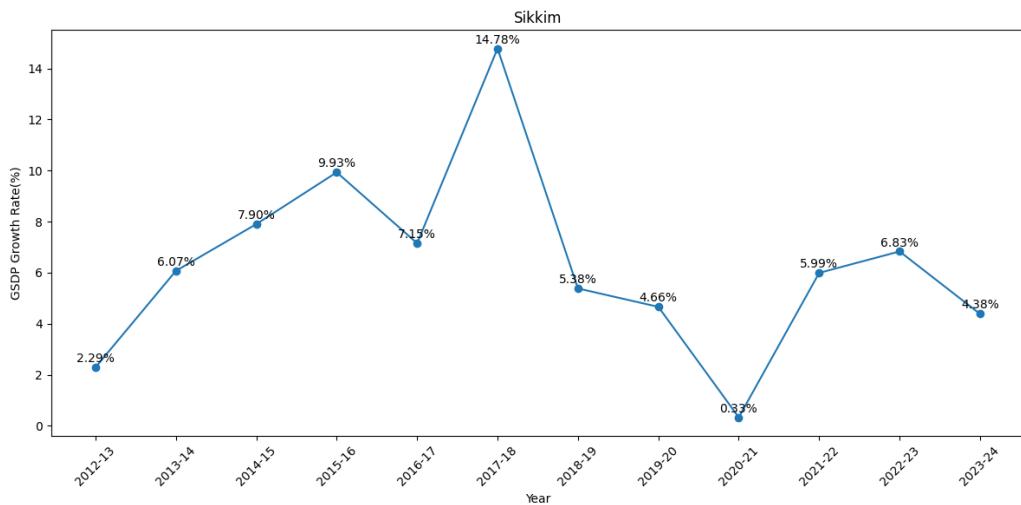


Figure 4.225: Trend of Growth Rate of GSDP of Sikkim

4.3.2.1.23 Tamil Nadu

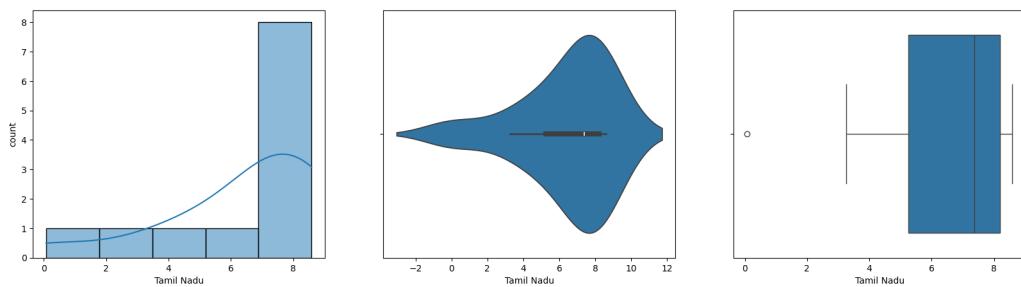


Figure 4.226: Data Distribution of Growth Rate of GSDP of Tamil Nadu

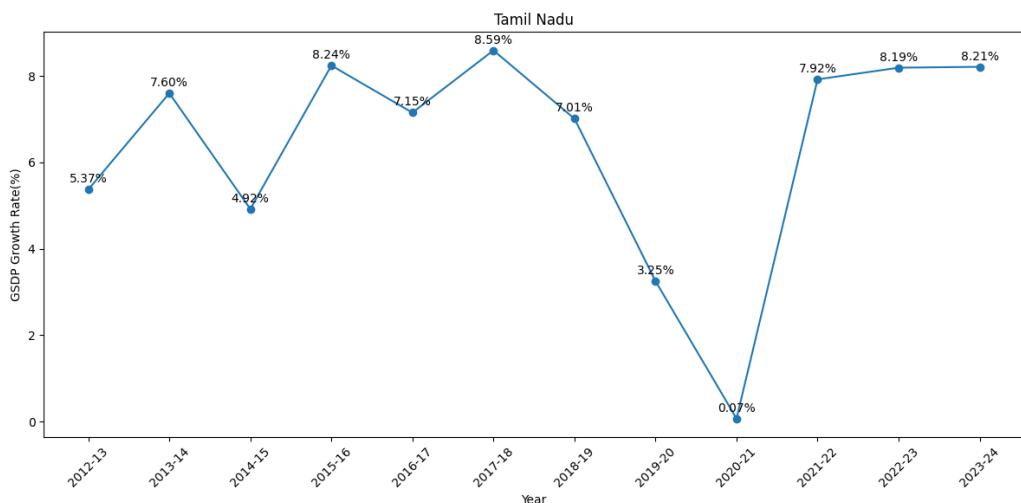


Figure 4.227: Trend of Growth Rate of GSDP of Tamil Nadu

4.3.2.1.24 Telangana

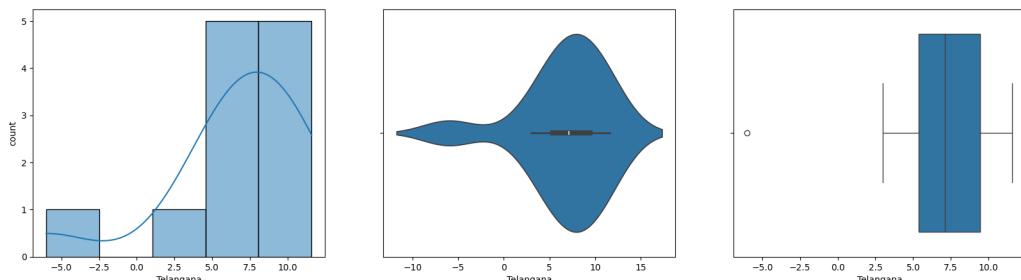


Figure 4.228: Data Distribution of Growth Rate of GSDP of Telangana

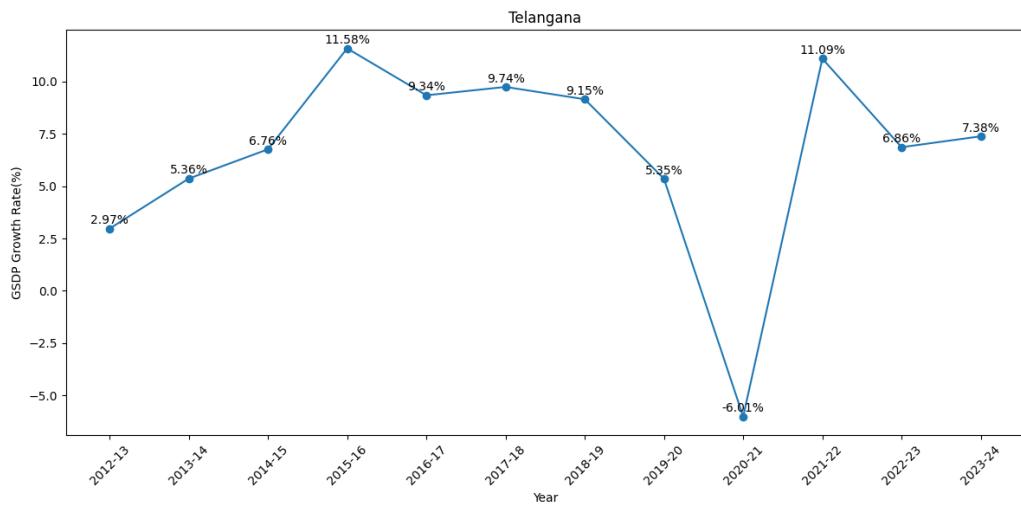


Figure 4.229: Trend of Growth Rate of GSDP of Telangana

4.3.2.1.25 Tripura

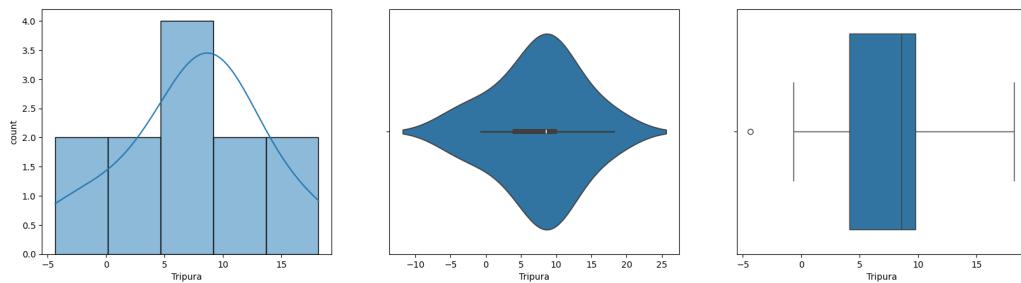


Figure 4.230: Data Distribution of Growth Rate of GSDP of Tripura

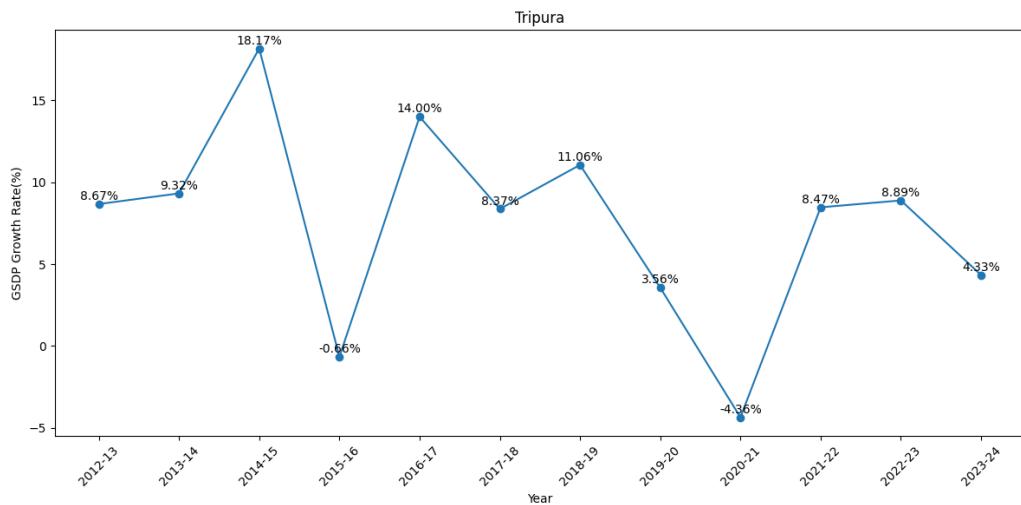


Figure 4.231: Trend of Growth Rate of GSDP of Tripura

4.3.2.1.26 Uttar Pradesh

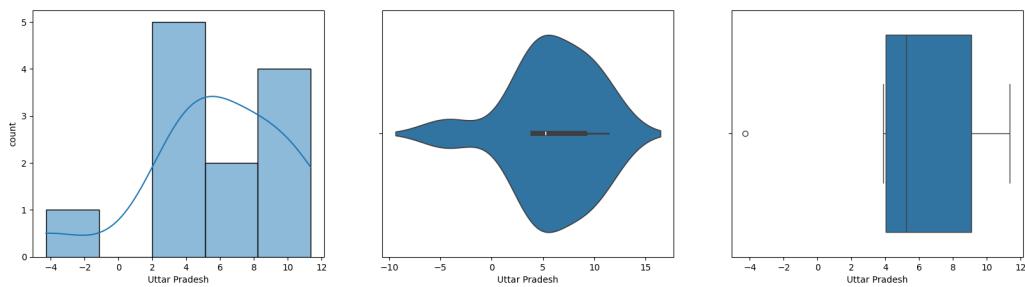


Figure 4.232: Data Distribution of Growth Rate of GSDP of Uttar Pradesh

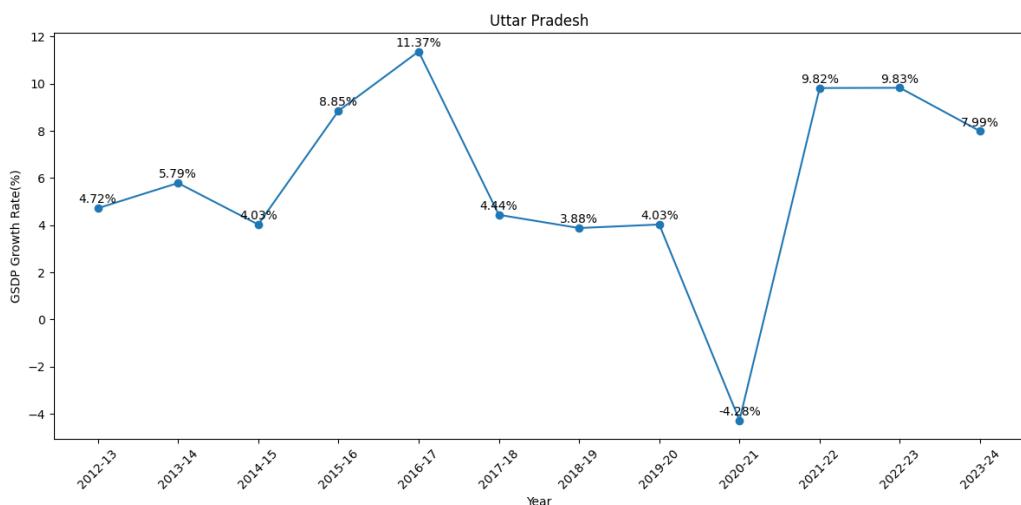


Figure 4.233: Trend of Growth Rate of GSDP of Uttar Pradesh

4.3.2.1.27 Uttarakhand

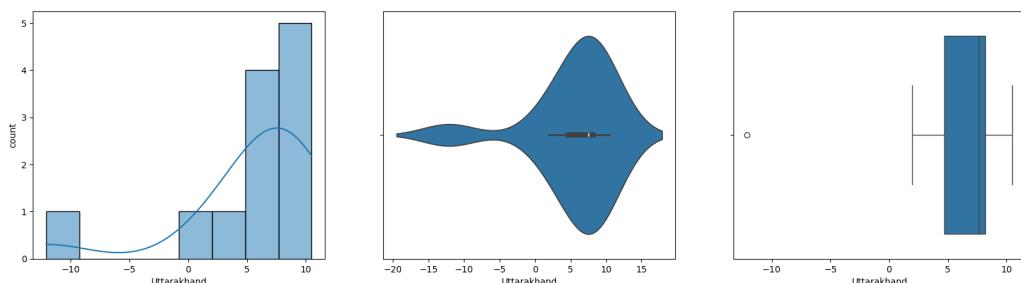


Figure 4.234: Data Distribution of Growth Rate of GSDP of Uttarakhand

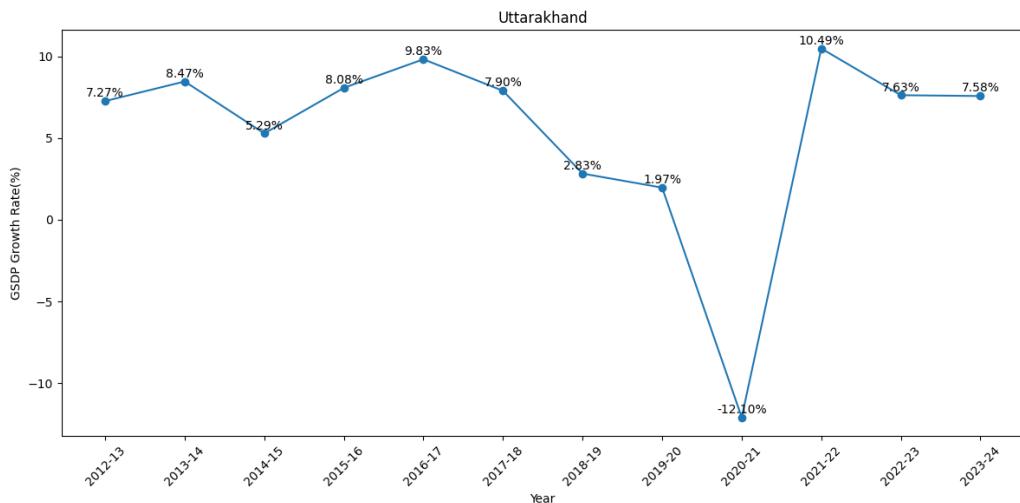


Figure 4.235: Trend of Growth Rate of GSDP of Uttarakhand

4.3.2.1.28 West Bengal

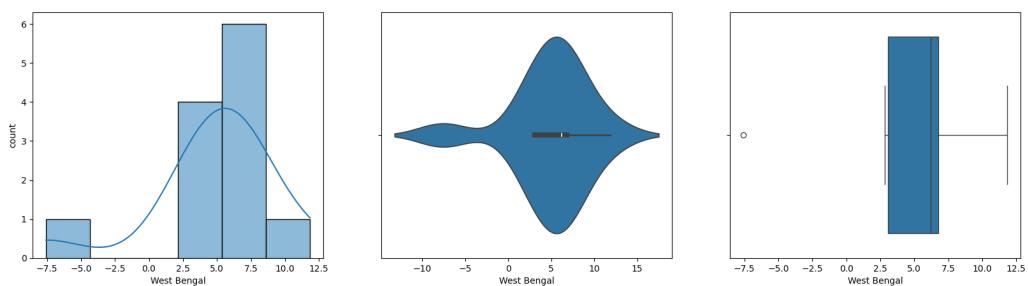


Figure 4.236: Data Distribution of Growth Rate of CSDP of West Bengal

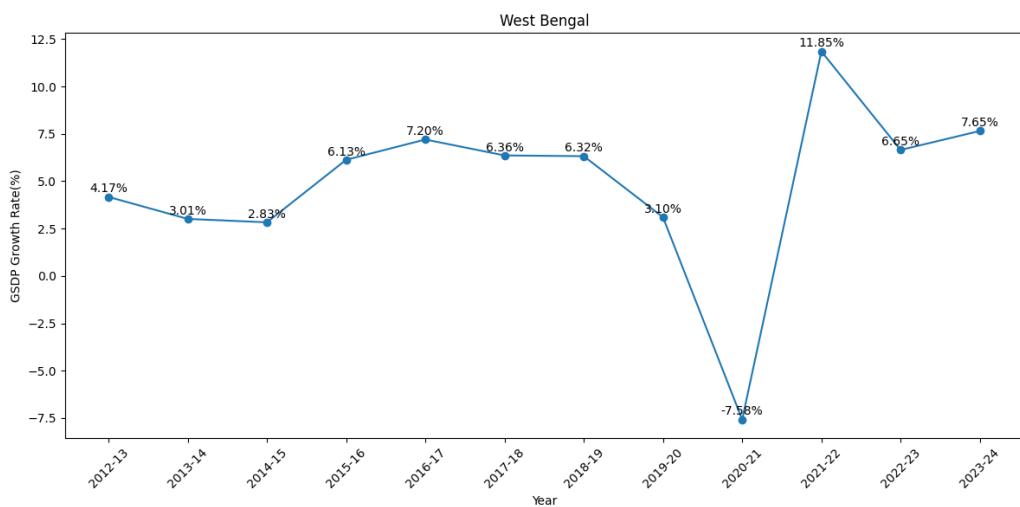


Figure 4.237: Trend of Growth Rate of GSDP of West Bengal

4.3.2.1.29 Andaman & Nicobar Islands

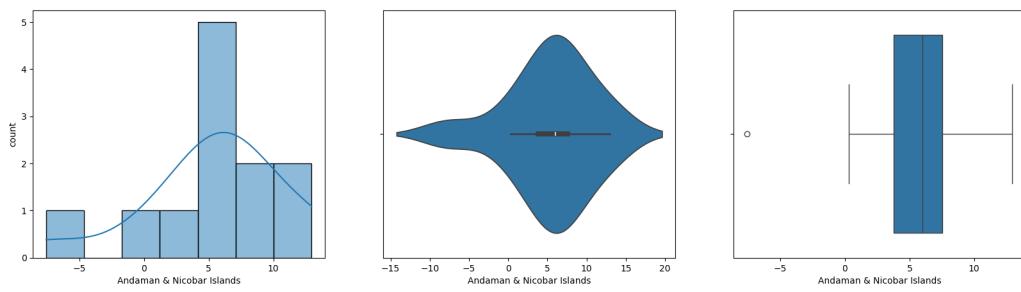


Figure 4.238: Data Distribution of Growth Rate of GSDP of Andaman & Nicobar Islands

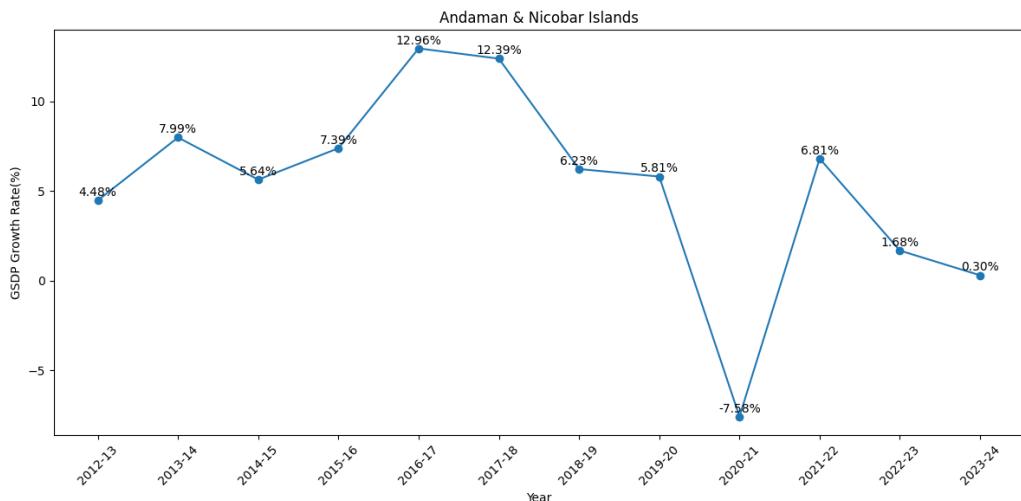


Figure 4.239: Trend of Growth Rate of GSDP of Andaman & Nicobar Islands

4.3.2.1.30 Chandigarh

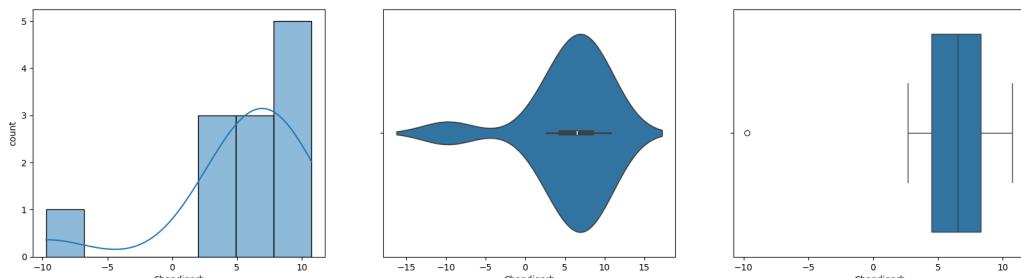


Figure 4.240: Data Distribution of Growth Rate of GSDP of Chandigarh

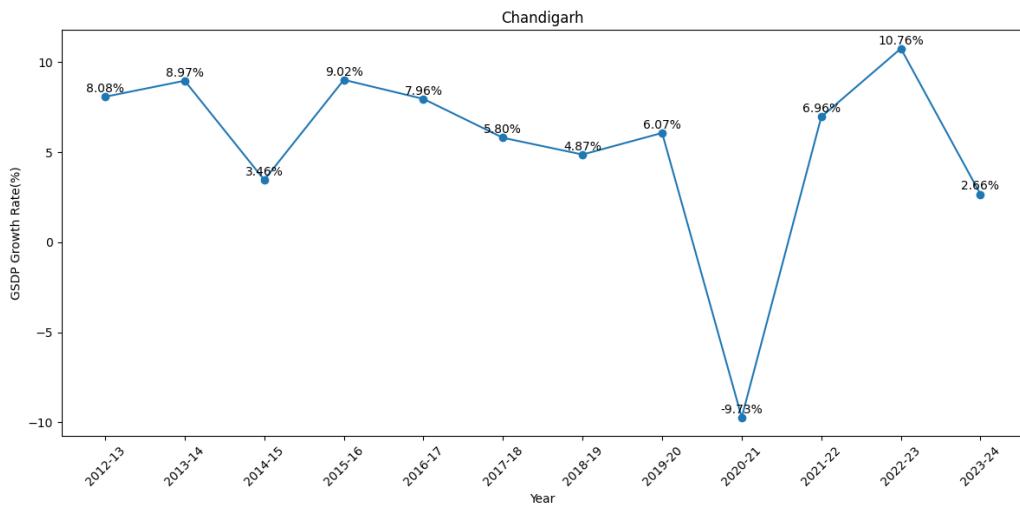


Figure 4.241: Trend of Growth Rate of GSDP of Chandigarh

4.3.2.1.31 Delhi

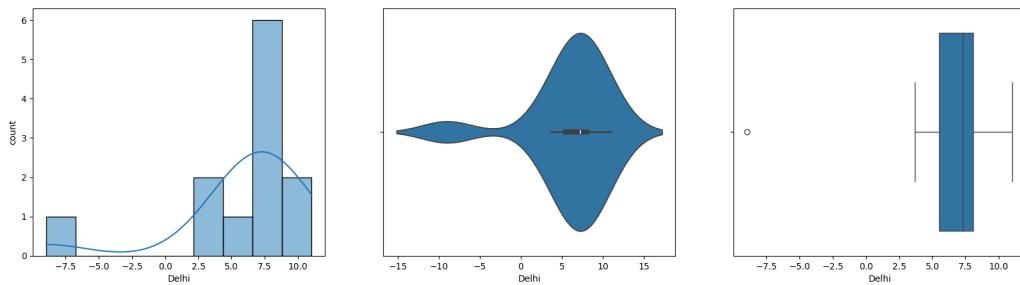


Figure 4.242: Data Distribution of Growth Rate of GSDP of Delhi

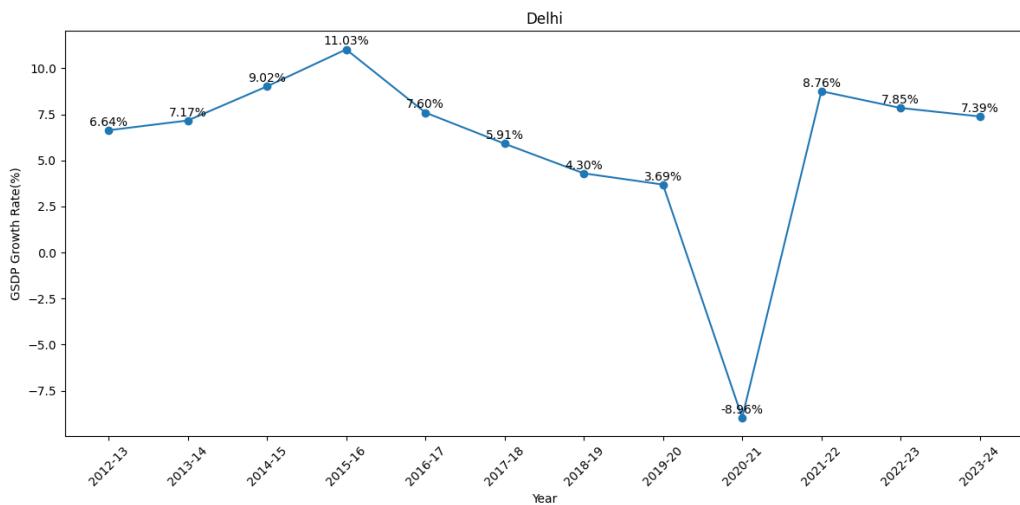


Figure 4.243: Trend of Growth Rate of GSDP of Delhi

4.3.2.1.32 Jammu & Kashmir-UT*

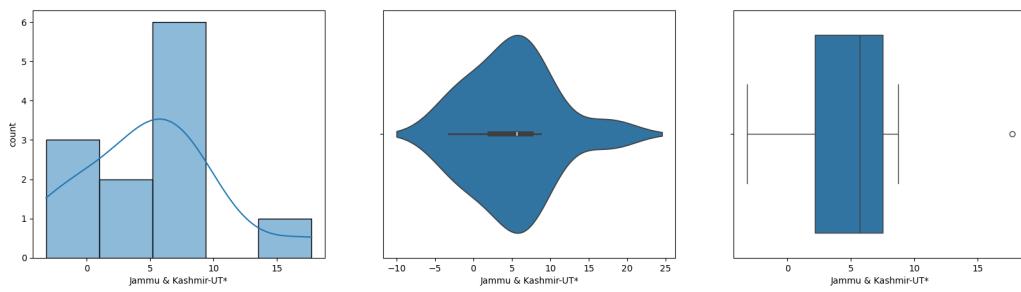


Figure 4.244: Data Distribution of Growth Rate of GSDP of Jammu & Kashmir-UT*

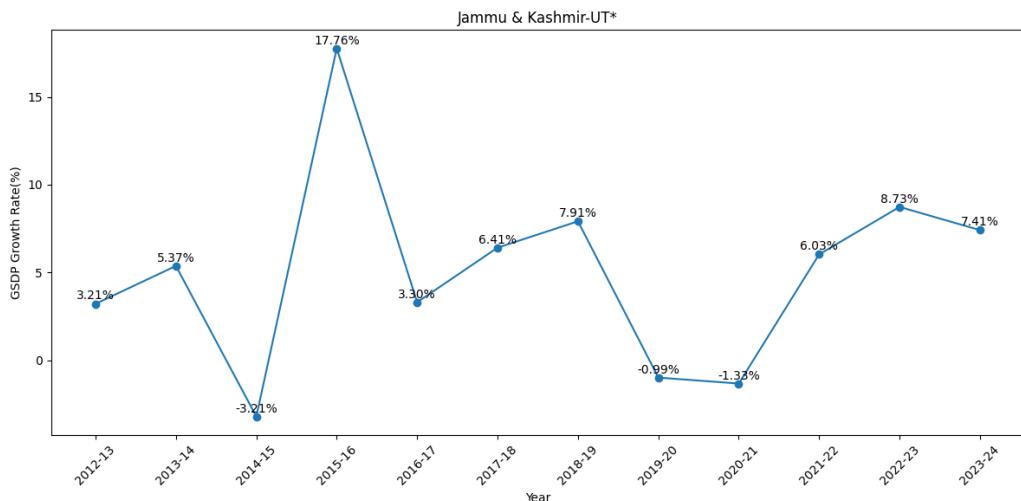


Figure 4.245: Trend of Growth Rate of GSDP of Jammu & Kashmir-UT*

4.3.2.1.33 Puducherry

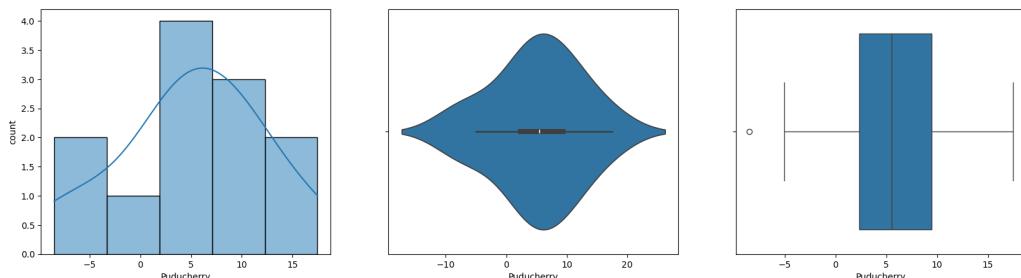


Figure 4.246: Data Distribution of Growth Rate of GSDP of Puducherry

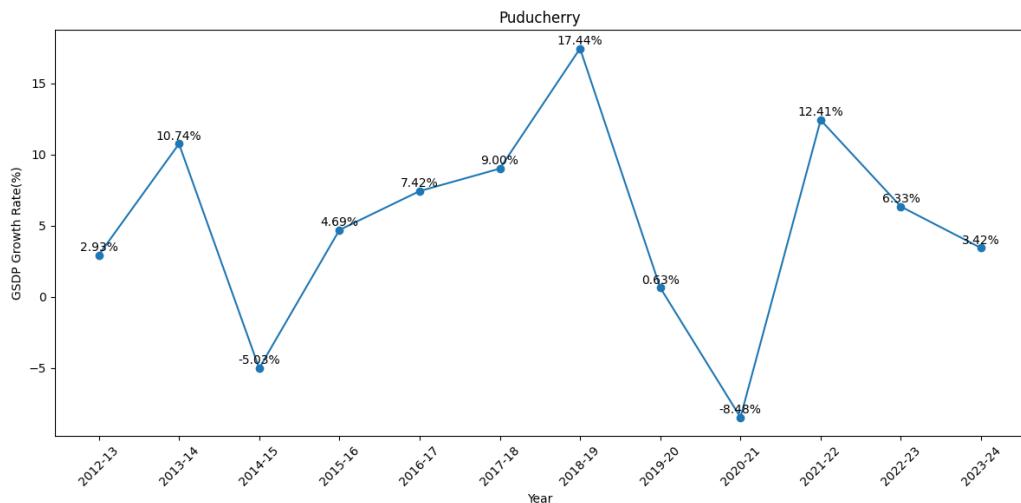


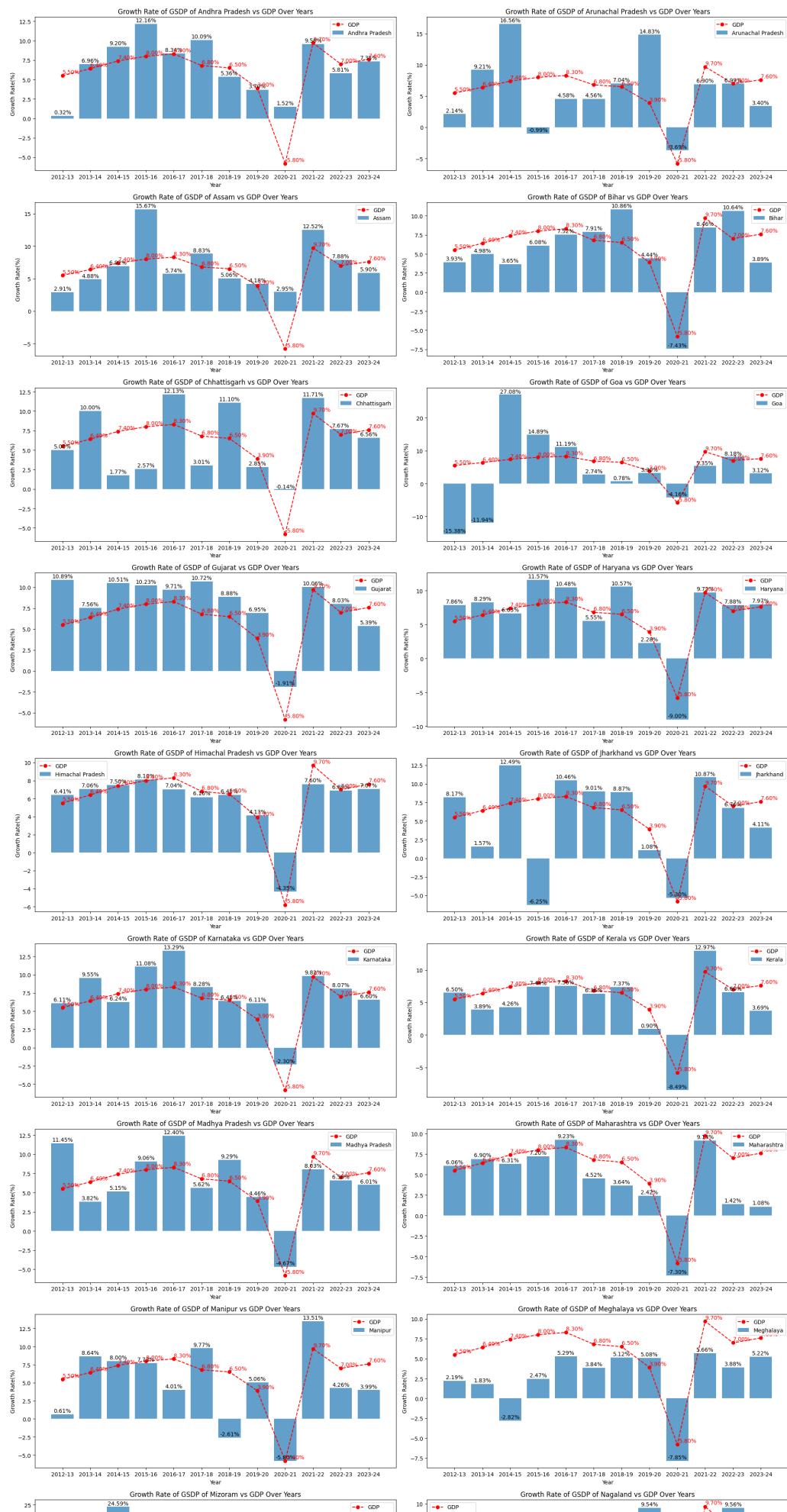
Figure 4.247: Trend of Growth Rate of GSDP of Puducherry

4.3.2.2 Bivariate Analysis

Here, we will analyze the relation between the growth rate of GSDP of two States of India.

4.3.2.2.1 Comparison with GDP

The comparison of the growth rate of GSDP of each State with respect to GDP growth is given below.





4.3.2.3 Multivariate analysis

Here, we will analyze the growth rate of GSDP of multiple States together.

4.3.2.3.1 Trend of Growth of GSDP of States of India

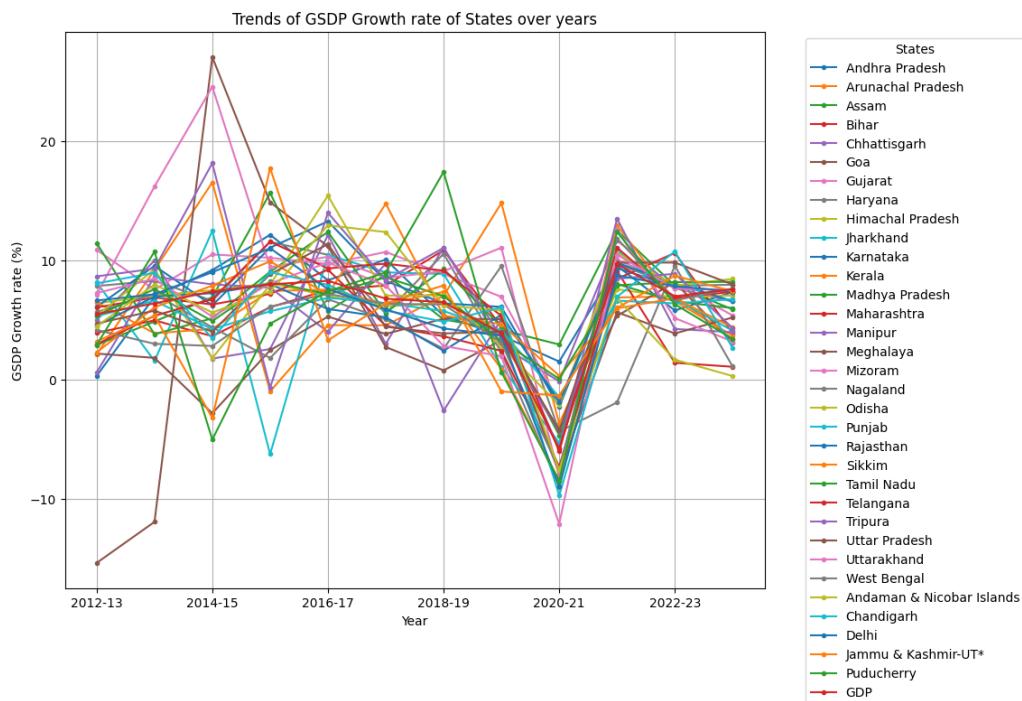


Figure 4.249: Growth of GSDP of Staes of India Over The Years

4.3.2.3.2 GSDP Growth Rate in 2023

Given below is a sorted representation of the 2023's GSDP Growth Rate of States of India.

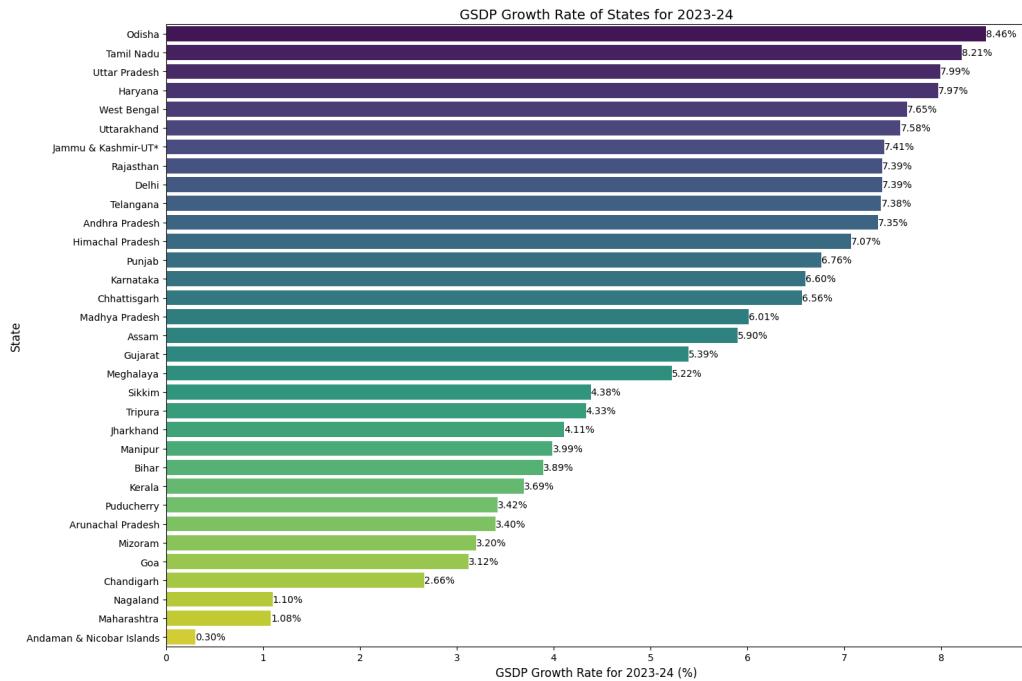


Figure 4.250: GSDP Growth Rate of States of India in 2023

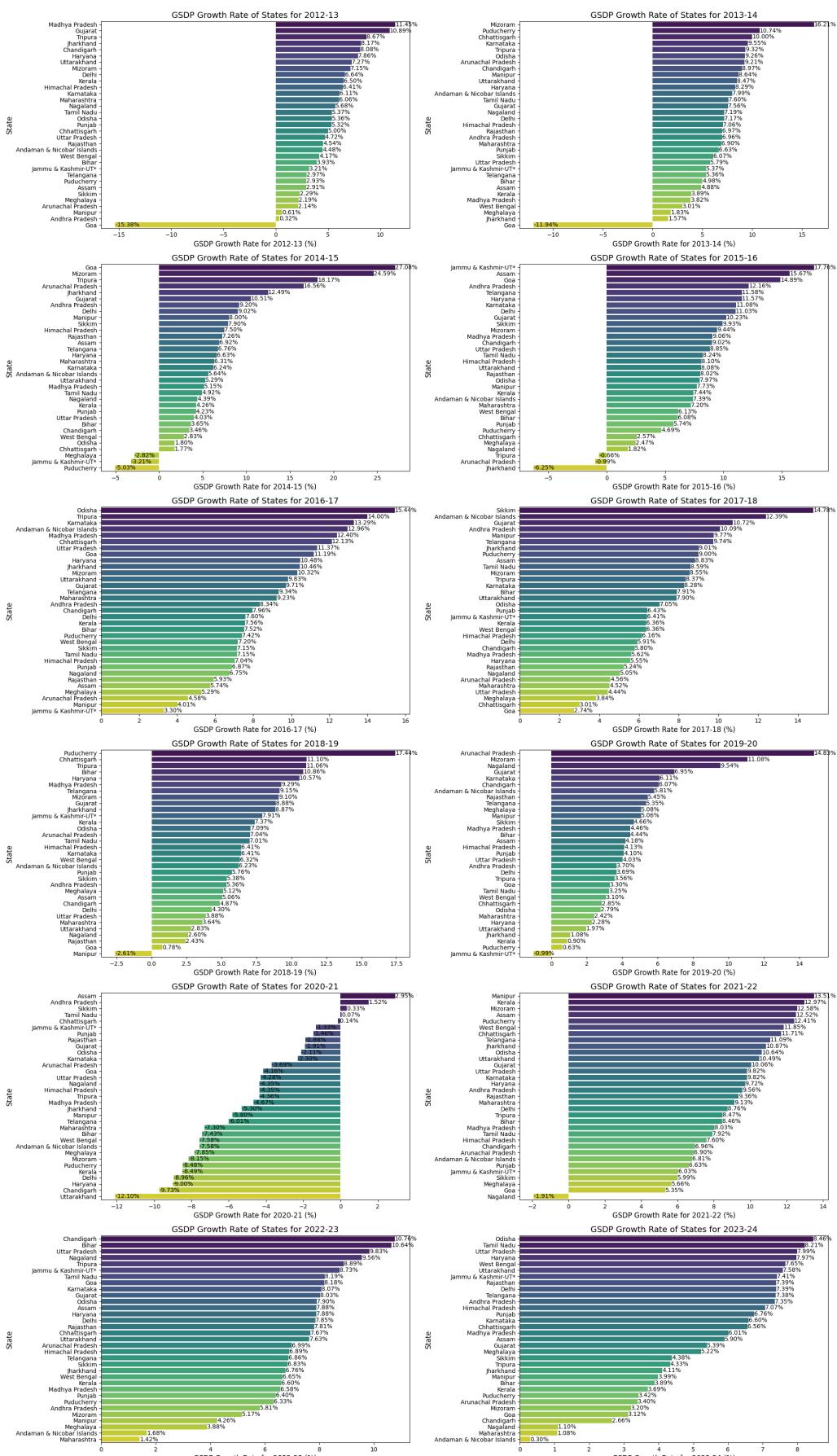
4.3.2.3.3 GSDP Growth Rate Over the Decade

Given below is a sorted representation of the Growth Rate of GSDP of States of India Over the Decade.



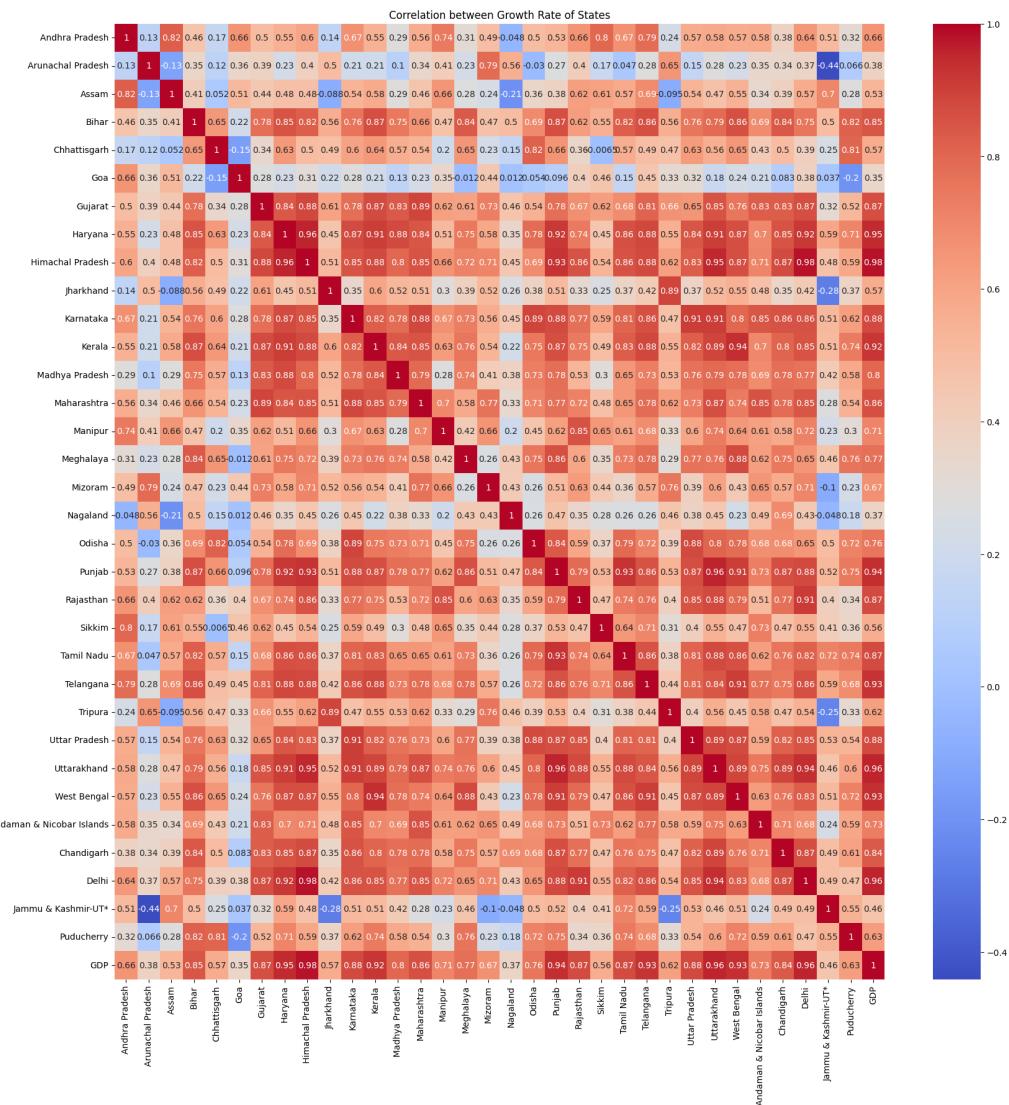
CHAPTER 4. VISUALIZATION

IT 462 EDA





4.3.2.3.4 Correlation of Growth Rate of GSDP of States



Chapter 5. Model Fitting

5.1 Contributions of Sectors

5.1.1 GVA

5.1.1.1 Linear Regression

5.1.1.1.1 Future Prediction

The plot given below has used linear regression (LR) to predict values for the contribution of various sectors till 2028.

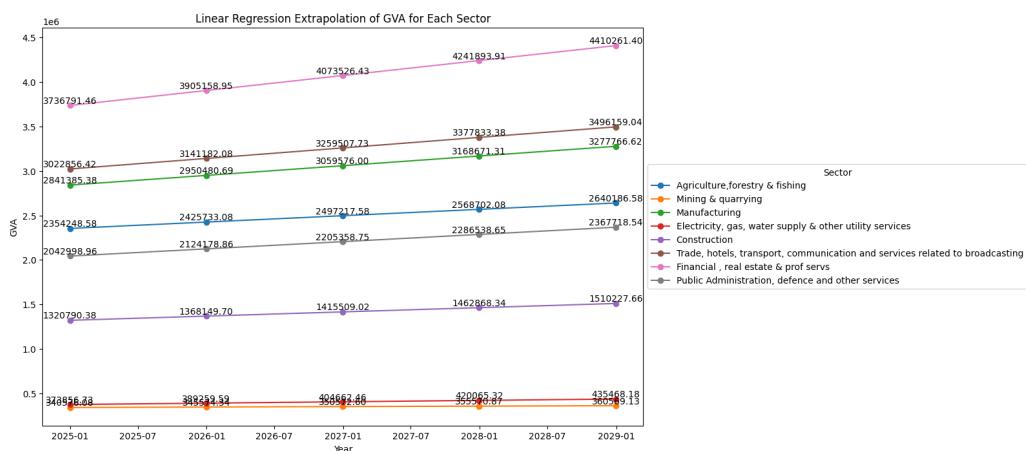


Figure 5.1: Predicted Contribution of Sectors till 2028

5.1.1.1.2 Accuracy Calculation

We have also predicted the contribution of the sectors from 2011 to 2023 using the same regression model and compared them with the actual data to calculate the accuracy of the model.

Residuals are the difference between the actual value and the predicted value.

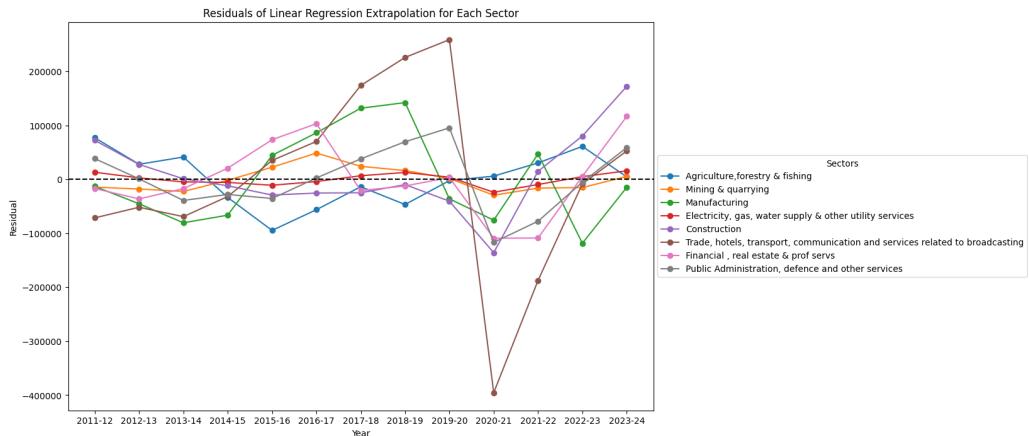


Figure 5.2: Residual of LR model for Contribution of Sectors in Terms of Absolute Value

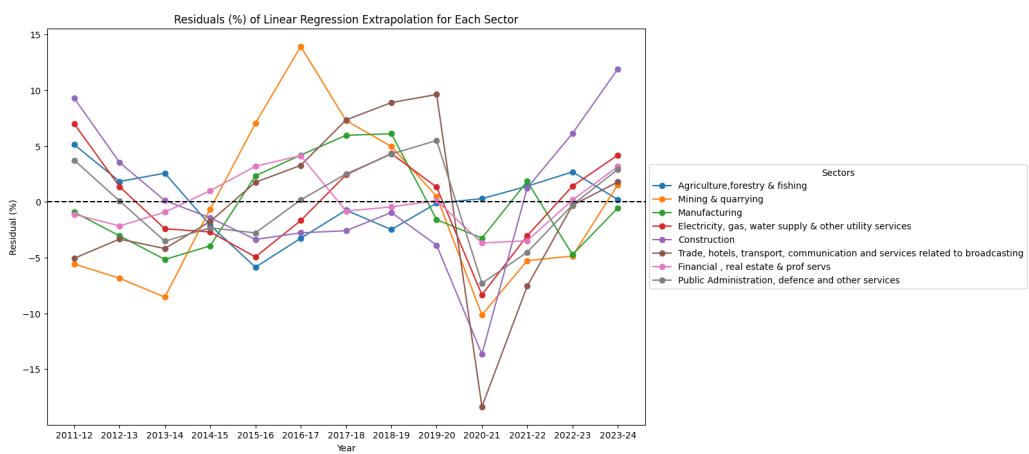


Figure 5.3: Residual of LR model for Contribution of Sectors in Terms of Percent

RMSE stands for 'Root Mean Square Error' and is calculated by the following formula

$$\sqrt{\frac{\sum(P_i - O_i)^2}{n}}$$

where P_i denotes the predicted value, O_i represents the observed value, and n is the total number of observations or data points [5].

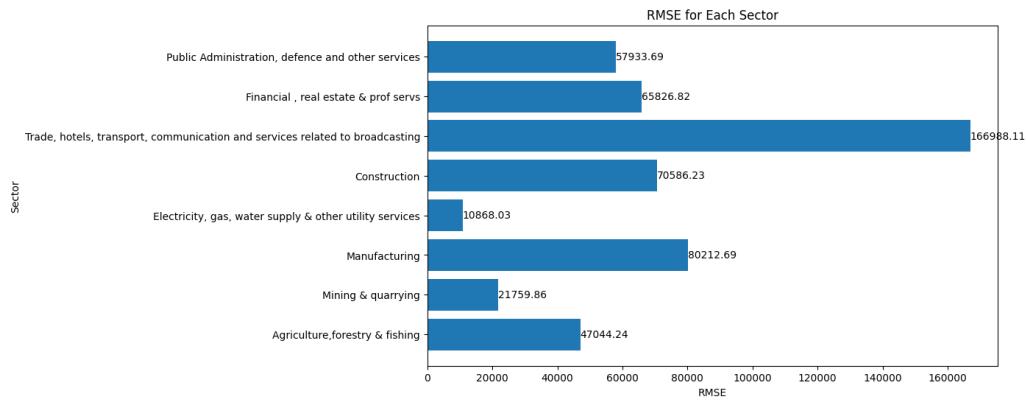


Figure 5.4: RMSE of LR model for Contribution of Sectors in Terms of Percent

R^2 is the regression coefficient and denotes the how efficient the model fitting is. Its value lies from 0 to 1. The closer the value to 1 the better the fit.

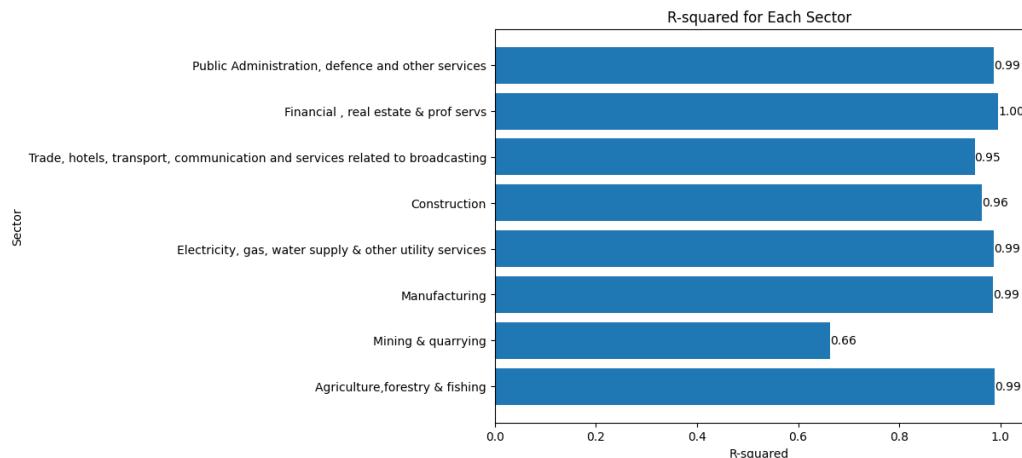


Figure 5.5: R^2 of LR model for Contribution of Sectors in Terms of Percent

5.1.1.2 Polynomial Extrapolation

5.1.1.2.1 Future Prediction

The plot given below has used polynomial extrapolation (PE) to the predict values for the contribution of various sectors till 2028.

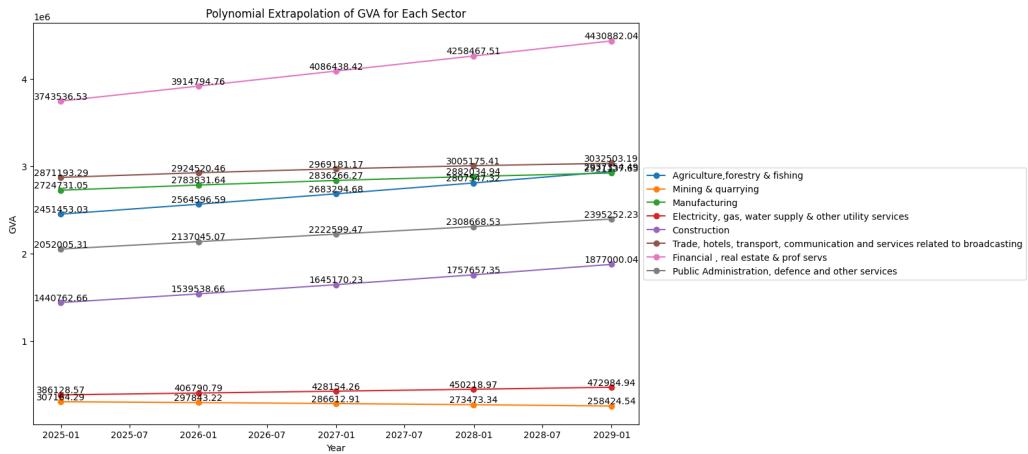


Figure 5.6: Predicted Contribution of Sectors till 2028

5.1.1.2.2 Accuracy Calculation

We have also predicted the contribution of the sectors from 2011 to 2023 using the same regression model and compared them with the actual data to calculate the accuracy of the model.

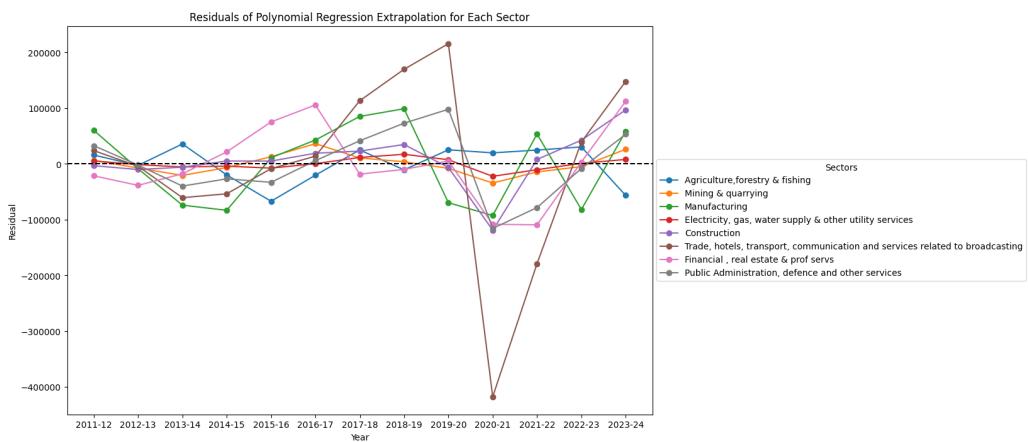


Figure 5.7: Residual of PE model for Contribution of Sectors in Terms of Absolute Value

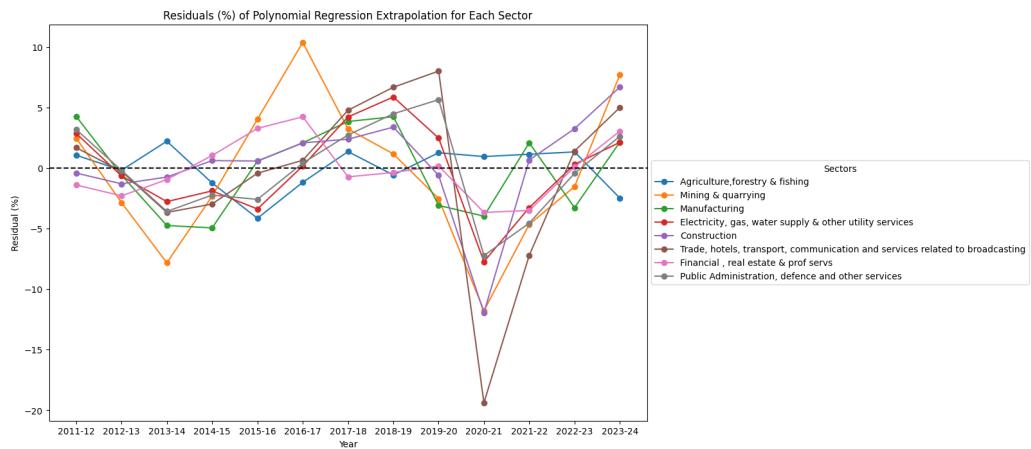


Figure 5.8: Residual of PE model for Contribution of Sectors in Terms of Percent

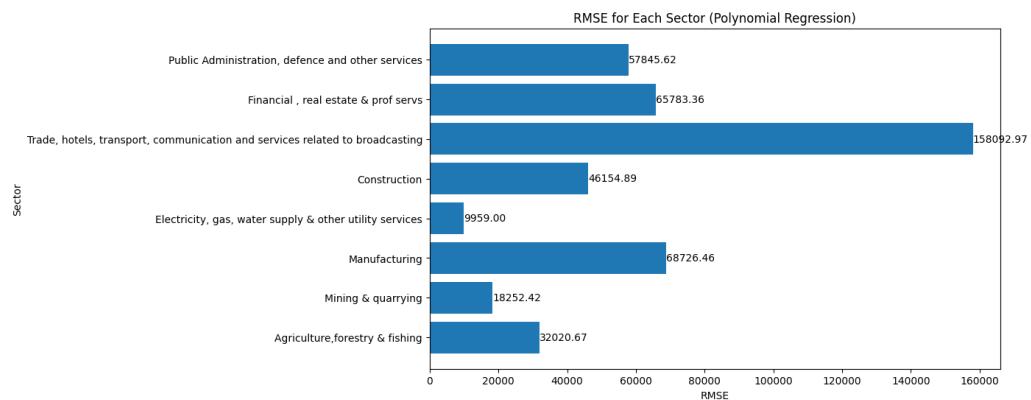


Figure 5.9: RMSE of PE model for Contribution of Sectors in Terms of Percent

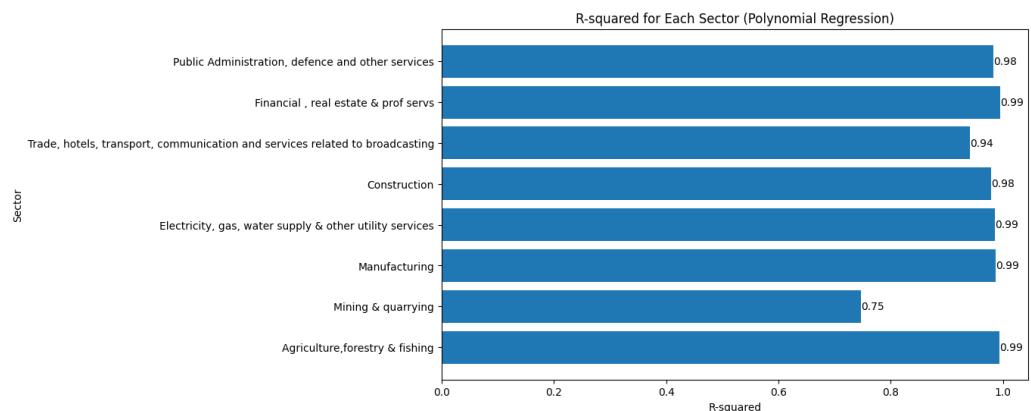


Figure 5.10: R^2 of PE model for Contribution of Sectors in Terms of Percent

5.1.1.3 Comparison of Models

Here we have compared the accuracy of both the models.



5.1.1.3.1 Value Comparison

The plot given below compares the predicted values obtained through linear regression and polynomial extrapolation.

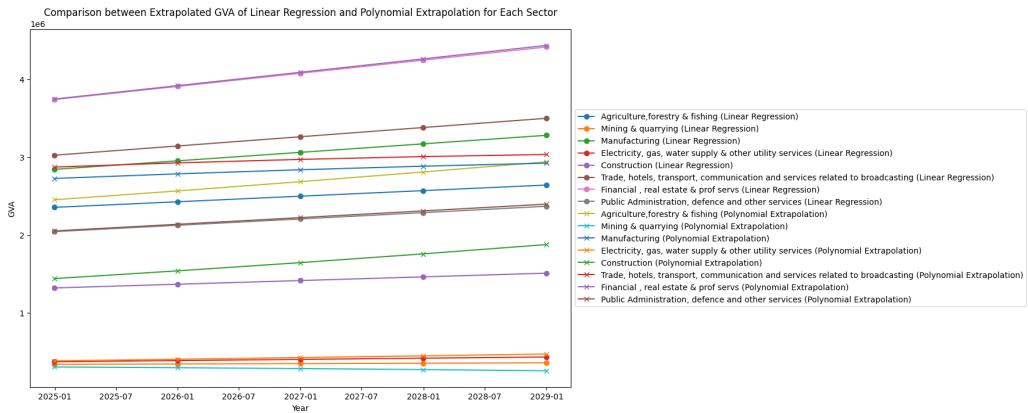


Figure 5.11: Comparison of Predicted Contribution of Sectors till 2028

5.1.1.3.2 Residual Comparison

The plot given below compares the residuals obtained through linear regression and polynomial extrapolation.

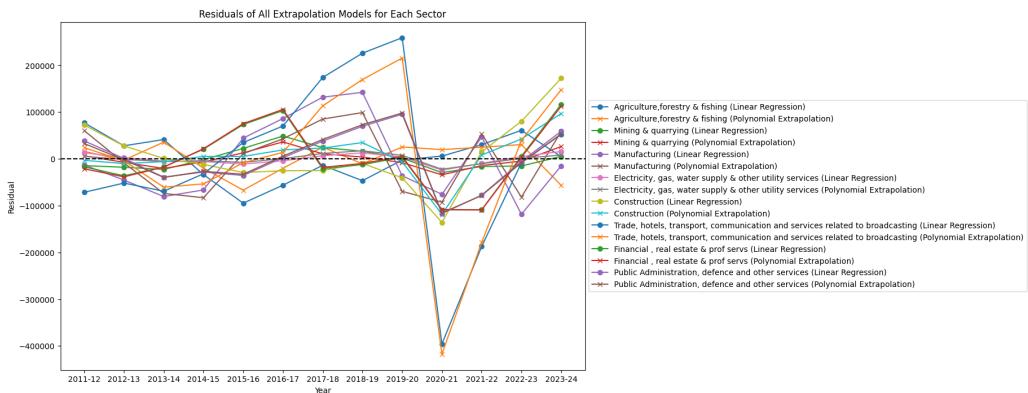


Figure 5.12: Comparison of Residuals for Contribution of Sectors in Terms of Absolute Value

5.1.1.3.3 RMSE Comparison

The plot given below compares the RMSE of both the models.

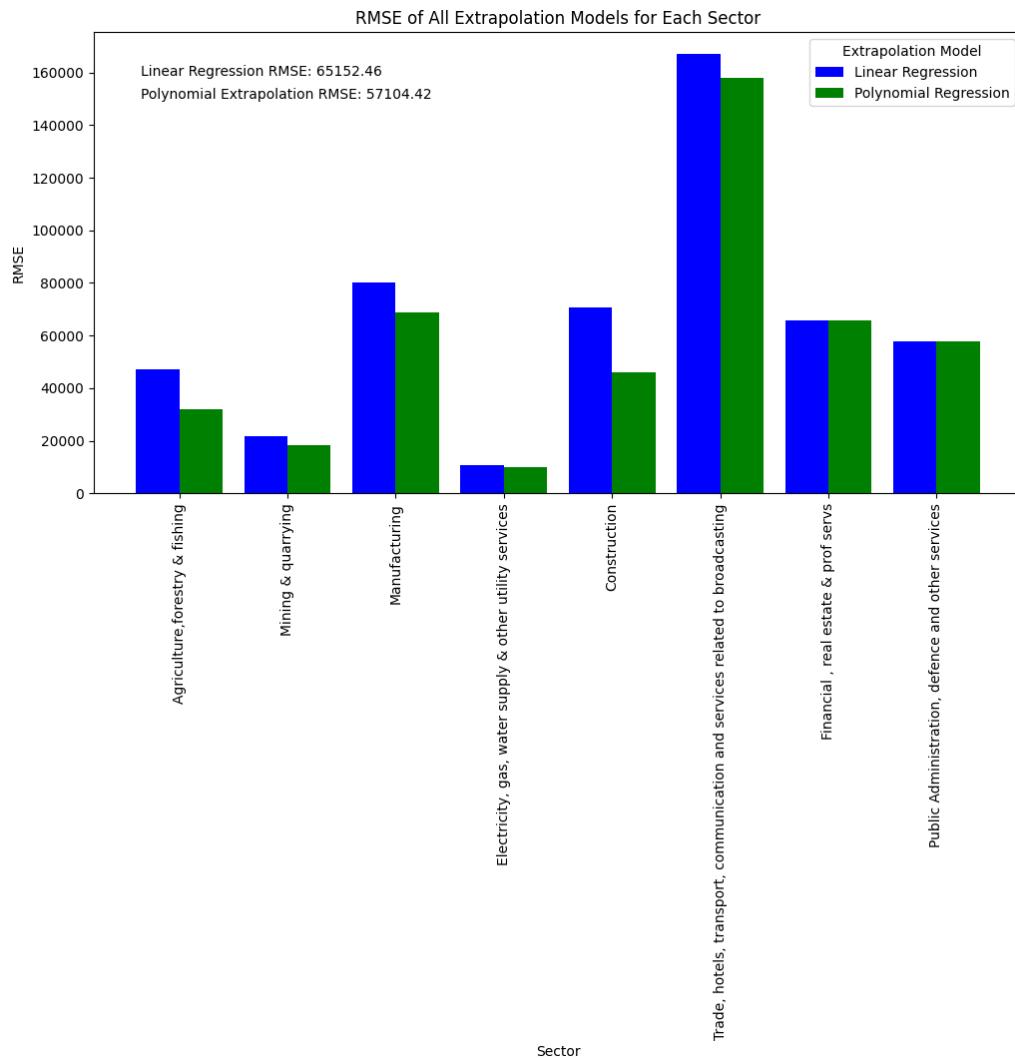


Figure 5.13: Comparison of RMSE for Contribution of Sectors in Terms of Absolute Value

5.1.1.3.4 R^2 Comparison

The plot given below compares the R^2 of both the models.

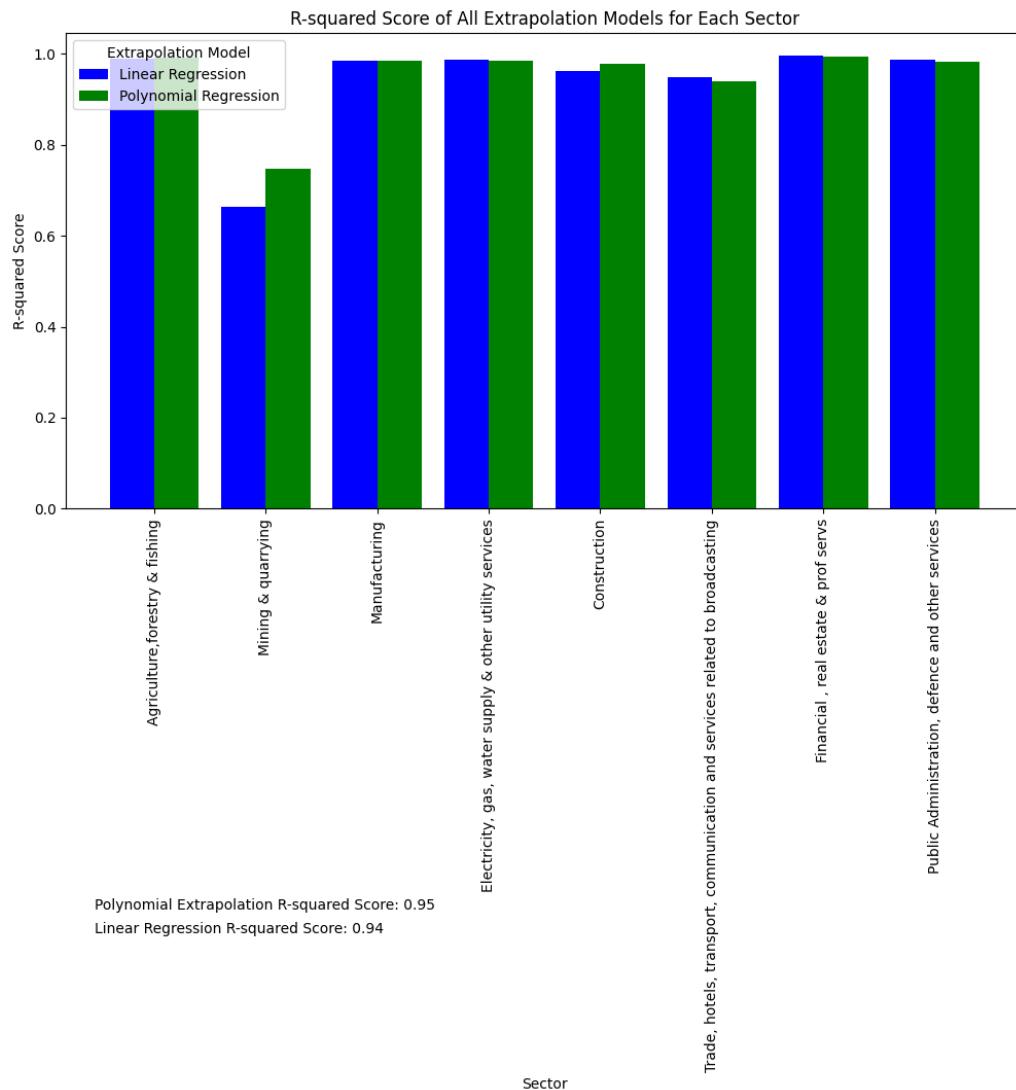


Figure 5.14: Comparison of R^2 for Contribution of Sectors in Terms of Absolute Value

5.2 Contributions of States

5.2.1 GSDP

5.2.1.1 Linear Regression

5.2.1.1.1 Future Prediction

The plot given below has used linear regression (LR) to predict values for the contribution of various sectors till 2029.

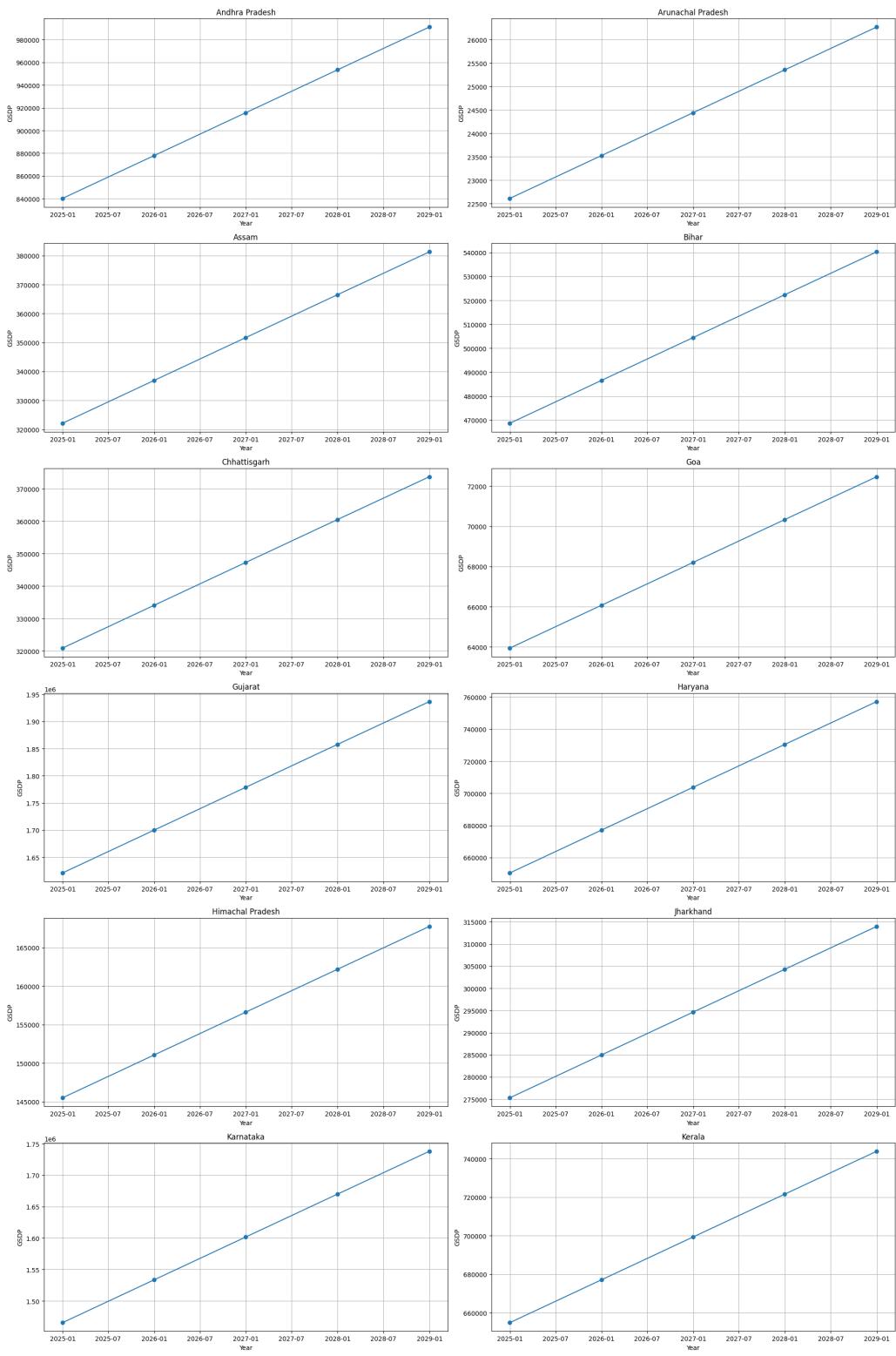


Figure 5.15: Predicted Contribution of States till 2029

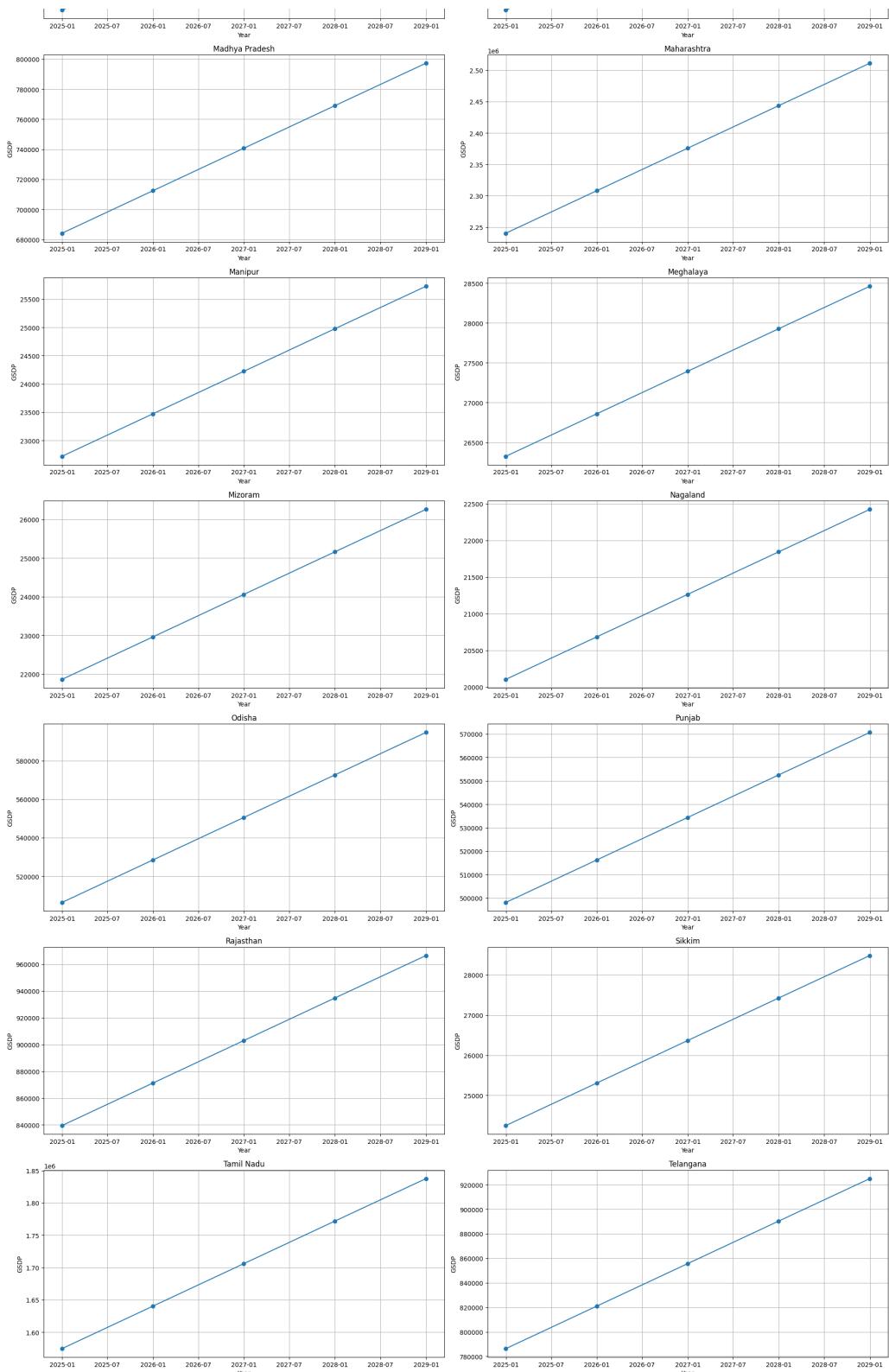


Figure 5.16: Predicted Contribution of States till 2029

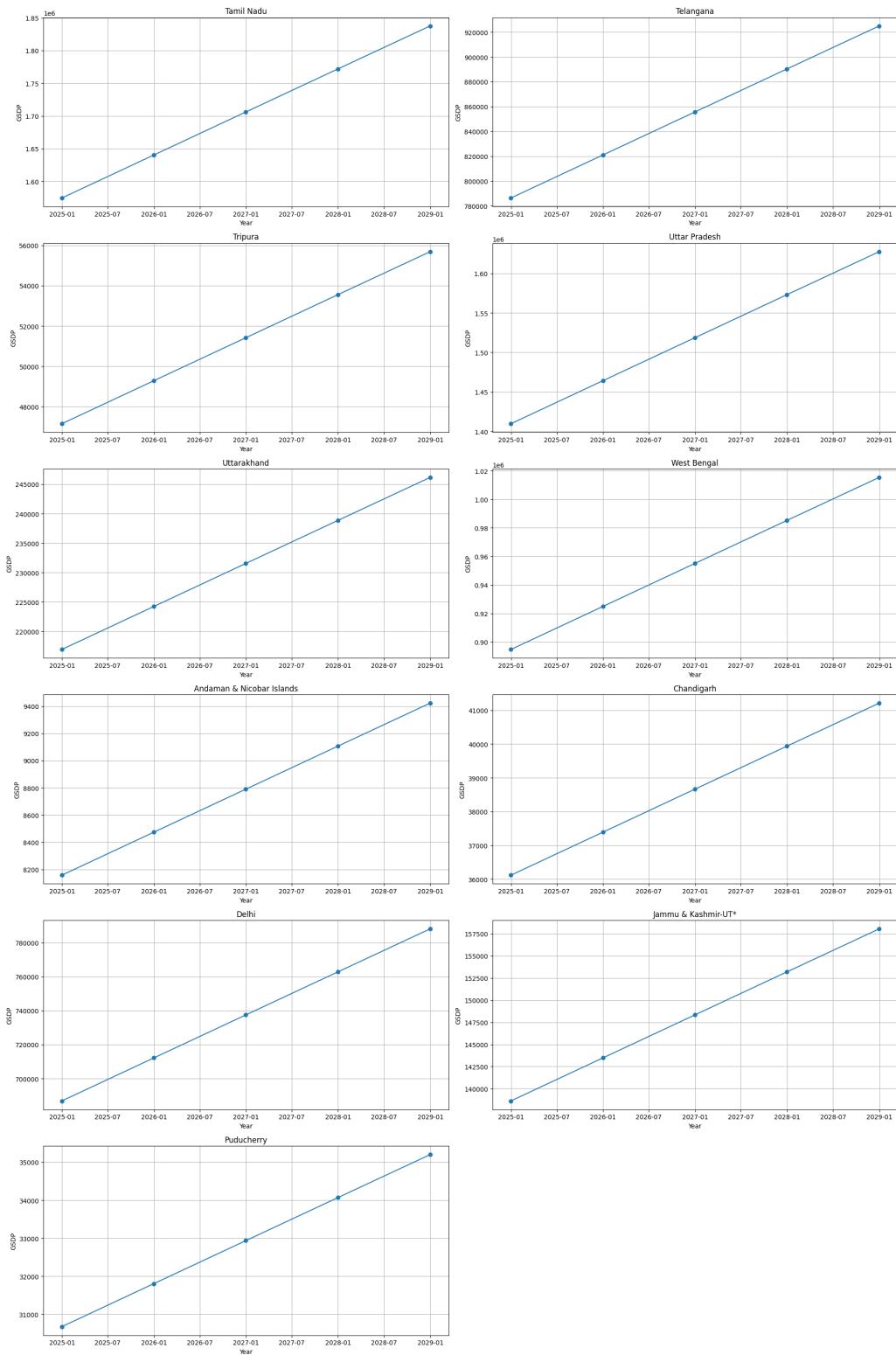


Figure 5.17: Predicted Contribution of States till 2029

5.2.1.1.2 Accuracy Calculation

We have also predicted the contribution of the sectors from 2011 to 2023 using the same regression

model and compared them with the actual data to calculate the accuracy of the model.

Residuals are the difference between the actual value and the predicted value.

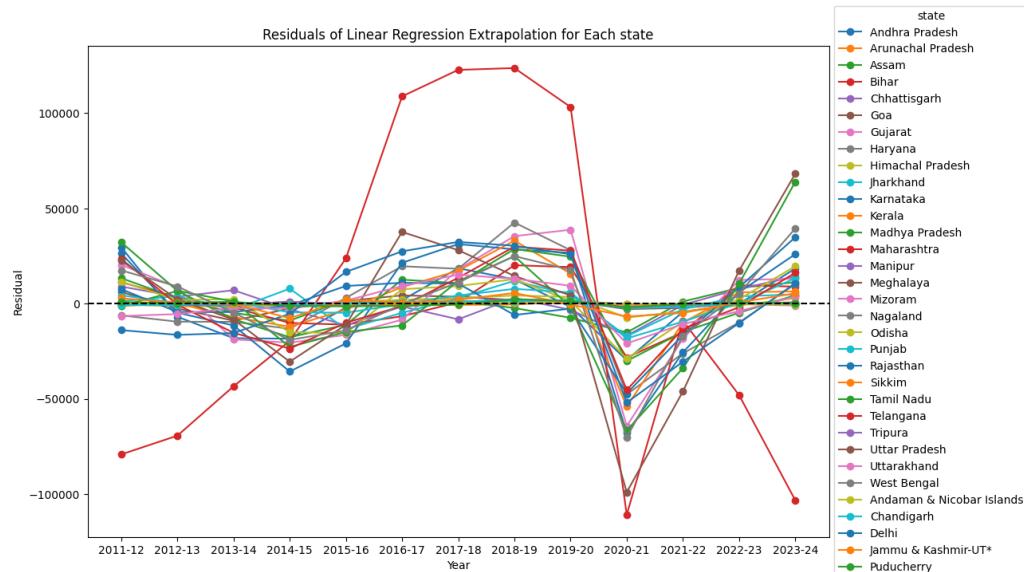


Figure 5.18: Residual of LR model for Contribution of States in Terms of Absolute Value

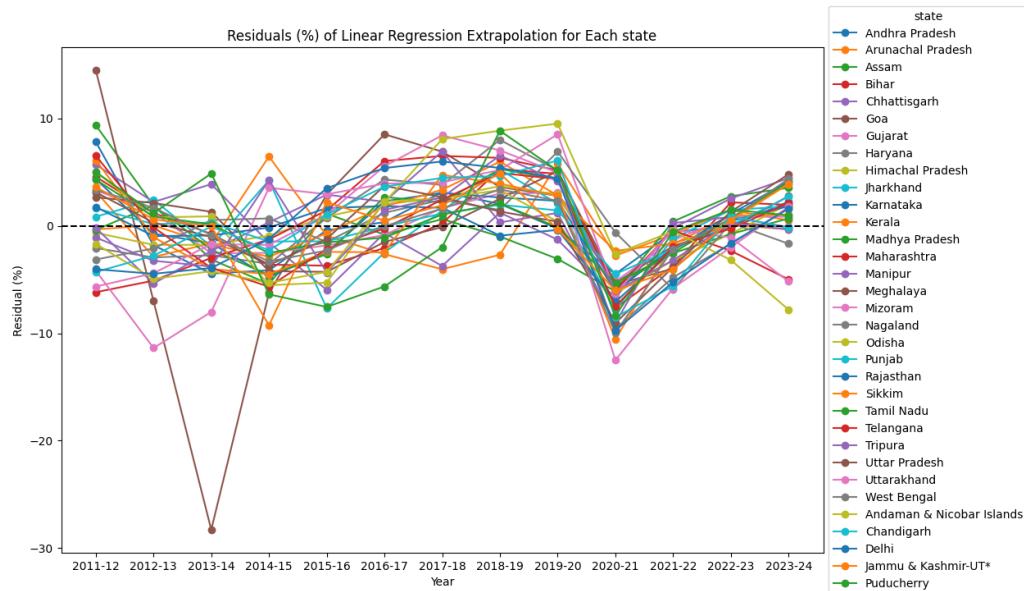


Figure 5.19: Residual of LR model for Contribution of States in Terms of Percent

RMSE stands for 'Root Mean Square Error' and is calculated by the following formula

$$\sqrt{\frac{\sum(P_i - O_i)^2}{n}}$$



where P_i denotes the predicted value, O_i represents the observed value, and n is the total number of observations or data points [5].

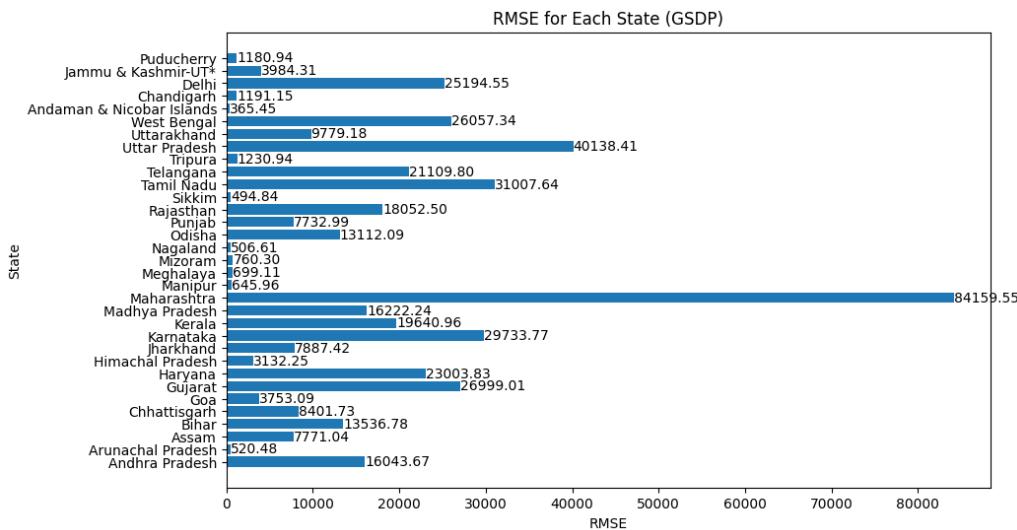


Figure 5.20: RMSE of LR model for Contribution of States in Terms of Percent

R^2 is the regression coefficient and denotes the how efficient the model fitting is. Its value lies from 0 to 1. The closer the value to 1 the better the fit.

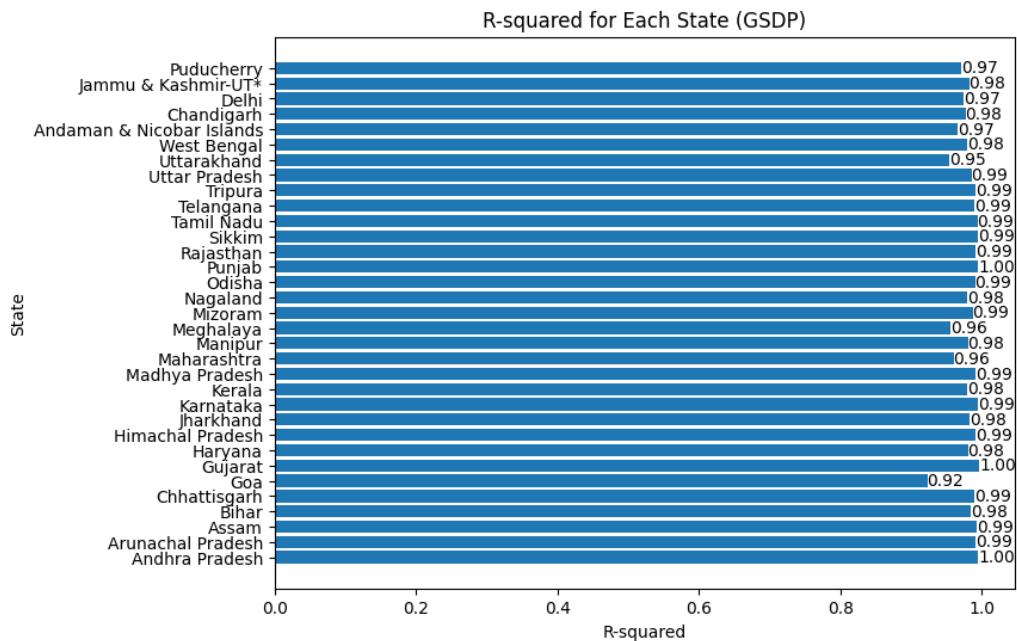


Figure 5.21: R^2 of LR model for Contribution of States in Terms of Percent



5.2.1.2 Polynomial Extrapolation

5.2.1.2.1 Future Prediction

The plot given below has used polynomial extrapolation (PE) to the predict values for the contribution of various sectors till 2029.

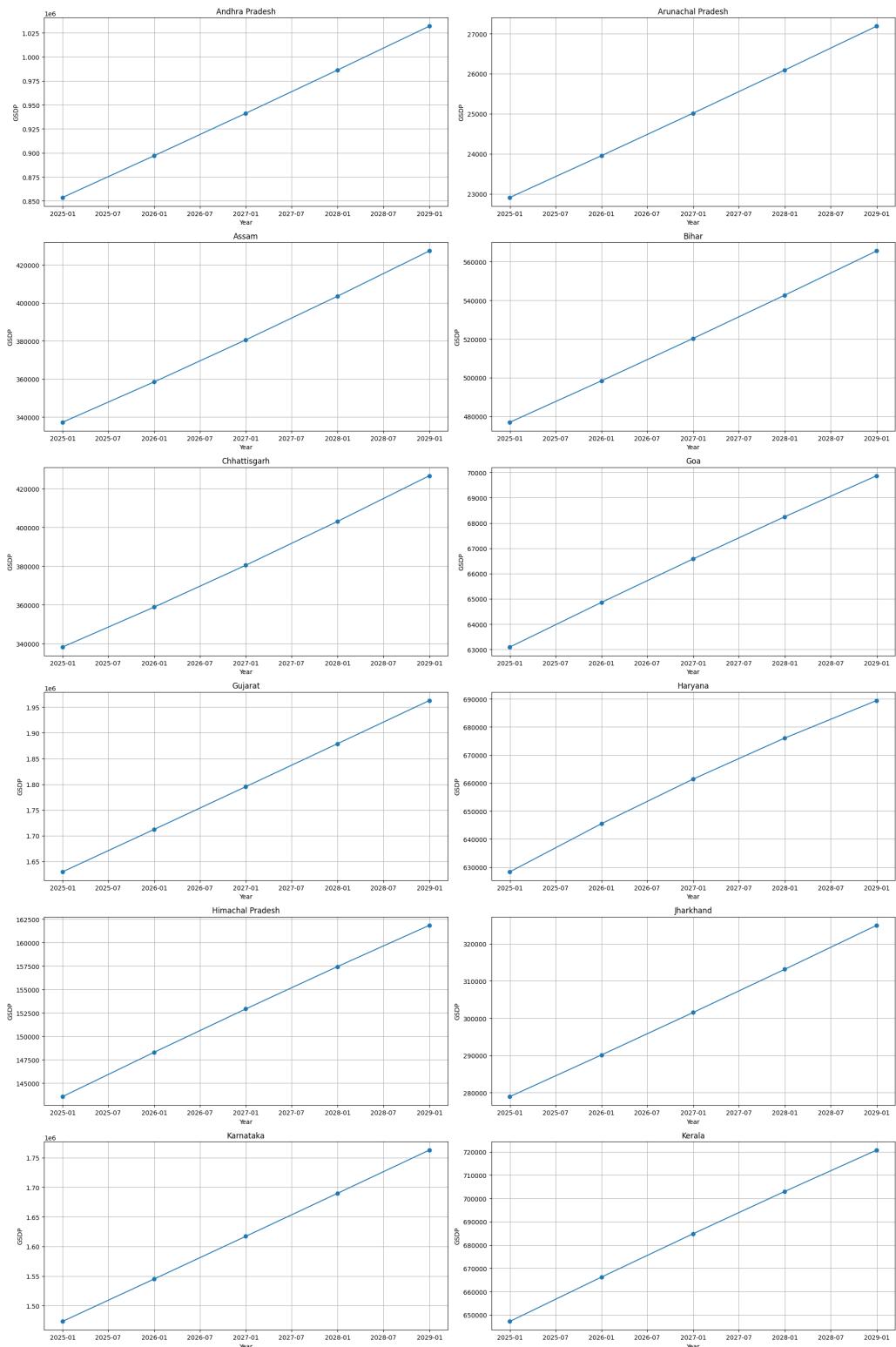


Figure 5.22: Predicted Contribution of States till 2029

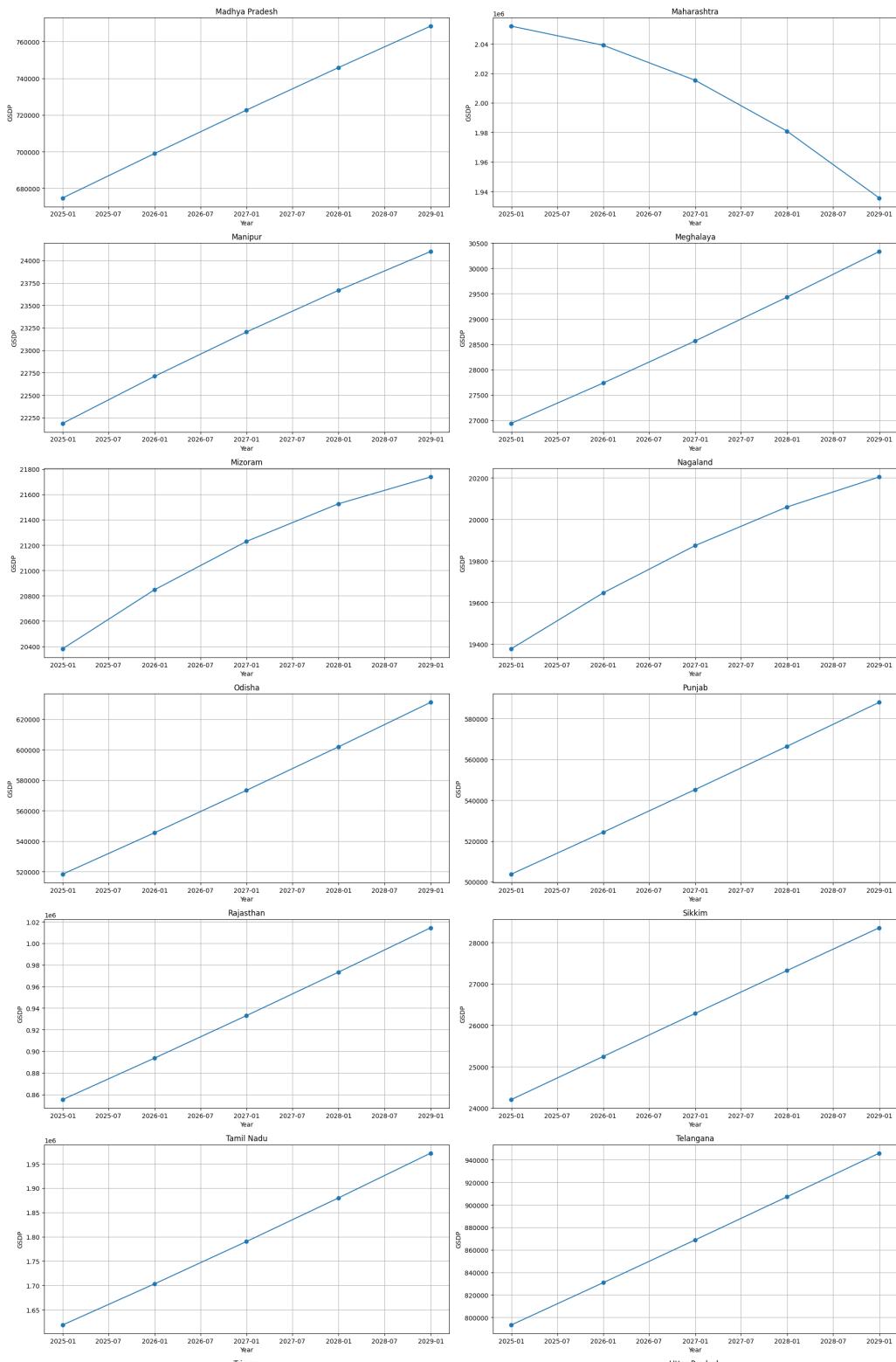


Figure 5.23: Predicted Contribution of States till 2029

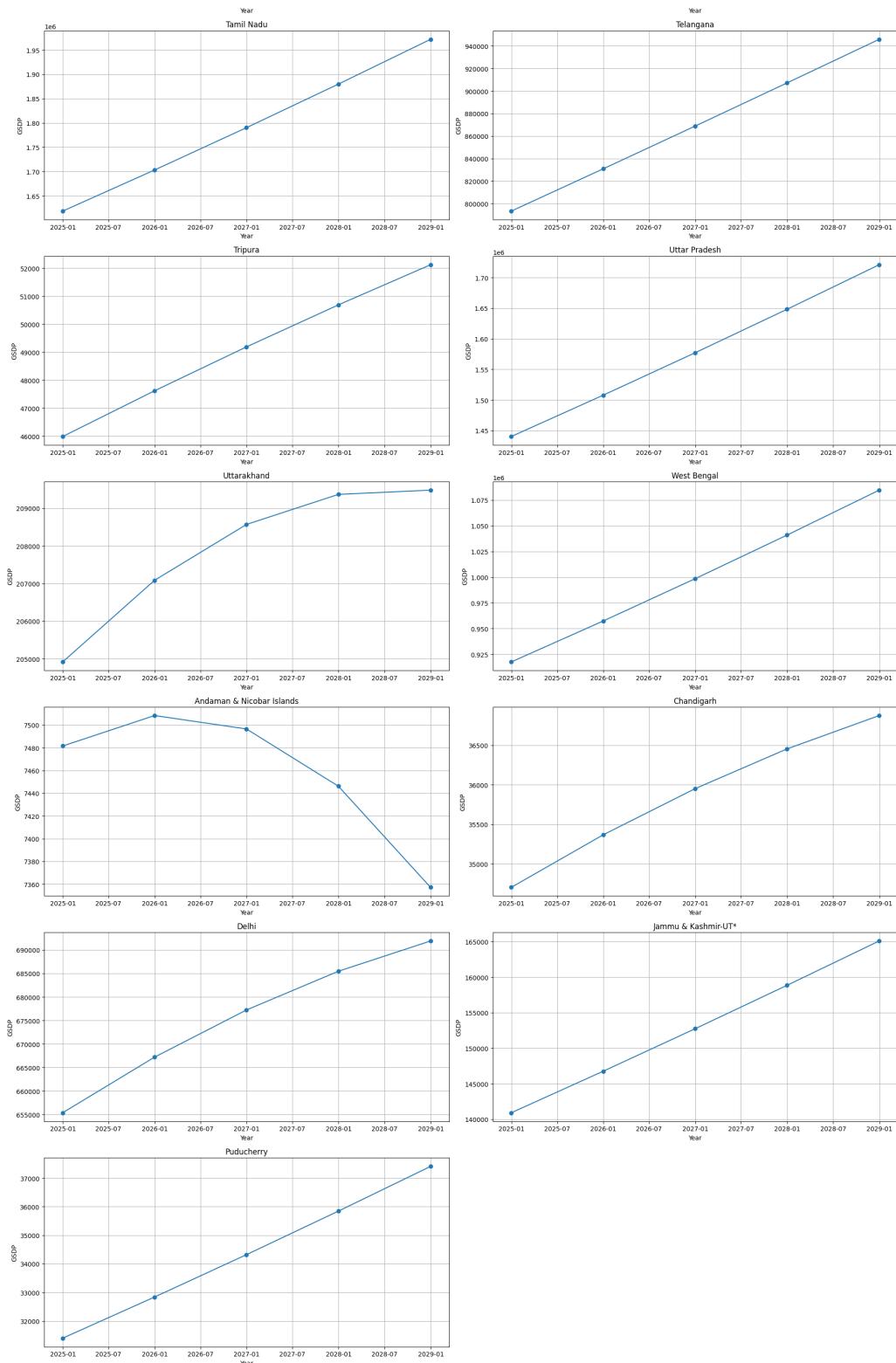


Figure 5.24: Predicted Contribution of States till 2029

5.2.1.2.2 Accuracy Calculation

We have also predicted the contribution of the sectors from 2011 to 2023 using the same regression

model and compared them with the actual data to calculate the accuracy of the model.

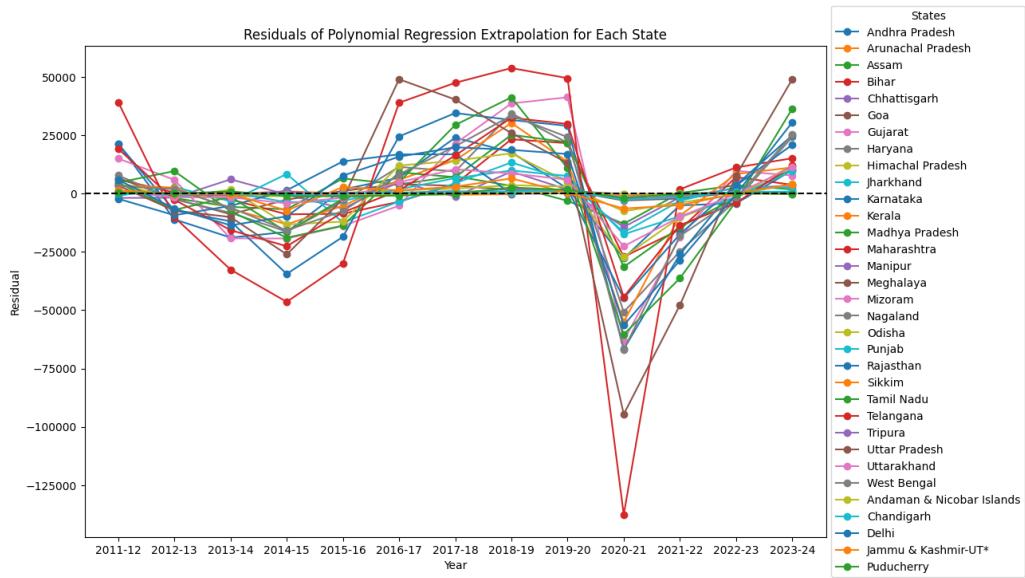


Figure 5.25: Residual of PE model for Contribution of States in Terms of Absolute Value

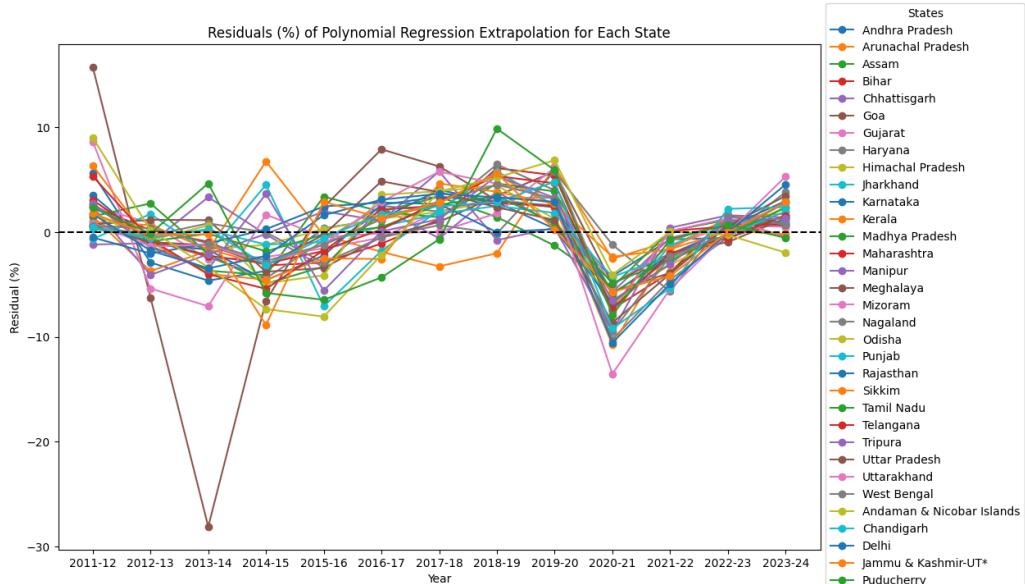


Figure 5.26: Residual of PE model for Contribution of States in Terms of Percent

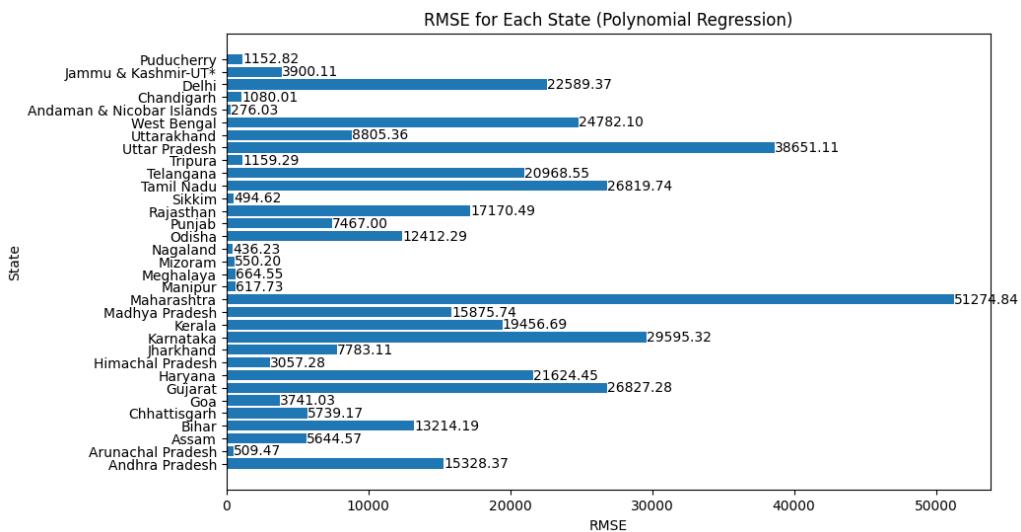
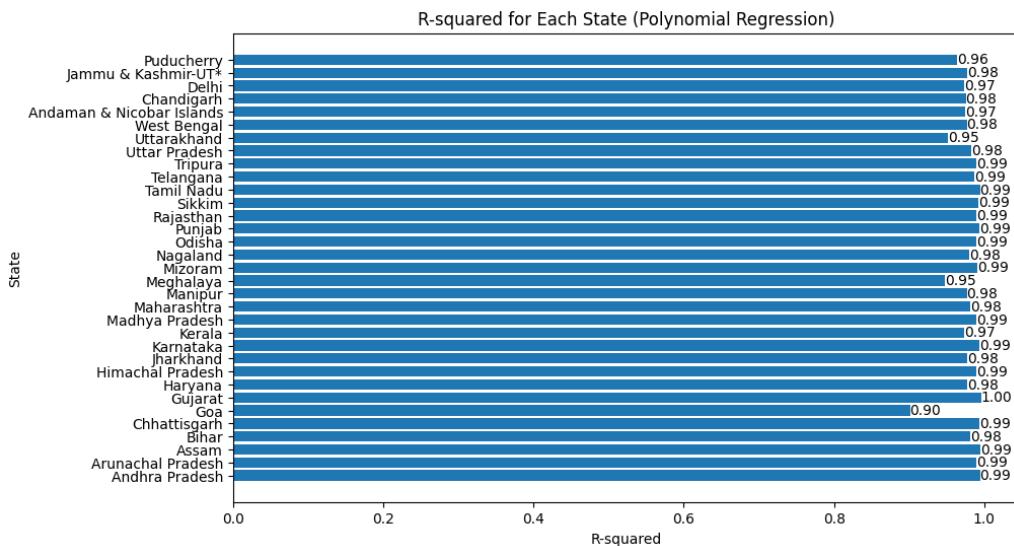


Figure 5.27: RMSE of PE model for Contribution of States in Terms of Percent

Figure 5.28: R^2 of PE model for Contribution of States in Terms of Percent

5.2.1.3 Comparison of Models

Here we have compared the accuracy of both the models.

5.2.1.3.1 Value Comparison

The plot given below compares the predicted values obtained through linear regression and polynomial extrapolation.

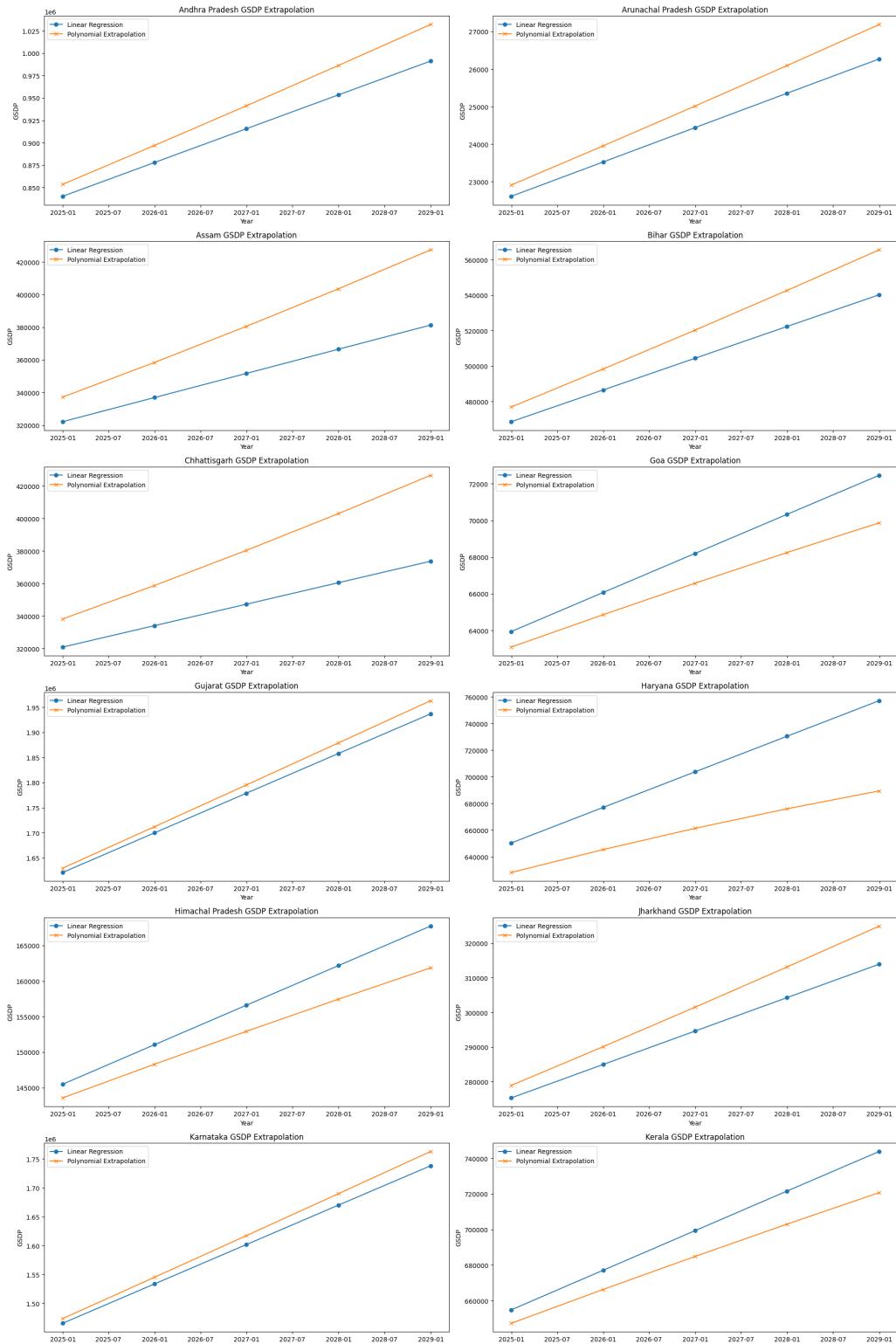


Figure 5.29: Comparison of Predicted Contribution of States till 2029

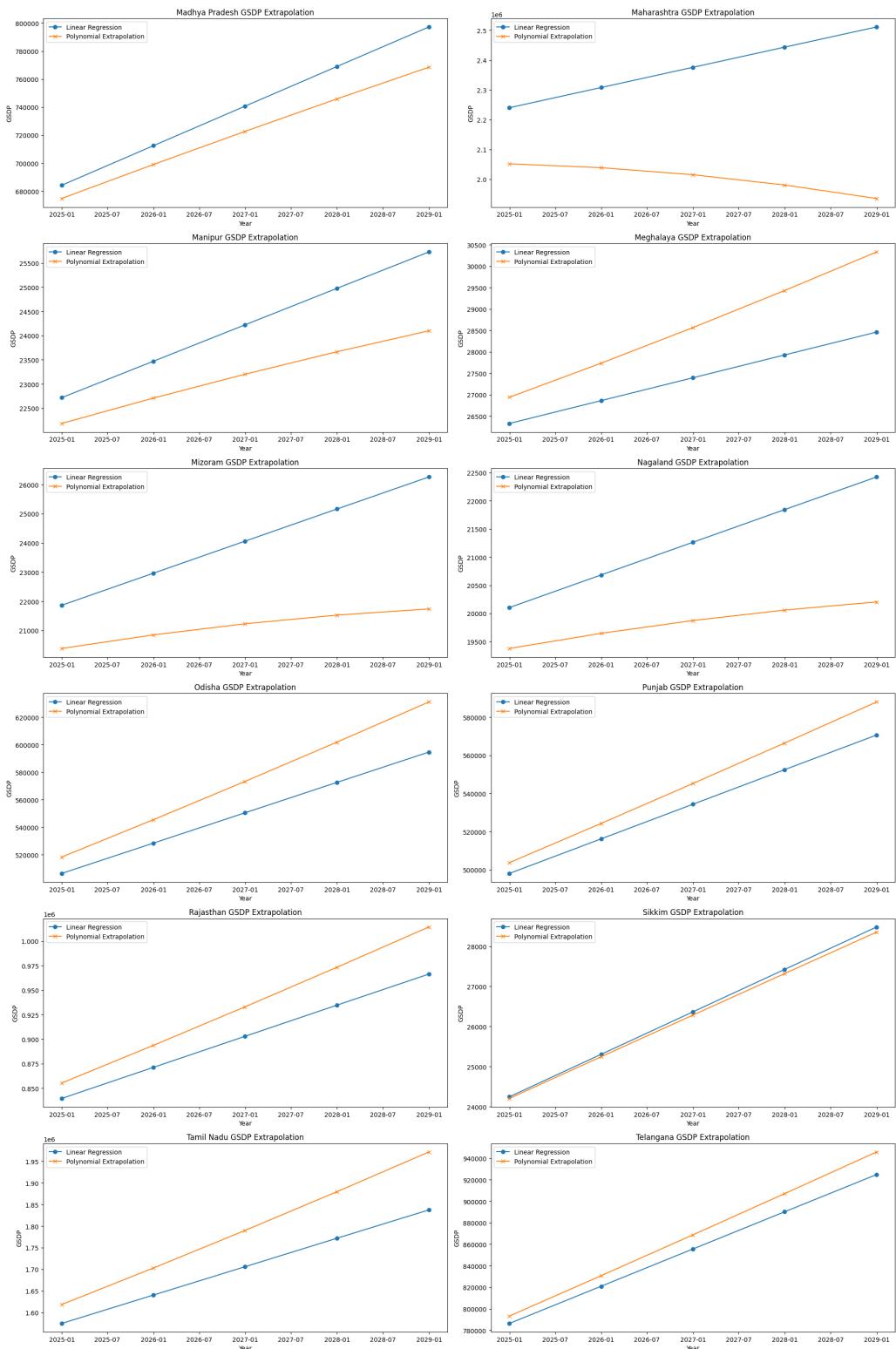


Figure 5.30: Comparison of Predicted Contribution of States till 2029

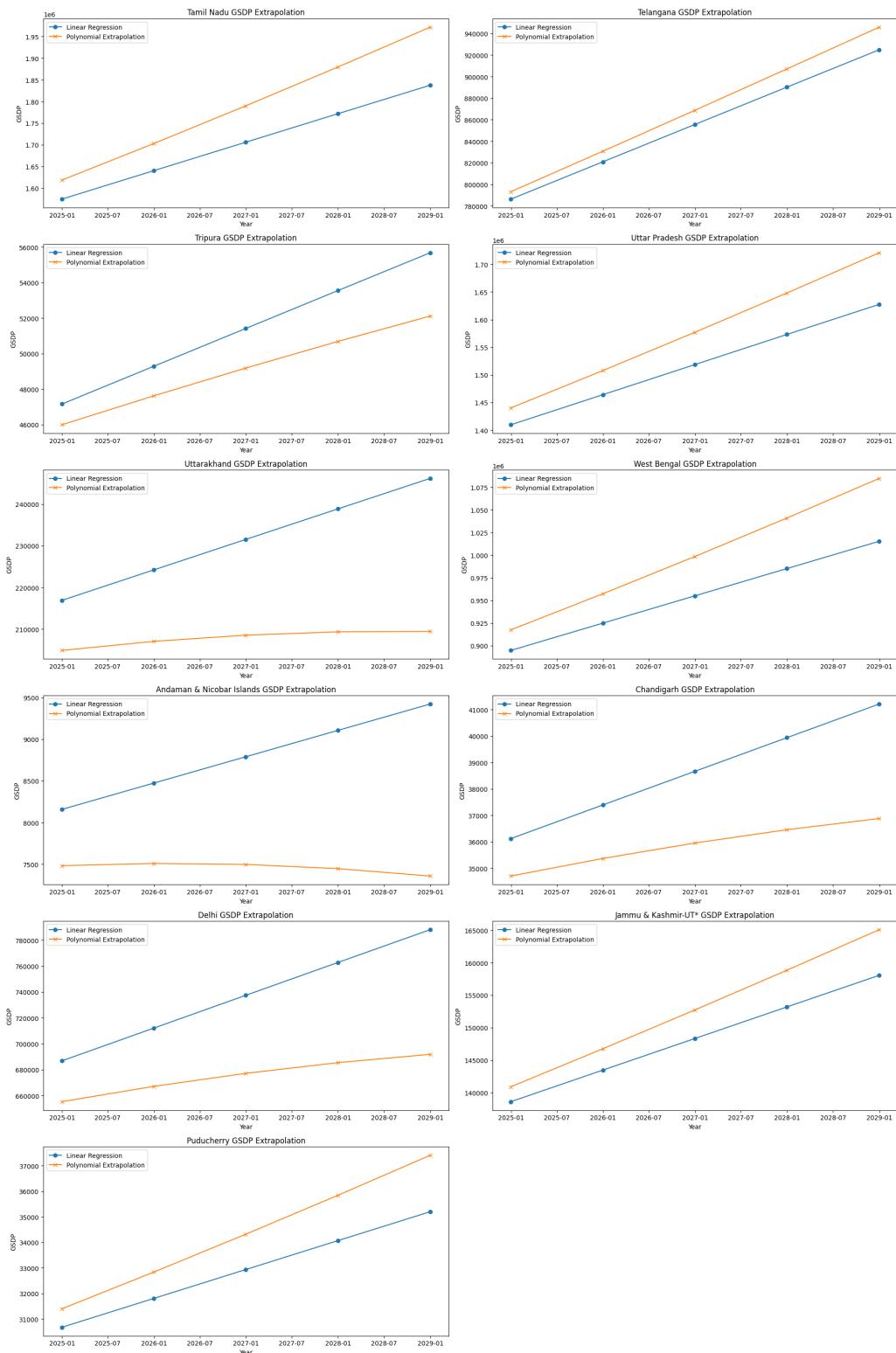


Figure 5.31: Comparison of Predicted Contribution of States till 2029

5.2.1.3.2 Residual Comparison

The plot given below compares the residuals obtained through linear regression and polynomial ex-

trapolation.



Figure 5.32: Comparison of Residuals for Contribution of States in Terms of Absolute Value

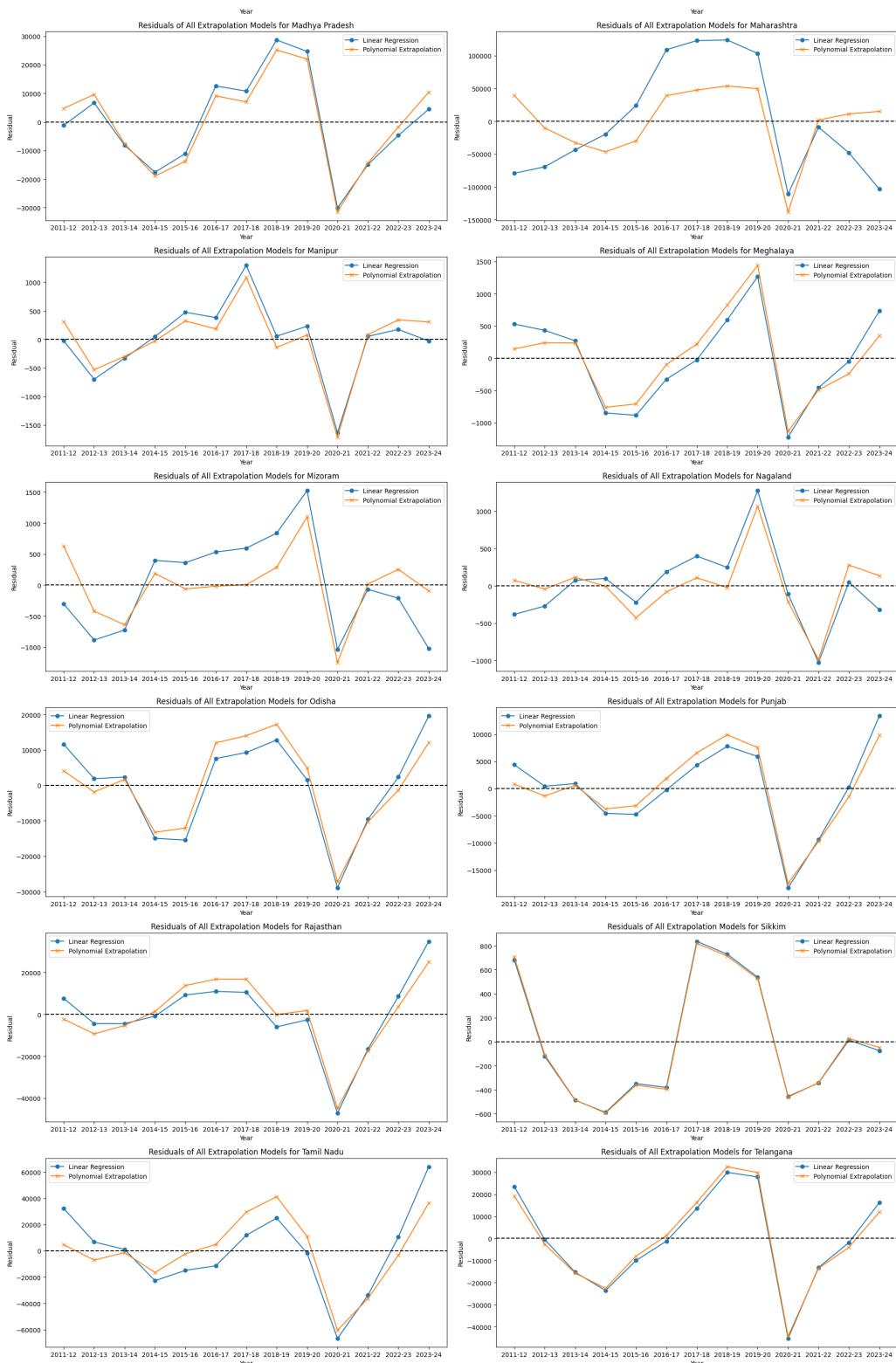


Figure 5.33: Comparison of Residuals for Contribution of States in Terms of Absolute Value

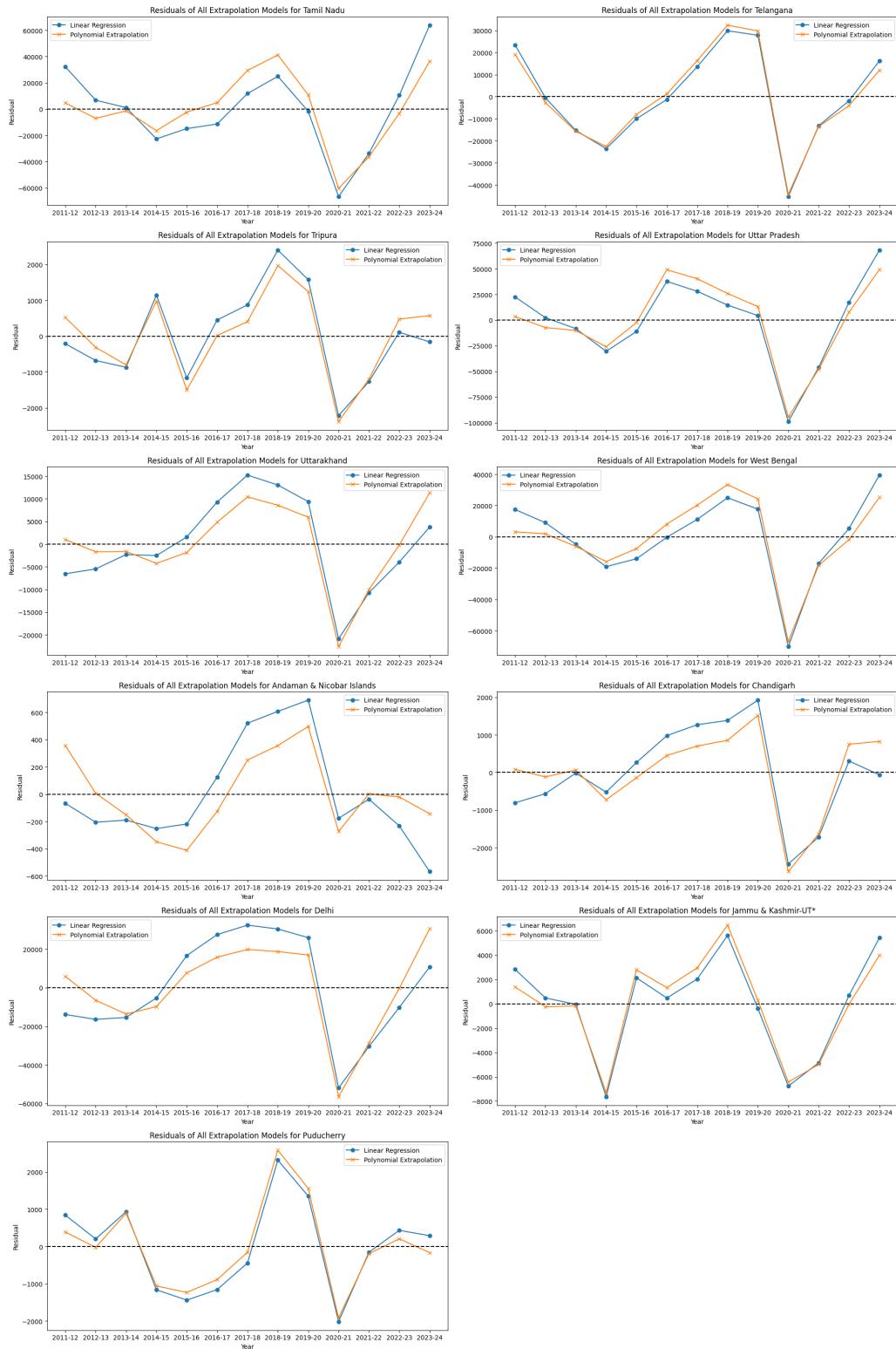


Figure 5.34: Comparison of Residuals for Contribution of States in Terms of Absolute Value

5.2.1.3.3 RMSE Comparison

The plot given below compares the RMSE of both the models.

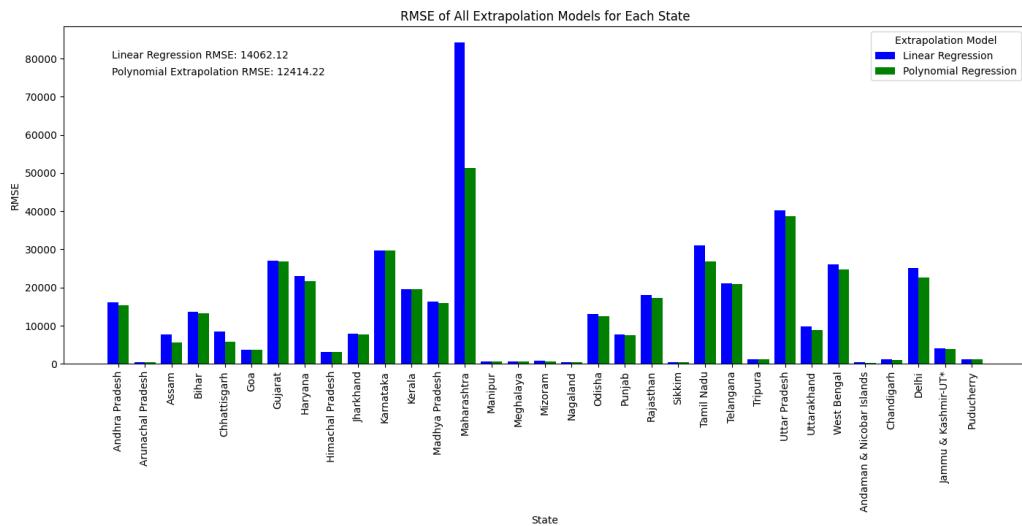


Figure 5.35: Comparison of RMSE for Contribution of States in Terms of Absolute Value

5.2.1.3.4 R² Comparison

The plot given below compares the R² of both the models.

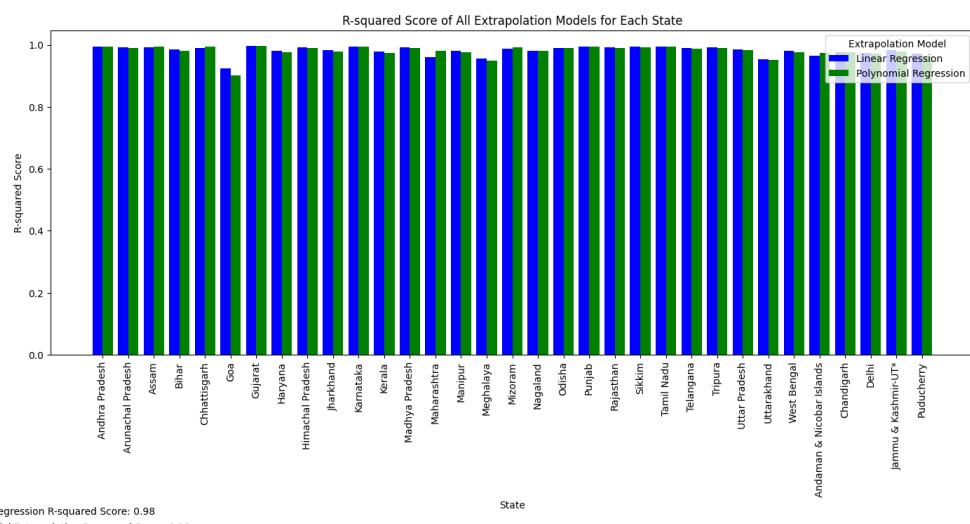


Figure 5.36: Comparison of R² for Contribution of States in Terms of Absolute Value

Chapter 6. Conclusion & future scope

6.1 Findings and Observations

6.1.1 World

1. United States had the highest GDP for the financial year 2022-23 with an amount of 254339.7 billion dollars. Its contribution over the years increased from 15049 billion dollars in 2011-12 to 254339.7 billion dollars in 2022-23.
2. South Africa had the lowest GDP for the financial year 2022-23 with an amount of 405.27 billion dollars. Its contribution over the years decreased from 417 billion dollars in 2011-12 to 405.27 billion dollars in 2022-23.
3. Contribution of India was 3416.65 billion dollars in the financial year 2022-23. Its contribution over the years increased significantly from 1675.62 billion dollars in 2010-11 to 3416.65 billion dollars. This significant growth propelled India into the ranks of the top five nations in terms of economic contribution, reflecting the country's rising economic stature and potential on the world stage.

6.1.2 Contribution of Sectors

1. The sector 'Financial, real estate & prof servs' had the highest contribution of 23.3% for the financial year 2023-24 with a contribution amount of Rs. 3684959 million crores. Its contribution over the years increased from 18.9% of the total GDP in 2011-12 to 23.3% in 2023-24.
2. The sector 'Mining & quarrying' had the lowest contribution of 2.2% for the financial year 2023-24 with a contribution of Rs. 340821 crores. Its contribution over the years increased from 3.2% in 2011-12 to 2.2% in 2023-24.
3. Considering the categories, the tertiary category had the highest contribution of 54.7% of the GDP with a contribution amount of Rs. 8662596 million crores. Its contribution over the years increased from 49% in 2011-12 to 54.7% in 2023-24.
4. The primary category had the lowest contribution of 16.6% with a contribution amount of Rs. 2628150 million crores. On the other hand, its contribution has decreased over the years from 21.7% in 2011-12 to 16.6% in 2023-24.
5. Considering the categories, the tertiary category was affected the most, while primary and secondary were almost constant with a slight increase.



6. By observing the comparison plots, we can see that there is a positive correlation between GVA, NNI, GNI, and NVA with GDP.

One of the major setbacks the economy faced was the COVID-19 Pandemic. Due to it

1. ‘Mining’, ‘Electricity, gas, water supply & other utility services’, ‘Construction’, ’Trade, hotels, transport, communication and services related to broadcasting’ and ‘Public Administration, defence and other services’ were the sectors that were affected, ‘Trade, hotels, transport, communication, and services related to broadcasting’ getting affected the most.
2. The contribution of this sector in 2019-20 was Rs. 2690060 million crores, while it fell down to Rs. 2153888 million crores in 2020-21, with a decrease of -19.9%.
3. Following were the contributions of each sector before and after the pandemic.
4. On the other hand, the contribution of the ‘Manufacturing’ sector increased possibly due to governmental initiatives such as ‘AtmaNirbhar Bharat’. This conclusion has been made by observing the data for three consecutive years.

Strangely, we also observed that the growth rate of almost every sector had a slight decrease in the year just before the Lok Sabha elections.

6.1.3 Contribution of States

1. Maharashtra had the highest contribution to the national GDP of 11.9% with a contribution amount of Rs. 2069199 million crores in 2023-24. Its contribution decreased from 14.8% in 2011-12 to 11.9% in 2023-24.
2. Andaman and Nicobar Islands had the least contribution in 2023-24, which was less than 0.1%. Its contribution was Rs. 7273 crores.
3. Considering zones, the South Zone had the highest contribution of 30.4% in 2023-24, and the Northeast Zone had the lowest contribution of 2.9% in 2023-24.
4. The pandemic affected the GSDP of each state, with Uttarakhand facing the highest decline of -12.10%.

6.1.4 Predictions

We performed linear regression and polynomial extrapolation to predict the future values of the contribution of various sectors and states. We further calculated the accuracies of these models and compared the parameters of both these models with each other and came to a conclusion that the ‘polynomial extrapolation’ method is more accurate and efficient than the linear regression model.

Given below are a few visualizations we made in PowerBI

6.1.4.0.1 GDP of G20 Countries

The plots given below show and compare GDP Growth of G20 Countries over the Decade.

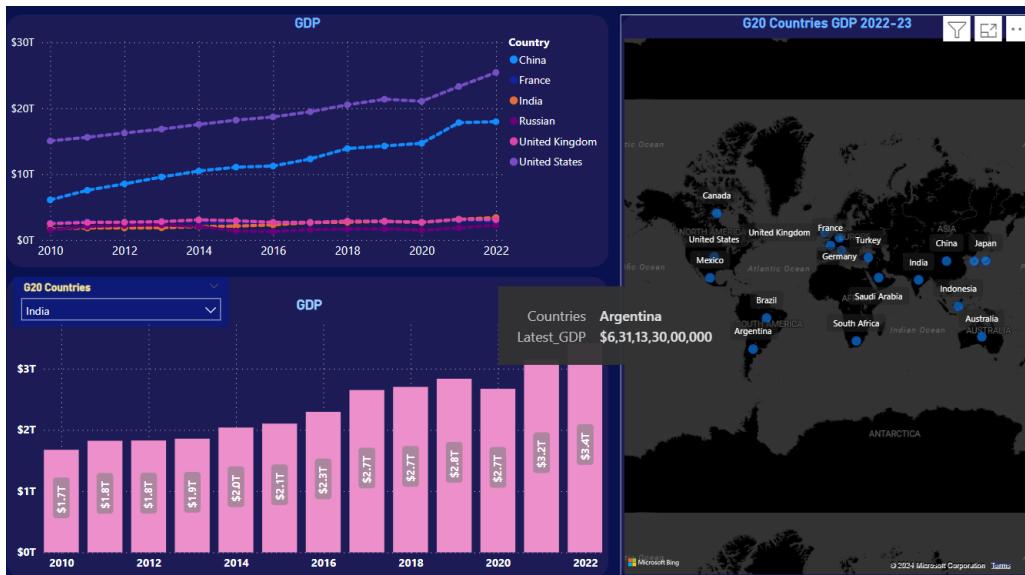


Figure 6.1: GDP Growth of G20 Countries and it's Trend Over the Decade

6.1.4.0.2 GVA

The plots given below show and compare the Growth and Growth Rate of GVA of various sectors over the Decade.



Figure 6.2: GVA Growth of various Sectors and it's Trend Over the Decade

6.1.4.0.3 GSDP Growth of States of India

The plots given below show and compare the contribution of various States of India in terms of GSDP Growth and its Growth Rate over the Decade.

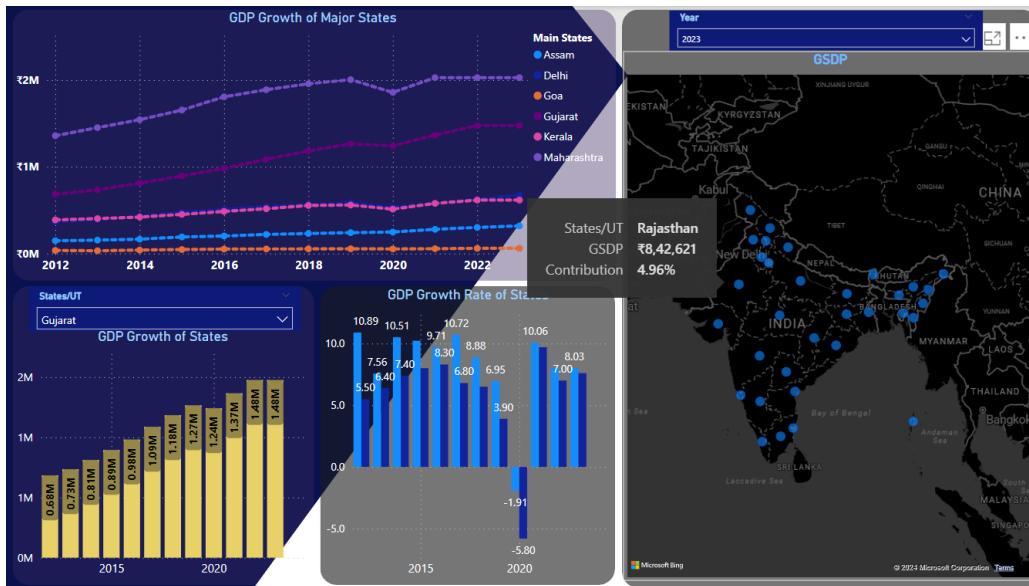


Figure 6.3: GSDP Growth and Growth Rate of States of India and it's Trend Over the Decade

6.2 Challenges

1. **Data Accuracy:** Ensuring precise data collection and maintaining data quality is essential, especially when dealing with extensive datasets covering national and state-level economies.
2. **Sectoral Analysis Complexity:** Analyzing sector-wise GDP poses challenges due to the complexity of sectors and their varied contributions to overall economic growth.
3. **Interpretation Amidst Dynamic Environments:** Interpreting GDP trends becomes challenging amidst rapidly changing global and domestic economic landscapes, requiring adaptable analytical frameworks.
4. **Multidisciplinary Approach:** Overcoming these challenges necessitates a multidisciplinary approach, integrating advanced analytical tools and collaborative efforts to extract meaningful insights.

6.3 Future plan

In the realm of GDP analysis, the future holds promising opportunities for leveraging insights into economic trends and sector-wise contributions. Our project delved into dissecting India's GDP at both national and state levels, providing a comprehensive understanding of each region's economic landscape. Moving forward, this analysis lays the groundwork for informed policy-making, investment strategies, and targeted development initiatives, ultimately driving sustainable growth and prosperity across the nation.



6.4 Conclusion

The project delves into multifaceted analyses of GDP growth within the G20 nations, offering insights into global economic trends and the underlying factors influencing their performances. By dissecting the contribution of various sectors to India's GDP growth, the study aims to illuminate the pivotal drivers of economic expansion while highlighting potential challenges faced by lagging sectors. Additionally, a thorough examination of India's GDP growth rate over the past decade provides valuable insights into periods of acceleration or deceleration, thus guiding policymakers in devising interventions to sustain economic momentum. Furthermore, the project scrutinizes the growth of Gross State Domestic Product (GSDP) across different Indian states, uncovering regional disparities and offering valuable data to shape policies geared towards fostering balanced development and reducing socio-economic inequalities. Through these comprehensive analyses, the project endeavors to provide actionable insights for policymakers, businesses, and stakeholders to navigate the complexities of global and national economic landscapes effectively.

Group Contribution

Mihir Paija (202101205)

- Data analysis on GVA and GDP of G20 countries.
- LaTeX Report.

Dhruv Shah (202101208)

- Data analysis on GVA Growth Rate and GDP of G20 countries.
- LaTeX Report.

Nikita Shah (202101209)

- Data analysis on GSDP Growth Rate and GDP of G20 countries.
- PowerBI Visualisation.

Rohan Mistry (202101231)

- Data analysis on GSDP and GDP of G20 countries.
- LaTeX Report.

Short Bio

1. **Mihir Paija** is a backend web developer with expertise in JavaScript, C++, and Python. He also possesses a foundational understanding of robotics with robot programming in ROS and Database Design. Beyond his technical pursuits, Mihir is an avid traveler and reader. He finds joy in analyzing the complexity of poetry and indulging in cinema. His reading interests encompass fiction, philosophy, and political ideologies, all explored with the mindset of a free thinker, and a firm belief in "carpe diem".

2. **Dhruv Shah** is an enthusiastic student pursuing a BTech in ICT from DAIICT. He has a passion for technology and enjoys brainstorming and exploring new ideas and concepts. He is also highly interested in web development. Beyond his academic pursuits, he has a love for cinema and content writing. He finds joy in creating engaging content and managing events, where he can bring people together. With a versatile skill set and a curious mind, he is eager to explore new opportu-

nities and make a positive impact in his field.

3. **Nikita Shah** is currently pursuing B.Tech in ICT at DA-IICT. She is a passionate web developer with a view of providing user friendly web-applications and keen interest in backend development. She has a keen interest in exploring new technologies. She is proficient in programming languages like C++ and JavaScript. Outside of work, she enjoys reading books and listening to music. This passion not only adds depth to her character but also enhances her creativity. With boundless energy and a drive to succeed, She is poised to make the most of her college experience and beyond.

4. **Rohan Mistry** is currently pursuing B.Tech in ICT at DA-IICT. He is a Full Stack Web Developer and Java Developer. With keen interest in backend, he likes to explore backend technologies. Apart from coding, he also likes to share his insights and enjoys travelling, playing cricket and gaining insights into the emerging tech trends.

References

- [1] Wikipedia. (n.d.). Gross domestic product. Retrieved from <https://en.wikipedia.org/wiki/Gross Domestic Product>
- [2] Wikipedia. (n.d.). Gross value added. Retrieved from <https://en.wikipedia.org/wiki/Gross Value Added>
- [3] Ministry of Statistics and Programme Implementation (MOSPI). (n.d.). State and National Dataset. URL: <https://mospi.gov.in/>
- [4] The World Bank. (n.d.). G20 countries' GDP. URL: <https://data.worldbank.org/indicator/NY.GDP.MKTP.PP.CD>
- [5] DeepChecks. (n.d.). RMSE. URL: <https://deepchecks.com/glossary/root-mean-square-error/>