Data Pre-processing

1 Missing Data Imputation

The number of weekly sales data should be the same in one loop for every departments. However, some departments may have missing data for several weeks. So, before prediction, we are supposed to implement the missing value. We use the historical average weekly sales of the store which this department belongs to as a replacement of the missing data.

2 Processing on data missing historical information

The number of weekly sales data should be the same in one loop for every departments. However, some departments may have missing data for several weeks. So, before prediction, we are supposed to implement the missing value. We use the historical average weekly sales of the store which this department belongs to as a replacement of the missing data.

Models

1 Simple Model

In Simple Model, we predict new sales by sales of around the same time of last year. We look at the same week, the week before, and the week after, and take their mean as our prediction of this year's week. If there is no historical data for a particular department in a particular store, we predict the sales by average sales history of the same entire store. If the week we are going to predict is a holiday, we use the exact sales of the same holiday week last year as our prediction.

2 Exponential Smoothing Forecast Model

In exponential smoothing forecast model, we recursively applying with exponential window functions to filter the noise in the time series data. We use ses() to fit a ets model with an adjustment of alpha parameter, which is a smoothing parameter for the level, to be 0.2 and leave beta, which is smoothing parameter for the trend, to be estimated by the machine. For the first a few observations, we set to use simple calculations to get the initial state values.

3 Default Time Series Model

Model3 is a simple time series model. When doing the prediction, the historical sales data are used to build a time series data with frequency of 52, which stand for one year. Instead of choosing ARIMA or Holt-Winters, we chose no extra models. So, this model can be applied to the whole prediction process because it does not have the request for more than two periods. However, this decision will trigger to an obvious defect. The accuracy of this model can be relatively low.

Model Comparison Based on Run Time and Accuracy

	Simple Model	**SES**	**Model3**
Error in each loop	2451.859	2451.859	4960.313
	2908.91	2908.91	4996.225
	2377.459	2377.459	5214.679
	2421.905	2421.905	5120.961
	2379.612	2379.612	5116.673
	2402.241	2402.241	5615.942
	4060.876	4060.876	10727.13

	2361.881	2361.881	5099.94
	5947.76	6145.207	18896.89
	6703.222	6735.277	15224.53
	2172.772	2172.772	7955.828
	5136.001	5136.001	12603.71
	2230.944	5356.168	4902.137
	3015.576	5501.213	5173.562
	2325.058	5359.662	4858.935
	2350.173	5409.973	4975.797
	2299.9	5401.882	4815.99
	2227.491	5711.196	5029.119
	4656.822	11387.06	10059.45
	2104.234	5221.378	4710.046
WMAE	3126.735	4545.126	7302.893
Runing Time	612.0757	1051.37	1561.306

We can find from the table that simple model is the fast and most accuarccy model.