**Sales Forecasting Python**

*By Ong Qiao Hui on 12 January 2024.*

# Dataset [Data.csv]

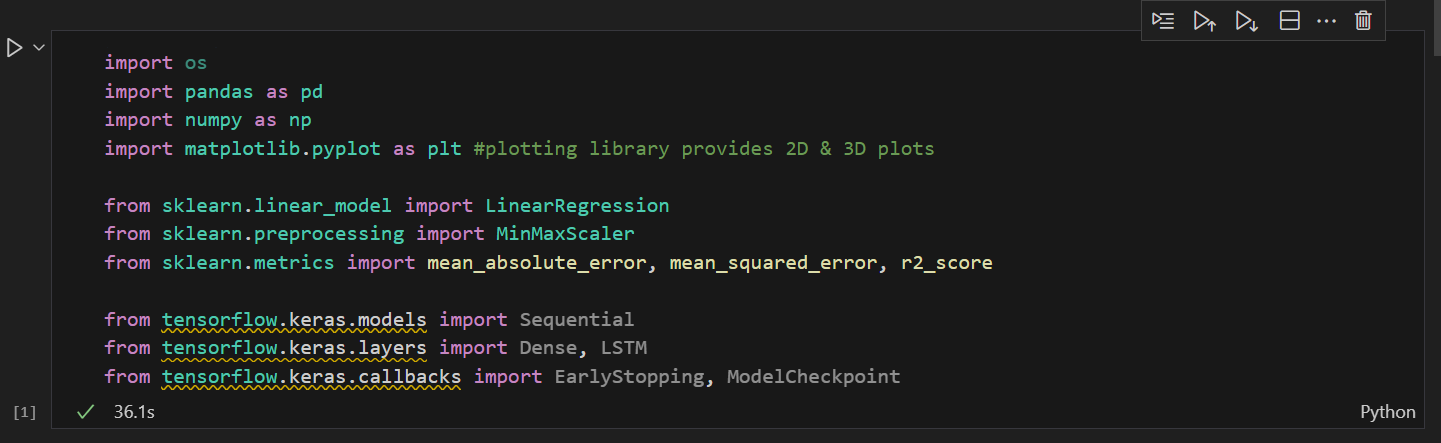
Daily historical data from July 2021 to January 2024.

* sales\_date– date in format dd/mm/yyyy
* sales\_amount– total sales for each day

## **Linear Regression**

**Goal: Get the predictions for monthly sales in future.**

1. Import all the library that would be used.



1. Read csv file and display the first 10 rows.

A screenshot of a computer

Description automatically generated

1. Check info from the dataset to know that does it contain null value and how much data is contained in the dataset.

A screenshot of a computer

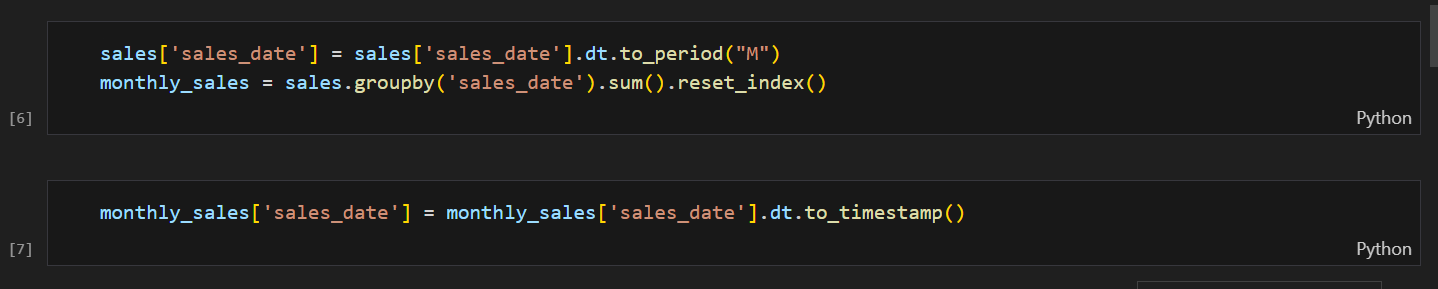
Description automatically generated

1. Converting data type of date from Object to DateTime.

A screen shot of a computer

Description automatically generated

1. Converting Date to a Month period, then sum the number of items in each month.



1. Visualization for the data.

A graph on a computer screen

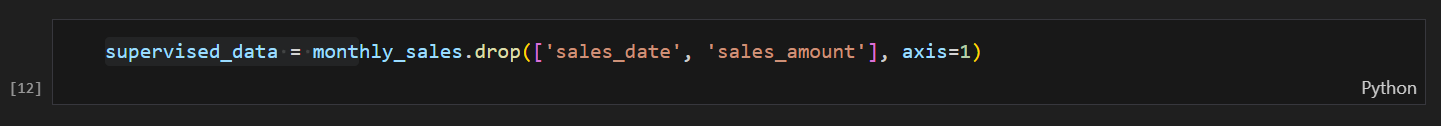
Description automatically generated

1. Call the difference on the sales columns to make the sales data stationary. (Difference between last month)

A screenshot of a computer

Description automatically generated

1. Dropping off sales\_date and sales\_amount because now only need Stationary data.



1. Prepare Supervised Data (Previous 12 months sales will be the input features, Next 12 months sales will be the output)

A black background with colorful text

Description automatically generated

1. Split data into Train & Test data.

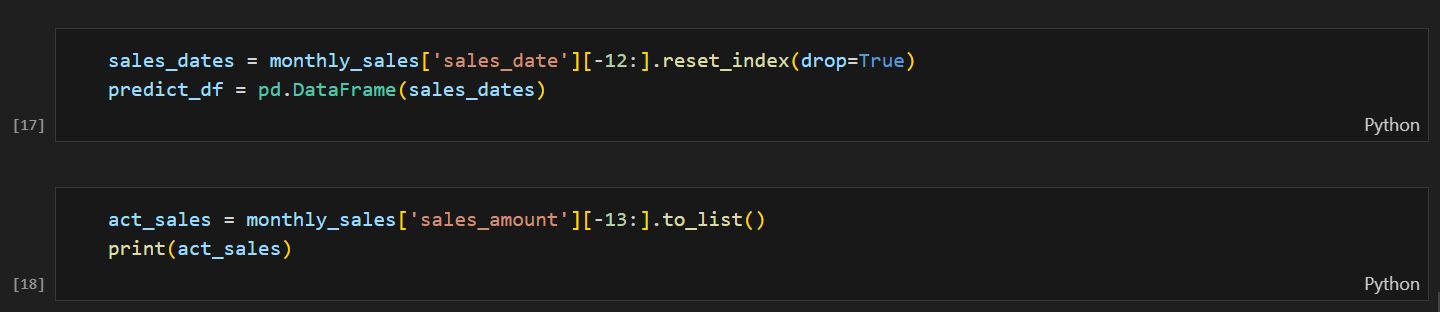
A screen shot of a computer

Description automatically generated

A screen shot of a computer

Description automatically generated

1. Make prediction data frame to merge the predicted sales prices of all trained algorithms.



1. Create Linear Regression Model and Predict Output.

A screenshot of a computer program

Description automatically generated

A screen shot of a computer program

Description automatically generated

1. Visualization of the prediction against the actual sales.

A screen shot of a computer screen

Description automatically generated

1. Get the predicted sales for every month.

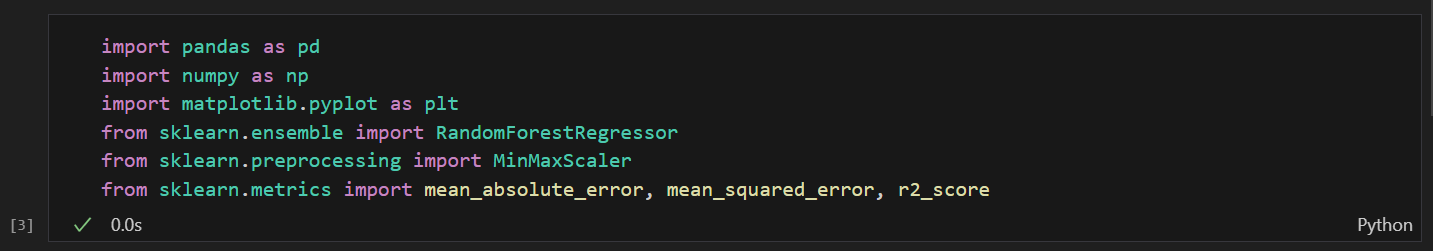
A screenshot of a computer

Description automatically generated

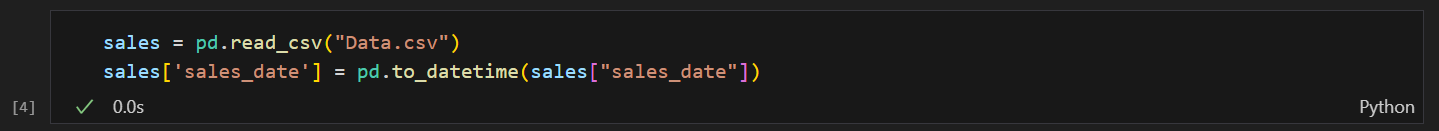
## **Random Forrest**

**Goal: Get the predictions for monthly sales in future.**

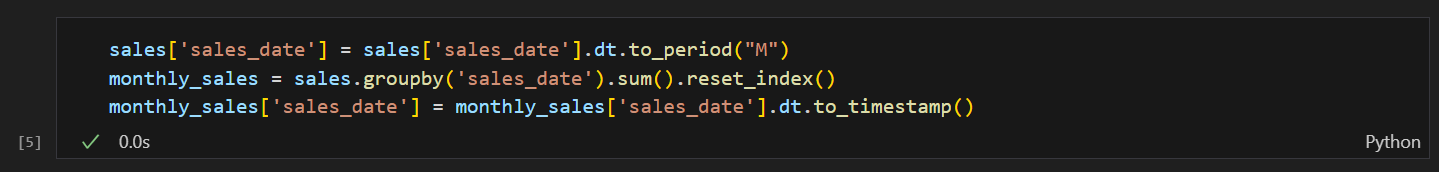
1. Import libraries.



1. Read dataset (.csv)



1. Converting Date to a Month period, then sum the number of items in each month.

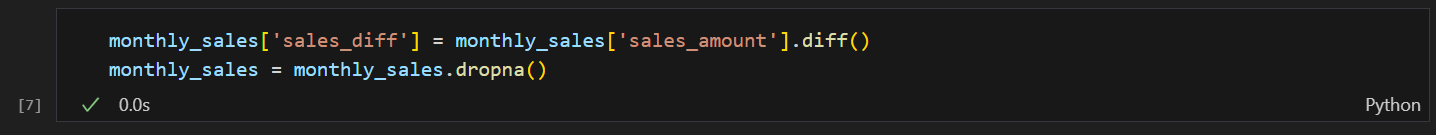


1. Visualize Monthly Customer Sales

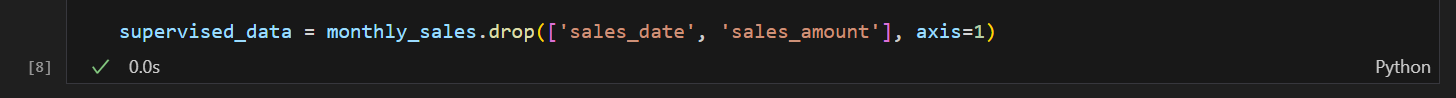
A screen shot of a graph

Description automatically generated

1. Call the difference on the sales columns to make the sales data stationary. (Difference between last month)



1. Dropping off sales\_date and sales\_amount because now only need Stationary data.

****

1. Prepare Supervised Data (Previous 12 months sales will be the input features, Next 12 months sales will be the output)

**A black screen with colorful text

Description automatically generated**

1. Split data into Train and Test data

A screenshot of a computer program

Description automatically generated

1. Create Random Forrest Regressor Model

A screenshot of a computer program

Description automatically generated

A screen shot of a computer

Description automatically generated

1. Visualization of the prediction against the actual sales.

A screen shot of a graph

Description automatically generated

## **XGBoost**

**Goal: Get the predictions for monthly sales in future.**

1. Import Libraries and read dataset.

A screenshot of a computer program

Description automatically generated

1. Converting date from Object to dateTime

A screen shot of a computer

Description automatically generated

1. Converting Date to a Month period, then sum the number of items in each month

A screen shot of a computer code

Description automatically generated

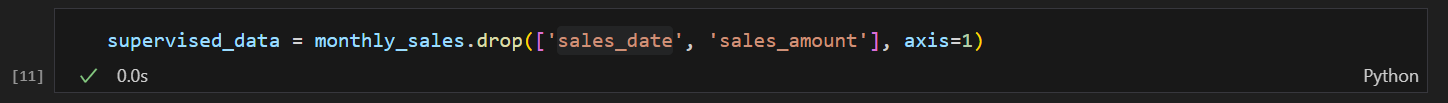
1. A screen shot of a graph

   Description automatically generatedVisualization
2. Call the difference on the sales columns to make the sales data stationary (Difference between last month)

A screenshot of a computer

Description automatically generated

1. Dropping off sales\_date and sales\_amount because now only need Stationary data



1. Prepare Supervised Data (Previous 12 months sales will be the input features, Next 12 months sales will be the output)

A computer code on a black background

Description automatically generated

1. Split data into Test and Train data

A screen shot of a computer program

Description automatically generatedA screen shot of a computer program

Description automatically generated

1. Create XGBRegressor Model

A screenshot of a computer program

Description automatically generated

A screen shot of a computer

Description automatically generated

1. Visualization of the prediction against the actual sales.

A screen shot of a computer

Description automatically generated

# **Dataset [Sales\_datasetA.csv]**

Daily historical data from June 2020 to December 2023.

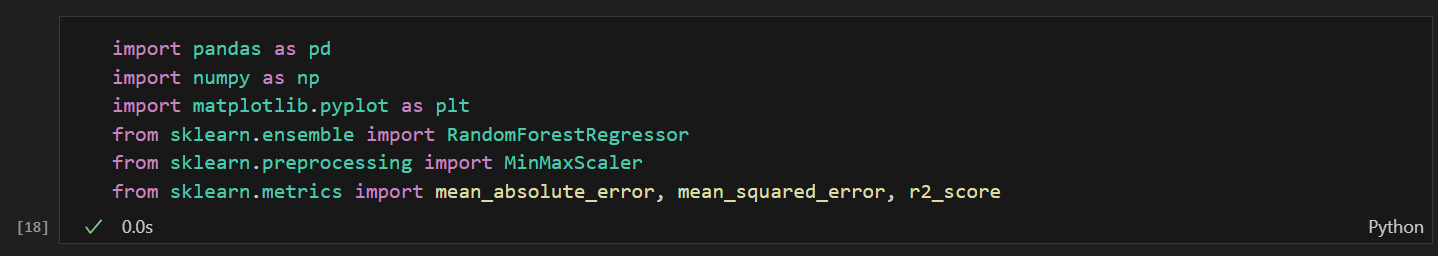
Columns will be used:

* Order\_Date – date in format yyyy-mm-dd
* Order\_TotalSales - total sales for each day

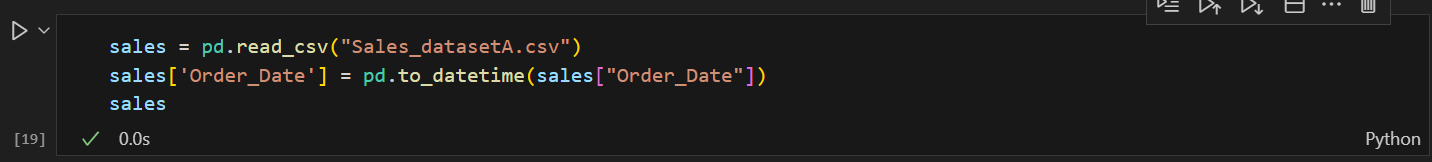
## **Random Forrest**

**Goal: Get the predictions for daily sales in future.**

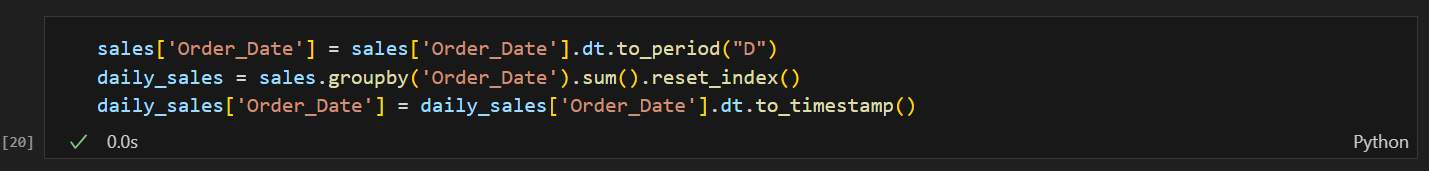
1. Import libraries.



1. Read dataset.



1. Converting Date to a Day period, then sum the number of items in each day.



1. A screen shot of a graph

   Description automatically generatedVisualization
2. Call the difference on the sales columns to make the sales data stationary and drop the unuse columns

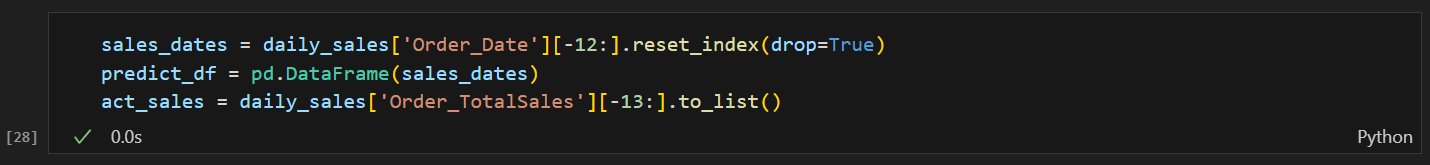
A screen shot of a computer

Description automatically generated

1. Prepare Supervised Data

A screen shot of a computer program

Description automatically generated



1. Create Random Forrest Regressor Model

A screenshot of a computer program

Description automatically generated

A screen shot of a computer

Description automatically generated

1. Visualization of the prediction against the actual sales.

A computer screen shot of text

Description automatically generated

A graph with blue lines

Description automatically generated

1. Show the predicted sales amount for each day.

A screenshot of a computer

Description automatically generated

## **Time Series (XGBoost)**

**Goal: Get the predictions for daily sales in future.**

1. Import libraries.

A black rectangular object with a black border

Description automatically generated

1. Read dataset and drop the unuse columns.

A screenshot of a computer

Description automatically generated

1. Visualization

A screen shot of a computer

Description automatically generated

1. Split data into Test and Train data

A screen shot of a computer screen

Description automatically generated

1. Create time series for the data

A screen shot of a computer program

Description automatically generated

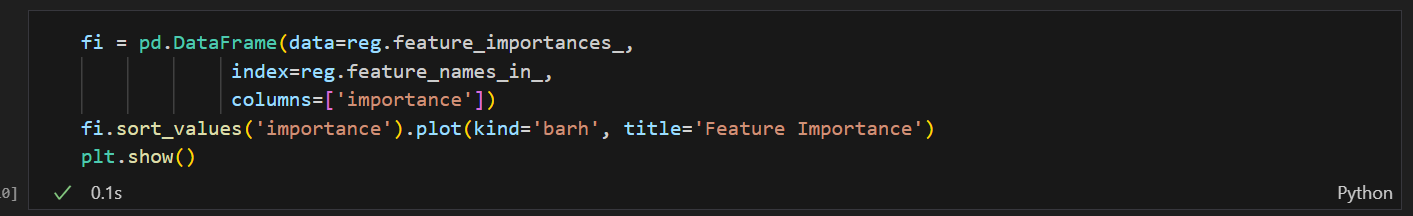
A graph of a graph with blue and black squares

Description automatically generated with medium confidence

1. Add features and target to check which features is important.

A screenshot of a computer program

Description automatically generated



A graph with blue squares

Description automatically generated

1. Prediction

A screen shot of a graph

Description automatically generated

1. Predict sales for 1 week

A graph with a line going up

Description automatically generated