five parts:  
eda: graphs and conclusion

fe:  
9 parts:  
clustering and remove outlier release and remove outliner event related features  
price related features  
time related features  
lag and rolling features quantile features aggregation features weather related features

model:  
5 models:

clsutering

lightgbm

xgboost

catboost

cnn

strategies:  
4 split strategies:  
by store  
by cluster  
by state + category  
by department  
2 weights strategies:  
add weight price  
add weight year related trend  
3 fold time series split strategies and replacement by store

final submission:  
build four models based on the smallest wrmsse in stage 1 and make prediction for stage 2