2013年11月11日

進捗状況

・矢吹先生のGitHub内でのtaroyabukiの活動のログをとって，Excelで表にしてみる．

・時系列データの解析をやってみる．の途中です．

・矢吹先生のGitHub内でのtaroyabukiの活動のログをとって，Excelで表にしてみる．

今回は矢吹先生のtaroyabukiでの活動のログを選びました．

具体的には2013年7月21日から2013年11月9日までのcommitとIssueのログを集めました．

関口のGitHub内にあるエクセルのファイル名「yabuki先生の活動のログ分析」を参照してください．

Commitは矢吹先生がコミットした日時を集めました．

Issueは矢吹先生がクローズした日時を集めました．

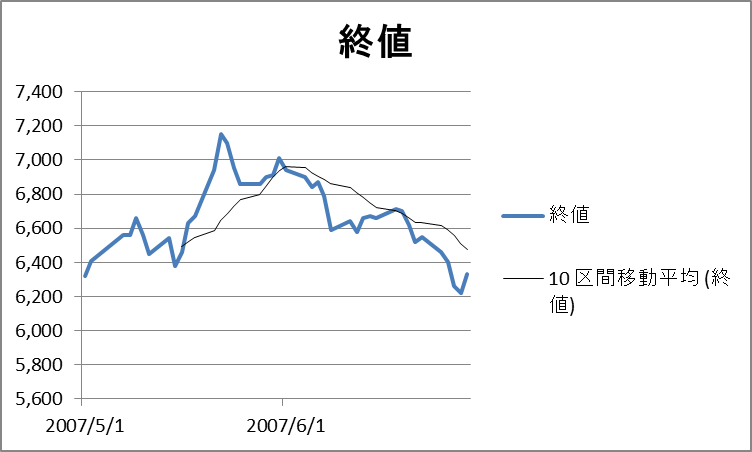
Issueは一か月を過ぎてしまうと月単位での表示でしか残ってなかったです．

以下の図のような感じです．



・時系列データの解析をやってみる．の途中です．

時系列解析の手法の習得の途中です．実際にRを導入し，データマイニングの講義で指定されたGDPデータと株価（終値）のデータを近似曲線の移動平均を区間5，10でグラフ化しました．そのエクセルデータをRでトレンドと周期変動，残差に分解するため，CSV形式で保存しましたが，Rの使い方がまだよく理解できていないので具体的な作業を教えていただけたらなと思っています．



質問事項

R i386 3.0.2とR x64 3.0.2のどちらをしようすればいいのか　A.どっちでもいい(データがすごく多い場合は64ビットがいいかもしれない)

R起動後，CSV形式のエクセルをどう読み取らせるのか

setwd("c:/R")

myData <- read.csv("GDPデータ.csv")

head(myData)

をどう入力？するのか

宿題

* 同じ結果（矢吹の行動履歴）をAPIで取得する．手作業の結果とあっているかを確認する．
* 授業用のメモ（<https://gist.github.com/taroyabuki/6793323>）を参考にする．
* 近藤 宏ほか『Excelでできるかんたんデータマイニング入門 』（同友館, 2008）や金 明哲『Rによるデータサイエンス - データ解析の基礎から最新手法まで 』（森北出版, 2007）の時系列解析のところを読んでおく．このほかにも，時系列処理関係の資料を探して，使えそうなテクニックを探しておく．