

ブートキャンプ° for Data Science

5回目講義資料

Tryeting Inc. ながえゆうき

アジェンダ

- イン트로ダクション
 - ラスト！エンディングに向けて…！
 - 宿題：時系列データの処理
- 早速やってみよう
 - twitterデータの取得
 - 前処理
 - 解析・可視化

- 前回までで、できるようになったこと
- 回帰・分類の機械学習は計算できる
- データ取得して、自分でモデル構築・予測までできる
- お客様に向けて、機械学習を用いたご提案ができる
- **データの前処理による精度の違いを説明できる**

今日の宿題

(デンソーさんの)株価予測しませんか？

- 株価取得から予測の流れ
- 株価のスクレイピング
- 株価モデリング
 - データ前処理
- 株価予測
- 時系列データの評価
 - 決定係数(R^2)
 - 売買シミュレーション

キーワード：データ前処理、関数、評価

- 株価取得…WEBで落ちてるもの使おうぜ：今回はこちら

株価データサイト k-db.com

トップ先物指数個別株統計

Google

あなたの履歴書に
Googleの公式認定証を。



デジタル
マーケティングの
知識やスキルを
無料で学べるe
ラーニング。



採用コストを抑えたいあなたへ

完全無料の採用支援ツール 

採用サイト作成

求人掲載

応募者管理ツール

完全無料! 詳しくはこちら >

当サイトは2017年12月末でサービス終了です。

■先物

日経平均先物やTOPIX先物などのデータ 4本値+出来高+売買代金 がCSVダウンロードできます。
セッションごとデータ 日中・夜間
各銘柄ごとの時系列データ 日夜・4時間足・1時間足・30分足・15分足・5分足

■指数

日経平均やマザーズ指数など株価指数の4本値データ がCSVダウンロードできます。
セッションごと 日足・前場・後場
各指数ごとの時系列データ 日足・前後場・1時間足・30分足・15分足・5分足

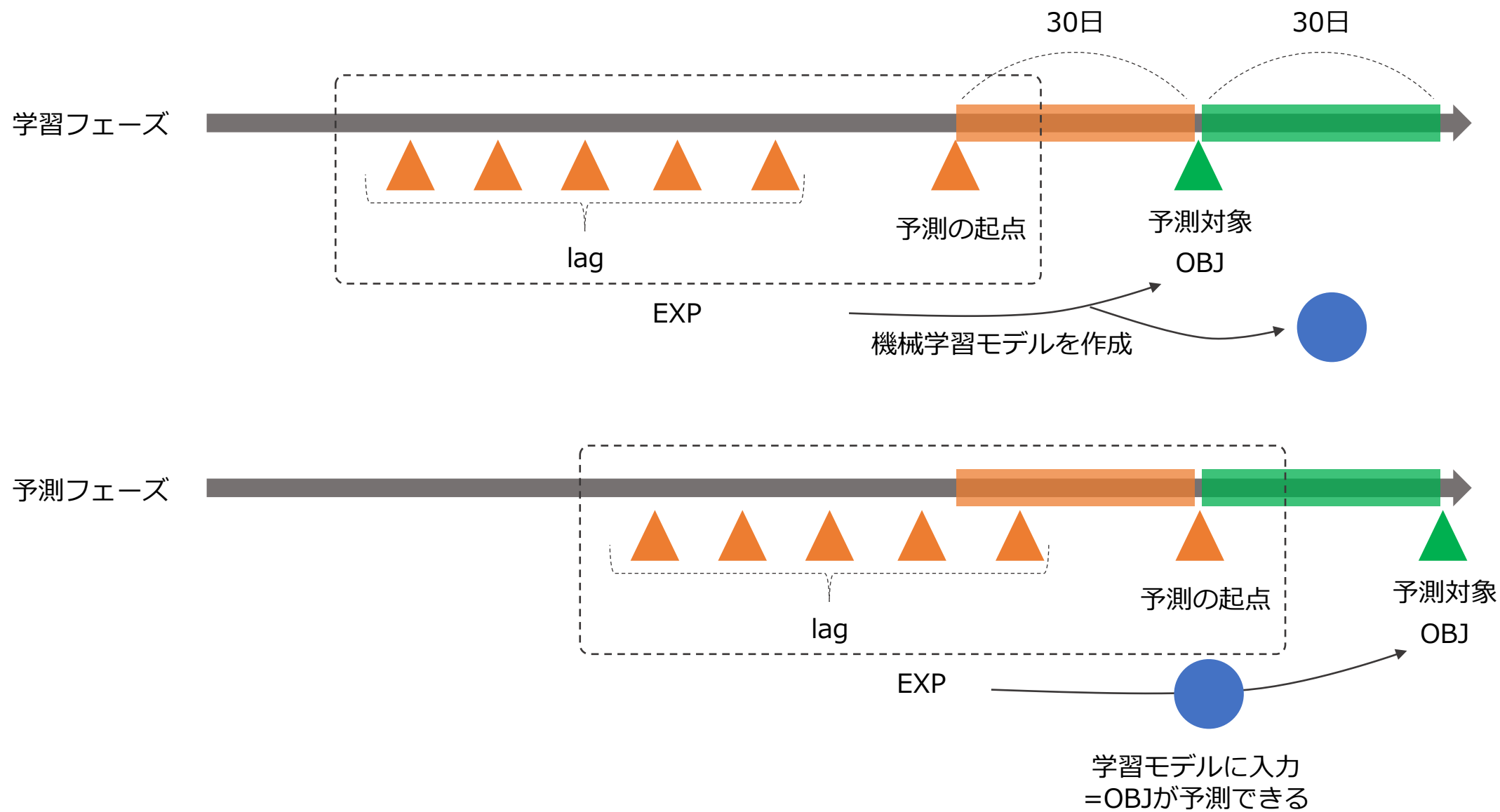
■個別株

個別株のデータ 4本値+出来高+売買代金 がCSVダウンロードできます。
セッションごとデータ 日足・前場・後場
各銘柄ごとの時系列データ 日足・前後場・1時間足・30分足・15分足・5分足

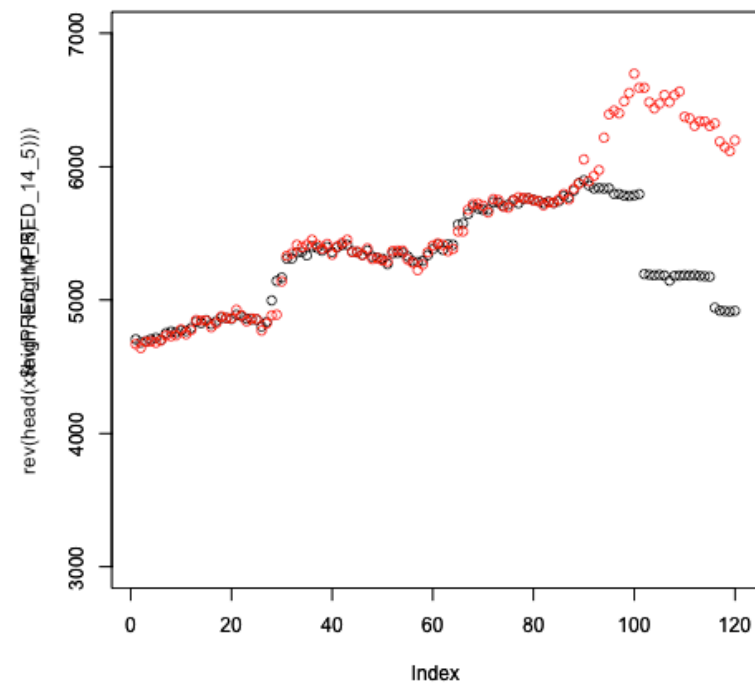
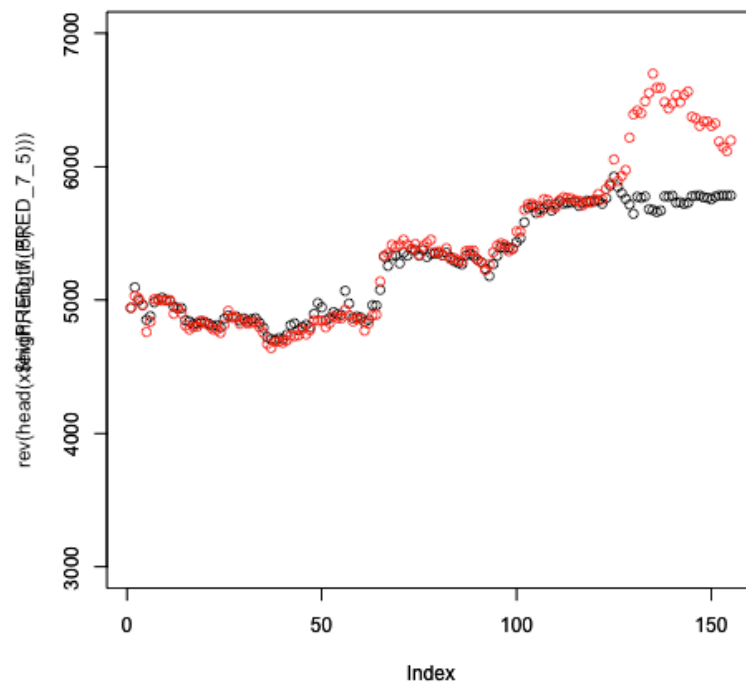
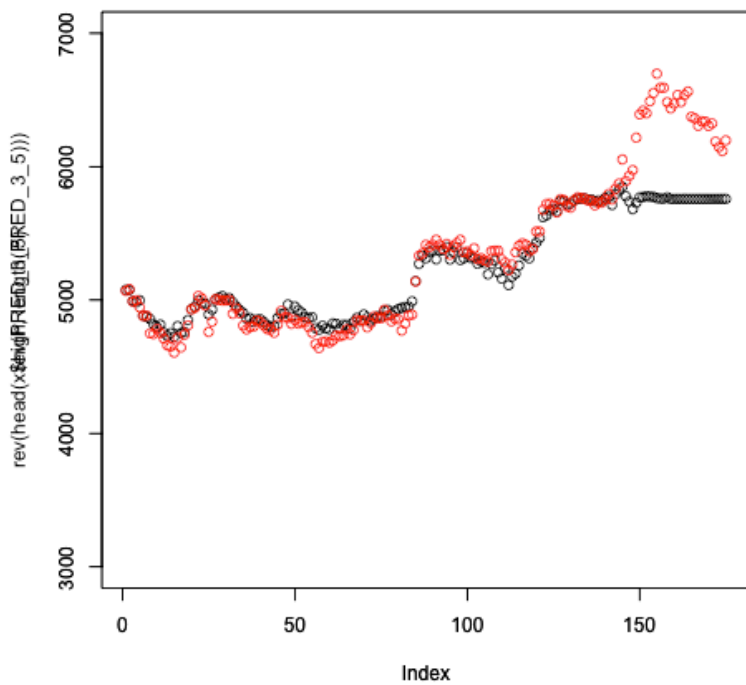
■市場統計

市場ごと・業種ごと の売買代金の集計や値上がり/値下がり銘柄数など日々の統計のデータです。
市場・業種を選択して、各市場業種ごとの 時系列表示 も可能です。

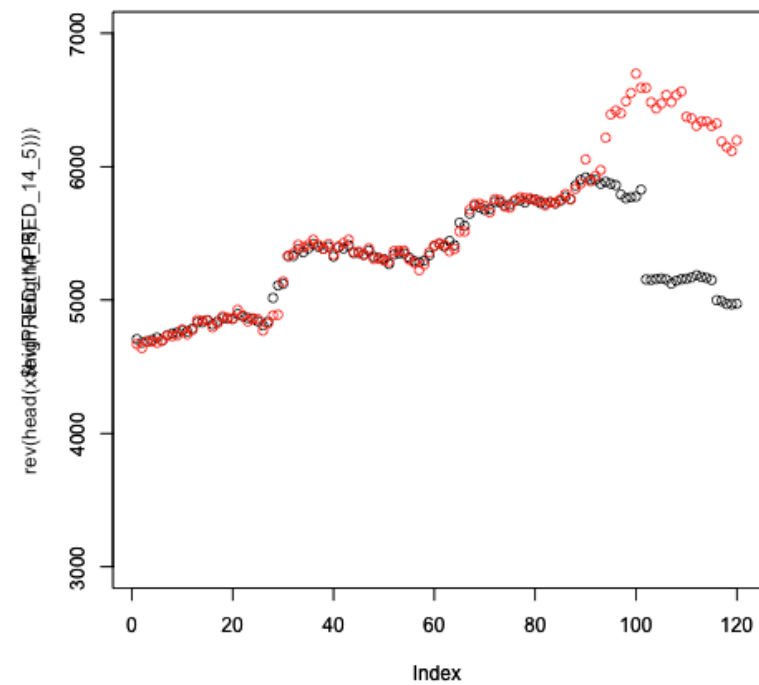
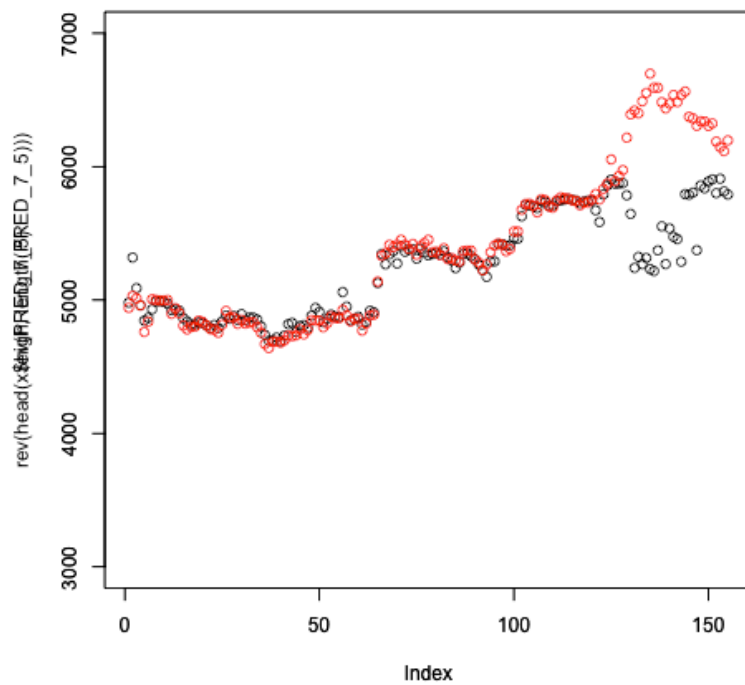
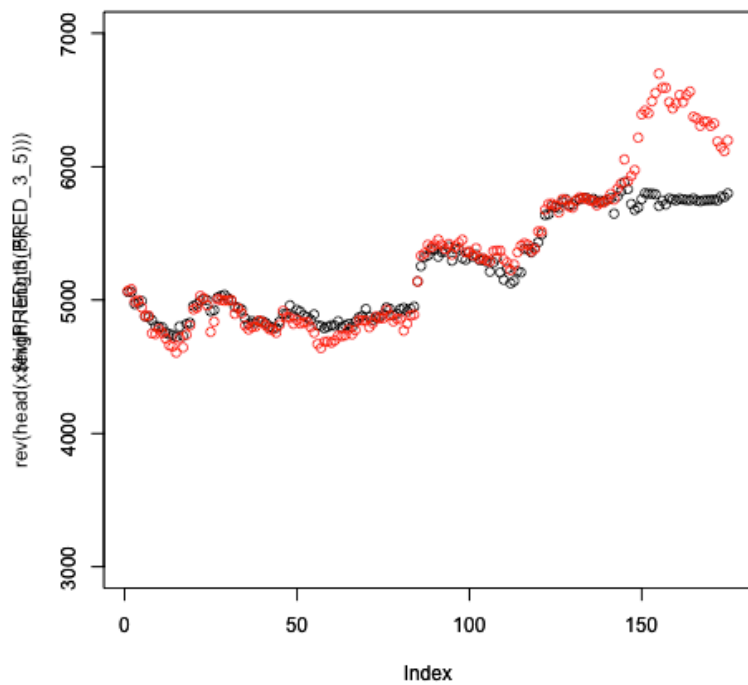
- 何やってるか



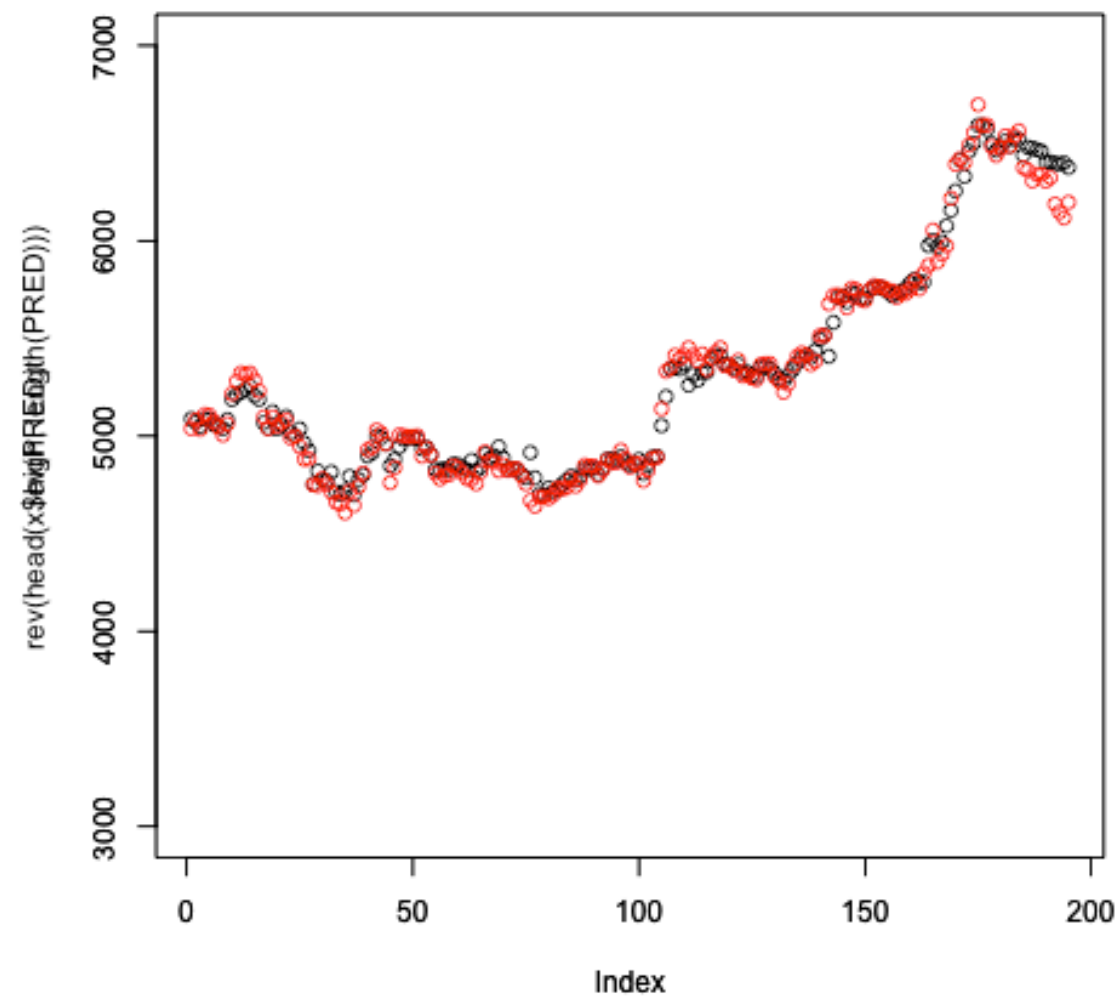
- 結果(単純に過去のデータを一緒に学習, 30日予測)



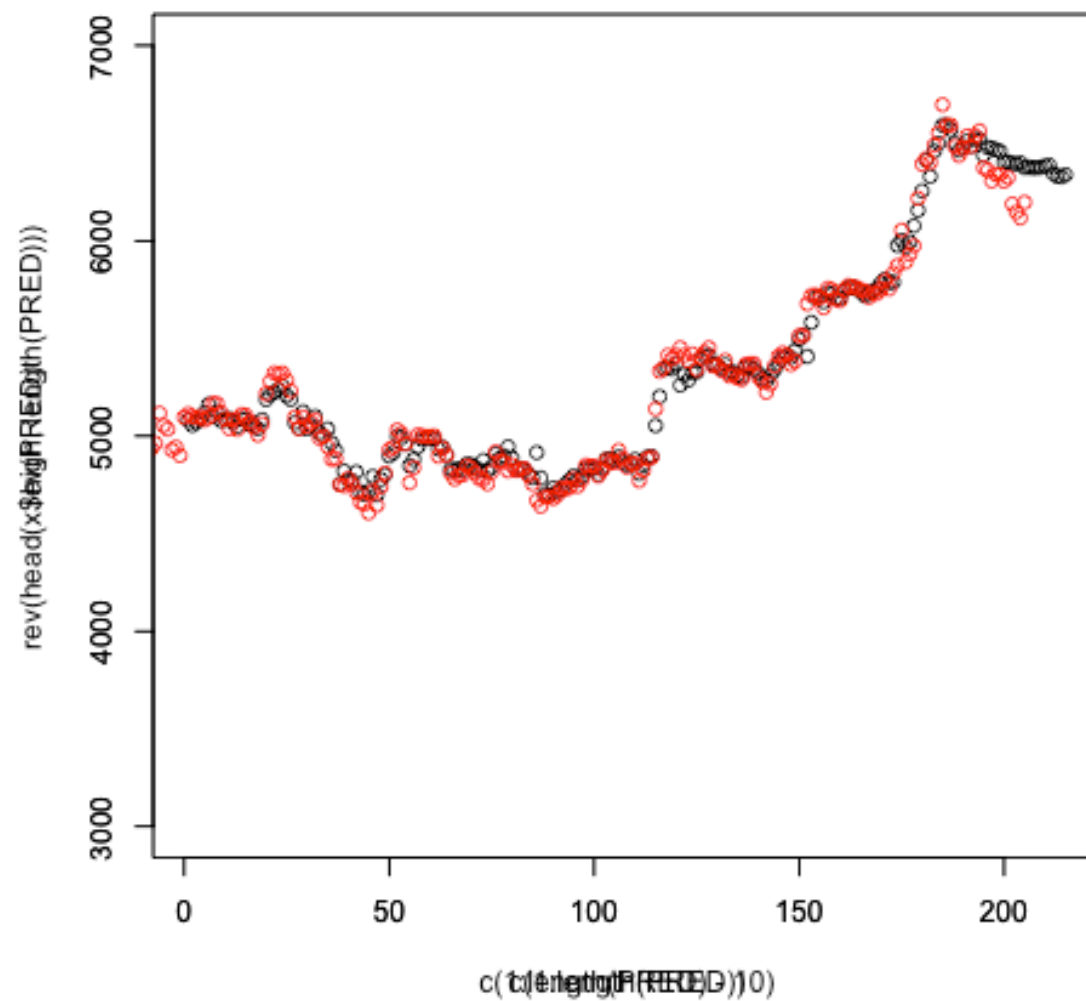
- 結果(過去データ+変化量, 30日予測)



- 結果(過去データ+変化量, 10日予測)



- ちなみに、未来予測：上がりそう？



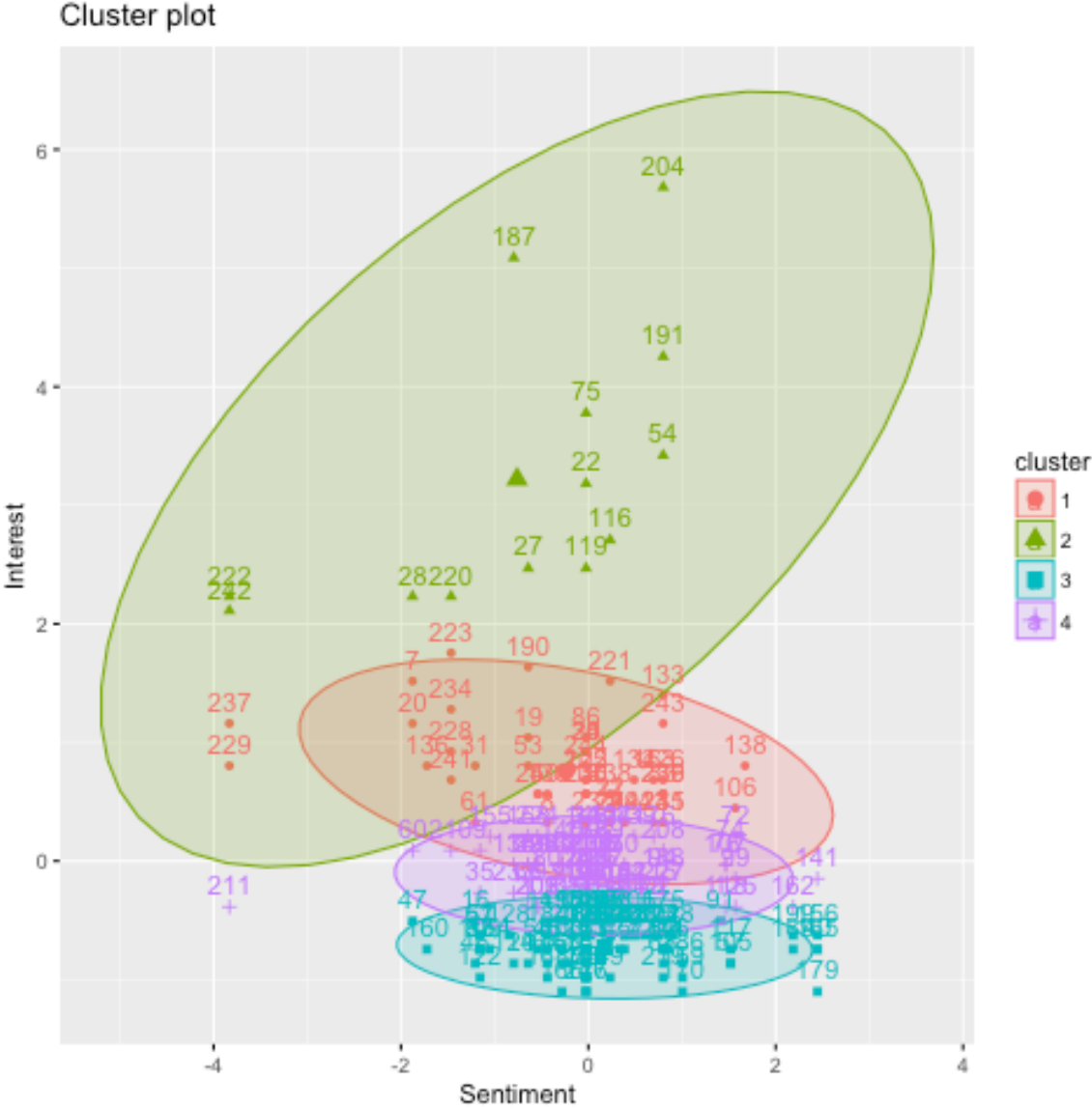
- 宿題

- デンソーの株価について、取引量も説明変数に入れて学習して評価
- パラメータを変えて、RMSEの小ささを競おう
- 図面も見せること

アジェンダ

- イン트로ダクション
 - ラスト！エンディングに向けて…！
 - 宿題：時系列データの処理
- 早速やってみよう
 - twitterデータの取得
 - 前処理
 - 解析・可視化

WSJ markets



ワシントンポスト

