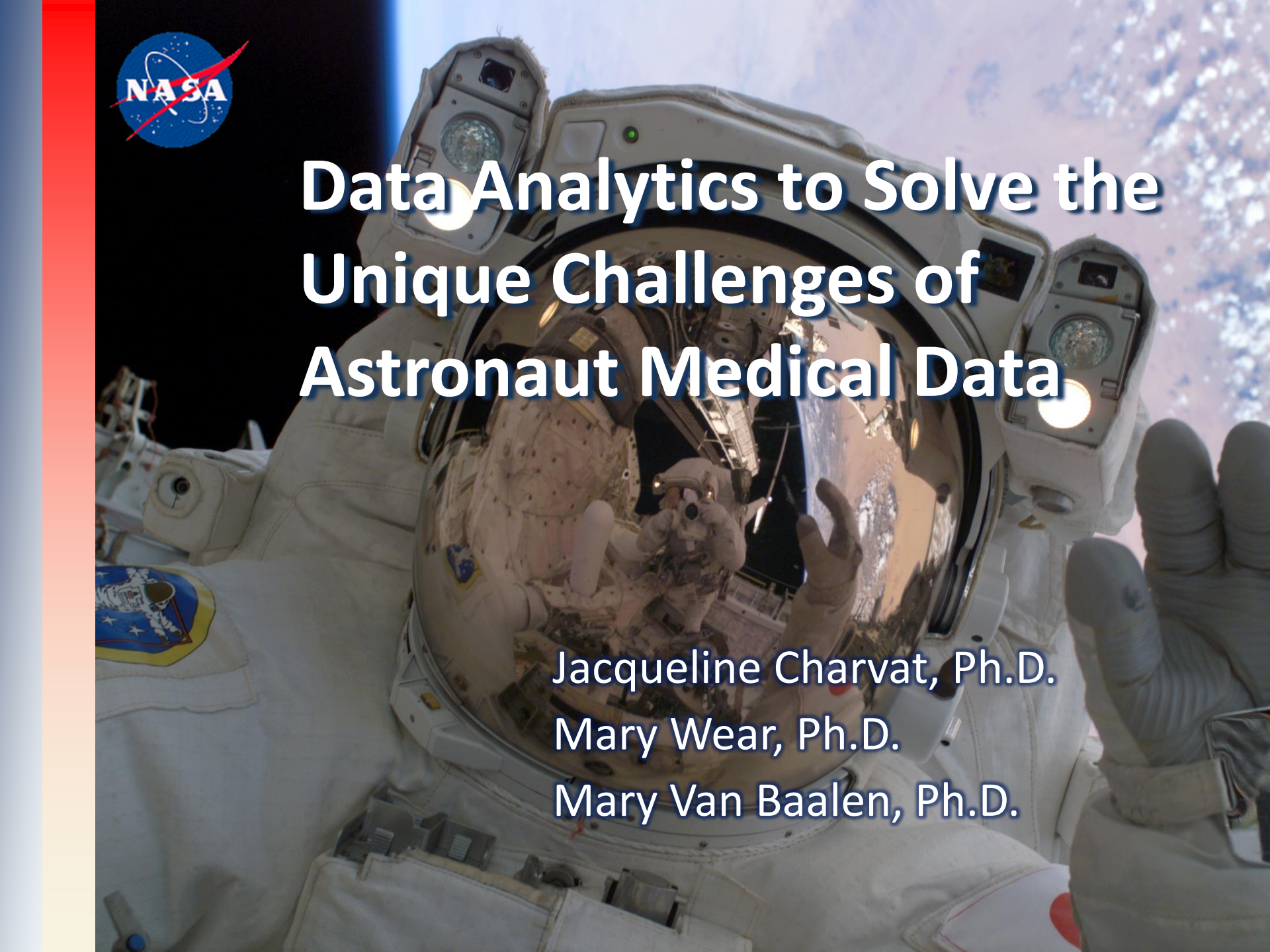
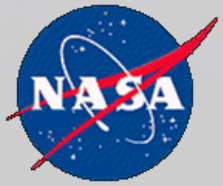


Data Analytics to Solve the Unique Challenges of Astronaut Medical Data

Jacqueline Charvat, Ph.D.
Mary Wear, Ph.D.
Mary Van Baalen, Ph.D.



- **Aim of LSAH is:**
  - Describe morbidity (acute and chronic) and mortality of astronauts
  - Determine whether unique occupational exposures encountered by astronauts are associated with increased risks of morbidity or mortality
  - Repository for all archived medical data
    - Provide data for internal and external requestors

  - Exposures
    - Spaceflight
      - Microgravity
      - Chemicals
      - Radiation
      - Closed loop life support system
      - Particulates
      - Lasers
      - Sunlight
  - High Performance Jet Training
  - Analog Training
    - NEEMO
    - Antarctica

Lifetime of Exposures

Spaceflight Exposures

Pre-NASA Exposures | ASCAN Training | Active Astronaut | Retirement

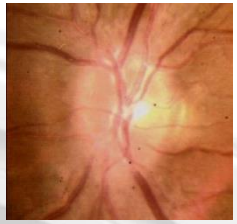
Terrestrial Exposures

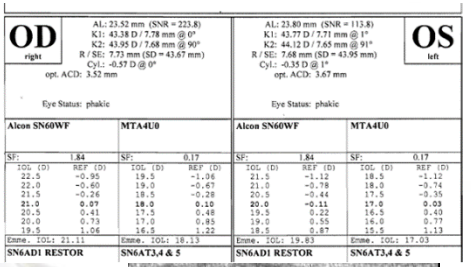
Challenges of Astronaut Data

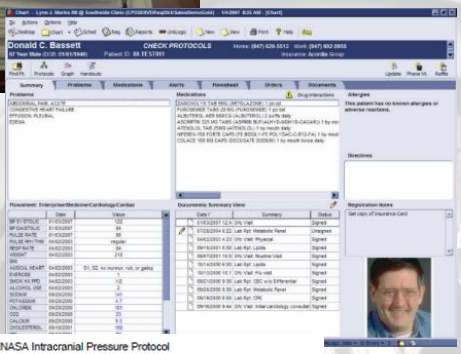
- Celebrity Effect
- Selection Bias and Healthy worker effect
- Cohort Effect
- Screening Bias
- Small N
  - Varied exposure to spaceflight
- Over-interpretation

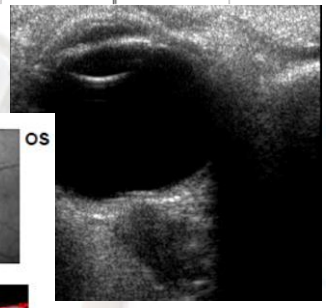
Data Compilation & Audit

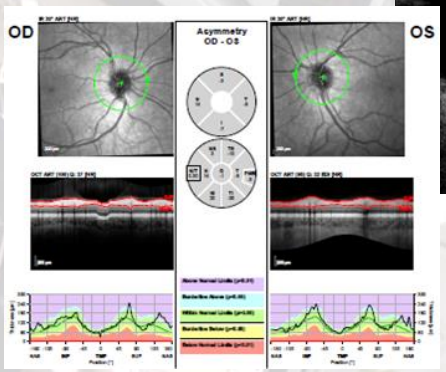
- Patient Medical Records
  - EMR
  - Hardcopy records
- Mission Medical Repository
  - Monitoring around spaceflight missions
- Individual Exposure Profile
- Training Records (NBL, survival, NOLS, etc.)
- Spacecraft Environmental Data
- Biospecimen Data
- Research Study Data
- “Other”Data











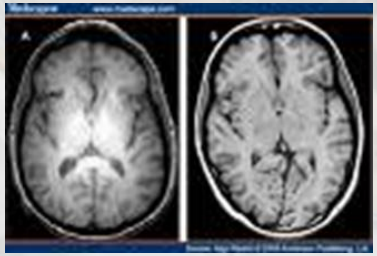
3 Tesla MRI NASA Intracranial Pressure Protocol

Technique:
3 Tesla MRI of the brain was obtained using isovolumetric T1 and T2 images of the whole brain; High resolution fat saturated , isovolumetric axial and oblique sagittal images of the orbits were obtained, along with axial FLAIR images of the brain.
Indication: Flight Evaluation

Findings:

Brain:
The ventricles and sulci are normal for the patient's age. The cisterns are within normal limits.
The sella turcica shows a partially empty sella configuration, without chiasmatic herniation.
No evidence of intracranial lesions. A non specific punctuate FLAIR hyperintensity is noted in the deep white matter of the right frontal lobe, unchanged since 2011.
Perivascular space visualization is the expected for the patient's age.
No extra-axial fluid collections are seen.
The paranasal sinuses are clear.

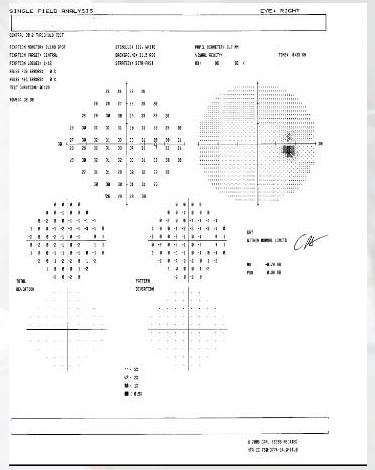
ORBITAL ANALYSIS:
Right eye
Left eye
Optic Nerve sheath diameter (ONSD) in cm axial 0.68 cm
0.75 cm
Optic Nerve sheath diameter (ONSD) in cm sagittal 0.68 cm
0.74 cm



Optic Nerve Diameter (OND) in cm axial 0.31 cm
0.31 cm
Optic Nerve Diameter (OND) in cm sagittal 0.30 cm
0.39cm
Optic Nerve sheath to nerve ratio (ON SNR) 2.19 ? 2.27
2.42 ? 1.90
Papilledema not present
Globe flattening grade 0-3 0
0
Optic nerve tortuosity grade 0-3 0
0
Optic Nerve hyperintensity 0-3 1
1

Impression:

1. Normal MRI of the brain without contrast.
2. No nerve sheath distention, optic globe flattening, or optic nerve hyperintensity to suggest increased intracranial pressure.
3. Normal sella.



11-28-05 pt picked copy of tri. Rx (ED)
5-5-08 pc. c pt c/o smudge in vision x weeks. not always there but in the center when it is. had exam c ophthal. in A2. told of some wrinkling. sched tomorrow DOA/OT (FV)

Complicating Factors: Data Harmonization

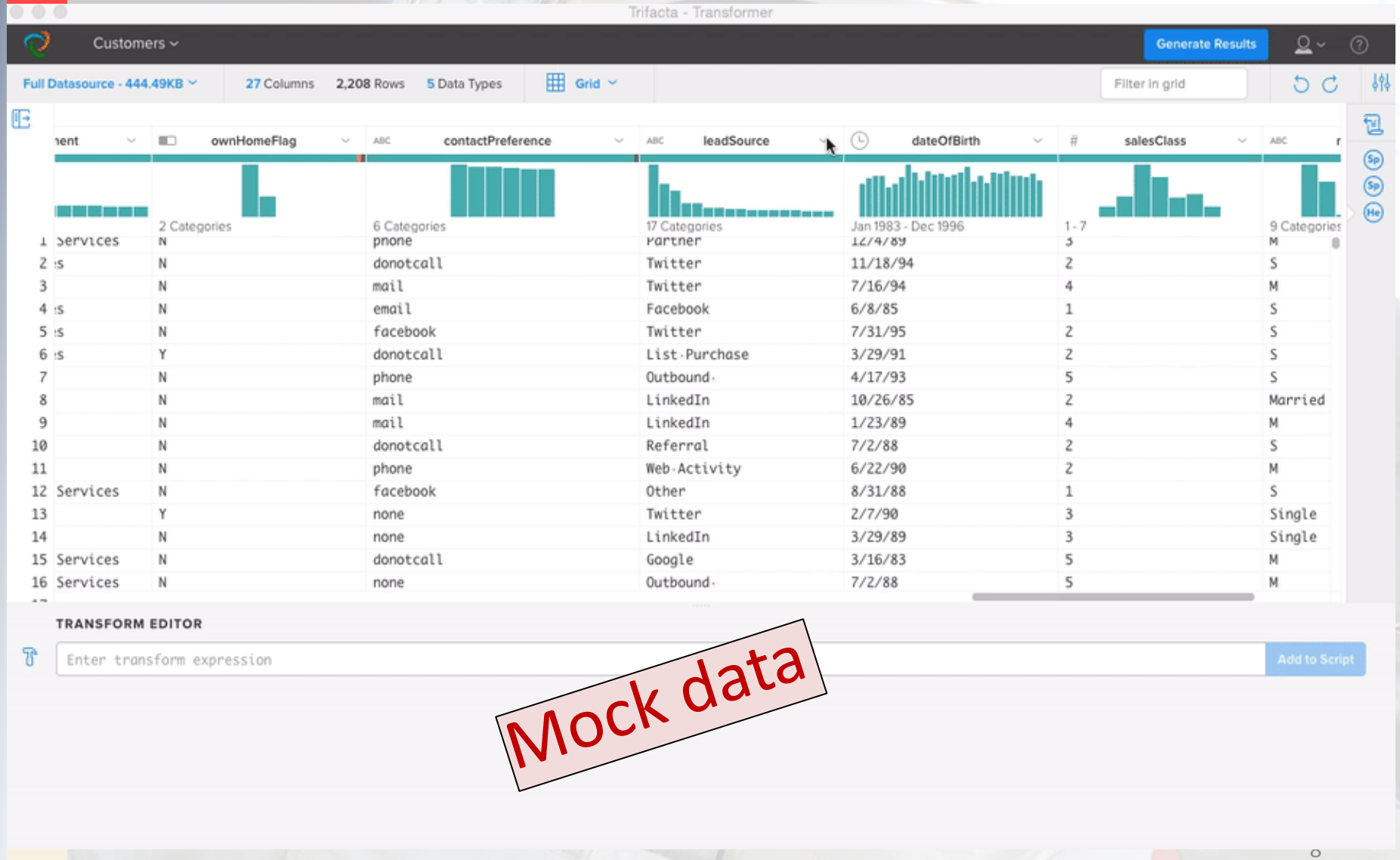
- Multiple sources collected for multiple reasons collected in multiple formats
  - Data entry into different places
- Dates
  - **Flight Day ≠ GMT ≠ MET ≠ Date of Entry**

IMPALA OVERVIEW

"Information Management Platform for Analytics & Aggregation"

- Platform to store, process, perform and share visualizations and advanced analytics of astronaut data to solve our data challenges
- Analyze multi-structured data, gain insights and make data driven decisions
- Self-service system to visualize large datasets.
- Share deep insights by collaborating in (near) real time across the board

IMPALA: Trifacta



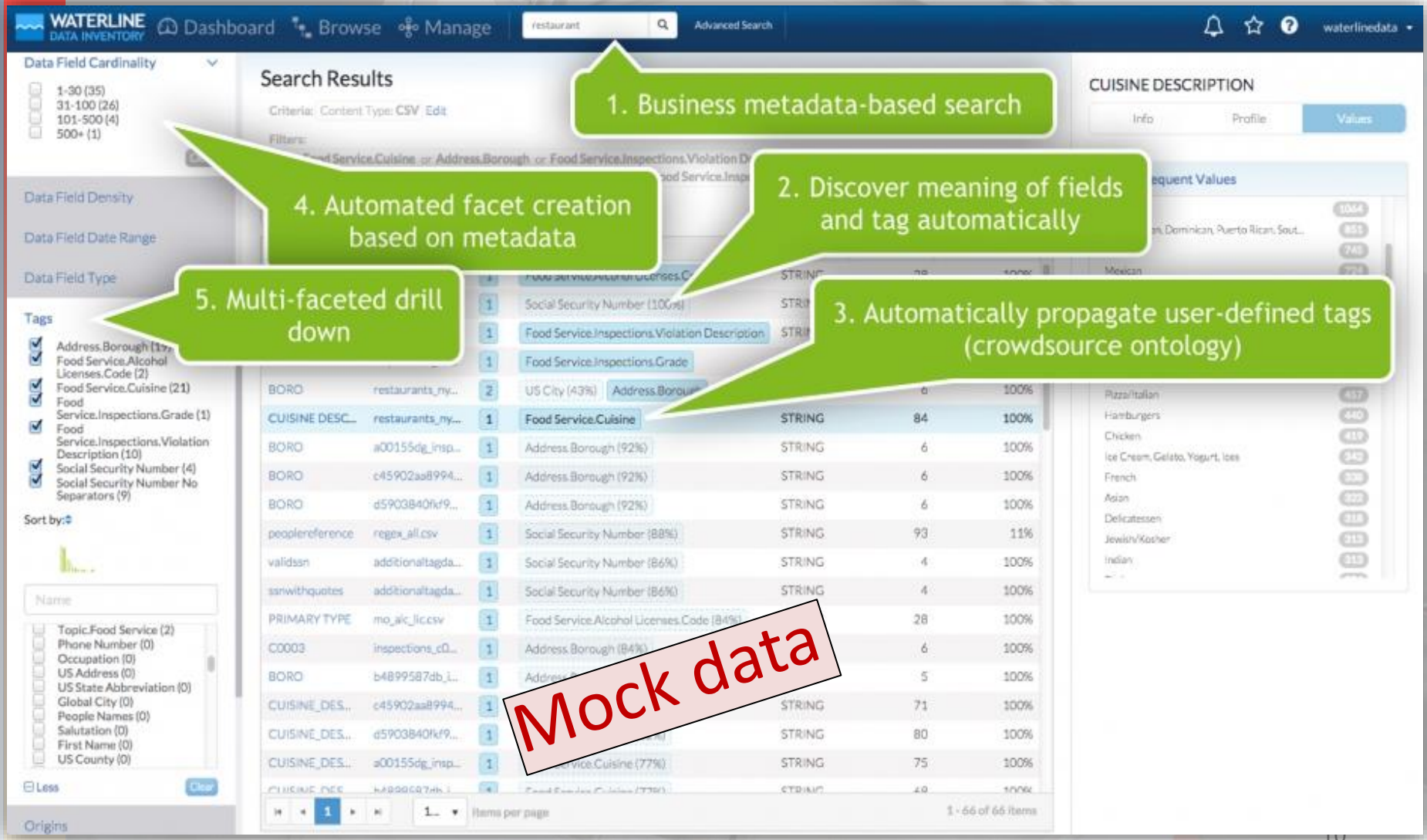
		ownHomeFlag	contactPreference	leadSource	dateOfBirth	salesClass	
1	ervices	N	phone	Partner	12/4/89	3	M
2	s	N	donotcall	Twitter	11/18/94	2	S
3		N	mail	Twitter	7/16/94	4	M
4	s	N	email	Facebook	6/8/85	1	S
5	s	N	facebook	Twitter	7/31/95	2	S
6	s	Y	donotcall	List Purchase	3/29/91	2	S
7		N	phone	Outbound	4/17/93	5	S
8		N	mail	LinkedIn	10/26/85	2	Married
9		N	mail	LinkedIn	1/23/89	4	M
10		N	donotcall	Referral	7/2/88	2	S
11		N	phone	Web Activity	6/22/90	2	M
12	Services	N	facebook	Other	8/31/88	1	S
13		Y	none	Twitter	2/7/90	3	Single
14		N	none	LinkedIn	3/29/89	3	Single
15	Services	N	donotcall	Google	3/16/83	5	M
16	Services	N	none	Outbound	7/2/88	5	M

Mock data

Complicating Factors: Metadata

- Who, what, why about our data
- Different types of data = different types of metadata
- Documentation is static, outdated or doesn't exist
- Historical context and description of changes across time largely absent
- Granularity differs across data types and across time
- Tags to find appropriate data largely absent

IMPALA: Waterline

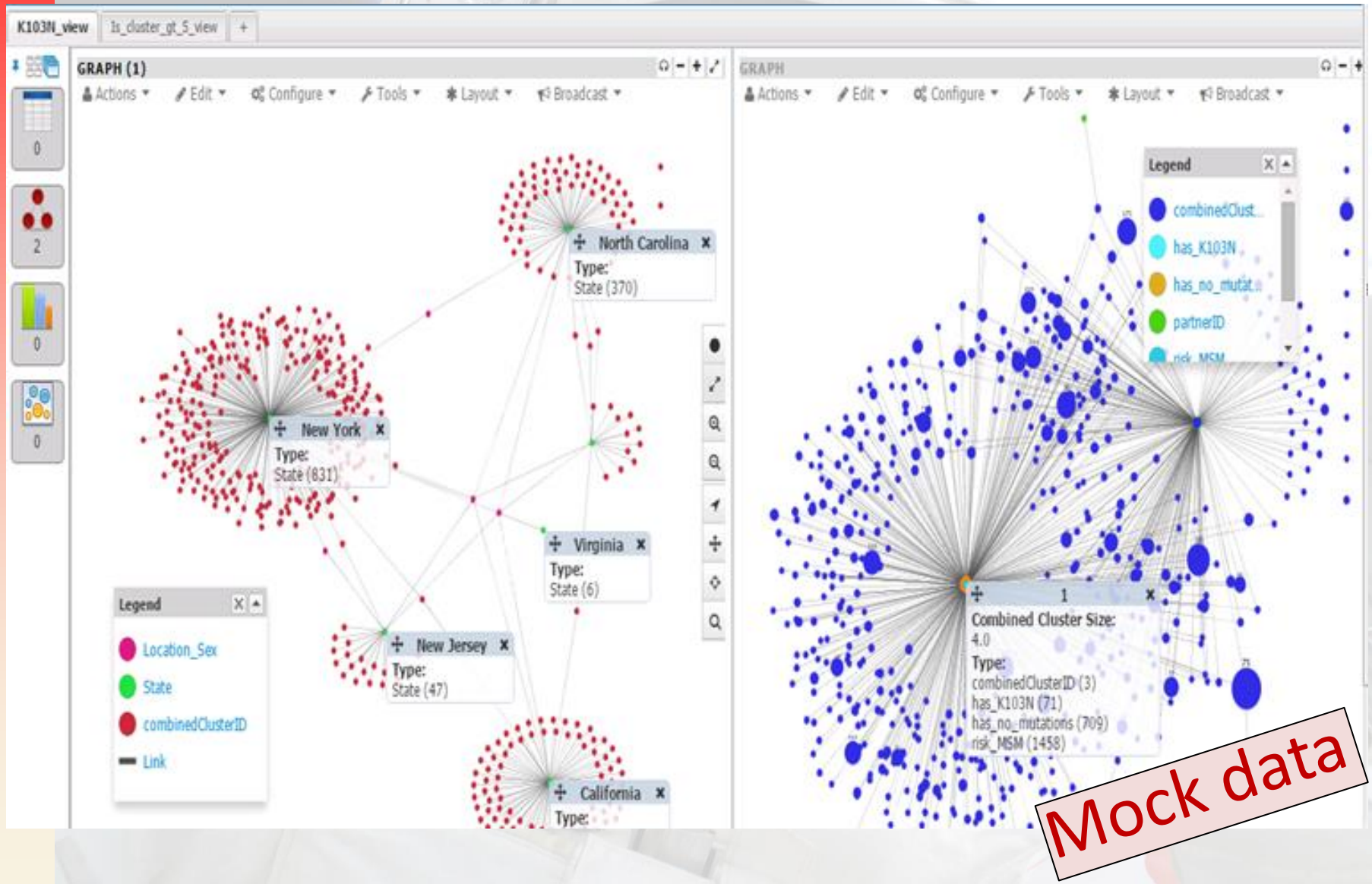


Complicating Factors: Data Analytics and Visualizations

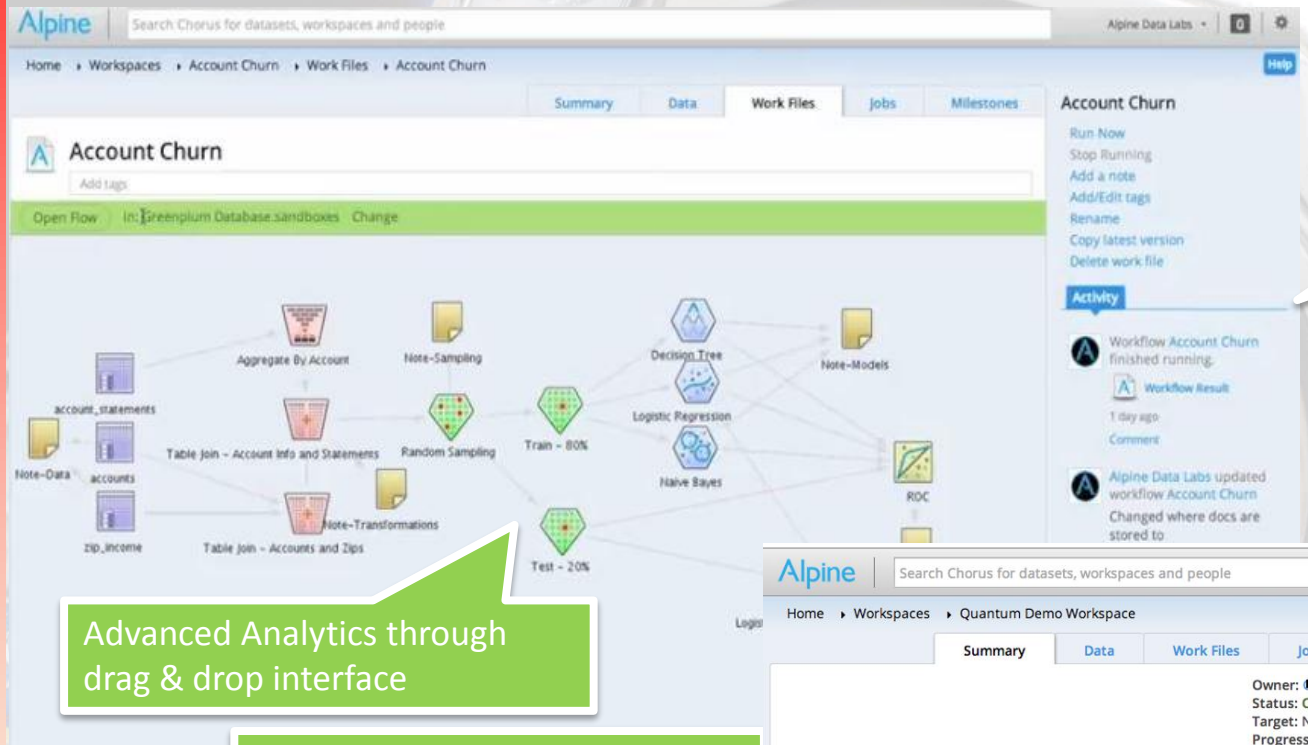

```
title 'QTc Change from Baseline by Week and Treatment';
proc sgplot data=QTcBand;
  format week qtcweek.;
  styleattrs datalinepatterns=(solid);
  band x=wk lower=L0  upper=L30 / fill legendlabel='Normal'
       fillattrs=(color=white transparency=0.6);
  band x=wk lower=L30 upper=L60 / fill  legendlabel='Concern'
       fillattrs=(color=lightgray transparency=0.6);
  band x=wk lower=L60 upper=L90 / fill legendlabel='High'
       fillattrs=(color=gray transparency=0.6) ;
  vbox qtc / category=week group=drug groupdisplay=cluster nofill;
  scatter x=wk y=QTc / group=drug name='a' nomissinggroup;
  text x=wk y=ylabel text=label / contributeoffsets=none;
  xaxistable risk / class=drug colorgroup=drug location=inside;
  refline 26 / axis=x;
  xaxis type=linear values=(1 2 4 8 12 16 20 24 28) valueshint
        min=1 max=29 display=(nolabel)
        colorbands=odd colorbandsattrs=(transparency=1);
  yaxis label='QTc change from baseline' values=(-120 to 90 by 30);
  keylegend 'a' / title='Treatment:' linelength=20;
run;
```

```
proc genmod data=resp descend;
   class id treatment(ref="P") center(ref="1") sex(ref="M")
      baseline(ref="0") / param=ref;
   model outcome=treatment center sex age baseline / dist=bin;
   repeated subject=id(center) / corr=unstr corrw;
run;
```

IMPALA: Visualizations



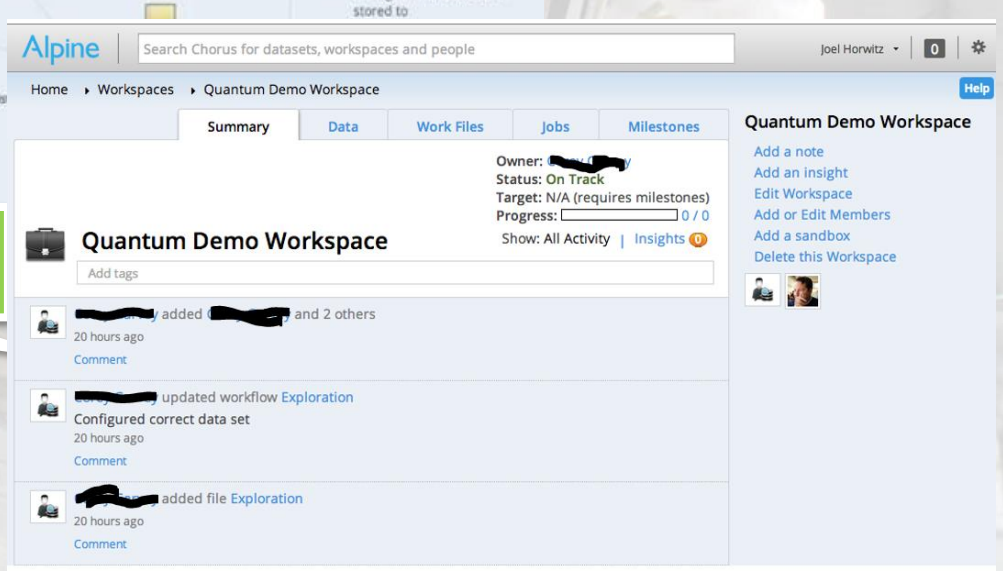
IMPALA Workflow
Alpine



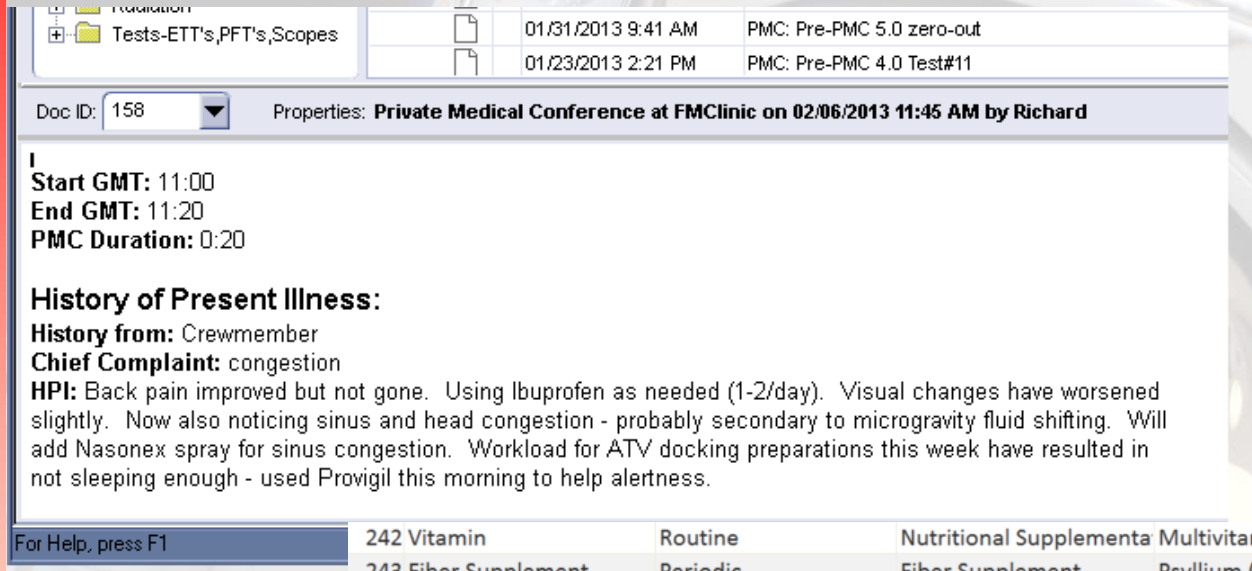
Native integration with R & Python scripts & executables

Advanced Analytics through drag & drop interface

Collaborate with team through a dedicated workspace



Complicating Factors: Text Data



242	Vitamin	Routine	Nutritional Supplementa	Multivitamin	9999	9999	Oral
243	Fiber Supplement	Periodic	Fiber Supplement	Psyllium (Metamucil)	9999	9999	Oral
244	Antilipemic	Routine	Hyperlipidemia	Atorvastatin (Lipitor)	9999	9999	Oral
245	Decongestant	Periodic	Other	Oxymetazoline (Afrin Nasal Spr	9999	9999	Nasal
246	Antihypertensive	Routine	Hypertension	Lisinopril (Zestril)	9999	9999	Oral
248	Analgesic	Periodic	Pain	Ibuprofen (Motrin)	9999	9999	Oral
249	Antibiotic/Corticosteroic	Periodic	Irritation	Tobramycin/Dexamethasone (T	9999	9999	Optic
250	Lubcricant	Periodic	Irritation	Carboxymethylcellulose OPHTH	.5	%	Optic
251	Analgesic	Periodic	Pain	Acetaminophen (Tylenol)	325	mg	Oral
252	Antibiotic/Corticosteroic	Periodic	Irritation	Tobramycin/Dexamethasone (T	9999	9999	Optic
253	Vitamin	Routine	Nutritional Supplementa	Multivitamin	9999	9999	Oral
254	Analgesic	Periodic	Congestion	Ibuprofen (Motrin)	600	mg	Oral
255	Decongestant	Periodic	Congestion	Oxymetazoline (Afrin Nasal Spr	9999	9999	N
256	Antibiotic	Periodic	Irritation	Ciprofloxacin OPHTH Oint (Cilo	9999		
257	Antibiotic	Periodic	Irritation	Ciprofloxacin OPHTH Soln (C			

Mock data

Complicating Factors: Text Data

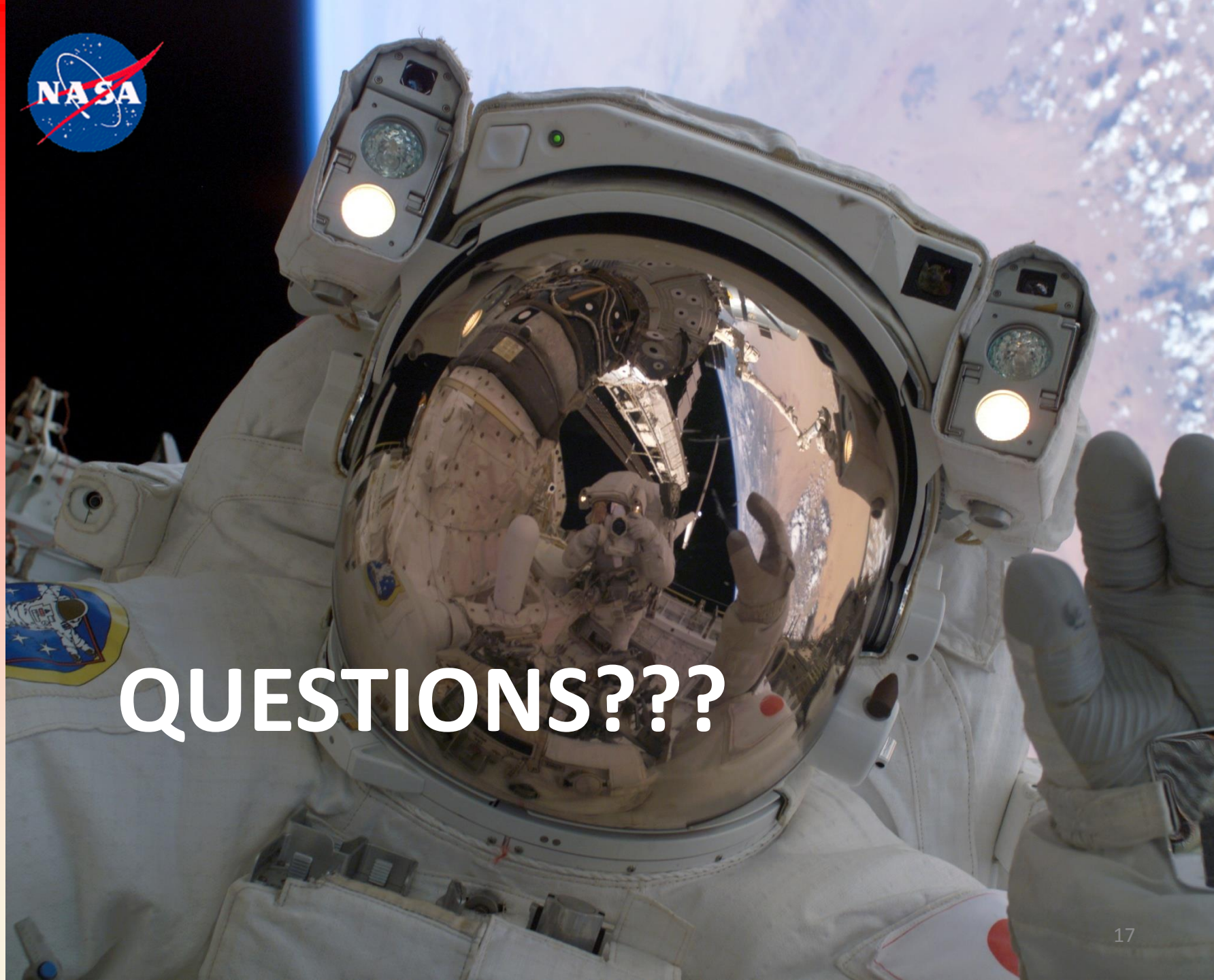
"CM WORE HIS PRIME TRAINING GLOVES AND REPORTED **'SORE' FINGERTIPS.** THE SIZING OF THE GLOVES WAS REPORTED AS ACCEPTABLE, WITH NO RESIZING REQUESTS. FOR THE TIME BEING, CM REQUESTED TO **REMOVE THE NAIL HARDENER** FROM HIS CREW PREFERENCE ITEMS. ALL OTHER SUIT FIT COMMENTS WERE FAVORABLE. "

Over 3000 pages of Sizing Comments Unable to be Analyzed

Mock data

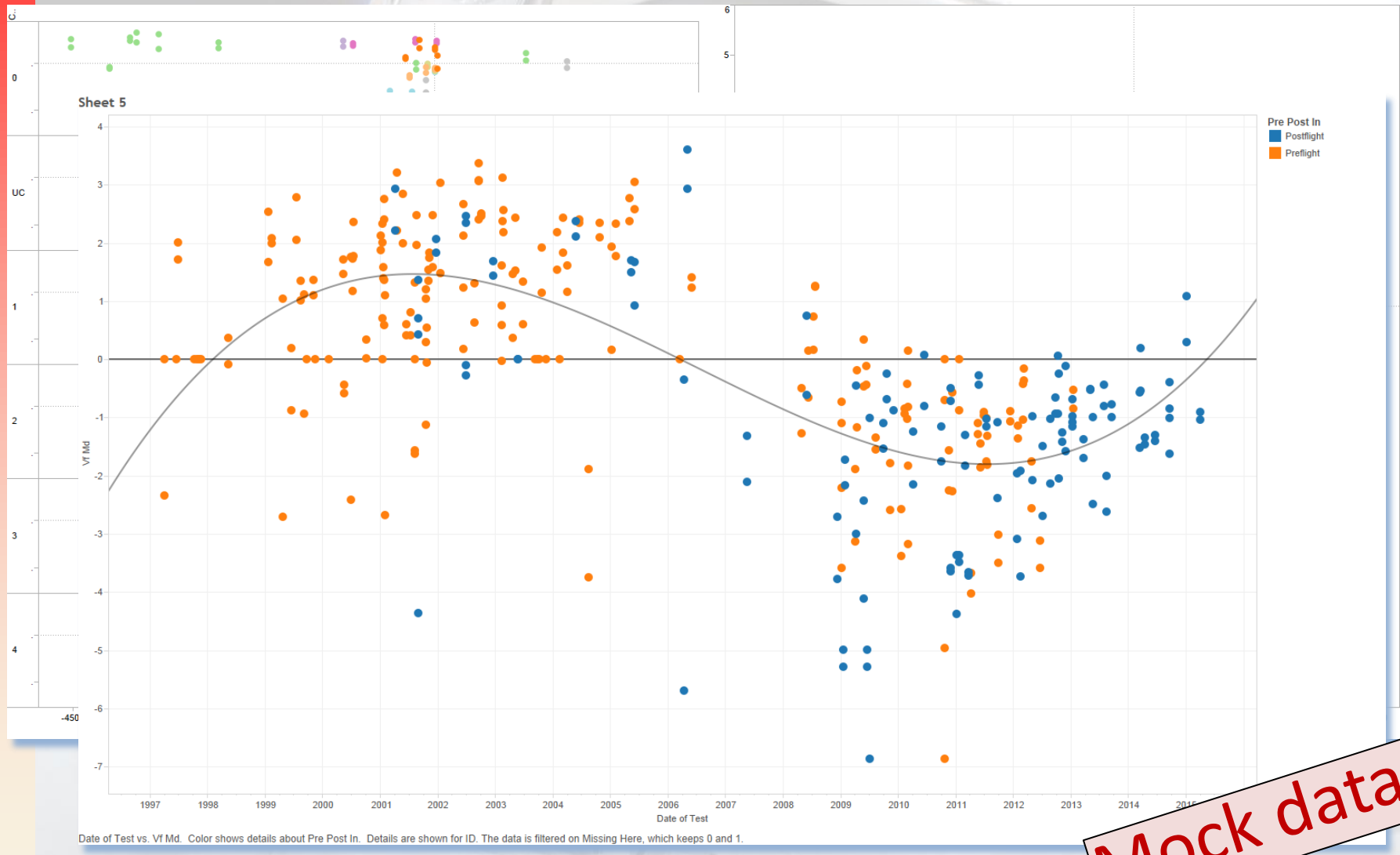
Summary

- Astronaut data and analytics of astronaut data is complex
- Creative and unique solutions are needed and some are currently being implemented
- Solutions for qualitative analysis including NLP are still needed


QUESTIONS???

Back Up

Trending and Analyses



Mock data