Online Hackathon On Data Driven Innovation For Citizen Grievance Redressal

- Prateek Khanna

DataSets

- Public Grievance details in CPGRAMS along with feedback details
- Public Grievance movement details across the organization in CPGRAMS
- Monthly Department-wise public grievance receipts and disposals from January 2016 to Octorber 2019
- Department-wise receipts, disposal and pendency of Public Grievance detailed statistics from 01.01.2016 to 01.11.2019
- Details of registered users of CPGRAMS as on 24.10.2019
- CPGRAMS Nodal Public Grievance Officers Details as on 30th October 2019

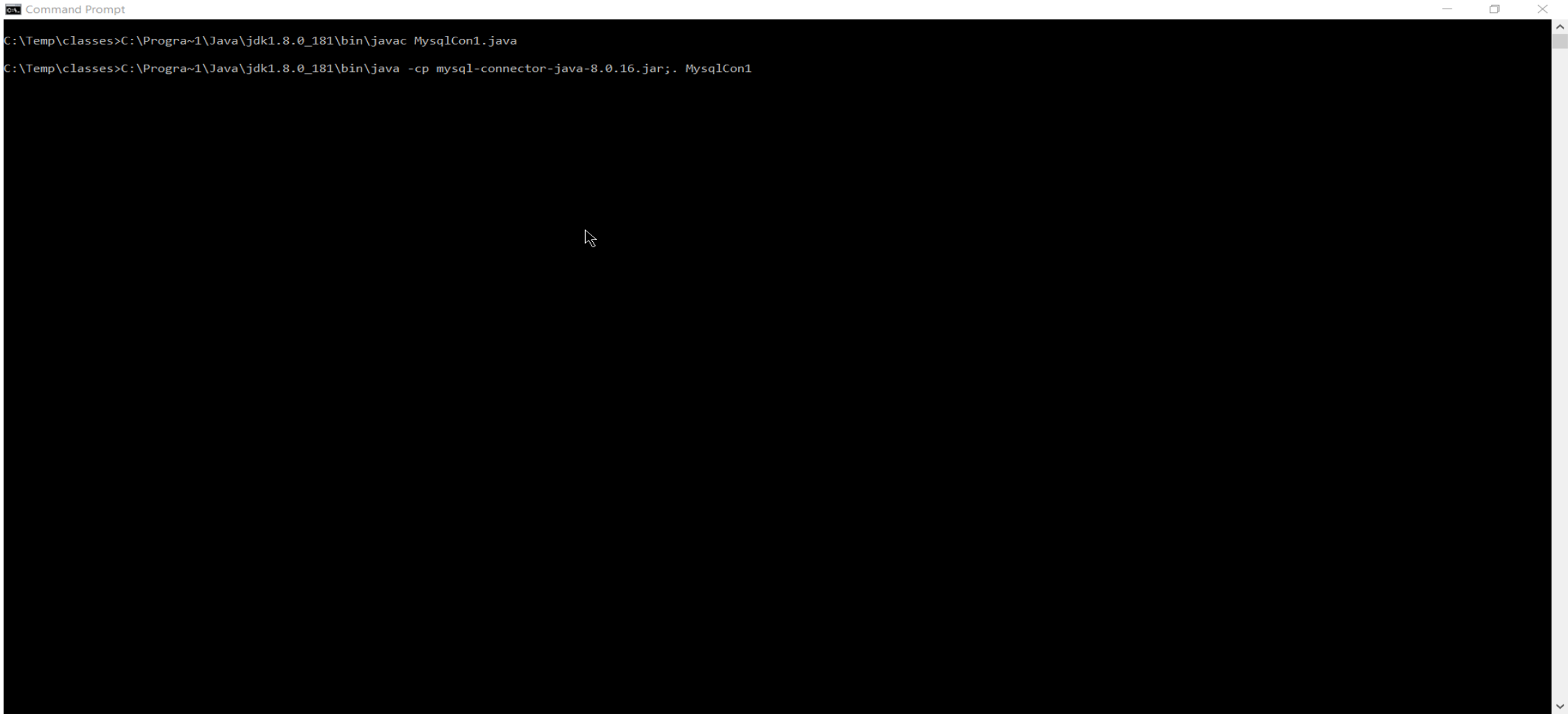
Predictive Models

- The implementation falls under the category of Application Predictive Models.
- It makes use of the shared datasets and predicts future trends for the relevant metrics.
- The time series data is analysed Auto Regressive Integrated Moving Average(ARIMA) methodology.
- Source Data is populated in a MySQL database and java code is used for generating relevant csv files.
- These csv files are then used as input to a R program to generate future predictions.

Setup

- Source code repository : https://github.com/solfinder/predictivemodels
- Deployment steps: https://github.com/solfinder/predictivemodels/blob/master/README.md
- Java Class : MySqlCon1
- R module : darpg.r
- Time Series Predictive Model : ARIMA
- Prediction data output format : csv files containing data for future months
- Demo : https://youtu.be/otq0YXMwv1g

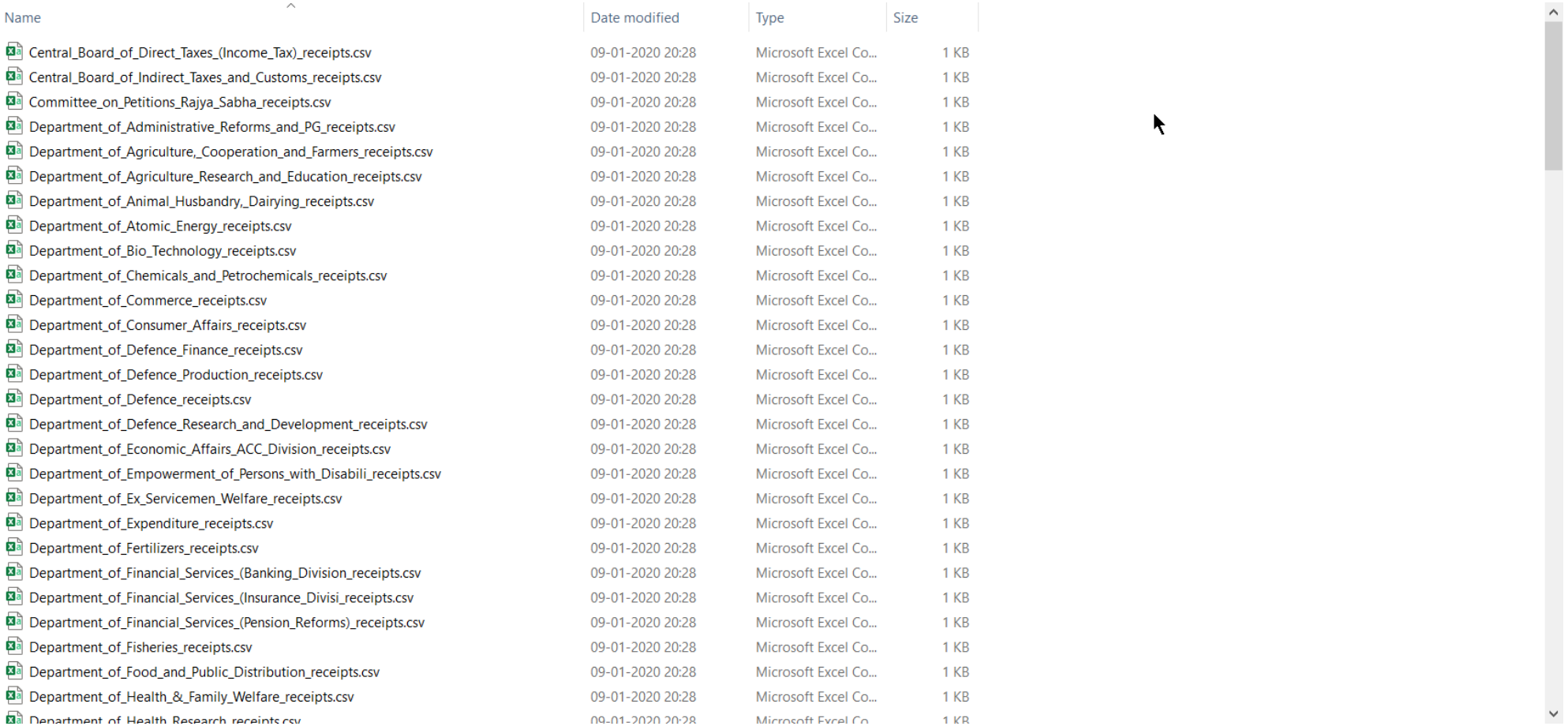
Predictive Models – Java Preprocessing



Predictive Models – Java Preprocessing

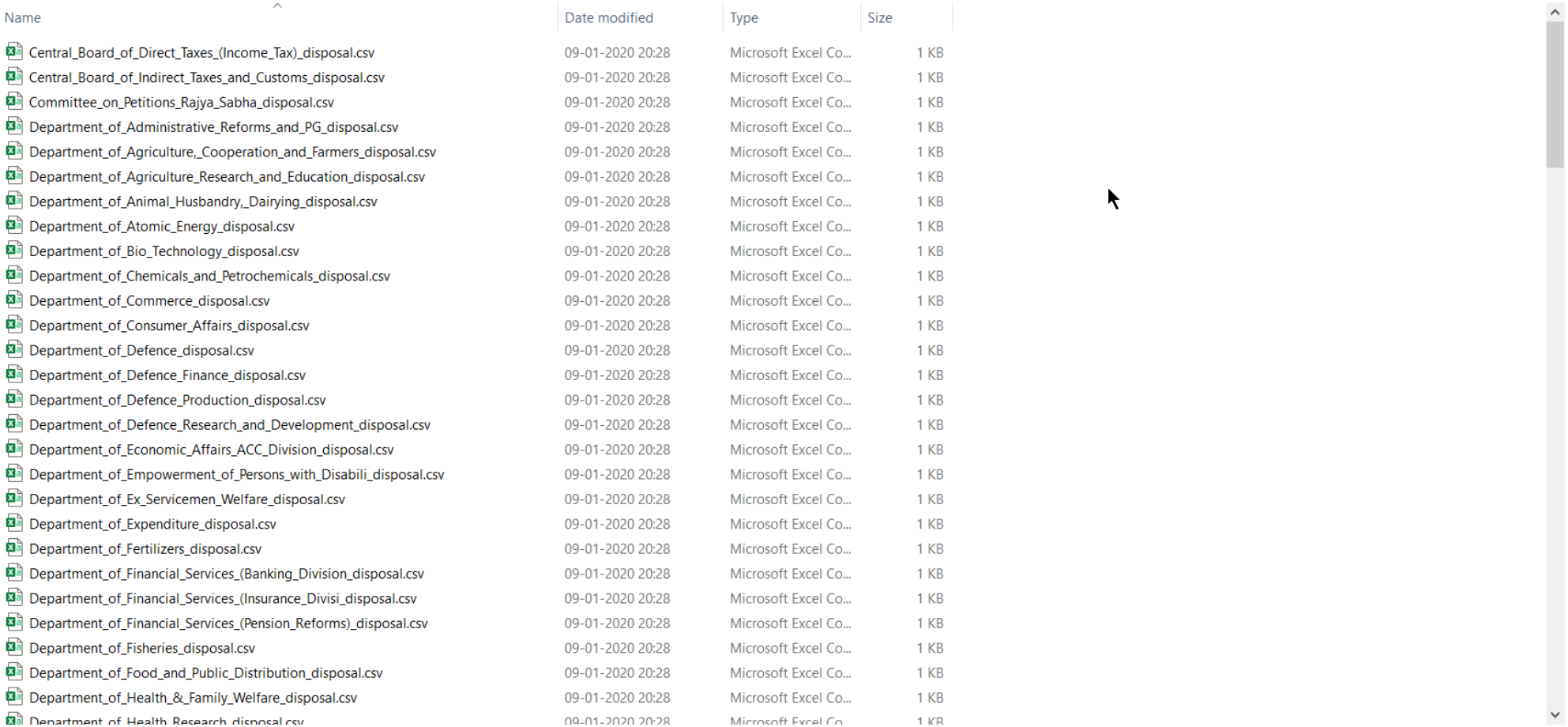
- Java Preprocessing Class MysqlCon1 connects to the database table containing raw data about grievance receipts and disposals and transforms it into a set of csv files organized by departments.
- Individual csv file is generated for each department for both receipts and disposal of grievances.
- These csv files are used as input for creating time series models for the respective departments.
- The csv files contain metrics from a start month to an end month.
- Prediction model computes the metrics for the subsequent months.

Predictive Models – Java Preprocessing



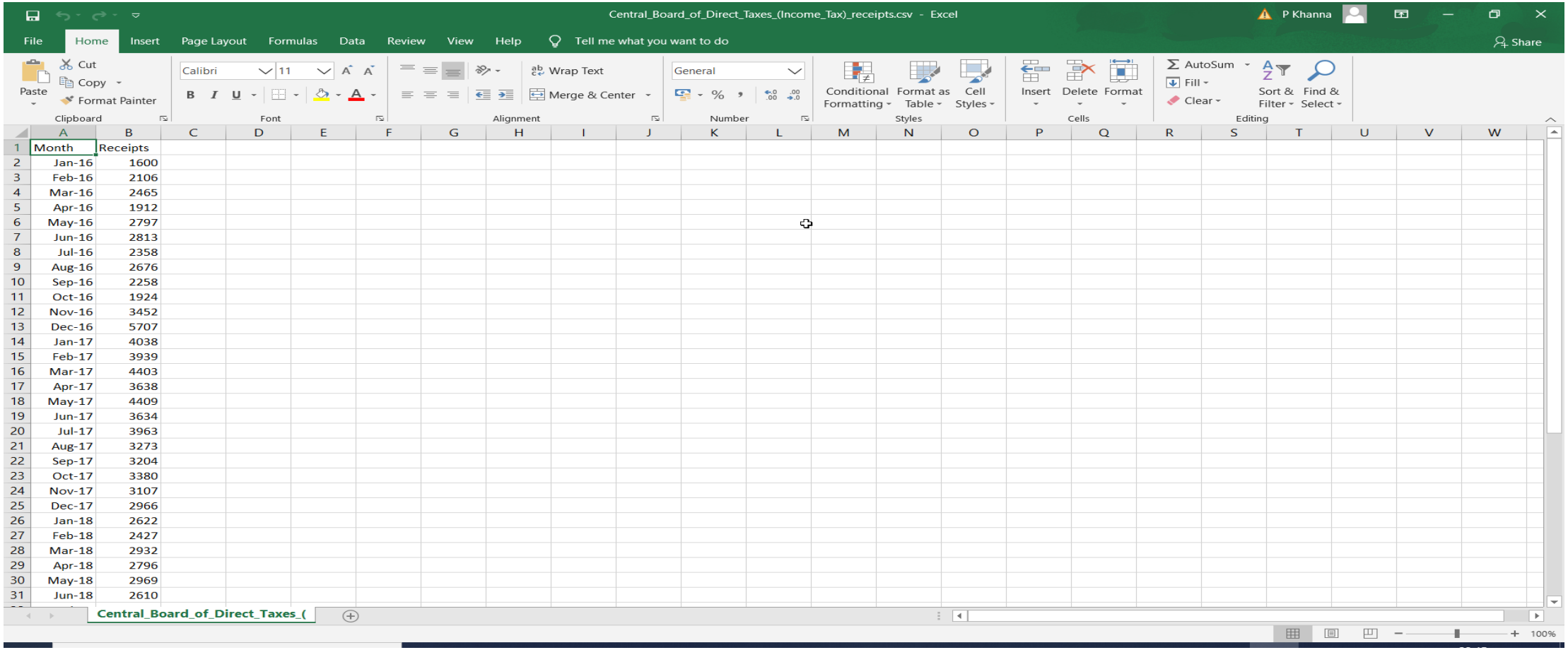
Name	Date modified	Type	Size
Central_Board_of_Direct_Taxes_(Income_Tax)_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Central_Board_of_Indirect_Taxes_and_Customs_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Committee_on_Petitions_Rajya_Sabha_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Administrative_Reforms_and_PG_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Agriculture,_Cooperation_and_Farmers_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Agriculture_Research_and_Education_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Animal_Husbandry,_Dairying_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Atomic_Energy_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Bio_Technology_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Chemicals_and_Petrochemicals_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Commerce_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Consumer_Affairs_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_Finance_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_Production_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_Research_and_Development_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Economic_Affairs_ACC_Division_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Empowerment_of_Persons_with_Disabili_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Ex_Servicemen_Welfare_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Expenditure_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Fertilizers_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Financial_Services_(Banking_Division_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Financial_Services_(Insurance_Divisi_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Financial_Services_(Pension_Reforms)_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Fisheries_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Food_and_Public_Distribution_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Health_&_Family_Welfare_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Health_Research_receipts.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB

Predictive Models – Java Preprocessing



Name	Date modified	Type	Size
Central_Board_of_Direct_Taxes_(Income_Tax)_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Central_Board_of_Indirect_Taxes_and_Customs_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Committee_on_Petitions_Rajya_Sabha_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Administrative_Reforms_and_PG_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Agriculture,_Cooperation_and_Farmers_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Agriculture_Research_and_Education_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Animal_Husbandry,_Dairying_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Atomic_Energy_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Bio_Technology_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Chemicals_and_Petrochemicals_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Commerce_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Consumer_Affairs_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_Finance_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_Production_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Defence_Research_and_Development_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Economic_Affairs_ACC_Division_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Empowerment_of_Persons_with_Disabili_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Ex_Servicemen_Welfare_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Expenditure_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Fertilizers_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Financial_Services_(Banking_Division_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Financial_Services_(Insurance_Divisi_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Financial_Services_(Pension_Reforms)_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Fisheries_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Food_and_Public_Distribution_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Health_&_Family_Welfare_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB
Department_of_Health_Research_disposal.csv	09-01-2020 20:28	Microsoft Excel Co...	1 KB

Predictive Models – Java Preprocessing



	A	B
1	Month	Receipts
2	Jan-16	1600
3	Feb-16	2106
4	Mar-16	2465
5	Apr-16	1912
6	May-16	2797
7	Jun-16	2813
8	Jul-16	2358
9	Aug-16	2676
10	Sep-16	2258
11	Oct-16	1924
12	Nov-16	3452
13	Dec-16	5707
14	Jan-17	4038
15	Feb-17	3939
16	Mar-17	4403
17	Apr-17	3638
18	May-17	4409
19	Jun-17	3634
20	Jul-17	3963
21	Aug-17	3273
22	Sep-17	3204
23	Oct-17	3380
24	Nov-17	3107
25	Dec-17	2966
26	Jan-18	2622
27	Feb-18	2427
28	Mar-18	2932
29	Apr-18	2796
30	May-18	2969
31	Jun-18	2610

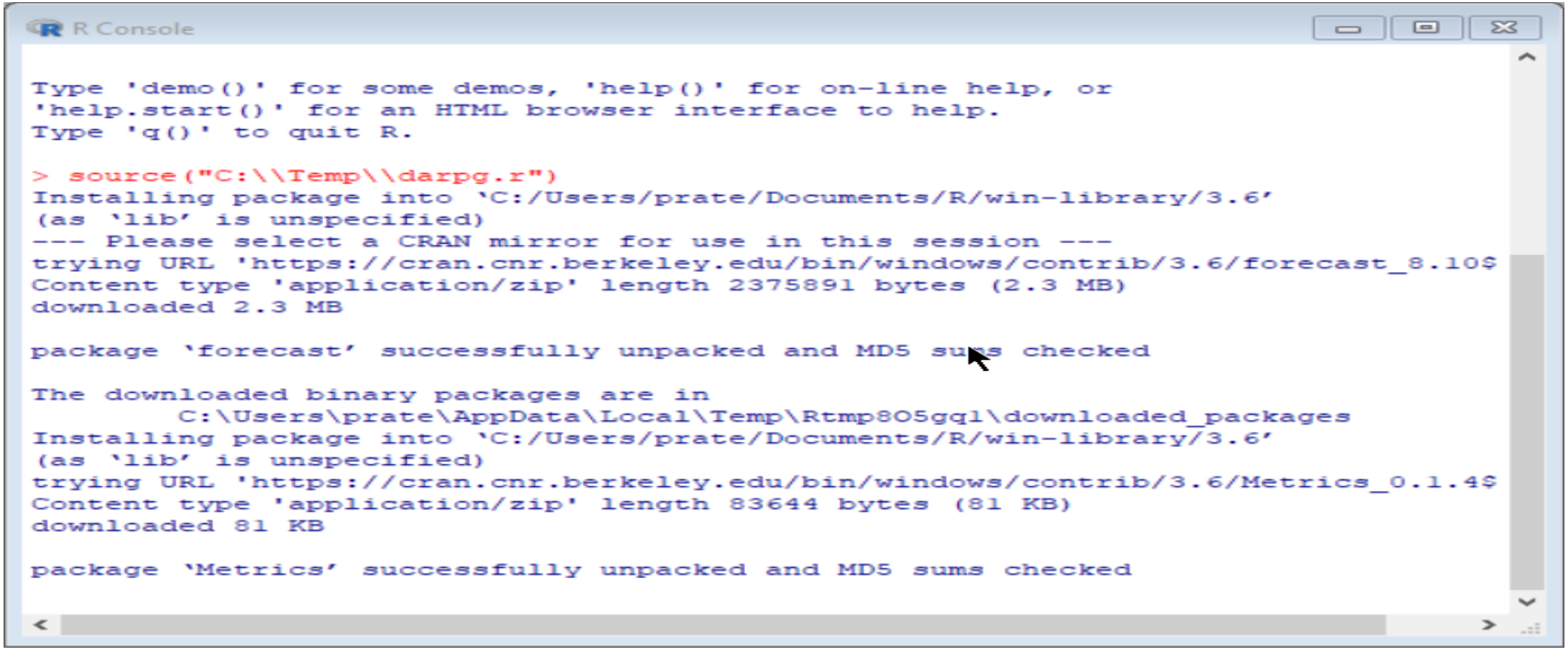
Predictive Models – R TimeSeries Prediction Model

RGui (32-bit) - [C:\Temp\darpg.r - R Editor]

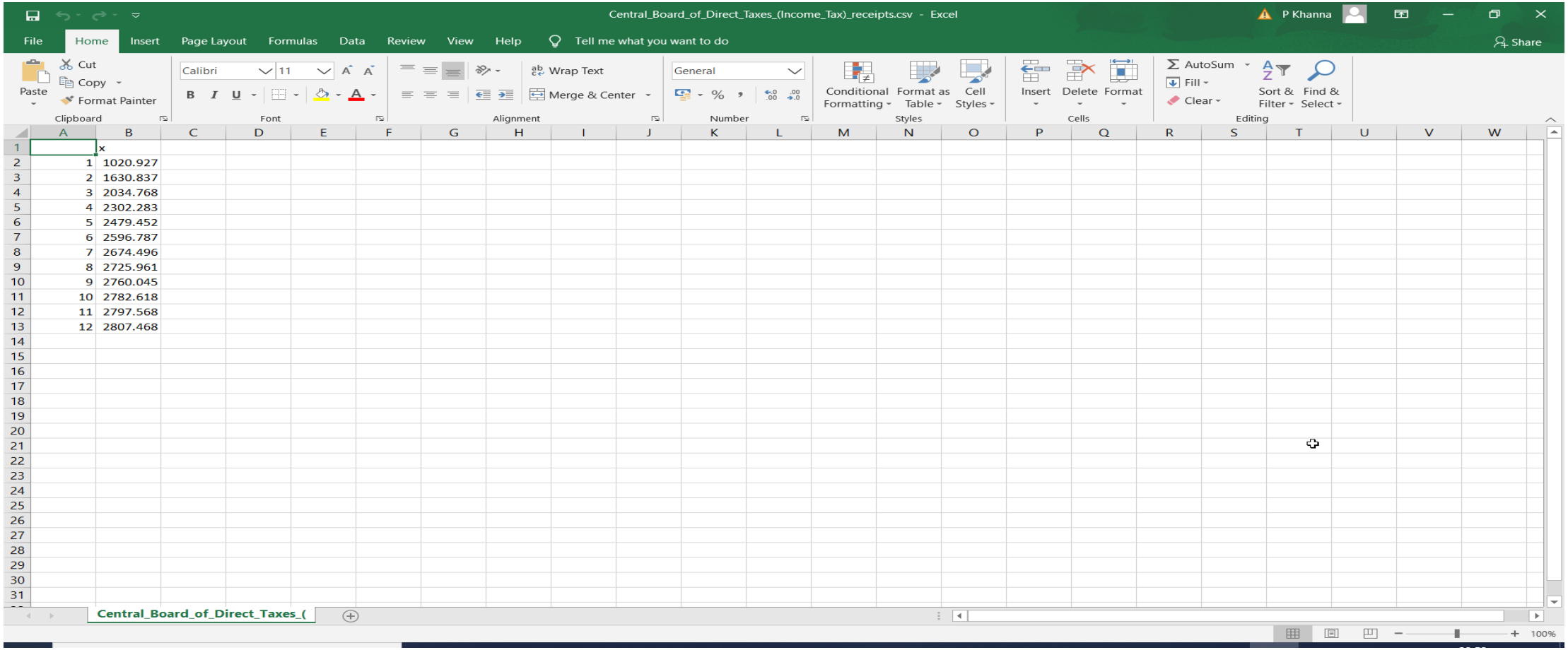
File Edit Packages Windows Help

```
# R Module for timeseries forecasting of received and disposed grievances
# Developed by Prateek Khanna

install.packages("forecast")
install.packages("Metrics")
# loading packages
library(forecast)
library(Metrics)
library(ggplot2)

#Default forecast period is set to 12
#Update this value to generate forecasts for different periods
forecast_period=12

data1 = read.csv("C:\\temp\\classes\\csv_receipts\\NITI_Aayog_receipts.csv")

nm <- list.files(path="C:\\temp\\classes\\csv_receipts")
for(filename in nm){
# reading data from csv file
data = read.csv(paste("C:\\temp\\classes\\csv_receipts\\",filename,sep=""))
#ggplot(data, aes(x=Month, y=Receipts)) + geom_point()
#ggsave(paste(paste("C:\\temp\\classes\\csv_receipts_graph\\",filename,sep=""),".png",sep=""),width = 25, height = 25)
# training data sets
train = data[1:nrow(data),]

# removing "Month" column
train$Month = NULL

# training model
model = auto.arima(train)

# model summary
sink(paste(paste("C:\\temp\\classes\\csv_receipts_predictionmodel\\",filename,sep=""),".txt",sep=""))
print(summary(model))
sink()

# forecasting
forecast = predict(model,forecast_period)

write.table(forecast$pred, file = paste("C:\\temp\\classes\\csv_receipts_prediction\\",filename,sep=""), sep = ",",col.names=NA)

#ggplot(data2, aes(x=Month, y=Receipts)) + geom_point()
#ggsave(paste(paste("C:\\temp\\classes\\csv_receipts_prediction_graph\\",filename,sep=""),".png",sep=""))
}

nm <- list.files(path="C:\\temp\\classes\\csv_disposal")
for(filename in nm){
```

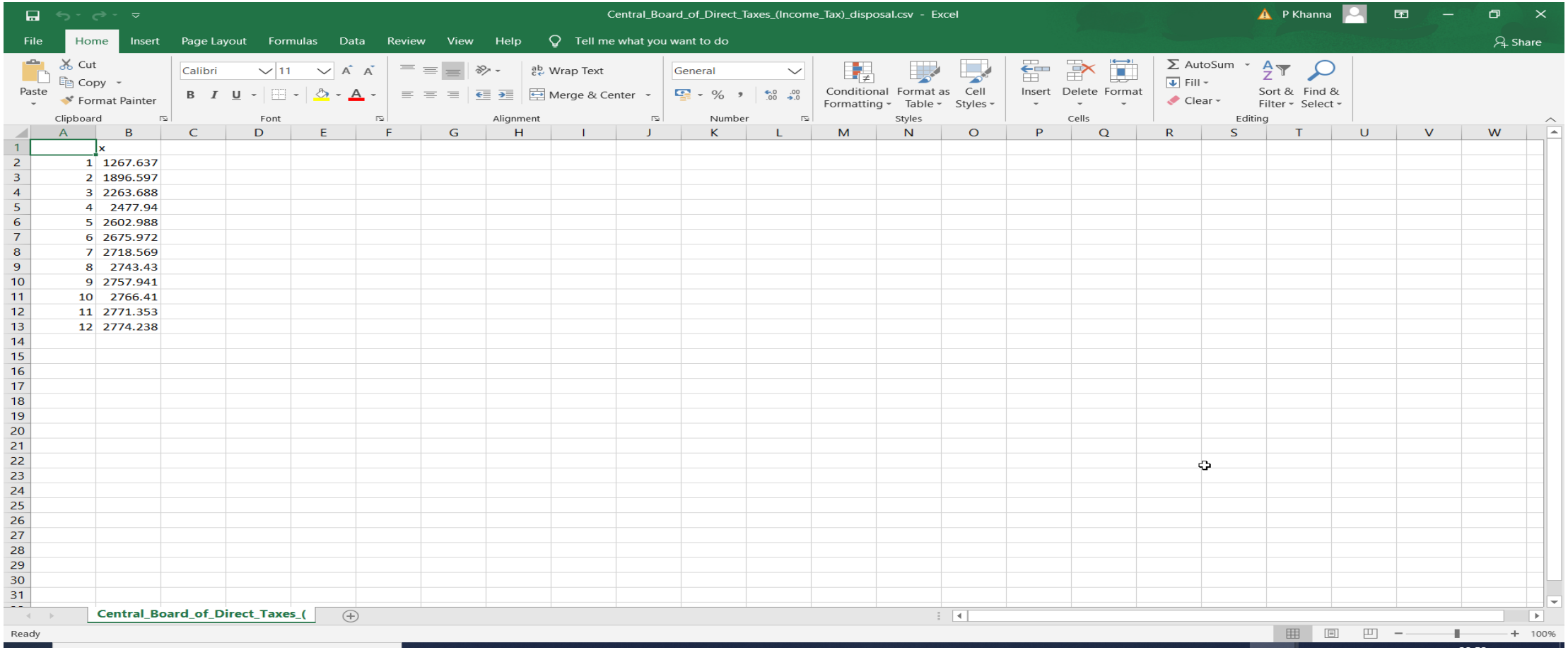
Predictive Models – R TimeSeries Prediction Model



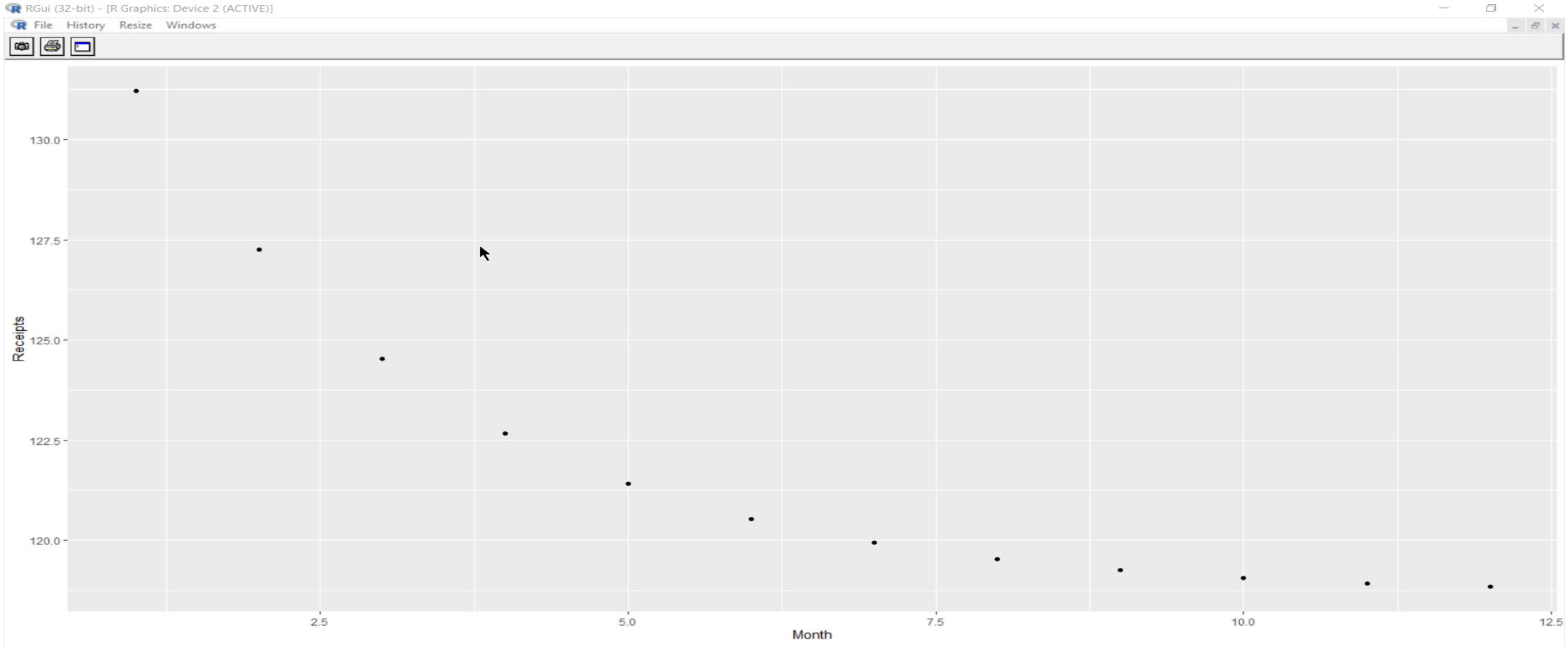
Predictive Models – R TimeSeries Prediction Example



Predictive Models – R TimeSeries Prediction Example



Predictive Models – R TimeSeries Prediction Example



Predictive Models – R TimeSeries Prediction Model Details

Department_of_Bio_Technology_receipts.csv.txt - Notepad

File Edit Format View Help

```
Series: train
ARIMA(0,0,0) with non-zero mean

Coefficients:
          mean
      20.9767
s.e.   1.6603

sigma^2 estimated as 121.4:  log likelihood=-163.68
AIC=331.36   AICc=331.66   BIC=334.88

Training set error measures:
                      ME     RMSE     MAE       MPE     MAPE      MASE
Training set -1.817377e-15 10.88735 7.92861 -30.49891 52.00143 0.7909777
                    ACF1
Training set 0.1981748
                      ME     RMSE     MAE       MPE     MAPE      MASE
Training set -1.817377e-15 10.88735 7.92861 -30.49891 52.00143 0.7909777
                    ACF1
Training set 0.1981748
```

Predictive Models – R TimeSeries Prediction Model Details

Central_Board_of_Direct_Taxes_(Income_Tax)_disposal.csv.txt - Notepad

File Edit Format View Help

```
Series: train
ARIMA(1,0,0) with non-zero mean

Coefficients:
          ar1       mean
      0.5836  2778.2818
s.e.  0.1454   328.5908

sigma^2 estimated as 901085:  log likelihood=-388.09
AIC=782.19   AICc=782.74   BIC=787.74

Training set error measures:
                    ME     RMSE      MAE       MPE     MAPE      MASE        ACF1
Training set 44.8403 928.8383 651.8077 -134.768 153.4398 0.9129238 -0.1436935
                    ME     RMSE      MAE       MPE     MAPE      MASE        ACF1
Training set 44.8403 928.8383 651.8077 -134.768 153.4398 0.9129238 -0.1436935
```

Conclusion

- The application makes use of ARIMA modelling in R to predict future datapoints.
- Details of the actual model characteristics for each series are also captured and reported for reference.
- Predicted data can be used for:
  - Optimum resource utilization
  - Reallocation of resources is required
  - Investment strategy formulation
  - Contingency planning