



Data Science to Patient Value (D2V)

UNIVERSITY OF COLORADO **ANSCHUTZ MEDICAL CAMPUS**

SEPTEMBER 2017

CU Anschutz Data Science Tour: From D2V to KDD

SETH RUSSELL, MS

About Seth

Github:

- <https://gist.github.com/magic-lantern>
- <https://github.com/magic-lantern/>

About Seth

BS Business Administration – Central Washington University

BS Computer Science – Central Washington University

MS Biomedical Informatics – University of Utah

Capstone Project: Identification and Automation of Quality Metrics for Radiology Reporting in
CT Pulmonary Angiography



About Seth

Systems Analyst/Programmer @ Central Washington University

Programmer Analyst @ University of California

Software Engineer & Engineering Team Lead @ Backcountry.com

Senior Software Engineer & Team Lead @ Intermountain Healthcare

Research Instructor @ Data Science for Patient Value University of Colorado Anschutz



About Seth

Intermountain Care Process Models: <https://goo.gl/wBE6ap>

Bacterial disease diagnosis and medication guide
Based on Intermountain Healthcare's Acute Otitis Media Care Process Model updated September 2013

Select the image that most closely resembles the patient's inner ear



Meets the following Does NOT meet the following

Has cloudy yellow fluid in the middle ear space PLUS at least 1 of the following:

1. Bulging ear drum
2. Otorrhea
3. Mild bulging with redness AND new pain

Severe AOM Mild or Moderate AOM

Severe Includes:

- Significant ear pain for > 48 hours OR
- Fever > 39°C Current temp: 40.00Cel OR
- Otorrhea

Diagnostic Criteria for Acute Otitis Media requiring antibiotics **not** met.

Recommend close follow-up but no antibiotics.

Radiology Reporting Tool + Information Retrieval A University of Utah BMI 6470 Project

14:31 User: Sam Jorgensen Patient: Pat Williams Age: 48y
Monday, Oct 7, 2015 MRN: 1025582 Sex: - Gender - Report: 2015.10.345AC Urgency: Normal

Indication
Clinical Indication (if applicable)

Procedure
- Select Modality - Area Imaged
Additional Technique information

Pending Reports
2 Reports started but not completed

Comparison
Comparison exams with dates (if applicable)

Recently Completed Reports
70 Completed reports in the last week

Findings
Detailed explanation

Messages

John Can you call me about the MRI...
Amber RE: Contextual Inquiry study...

Conclusion
Summary

Radiologic Assistant Results
No results found - you must fill out at least 1 field for results to appear, or provide more appropriate terms.

About Seth

Patient Engagement Platform

Intermountain Alta View Hospital

Sara (Nurse), Ela (Housekeeper), 10:37:12 AM

Day/Date: Thursday, Sep 17 Room: Alta View Room 2147 Room Phone: 801-501-4647 Diet: No Restrictions Assist Level: Independent

Patient Name: [] Family Contact: [] Family Phone: [] Alternate Contact: [] Alternate Phone: []

Your Care Team

Physician: [] Therapist: [] Therapist: [] Therapist: []

Discharge Planning

Disch. Date: Friday, Sep 18 Discharge Location: []

Goals for Today

A- A- Ti- B I S U = = = = = = = = = =

When can I go home?

Delete Edit Cancel How do I care...

Pain Medications

Please call for pain medication

Pain/Comfort Goal: [] 0 1 2 3 4 5 6 7 8 9 10

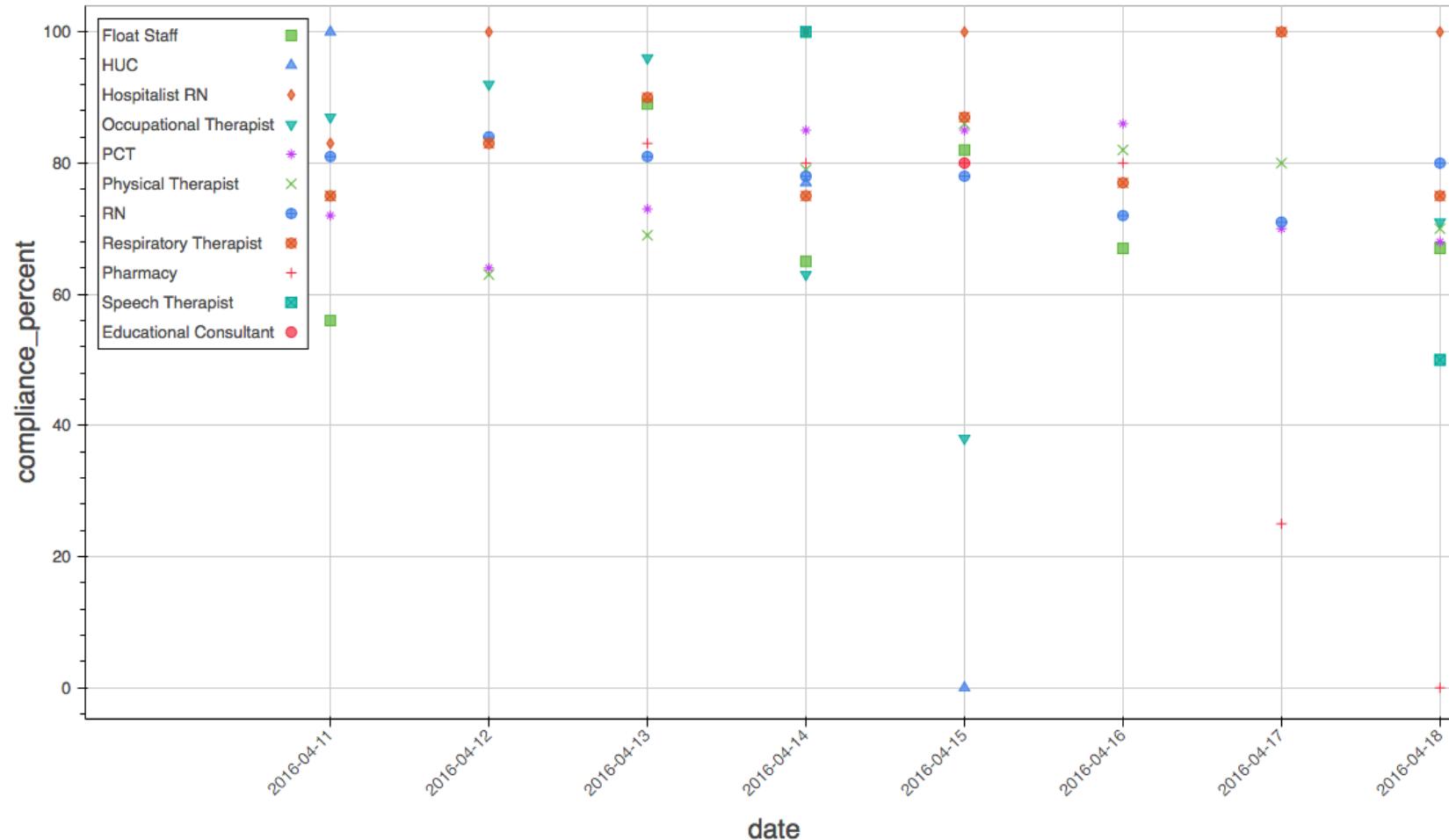
Pain Medication Last Dose Next Dose

Double tap text area to edit

Night Mode Scribble Board Clear All

About Seth

Real-time Location Services – Hand Hygiene Compliance



Data Science to Patient Value



School of Medicine

UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS

A-Z Index | Find a Health Care Provider

Contact Us | Maps and Parking

EDUCATION

APPLY NOW

RESEARCH

PATIENT CARE

DEPARTMENT|CENTER|INSTITUTE

DEAN'S OFFICE

FACULTY

[Home](#) > [Research](#) > **Transformational Research Funding**

Consortium for
Fibrosis Research
and Translation

GI & Liver Innate
Immune Program

Human
Immunology and
Immunotherapy
Initiative

RNA Bioscience
Initiative

D2V

Transformational Research Funding CU School of Medicine

(Jan. 27, 2016) At his first State of the School address on Wednesday University of Colorado School of Medicine Dean John J. Reilly, Jr., MD, named five recipients of Transformational Research Funding awards.

Last fall, Reilly announced a competitive process for applicants to seek funding for proposals that would position the University of Colorado School of Medicine as a leader in cutting-edge and emerging fields, attract extramural funding, help recruit and retain outstanding faculty, enhance education and training, and positively impact human lives and society in Colorado, the nation and the world.

The proposals were selected by an external review committee.

The five selected proposals are:

- Data Science to Patient Value (D2V), which aims to build infrastructure, develop methods and establish implementation pathways to prepare for population health.

The team leaders are Jean Kutner, MD, MPH, professor of medicine, chief medical officer for University of Colorado Hospital and associate dean for clinical affairs for the School of Medicine; Michael Ho, director of the Denver VA Center of Innovation; Lisa Schilling, MD, MSPH, professor of medicine and medical director of the Office of Value Based Performance; and Michael Kahn, MD, professor of pediatrics, director of informatics, Children's Hospital Colorado, and interim director of Health Data Compass.



Data Science to Patient Value (D2V)

UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS

Data Science to Patient Value



Data Science to Patient Value (D2V)

UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS

ABOUT D2V

SCIENTIFIC CORES

RESEARCH

EVENTS

TRAINING & EDUCATION

A-Z Index | Find a Health Care Provider
Contact Us | Maps and Parking



Data Science to Patient Value (D2V)

Data Science to Patient Value (D2V) is a multidisciplinary research initiative that focuses on Big Data methods, their applications to medicine and health care delivery, and ultimately, the achievement of high value, patient-centered health care. D2V accomplishes this by creating an environment of collaboration and innovation in which the top minds in data science, health care delivery and health services research work closely with patients and other stakeholders to tackle important problems. D2V places a priority on using big data in translation and dissemination work (i.e., implementation science), ensuring that D2V 'products' – be they analytic methods, care pathways or other tools – achieve maximal impact in the real world.

News and Events:

D2V Spotlight: [Dr. Vinay Kini](#)

Next D2V Seminar:
[Tuesday, September 19](#)

D2V Announces Project Topics for Research Track Students

[more...](#)

Leverage data to improve value in medically complex patients

<http://cud2v.org/>

Follow Us:

- [Subscribe](#) to our e-mail list
- Twitter: [@cuD2V](#)
- Facebook: [cuD2V](#)
- LinkedIn: [D2V](#)



Data Science to Patient Value (D2V)

UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS

Data Science to Patient Value – Scientific Cores

- Patient and Systems Value
- Data and Informatics
- Analytics
- Stakeholder Engagement and Governance
- Training and Education
- Dissemination and Industry Collaboration



Data Science to Patient Value

Data Science to Patient Value Scholars Program

<https://goo.gl/on8PGT>

Research Track School of Medicine

<https://goo.gl/hnPTFX>

D2V Pilot Project Funding Opportunity

<https://goo.gl/NvKopi>



Data Science to Patient Value (D2V)

UNIVERSITY OF COLORADO **ANSCHUTZ MEDICAL CAMPUS**



I am currently a Research Fellow in the Data Science to Patient Value (D2V) Scholar Program in the School of Medicine at the University of Colorado Anschutz Medical Campus.

Previously I held positions as a Postdoctoral Fellow in the Computational Bioscience Program at the University of Colorado Anschutz Medical Campus, and as an Associate Professor in the School of Information Systems and Applied Technologies at Southern Illinois University. Before becoming an academic, I worked in industry in various capacities including systems analyst, software engineer, computer programmer, and system administrator.



Data Science to Patient Value (D2V)

UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS

D2V Analytics Core

Analytics Environment to built on common infrastructure using open tools and to promote reproducible research

	Non-PHI	PHI
Analytics services	D2V, CBC/BERD, departmental biostatisticians	
Analytics infrastructure (HPC workloads)	Summit (CU Boulder)	TICR/ROSALIND
Analytics infrastructure (non-HPC workloads)	Personal computers, departmental servers	Compass Eureka
Data providers	Various sources	Compass Data Warehouse



- OS Updates ??

- "Local" repo

Mirror

Anacoda

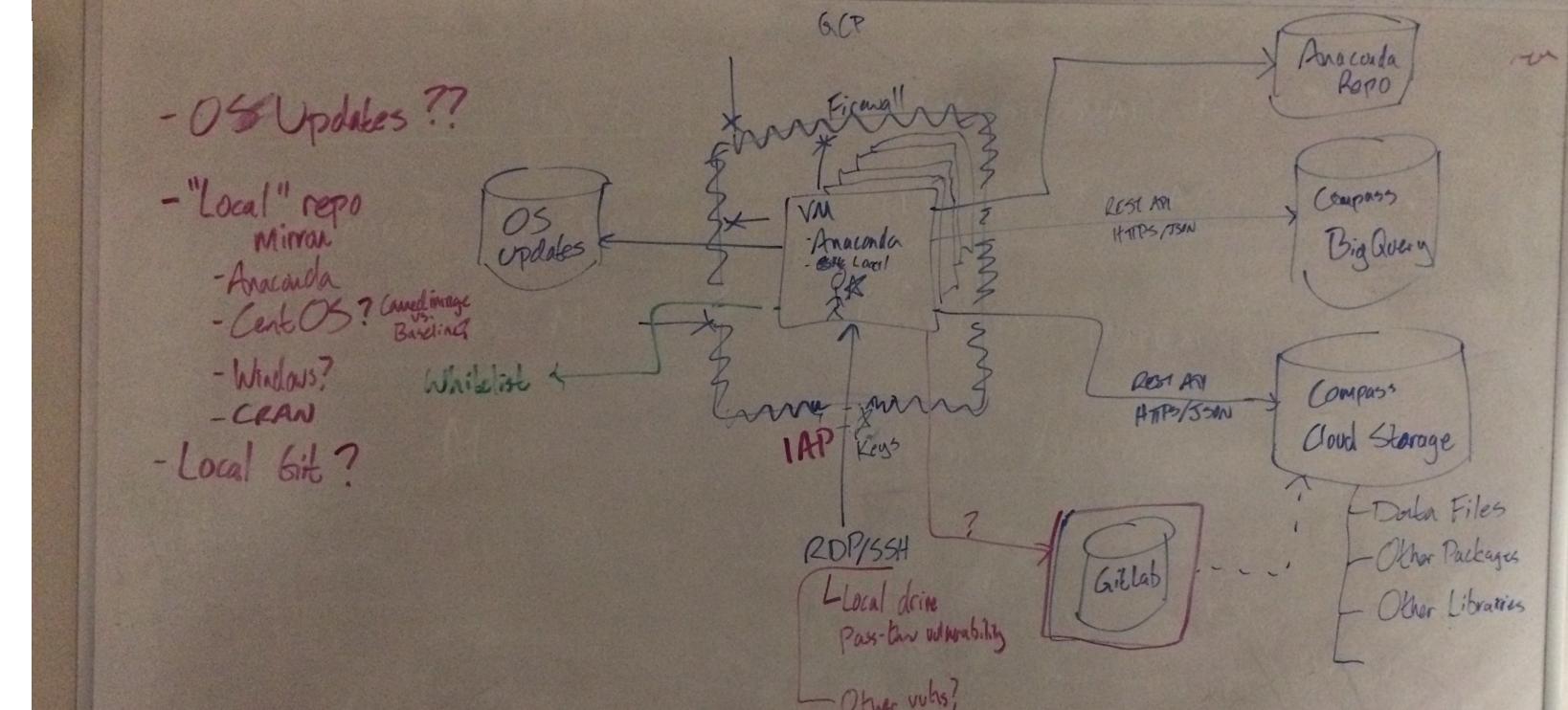
CentOS?

CentOS image
Baseline

Windows?

CRAW

- Local Git ?

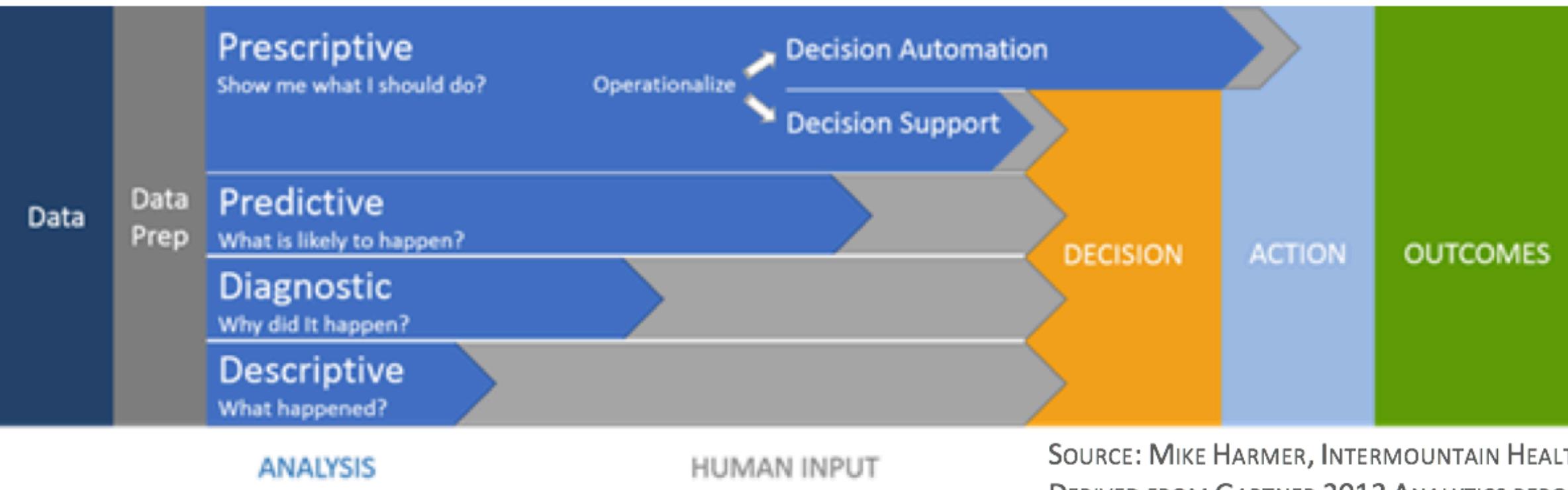


<http://healthdatacompass.org>

<https://github.com/magic-lantern/biostatistics-seminar-2017>



Data Science to Patient Value (D2V)
UNIVERSITY OF COLORADO ANSCHUTZ MEDICAL CAMPUS





Analytics Core: Guidance for Funded Projects

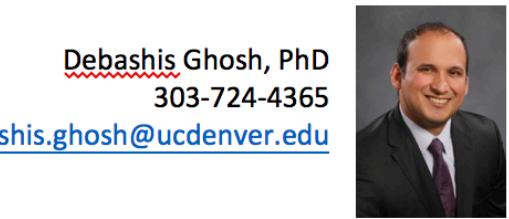
Leverage data to improve value

This Core is designed to fill gaps in current data analytic capabilities on campus. The initial focus is on applying the analytic tools of machine learning, health econometrics, and causal inference to analyses of observational data. In particular, we will develop and apply expertise in analyzing big, complex, and messy data such as those emerging from Health Data Compass and the D2V Informatics core. We envision the D2V Analytics Core as being a collaborative group that has unique expertise in these types of analyses, but which also builds strong bridges to other analytic groups on campus in support of D2V projects.

Analytics Core Leads



Tell Bennett, MD, MS
303-724-8661
tell.bennett@ucdenver.edu



Debashis Ghosh, PhD
303-724-4365
debashis.ghosh@ucdenver.edu

Services available

We offer limited engagement consultancy to all D2V funded projects. We can review each project's aims, methods, and provide high level analytical guidance.

We may offer additional assistance with analytic and computing needs – or pull in other cores as appropriate - for research that will develop or extend big data methods and tooling or involve multi-organizational collaborations.

The goal of healthcare analytics should be to improve an outcome in a desired area. The diagram shown at the bottom of this page depicts the process by which analytics can affect outcomes. An analysis or analytics tool can provide value at multiple time points. The closer an analytics process is to affecting outcomes, the more mature and robust it should be.

Some key analytics questions to ask regarding your research project:

- What is the specific outcome you want to

https://github.com/CUD2V/analytics_examples

CUD2V / analytics_examples

Code Issues Pull requests Projects Wiki Settings Insights

Branch: master analytics_examples / synpuf_exploration.ipynb Find file Copy path

 magic-lantern Adding some more documentation sections ed85268 2 hours ago

1 contributor

567 lines (566 sloc) | 17.9 KB Raw Blame History

Setup and helper functions

```
In [ ]: import pandas

from IPython.core.display import display, HTML
display(HTML("<style>.container { width:95% !important; }</style>"))

import psycopg2 as pg
import pandas.io.sql as psql

from bokeh.io import output_notebook, show
output_notebook()
```

```
In [ ]: # this is to hide useless errors if using OAuth with BigQuery
import logging
logging.getLogger('googleapiclient.discovery_cache').setLevel(logging.CRITICAL)
# don't want to be messaged about future warnings as I'm not explicitly calling code that is being
warned about
import warnings
warnings.simplefilter(action='ignore', category=FutureWarning)

# set this to either postgres or bigquery ####
datasource = 'bigquery'
```



KDD 2017 Highlights

- About 1700 attendees
- Paper acceptance rate ~9%
- Poster acceptance rate ~10%
- Workshops/Tutorials acceptance rate ~60%
- \$540k in sponsorships



KDD 2017 Highlights

- User voted best video (paper summary) : https://youtu.be/G-omu_ki7YM
- Videos of all presentations not yet available. Should be available on youtube in the next few weeks- search “kdd2017”
- Accepted Papers Available at <http://www.kdd.org/kdd2017/accepted-papers> many have short video introduction/overview



KDD 2017 Highlights

Trends

- Machine Learning can be applied to any domain
- Hyperparameter Tuning
- How to scale to handle larger data, more CPU needs
- Working to migrate from supervised learning methods to unsupervised methods
- The real bottle neck to analysis is people
- Research problems
- Healthcare specific topics
 - Relevant Paper: <http://www.kdd.org/kdd2017/papers/view/patient-subtyping-via-time-aware-lstm-networks>



KDD 2017 Highlights

Sub-Conference: Big data analytics-as-a-Service: Architecture, Algorithms, and Applications in Health Informatics

Top Challenges

- Privacy & Security
- Compliance with regulations
- Mobile Computing
- Sharing of data
- Lack of Integration between clinical and administrative systems
- Operational Analytics
- Lack of analytics or people with analytics skills

Sub Conference Papers: <http://bigdas.org/callforpapers/>

Videos should be posted to YouTube along with rest of conference

