
SOFTWARE METAPAPER

eAnalytics: Dynamic Web-based Analytics for the Energy Industry

Paul Govan

Texas A&M University, US

Corresponding author: Paul Govan
(pgovan1@aggienetwork.com)

eAnalytics is a web application built on top of R that provides dynamic data analytics to energy industry stakeholders. The application allows users to dynamically manipulate chart data and style through the Shiny package's reactive framework. eAnalytics currently supports a number of features including interactive datatables, dynamic charting capabilities, and the ability to save, download, or export information for further use. Going forward, the goal for this project is that it will serve as a research hub for discovering new relationships in the data. The application is illustrated with a simple tutorial of the user interface design.

Keywords: R; energy; data analytics; visualizations; shiny; shinydashboard

(1) Overview

Introduction

Energy companies are interested in creating statistical models in order to benchmark performance [1]. Past work has often resulted in proprietary models (see, for example, Fisher et al, [2] and Lee et al. [3]). Fortunately (or unfortunately depending on one's perspective), the energy industry happens to be one of the most heavily regulated industries in the US, with much of the data that is collected by agencies being freely available online. The problem, however, is that the data is extremely difficult to find and understand. As a result, this potentially powerful source of information is at best underused, and at worst undiscovered.

eAnalytics is a free and open-source data analytics web application for energy industry stakeholders. The main motivation for developing eAnalytics was to gather and publish this information in order to spur interest in the research community. The secondary goal of this project was to provide a working example of how federal agencies could improve their database management systems. To the best of the author's knowledge, eAnalytics has the largest free and open-source database of US energy project information. The application's current features include allowing users to explore the data in greater detail, take an overview of different industry segments, measure key performance indicators (KPIs), and identify changes in the industry over time. Going forward, the goal for this project is that it will serve as a research hub for discovering new relationships in the data.

eAnalytics is built around the `energyr` [4] R [5] package of data published by the United States Federal Energy

Regulatory Commission (FERC) (www.ferc.gov). `energyr` contains several datasets for different industry segments:

- Electric: electric company financial information
- Gas: natural gas company financial data
- Hydropower: hydropower plant data
- LNG: LNG plant data
- Oil: oil company financial data
- Pipeline: natural gas pipeline project data
- Storage: natural gas storage field data

Implementation and architecture

eAnalytics is built using the Shiny [6] framework for developing web applications in R. The structure of a Shiny app consists of two primary components: (1) a user-interface script that controls the layout and appearance of the app and (2) a server script that contains instructions for the computer to build the app. The Shiny package contains multiple layout templates or the ability to build the user-interface from html content. eAnalytics is designed using the `shinydashboard` [7] package, which is a theme on top of Shiny for creating dashboard pages. The application is organized into a number of dashboard tabs based on the current features, which are discussed in the following section.

The application also employs the `htmlwidgets` [8] framework for binding JavaScript data visualizations in R. `htmlwidgets` create R bindings to JavaScript libraries, which allows these widgets to be embedded in different environments, including Shiny web applications. eAnalytics depends on the following `htmlwidgets`: `plotly` [9], `leaflet` [10], `googleVis` [11], and `DT` [12] packages.

eAnalytics is available on the Comprehensive R Archive Network (CRAN) as an R package at <https://cran.r-project.org/web/packages/eAnalytics/>. Full documentation and working examples are available at <https://github.com/paulgovan/eAnalytics>. Issues or requests may be filed at <https://github.com/paulgovan/eAnalytics/Issues>. To install the package in R:

```
install.packages("eAnalytics")
```

To install the latest development version:

```
devtools::install_github('paulgovan/eAnalytics')
```

To launch the app:

```
eAnalytics::eAnalytics()
```

eAnalytics currently contains a number of features including:

- Home: an introduction to the app
- Profile: take an overview of the industry
- Performance: measure key performance indicators (KPIs)
- Trends: identify changes in the industry over time
- Explorer: discover new relationships in the data
- Data: explore the data in greater detail

These features are illustrated in more detail in the following examples.

Illustrated Examples

Launching the app first brings up the Home tab, which is basically a landing page that gives a brief introduction to the app and includes three value boxes for the current number of projects, companies, and facilities in the database. **Figure 1** shows the Home tab as of this writing.

The Profile tab contains a number of interactive maps with information about facilities for the selected industry. **Figure 2** shows the Profile tab for the Natural Gas Industry.

Multiple options are currently available for customizing the maps. Choose a preferred size or color variable in the movable well panel, select from different basemaps via the lower-right control, and click on a specific facility to view additional information.

The Performance tab tracks a number of Key Performance Indicators (KPIs) for the selected industry. **Figure 3** shows the Performance tab for the Natural Gas Industry.

The Trends tab contains multiple interactive time-series charts of financial information for the selected industry. **Figure 4** shows the Performance tab for the Electric industry.

The time-series chart in the Trends tab is linked to the data table shown in the Data tab (see **Figure 6**). Searching, filtering, and sorting the data in the data table will automatically update the time-series chart with the selected data.

The Explorer tab contains a dynamic motion chart for exploring several indicators over time. **Figure 5** shows the Explorer tab for the Natural Gas Industry.

The Data tab contains interactive datatables of information for the selected industry. The data can be searched, filtered, and sorted as required. The selected data can then be copied to the clipboard, downloaded to a csv or pdf file, or sent to a local printer. **Figure 6** shows the Data tab for the Hydropower industry.

Quality control

eAnalytics has been tested in modern web browsers, including Google Chrome, Safari, Firefox, and IE10+. The application is available on the Comprehensive R Archive Network (CRAN) as an R package at <https://cran.r-project.org/web/packages/eAnalytics/>. Full documentation and working examples are available at <https://github.com/paulgovan/eAnalytics/>. Issues or requests may be filed at <https://github.com/paulgovan/eAnalytics/Issues>.

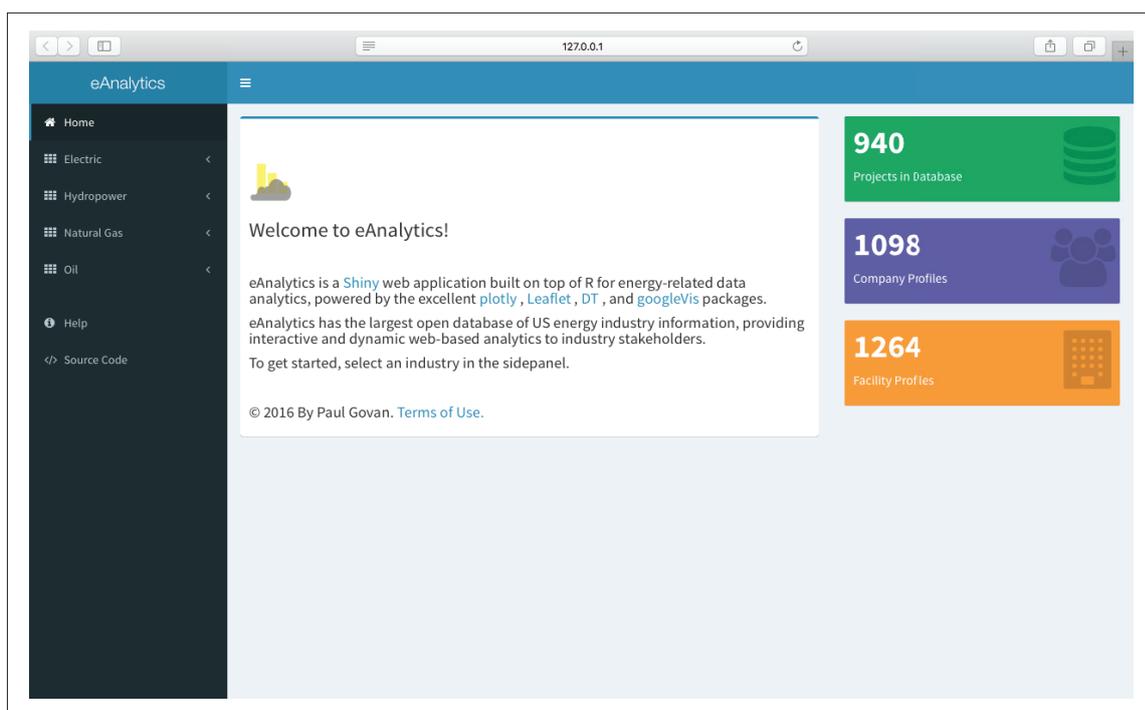


Figure 1: The Home Tab.

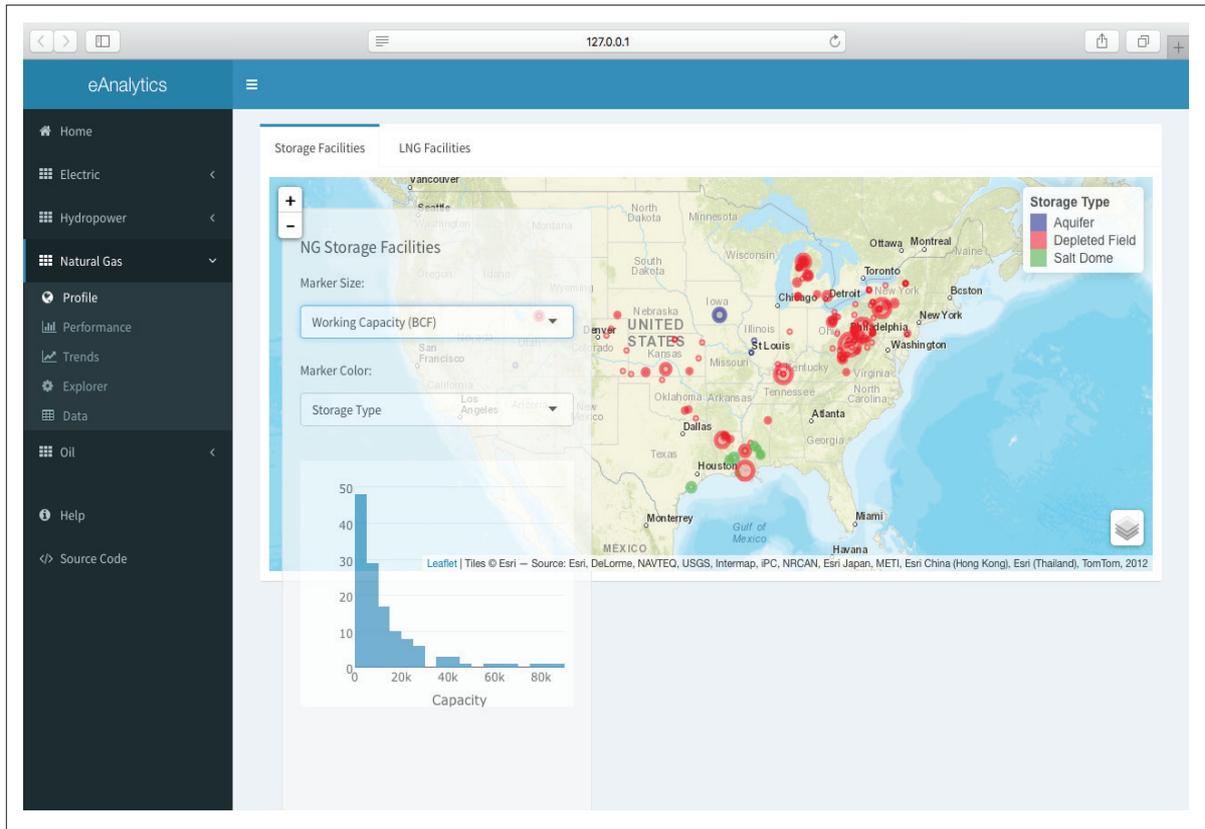


Figure 2: Profile Tab for Natural Gas Industry.

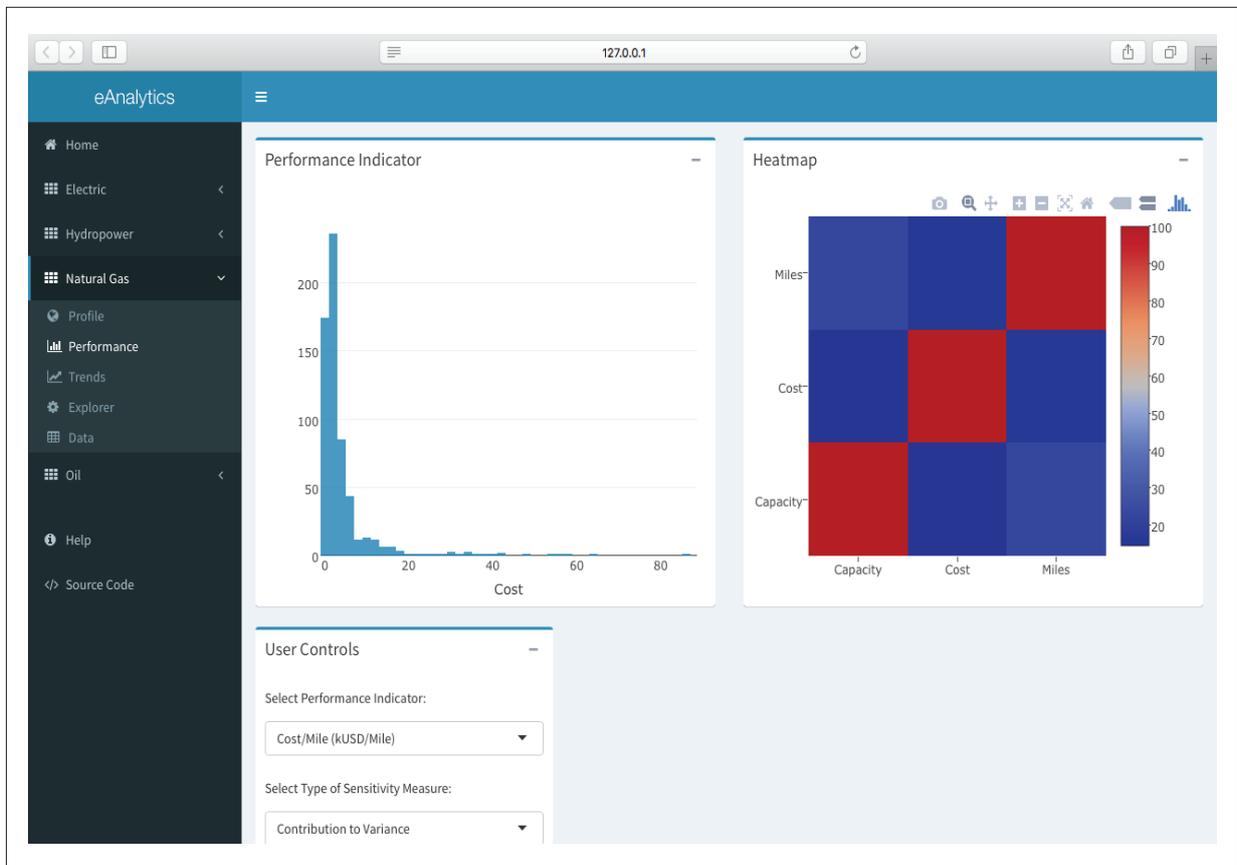


Figure 3: Performance Tab for Natural Gas Industry.

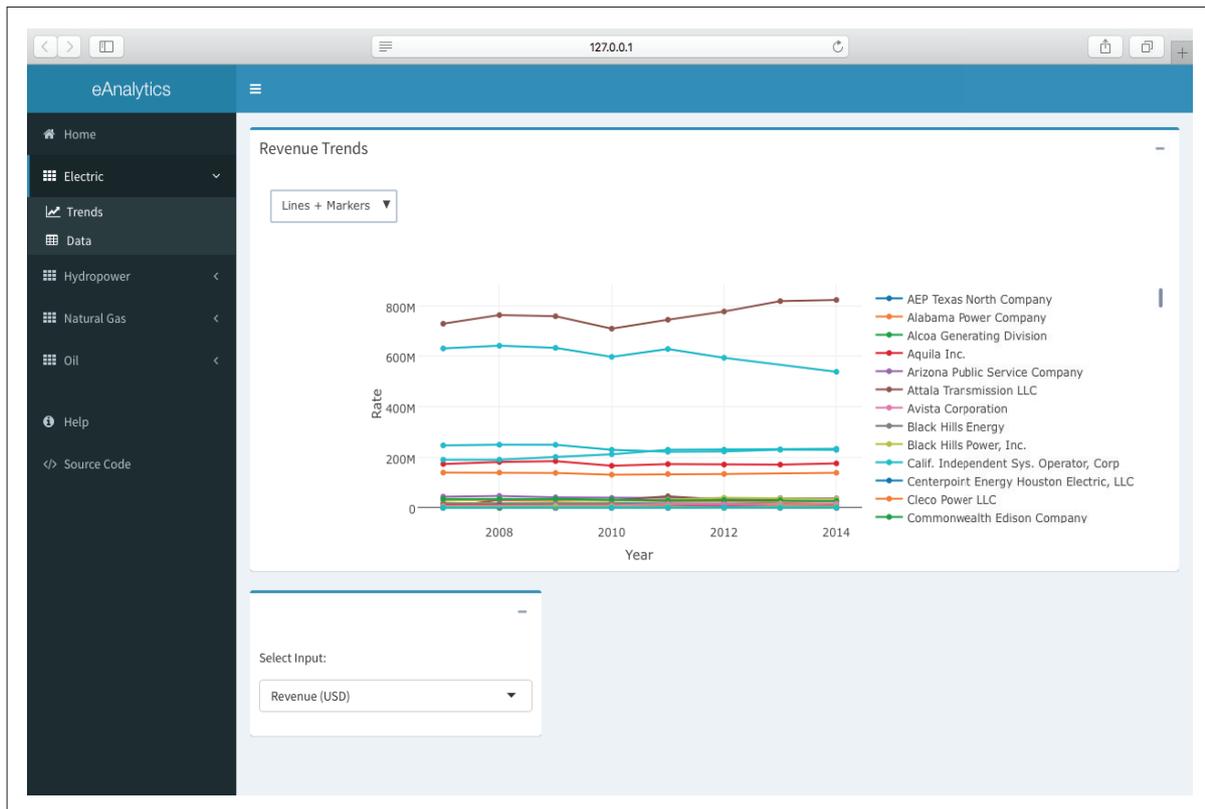


Figure 4: Trends Tab for the Electric Industry.

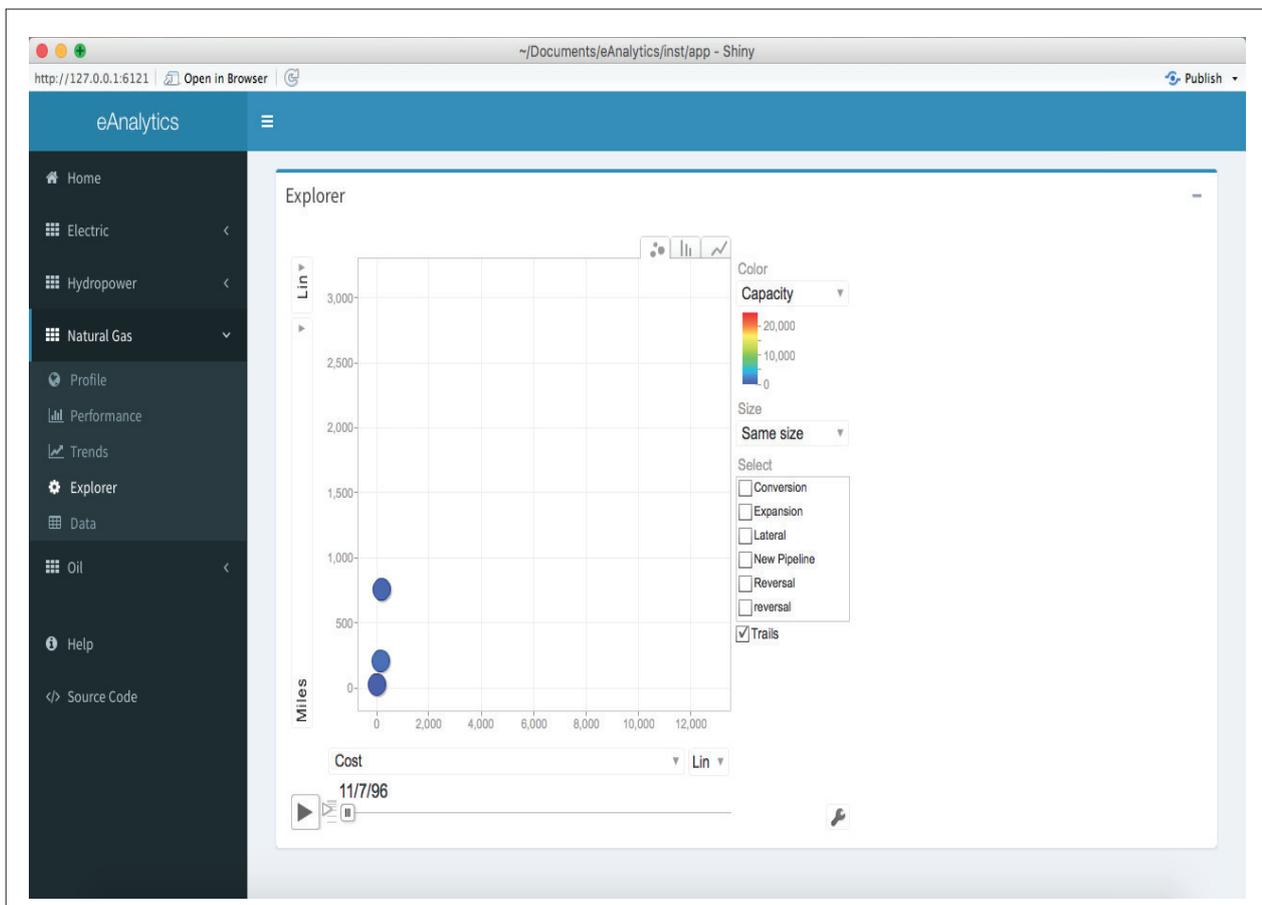


Figure 5: Explorer Tab for Natural Gas Industry.

The screenshot displays the 'Hydropower Facility Data' table within the eAnalytics application. The table contains the following data:

Number	Name	Expiration	Issued	Status	Capacity (KW)	Company	Waterway	State
2088	SOUTH FEATHER POWER	3/31/09	7/21/52	Expired	104100	SOUTH FEATHER WATER AND POWER AGENC	SOUTH FORK FEATHER RIVER	CA
2086	VERMILION VALLEY	8/31/03	9/29/53	Expired	0	SOUTHERN CALIFORNIA EDISON CO (CA)	MONO CREEK	CA
2107	POE	9/30/03	10/26/53	Expired	142830	PACIFIC GAS AND ELECTRIC CO (C)	NORTH FORK FEATHER RIVER	CA
2082	KLAMATH	2/28/06	1/28/54	Expired	161338	PACIFICORP (OR)	KLAMATH RIVER	OR
2105	UPPER NORTH FORK FEATHER RIVER	10/31/04	1/24/55	Expired	342628	PACIFIC GAS AND ELECTRIC CO (C)	NORTH FORK FEATHER RIVER	CA
2174	PORTAL	3/31/05	4/19/55	Expired	10800	SOUTHERN CALIFORNIA EDISON CO (CA)	RANCHERIA CREEK	CA
1971	HELLS CANYON	7/31/05	8/4/55	Expired	1166900	IDAHO POWER CO (ID)	SWAKE RIVER	OR
2100	FEATHER RIVER (PS&CON)	1/31/07	2/11/57	Expired	762850	CALIFORNIA DEPT-WTR RESOURCES (CA)	FEATHER RIVER	CA
2085	MAMMOTH POOL	11/30/07	12/30/57	Expired	150938	SOUTHERN CALIFORNIA EDISON CO (CA)	SAN JOAQUIN RIVER	CA
2197	YADKIN	4/30/08	5/1/58	Expired	221440	ALCOA POWER GENERATING INC.	YADKIN RIVER	NC

Figure 6: Data Tab for Hydropower Industry.

(2) Availability

Operating system

eAnalytics a platform-independent software package, compatible with modern web browsers (IE 10+, Google Chrome, Firefox, Safari, etc.).

Programming language

R

Additional system requirements

None

Dependencies

eAnalytics imports a number of R packages: plotly, dplyr, DT, energyr, googleVis, leaflet, shiny, shinydashboard.

List of contributors

Paul Govan, Author and Creator

Software location

Archive

Name: Zenodo

Persistent identifier: <http://dx.doi.org/10.5281/zenodo.165177>

Licence: Apache

Publisher: Paul Govan

Version published: v0.1.3

Date published: 7/11/16

Code repository

Name: GitHub

Identifier: <https://github.com/paulgovan/eAnalytics>

Licence: Apache

Date published: 7/11/16

Emulation environment (if appropriate)

Name: N/A

Identifier: N/A

Licence: N/A

Date published: N/A

Language

English

(3) Reuse potential

This project began at a mid-sized energy company that was looking for a way to model and benchmark project performance. The simple idea was that this information was undervalued and that the company could gain a strategic advantage by identifying areas of improvement and potential paths to growth. The result was a simple and adaptable tool that was easy to update and maintain. Now that this application is public, it could provide a strategic advantage to other organizations in an otherwise competitive industry.

eAnalytics is an open source project, with the goal of spurring growth in the research community of energy infrastructure. Researchers are encouraged to share and

adapt the package as required, with the only request being to pay-it-forward by sharing future insights. The information published in this package serves a wide range of industry segments, and is, therefore, applicable to a wide range of scientific enquiries. Generally speaking, this project helps answer basic questions about the state of energy in the US, such as how energy is generated and stored and how it is distributed and managed. These questions are important to a number of fields, from engineering to economics.

Competing Interests

The author has no competing interests to declare.

References

1. **Govan, P** and **Reinschmidt, K** 2013 Benchmarking Natural Gas Pipeline Projects. In *Pipelines 2013: Pipelines and Trenchless Construction and Renewals—A Global Perspective* (pp. 1532–1542). ASCE. Vancouver. DOI: <http://dx.doi.org/10.1061/9780784413012.143>
2. **Fisher, D, Miertschin, S** and **Pollock Jr, D R** 1995 Benchmarking in construction industry. *Journal of management in engineering*, 11(1): 50–57. DOI: [http://dx.doi.org/10.1061/\(ASCE\)0742-597X\(1995\)11:1\(50\)](http://dx.doi.org/10.1061/(ASCE)0742-597X(1995)11:1(50))
3. **Lee, S H, Thomas, S R** and **Tucker, R L** 2005 Web-based benchmarking system for the construction industry. *Journal of Construction Engineering and Management*, 131(7): 790–798. DOI: [http://dx.doi.org/10.1061/\(ASCE\)0733-9364\(2005\)131:7\(790\)](http://dx.doi.org/10.1061/(ASCE)0733-9364(2005)131:7(790))
4. **Govan, P** 2016 energyr: Data Published by the United States Federal Energy Regulatory Commission. R package version 0.1.0. <https://github.com/paulgovan/energyr>.
5. **R Core Team** 2014 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
6. **Chang, W, Cheng, J, Allaire, J, Xie, Y** and **McPherson, J** 2015 shiny: Web Application Framework for R. R package version 0.11.1.
7. **Chang, W** 2015 shinydashboard: Create Dashboards with 'Shiny'. R package version 0.5.1.
8. **Vaidyanathan, R, Cheng, J, Allaire, J, Xie, Y** and **Russell, K** 2014 htmlwidgets: HTML Widgets for R. R package version 0.3, 2.
9. **Sievert, C, Parmer, C, Hocking, T, Chamberlain, S, Ram, K, Corvellec, M** and **Despouy, P** 2016 plotly: Create Interactive Web Graphics via 'plotly.js'. R package version 3.6.0. <https://CRAN.R-project.org/package=plotly>.
10. **Cheng, J** and **Xie, Y** 2016 leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library. R package version 1.0.1. <https://CRAN.R-project.org/package=leaflet>.
11. **Gesmann, M** and **de Castillo, D** 2011 Using the Google visualisation API with R. *The R Journal*, 3(2): 40–44.
12. **Xie, Y** 2016 DT: A Wrapper of the JavaScript Library 'DataTables'. R package version 0.2. <https://CRAN.R-project.org/package=DT>.

How to cite this article: Govan, P 2016 eAnalytics: Dynamic Web-based Analytics for the Energy Industry. *Journal of Open Research Software*, 4: e45, DOI: <http://dx.doi.org/10.5334/jors.144>

Submitted: 27 August 2016 **Accepted:** 08 November 2016 **Published:** 30 November 2016

Copyright: © 2016 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.