



***Forecasting Economic Indicators of Andhra Pradesh & Telangana for
Central Budget Planning***

GROUP-A2

PGID	STUDENT NAME
61910086	Ankana Dhar
61910177	Sri Harsha Bandi
61910261	Sireesha Tekuru
61910747	W Snigdha Rao
61910552	Ujjwal Vasisht
61910755	Susruta Meka



EXECUTIVE SUMMARY

Business Problem

The ministry of Finance possesses a lot of data with respect to multiple financial indicators of all the states. A forecast of these indicators for the next 5 years, will guide the ministry in planning the fiscal budget for various states.

The dataset in usage has the data of all 28 states, however we have chosen to construct forecasting models for the state of Andhra Pradesh. However, it should be noted that, due to the separation of Telangana from Andhra Pradesh in 2014, the forecasts will be an aggregation of both the states.

Client: Office of Finance Minister of Government of India

How the forecasts will be used: This forecasting model will help the government predict various financial indicators and accordingly perform fiscal planning for the state of Andhra Pradesh and Telangana for the time 2016-2020 (5 years).

Forecasting Problem and data

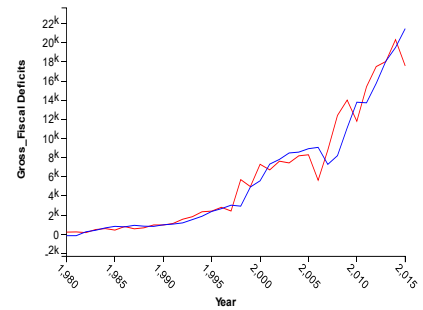
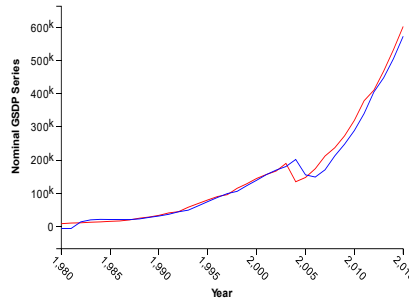
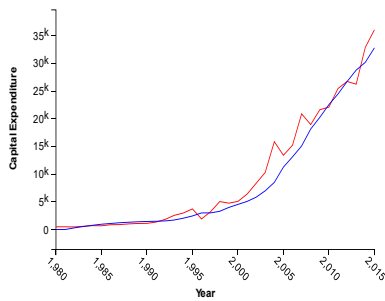
To forecast the fiscal indicators, we have 6 datasets each one related to one indicator. The data is available from 1980-81 to 2015-16. Below is the list of fiscal indicators we are trying to forecast:

1. Nominal Gross state domestic product (GSDP)
2. Gross Financial deficit
3. Capital Expenditure
4. Social sector expenditure
5. Revenue deficit
6. Own tax revenues

Forecasting Model

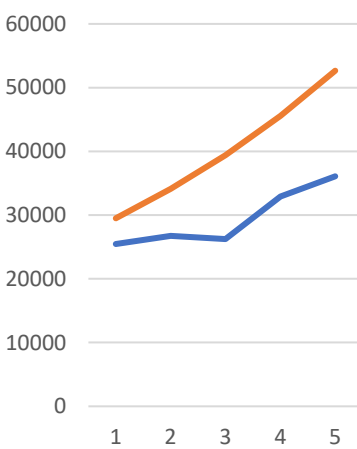
Double Exponential Smoothing and Linear Regression models performed optimally for the datasets in the project. The data sets are aggregated on an annual level and do not show substantial seasonality, only increasing trend.

Double Exponential Smoothing:

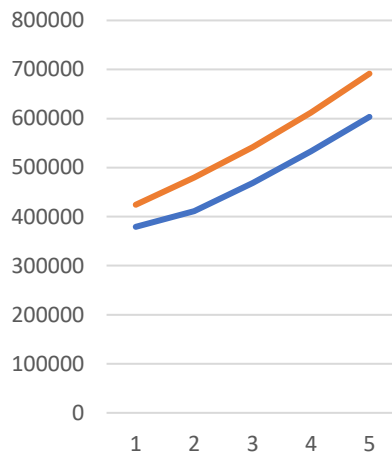


Linear Regression:

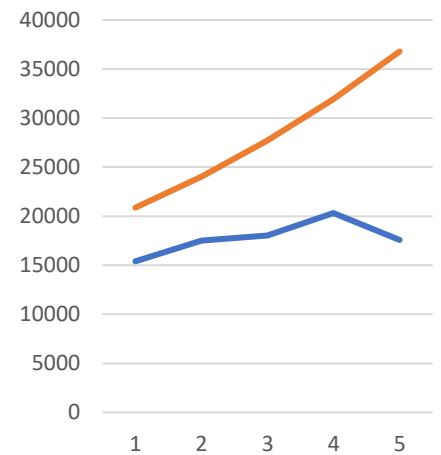
Capital Expenditure



GSDP



Fiscal deficit



Recommendations:

- Only time is the independent variable. Need other external variables for better forecasting the indicators
- No seasonality in the time series of the indicators. Quarterly data would have helped to see whether seasonality is present.
- It could be helpful to use some of the relevant data sets as predictor variable in building the forecasting model for other indicators. Many of the indicators are interdependent of each-other, so we need to know how the factors interact with each other in the regression model too.

- During application of this model, separate prospective CAGR of Population density, general allocation funds, financing surplus budget, revenue, and other macro-economic indicators of Telangana and AP should be factored in the model. Also, a roll forward approach for training data set should be used every period so that the model keeps improving with real time data.

TECHNICAL SUMMARY

Dataset and Data Processing

Original Dataset

The original dataset from Kaggle had eight indicators pertaining to all the states across the years, in 8 different files. The eight fiscal indicators include nominal GSDP, gross fiscal deficit, capital expenditure, social sector expenditure, revenue expenditure, revenue deficits, own tax revenues and aggregate expenditure.

Data processing

The data downloaded from Kaggle was processed as below:

- ▶ Out of the eight indicators, five have been selected to perform the forecasting since the other 3 variables (revenue expenditure, revenue deficit and aggregate expenditure) are inter-related.
- ▶ The data for Andhra Pradesh and Telangana is extracted from each of the five files and is tabulated with year and five economic indicators.

NOTE: Since the state separation happened in 2014, the data of both Telangana and Andhra Pradesh are summed for years 2015 and 2016. Thus, the forecast depicts the fiscal variables for the combined state of Telangana and Andhra Pradesh. However, we can include external information from government sources (AP and Telangana split revenues in the ratio 48:52 for the first 10 years) to split the revenue and expenditures for AP and Telangana.

Data Description

Below are the fields in the prepared dataset:

- ▶ Year (Ranges from 1980-81 to 2015-16)
- ▶ Capital expenditure

- ▶ Nominal GSDP series
- ▶ Own tax revenues
- ▶ Social sector expenditure
- ▶ Gross fiscal deficits

The data covers the past 35 years (Yearly data) and looks to be good enough to forecast for the next 5 years.

Forecasting methods used

From the time-series plots shown previously, all the indicators have only trend and no seasonality. This narrowed down the choice of forecasting methods to double exponential and regression models as depicted in the table below:

Forecasting Method	Descriptions	Can we use?
Simple Exponential Smoothing	Used for forecasting series that have no trend or seasonality	NO
Double Exponential Smoothing	Used for forecasting series that have trend	YES
Holt-Winter's Exponential Smoothing	Used for forecasting series that have both trend and seasonality	NO
Regression Model	Used for forecasting series that have global trend	YES

Indicators¹

1. Capital Expenditure
2. Own Tax Revenues
3. Nominal GSDP Series
4. Social Sector Expenditure
5. Gross Fiscal Deficits

¹ The time series plots of these indicators Exhibit 1

Data Partitioning

We partitioned the data to forecast the fiscal variables for 5 years as below:

Forecasting Horizon = 5 years; Seasonal Number = 0; Validation Period = 5

36 data points = 31 data points & Validation Period = 5 data points

Naïve Model and benchmarking

A naïve forecast has been performed on the data, Naïve forecast being the same value as that of the previous year and below is the prediction output.²

Double Exponential Smoothing

We have tried 4 and 3 values of alpha and beta respectively to find the best prediction model as shown in exhibit 7. The plots of Predicted vs Actual Capital Expenditure, Nominal GSDP, and Gross Fiscal Deficits³ for the best model are attached in the appendix.

Linear Regression with Exponential trend

We used regression model with linear trend and the forecasted output is not encouraging. We then used regression model with exponential trend to forecast the indicators and the regression Summary for the Indicator's Errors⁴ is attached in the appendix.

Best model to forecast

Comparing the accuracy of the 3 models used for forecasting, we get the following summary that determines the best model (considering least MAPE) to forecast each of the variables. The highlighted columns represent the best models for the respective indicators.

MAPE	Capital Expenditure	Nominal GSDP Series	Tax Revenues	Social Sector Expenditure	Gross Fiscal Deficits
Naïve	23.59%	31.30%	17.68%	20.49%	33.08%
Exponential Smoothing	6.15%	15.55%	23.87%	9.03%	12.02%
Linear Regression	35.57%	14.71%	38.72%	16.41%	58.56%

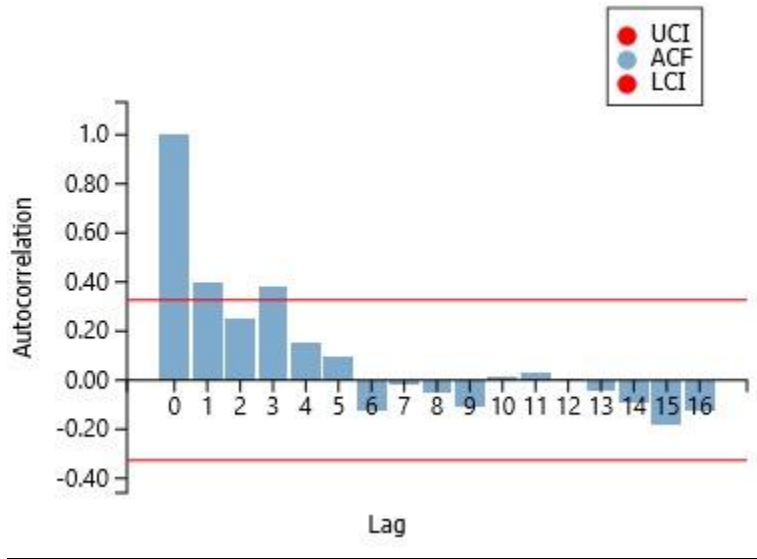
² The Naïve model Summary in Exhibit 2

³ The Plots in Exhibit 5

⁴ Table in Exhibit 6

Autocorrelation

The ACF plot for the residuals in the exponential smoothing method is shown below. There is no correlation observed between the residuals.



APPENDIX

EXHIBIT 1

Time series plot of individual indicators

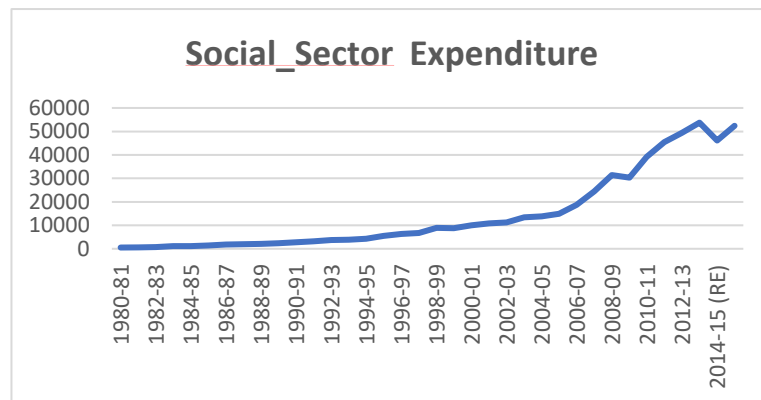
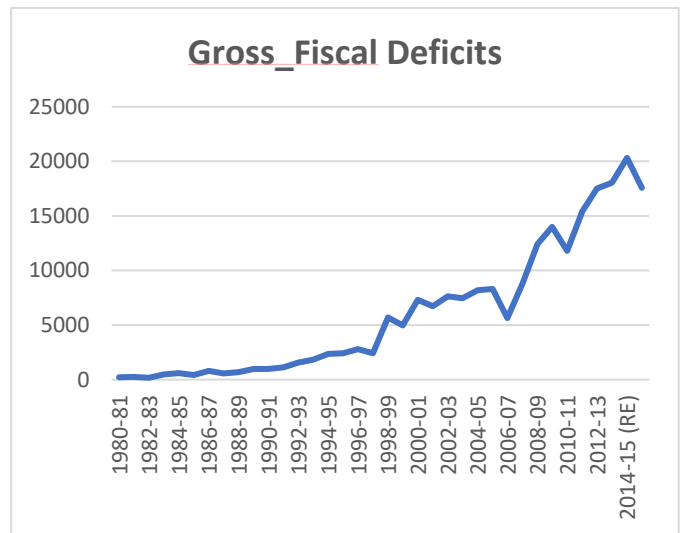
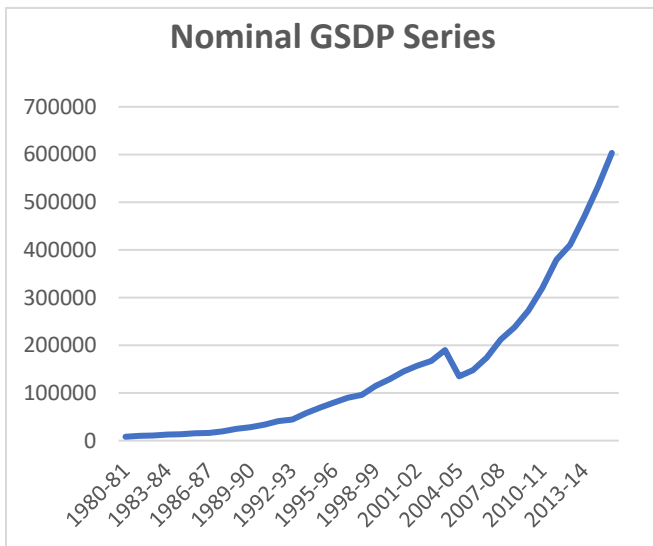
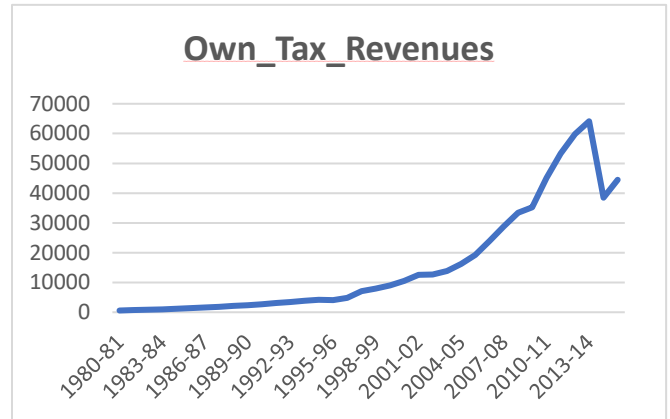
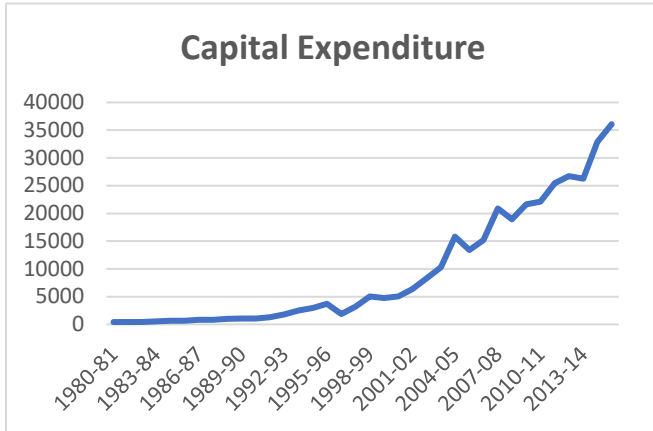


EXHIBIT 2

Naïve model Summary

	Capital Expenditure	Nominal GSDP Series	Own Tax Revenues	Social Sector Expenditure	Gross Fiscal Deficits
MAE	7388.568	159132.526	9848	10306	5970
Average Error	7388.568	159132.526	6896	10306	5970
MAPE	23.59216%	31.30259%	17.68460%	20.49149%	33.07675%
RMSE	8512.687351	178710.1966	11735.77096	10823.71378	6172.527845

EXHIBIT 3

$\alpha, \beta,$ and Γ values for the indicators for Double exponential smoothing

Indicators	Alpha	Beta	MAPE
Capital Expenditure	0.2	0.3	6.15
Nominal GSDP Series	0.7	0.3	15.55
Own_Tax_Revenues	0.5	0.1	23.87
Social_Sector Expenditure	0.7	0.15	9.03
Gross_Fiscal Deficits	0.5	0.3	12.02

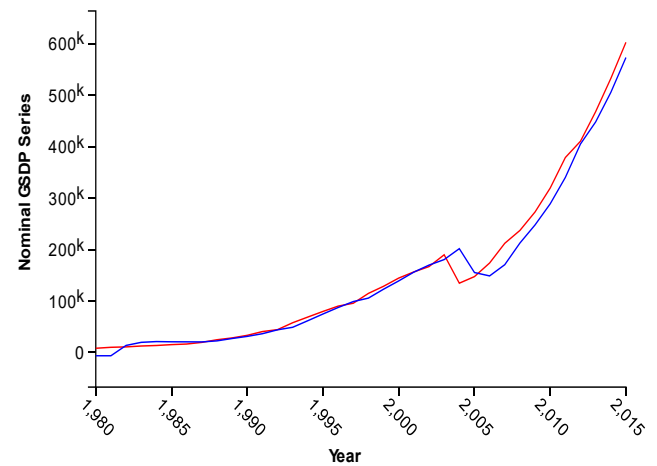
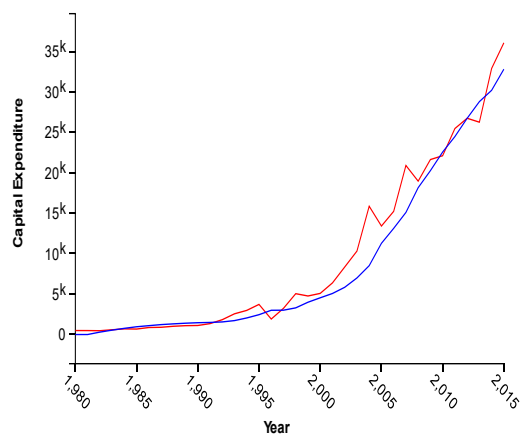
EXHIBIT 4

Data Set

Year	Capital Expenditure	Nominal GSDP Series	Own Tax Revenues	Social Sector Expenditure	Gross Fiscal Deficits
2016	35739.35592	649725.8429	51252.91577	55068.17439	20509.08519
2017	38009.22502	704949.185	53153.79418	57777.33435	21490.19325
2018	40279.09411	760172.5271	55054.6726	60486.49432	22471.3013
2019	42548.96321	815395.8693	56955.55102	63195.65428	23452.40935
2020	44818.83231	870619.2114	58856.42944	65904.81425	24433.5174

EXHIBIT 5

The predicted vs actual plots using exponential smoothing



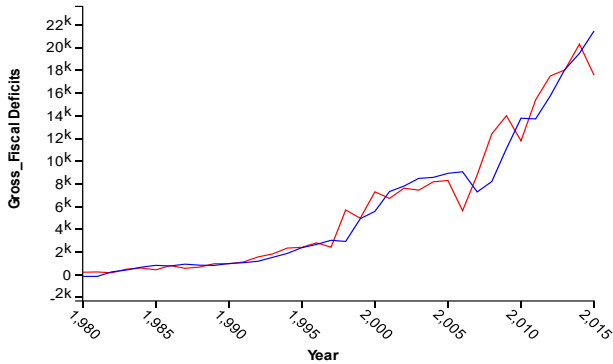


EXHIBIT 6

Error Summary for Linear Regression model with Exponential Trend

Indicators	MAE	Average Error	MAPE
Capital Expenditure	10754.22	-10754.2	0.355682
Nominal GSDP Series	70738.28	-70738.3	0.147114
Own Tax Revenues	16879.37	-10863.3	0.387175
Social Sector Expenditure	8036.148	-5707.92	0.16409
Gross Fiscal Deficits	10497.6	-10497.6	0.585627

EXHIBIT 7:

Double Exponential Smoothing (Various Alphas and Betas)

Beta = 0.15

MAPE/Alpha	Capital Expenditure	Nominal GSDP Series	Own_Tax_Revenues	Social_Sector Expenditure	Gross_Fiscal Deficits
0.7	9.64	21.76	25.83	9.03	17.3
0.5	9.06	25.87	25.51	11.37	16.6
0.2	14.5	35.09	24.4	28.1	22.2
0.1	31.3	42.62	35.3	43.26	31.1

Beta = 0.3

MAPE/Alpha	Capital Expenditure	Nominal GSDP Series	Own_Tax_Revenues	Social_Sector Expenditure	Gross_Fiscal Deficits
0.7	9.19	15.55	28.96	10.12	14.3
0.5	7.75	19.83	28.59	10.48	12.02
0.2	6.15	32.98	27.28	17.59	18.3
0.1	16.13	36.63	25.82	32.71	20.97

Beta = 0.1

MAPE/Alpha	Capital Expenditure	Nominal GSDP Series	Own_Tax_Revenues	Social_Sector Expenditure	Gross_Fiscal Deficits
0.7	11.51	24.32	24.11	10.9	19.18
0.5	11.43	28.25	23.87	15.42	18.92
0.2	20.84	38.26	24.96	33	25.93
0.1	37.77	45.89	41.19	47.79	36.27

