

 **DATUM STUDIO**

BIと機械学習のコラボレーション

自己紹介

- 山口 将央
 - DATUM STUDIO株式会社
 - データアナリスト
- 専門
 - Python・Rでの分析
 - 分析基盤構築
- 最近の活動
 - Tableauの本をかきました。(共著)
 - <https://www.amazon.co.jp/dp/4798050261/>



- 本題に入る前に、皆様はどんなBIツールを活用していますか？

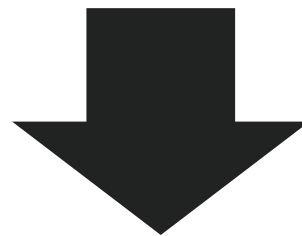


etc....

BIツールとは何か

BIツールとは？

- **ビジネスインテリジェンス**（英: Business Intelligence、**BI**）は、
- 経営・会計・情報処理などの用語で、企業などの組織のデータを、収集・蓄積・分析・報告することで、経営上などの意思決定に役立てる手法や技術のこと。



データによる意思決定を速くするためのツール

BIツールの種類

- 右の図は、米ガートナー者のMagic Quadrant (MQ)になります。
- 横軸に「Completeness of Vision (開発思想やアーキテクチャなどの先進性、完成度)」
- 縦軸に「Ability of Execute (市場シェアなど、企業としての実行力)」



「Magic Quadrant for Business Intelligence and Analytics Platforms」 (2016年) 出典: As of February 2016

BIツールの種類

- 米ガートナーのMQに乗っているツール

ポジション	製品名
Leader	Tableau
Leader	Qlik
Leader	Microsoft
Visionaries	Alteryx
Visionaries	SAS
Visionaries	SAP
Visionaries	Pentaho
Nitch Players	Domo
Nitch Players	Salesforce

- 米ガートナーのMQに乗っていないツール

Data Studio(Google、β版無料)

QuickSight(AWS、有料)

Re:Dash(OSS、一部有料あり)

Superset(OSS、無料)

- 今回は、Tableau・Data Studio・Re:Dashの3つを比較。



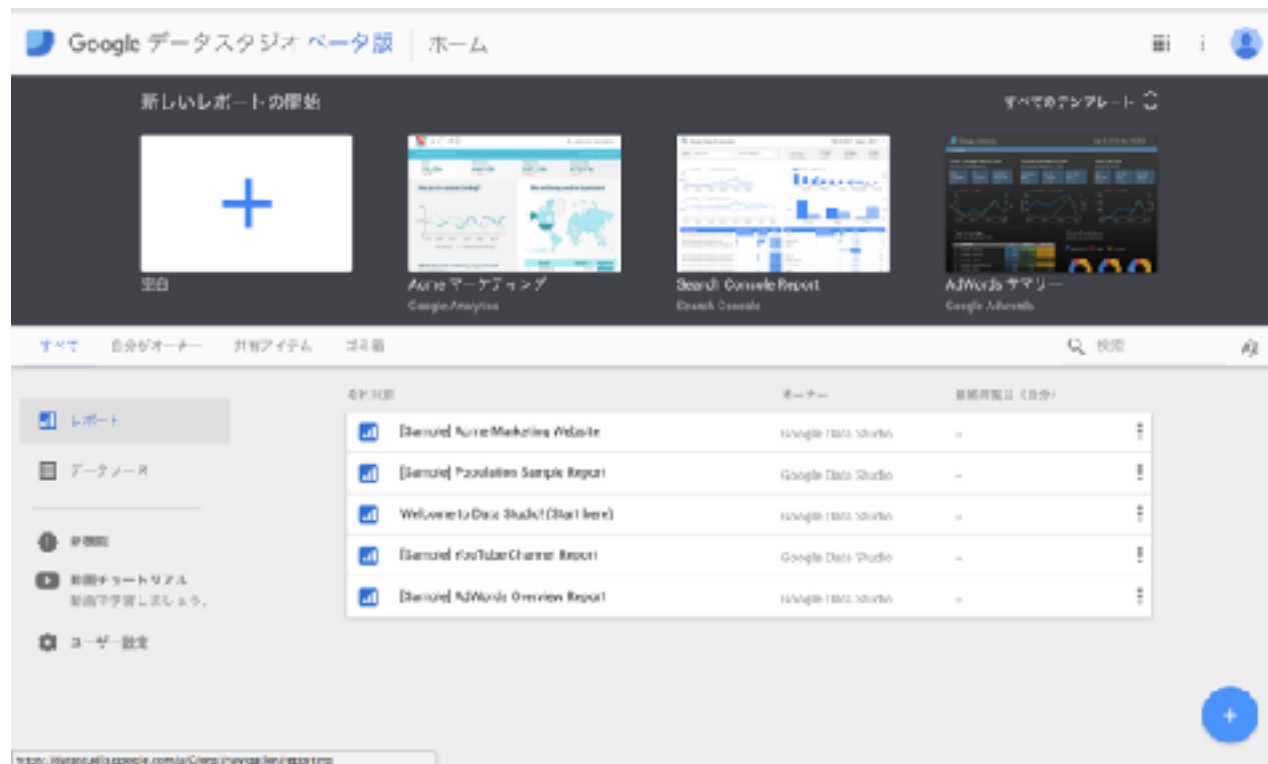
今回は、3つのBIを

「価格」、 「分析」、 「共有」

の3つの軸で比較してみたいと思います。

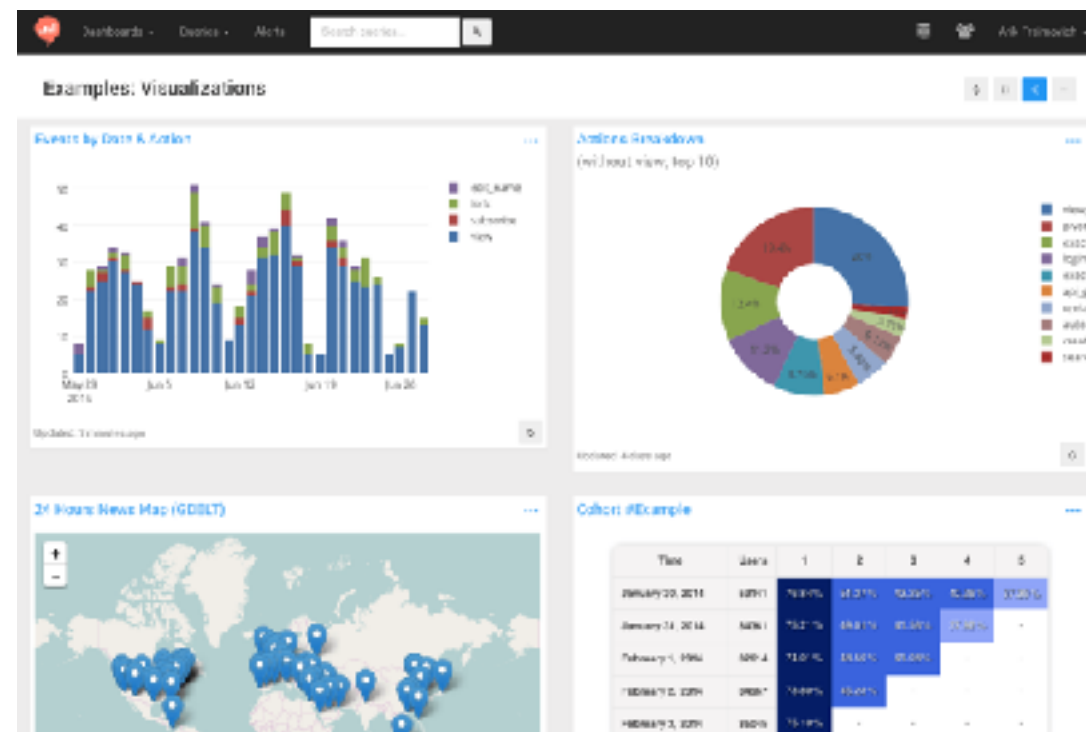
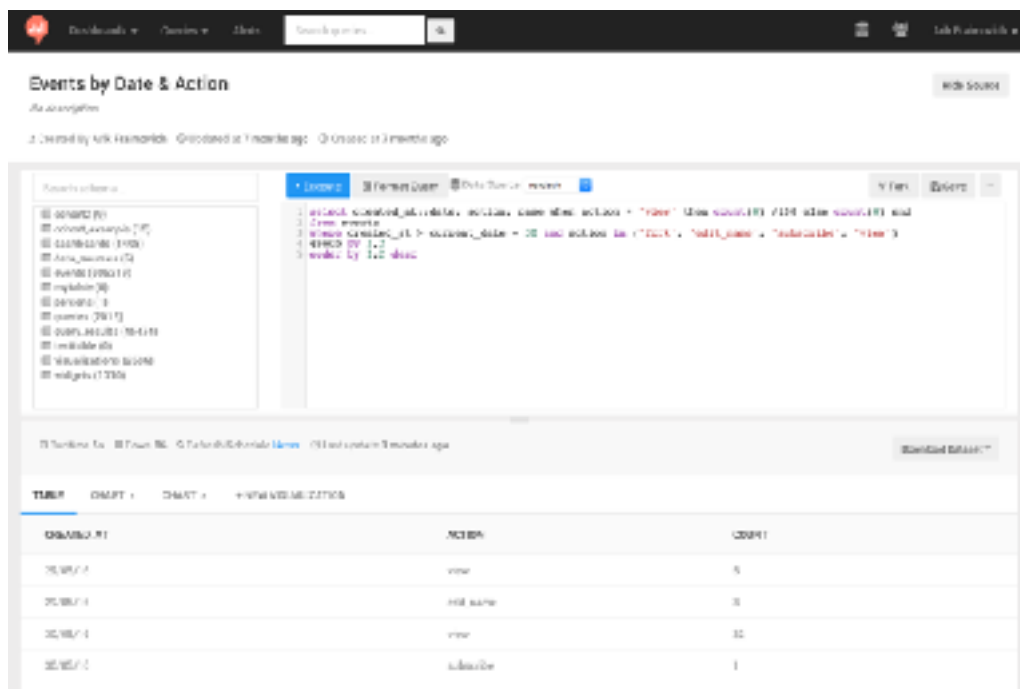
各BIの特徴 Data Studio

- 価格: 無料
- 分析: インタラクティブな分析には向かない
- 共有: 共有が簡単
 - (共有可能なURLを作成した共有する人のメールアドレスを入れるだけ)



各BIの特徴re:dash

- 価格: 無料(自分で立てるのであれば、別途サーバー費用必要)
- 分析: SQLベースなので、SQLがかけられる人であれば、分析までできる
- 共有: 共有が簡単
 - Google Apps認証あるので、gmailのアドレスがあればOK。

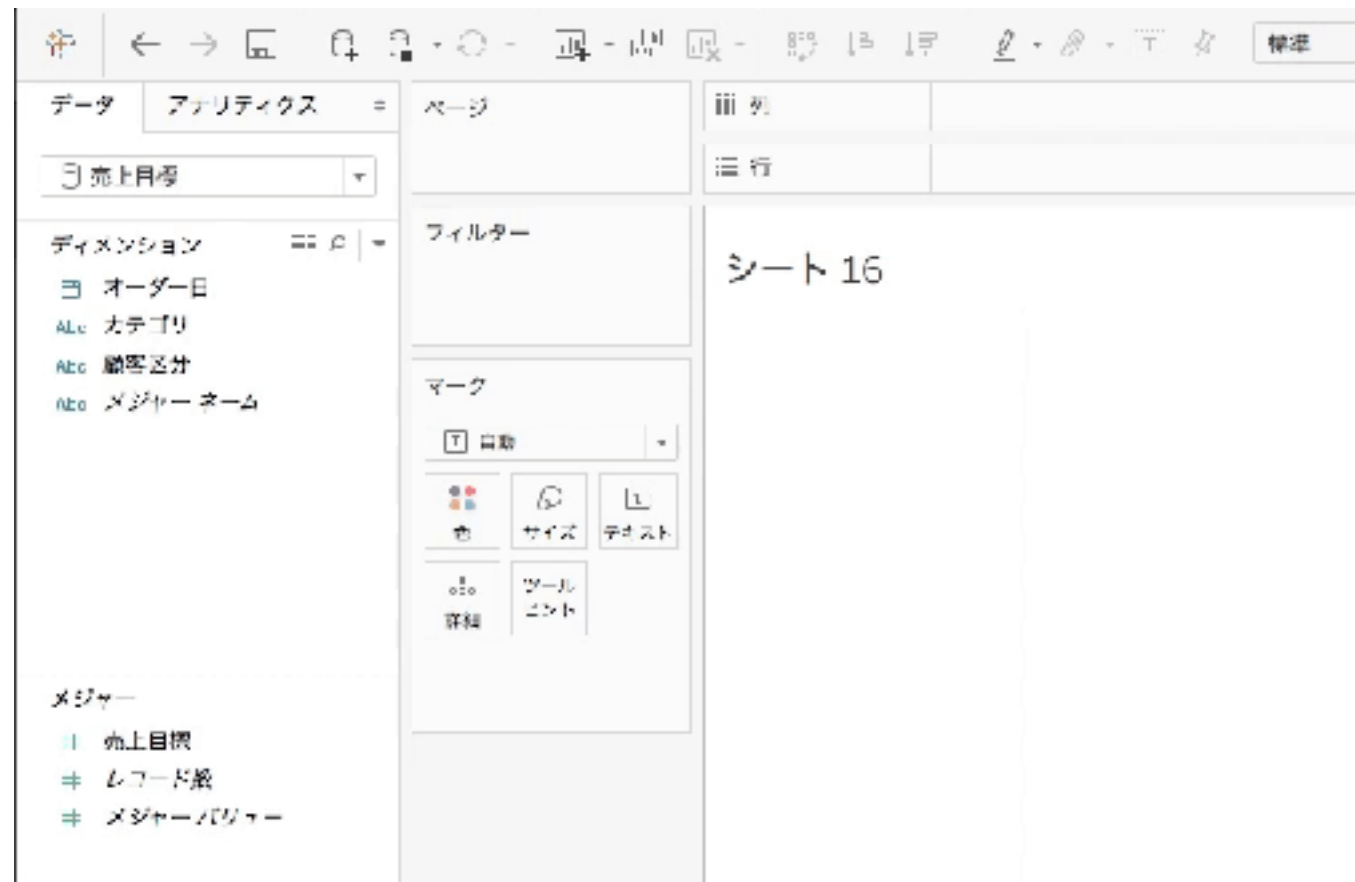


- Tableauは、今回紹介するなかで唯一有料のツールになります。

Tableau Desktop	
Personal エディション	Professional エディション
51,000 円	102,000 円
1ユーザーあたり年間(税 抜) (請求は1年ごと)	1ユーザーあたり年間(税 抜) (請求は1年ごと)
無料トライアルを始める	無料トライアルを始める

各BIの特徴 Tableau 分析

- ドラッグ&ドロップの操作で、データを可視化できる



- 共有方法

1. Tableau Readerを見る人に入れてもらうか
2. PDF・イメージでの共有
3. Tableau Serverを導入

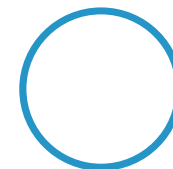
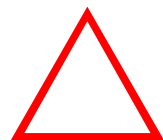
Tableau Serverまで導入しないと、
レポートの自動配信までは難しい

各BIの特徴

■ Tableau・Data Studio・Re:Dashの3つの比較



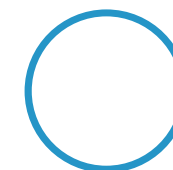
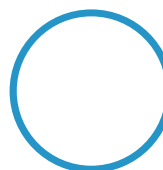
価格



分析



共有



- 個人的には、データを可視化するだけなら、Data Studio、re:dashで全く問題ないと思っております。
- ただ、分析をするのであれば、Tableauが一番使いやすいいと考
えております。

なぜか

- ドラッグ&ドロップで、データを可視化しながら、深掘りができる。
 - データ分析、特に探索的データ分析では、試行錯誤が発生し、行ったり来たりの作業が頻発するため。
- Python連携 & R連携ができる。
 - Python・Rのコードを書くことで機能拡張ができる。

今回は、Python連携のデモをします。

ここからは、Tableauを使用します。

- Tabpyの辛かった点

- パラメーターのPython渡し方が想定していたものと違った。

A. 1つの値だけかと思っていたらデータの数分はlistで来ていた。

- デバックばしづらい

- A. 書いたスクリプトでエラーがでてても、tabpy側のエラーしか見えない。

以上になります。