From questions to interventions

Towards an integrated model for using paradata to monitor and improve fieldwork.

Kyle Fennell and Tom W. Smith

Presented at 1st DASISH Quantitative Workshop
December 2012

NORC at the UNIVERSITY of CHICAGO

Outline

NORC at the UNIVERSITY of CHICAGO

1. Introduction to NORC
2. Paradata basics
3. Capturing paradata
4. Managing paradata
5. Using paradata
6. Recommended framework

Topic 1

NORC at the UNIVERSITY of CHICAGO

Introduction to NORC

NORC

- Founded in 1941, NORC's mission is to conduct high-quality social science research in the public interest. Our work is grounded in a commitment to research excellence, innovation, dissemination of data and findings, and collegiality.
- Our capabilities include research design, data collection, analysis, and dissemination.
- Visit http://www.norc.org for more details.

Who we are

Senior Fellow Tom W. Smith directs NORC's Center for the Study of Politics and Society. Since 1980, he has served as Director of the General Social Survey (GSS), one of NORC's most visible projects and one of the nation's most heavily utilized datasets. He is also co-founder of the International Social Survey Program (ISSP), former Secretary General of the ISSP, and currently serving on the ISSP Standing and Methodology Committees.

Kyle Fennell is Associate Director of Field Operations at NORC. Since he was hired in 2002, Fennell has participated in every aspect of in-field survey operations including recruiting, training, and management of field resources. Fennell has developed processes and systems for monitoring field interviewer recruiting, cost, and production which are essential to the efficient management of NORC’s large-scale field studies. Fennell's expertise includes staffing, survey management, and the use of paradata as a decision making tool.

Selected NORC Studies

NORC at the UNIVERSITY of CHICAGO

General Social Survey (GSS)	National Longitudinal Survey of Youth 1979 (NLSY79)	National Longitudinal Survey of Youth 1997 (NLSY97)	Survey of Consumer Finances (SCF)
National Survey of Early Care and Education (NSECE)	National Social Health and Aging Project (NSHAP)	Residential Energy Consumption Survey (RECS)	National Immunization Survey (NIS)
Survey of Doctorate Recipients (SDR)	Survey or Earned Doctorates (SED)	Data Enclave	International Projects

Topic 2

NORC at the UNIVERSITY of CHICAGO

Paradata basics

What are Paradata?

NORC at the UNIVERSITY of CHICAGO

- "Paradata are automatic data collected about the survey data collection process captured during computer assisted data collection, and include call records, interviewer observations, time stamps, keystroke data, travel and expense information, and other data captured during the process." (Kreuter, Couper, and Lyberg, 2010)

What are Paradata, 2

NORC at the UNIVERSITY of CHICAGO

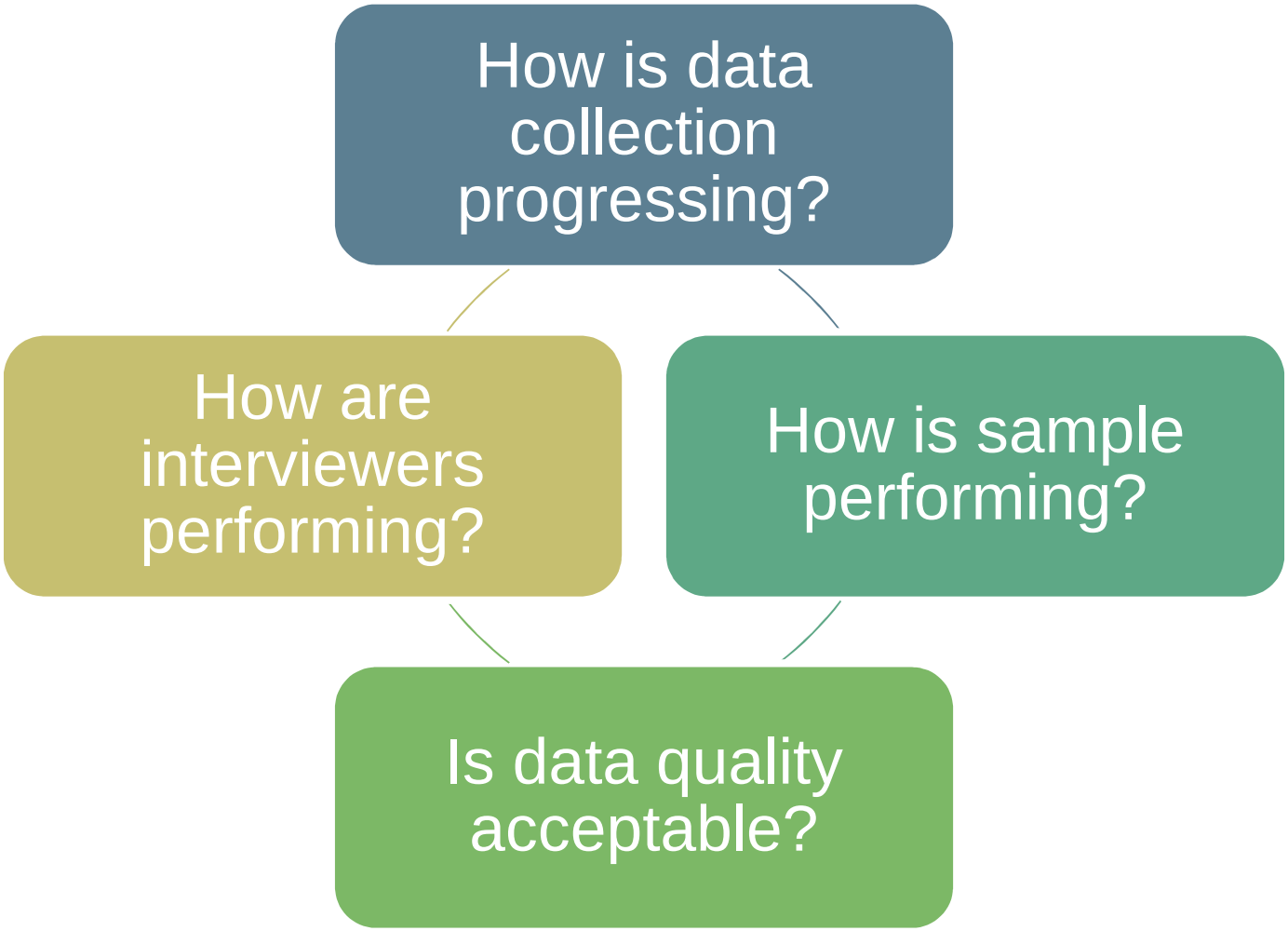
Micro and Macro

Administrative, Procedural, and Observational

Byproduct and by design

Source: JANS, ET AL.(2010)

Questions Paradata Help Answer



Topic 3

NORC at the UNIVERSITY of CHICAGO

Capturing Paradata

Interviewer Generated Paradata

NORC at the UNIVERSITY of CHICAGO

- Interviewers generate paradata when they
  - Make observations about Housing Units
  - Select dispositions
  - Enter comments during interviews
  - Write records of contact (ROCs)
- Interviewer generated paradata often
  - Require labor to produce
  - Include errors and missing data
  - CAI systems can reduce error and increase efficiency

Observational Paradata

NORC at the UNIVERSITY of CHICAGO

Use interviewer observations to validate sample design

Use post-interview debriefing comments to help improve data quality

Link interviewer comments, GPS data, and pictures to improve data quality

Computer Generated Paradata

- Computer generated paradata
  - Can be captured anytime an “event” occurs.
    - Time stamps for entering and exiting screens in a questionnaire
    - Recordings which begin and end at set intervals
    - Off path data trail
    - Coordinates from GPS devices
  - Can be derived data from interviewer data (most advanced disposition, length of comments,….)
  - Require planning and investment prior to data collection
  - Can overwhelm staff

Paradata and PAPI

NORC at the UNIVERSITY of CHICAGO

It is possible to collect paradata during PAPI efforts

- Use an IVR system so that interviewers can call in reports
- Have interviewers provide verbal progress reports to managers
- Data enter paper contact logs along with questionnaire data.
- Pair online case management system with paper questionnaires

Recommendations

NORC at the UNIVERSITY of CHICAGO

- Think about capturing paradata when designing systems
- Automate as much as possible and use structured forms for everything else.
- Estimate the cost of collecting and using paradata
- Don’t attempt to be too precise, but gather as much detail as you can
- Pretest the paradata collection tool and process for retrieving data from interviewers
- Train staff on paradata collection and monitor quality from start of effort.

Topic 4

NORC at the UNIVERSITY of CHICAGO

Managing paradata

Storage

NORC at the UNIVERSITY of CHICAGO

- Secure locations for raw data and for data files needed for analysis
- Secure enough space for the entire field period
- Be prepared for case resets
- Set up data access rights to protect data while making approved use as easy as possible.

Cleaning and processing

NORC at the UNIVERSITY of CHICAGO

Check data before using reports

Expect errors and missing data

Correct errors in the source file

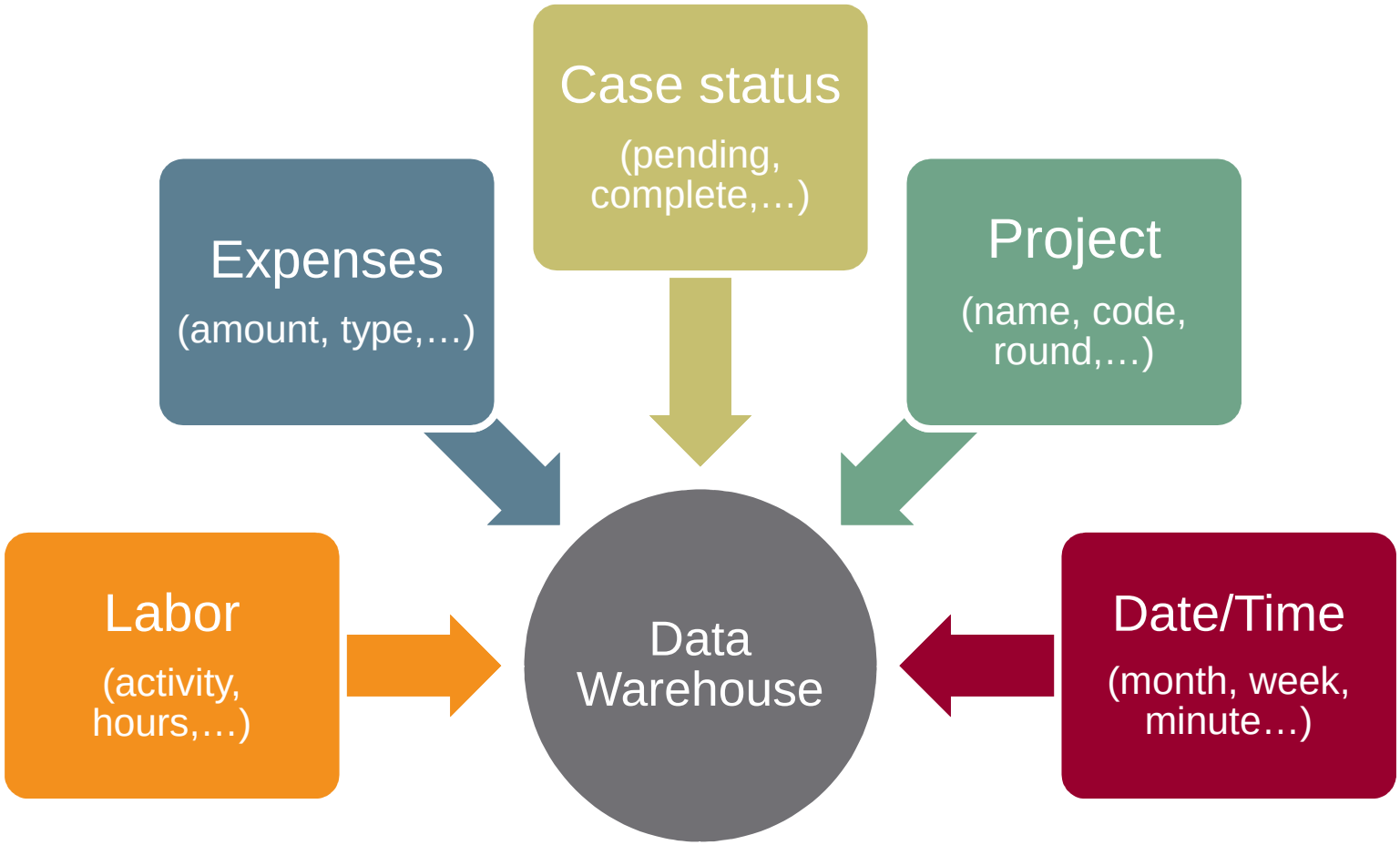
Recode to reduce noise

Topic 5

NORC at the UNIVERSITY of CHICAGO

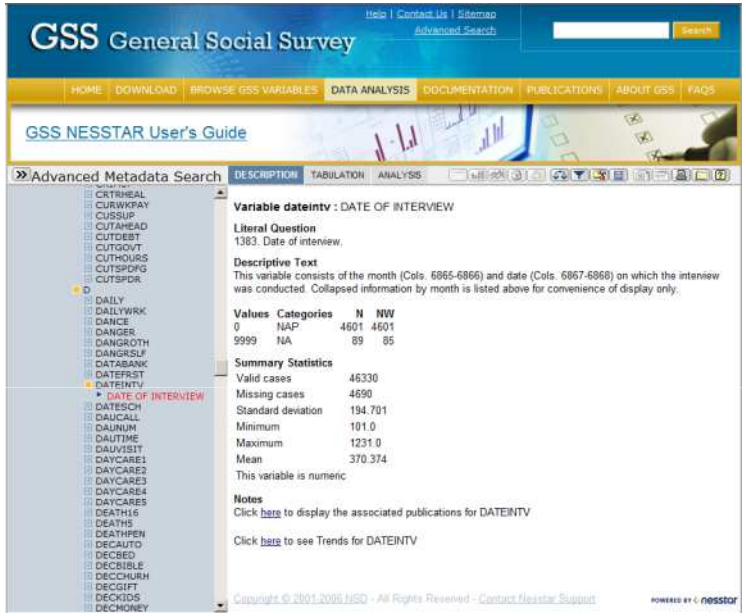
Using paradata for analysis and project management

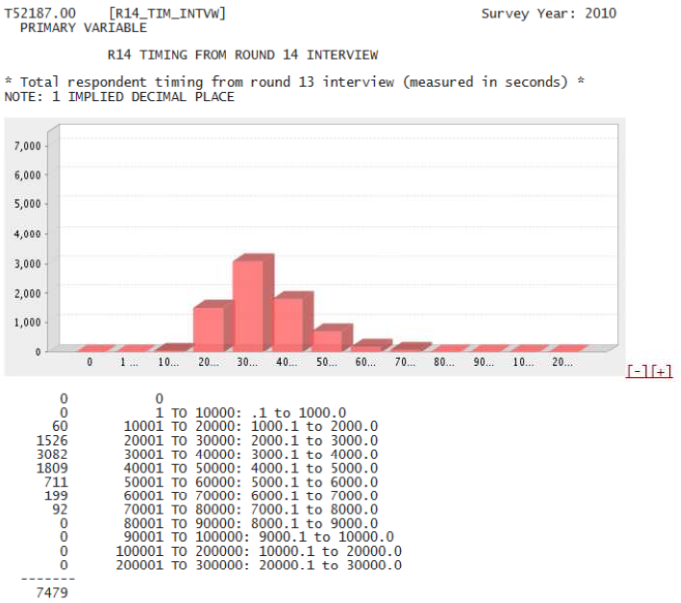
NORC Data Warehouse



Lesson: It is possible to harmonize paradata across projects

Paradata in public data files





GSS: http://www3.norc.org/GSS+Website/Data+Analysis/

NLSY: https://www.nlsinfo.org/investigator/

Lesson: Paradata can be integrated into public use data

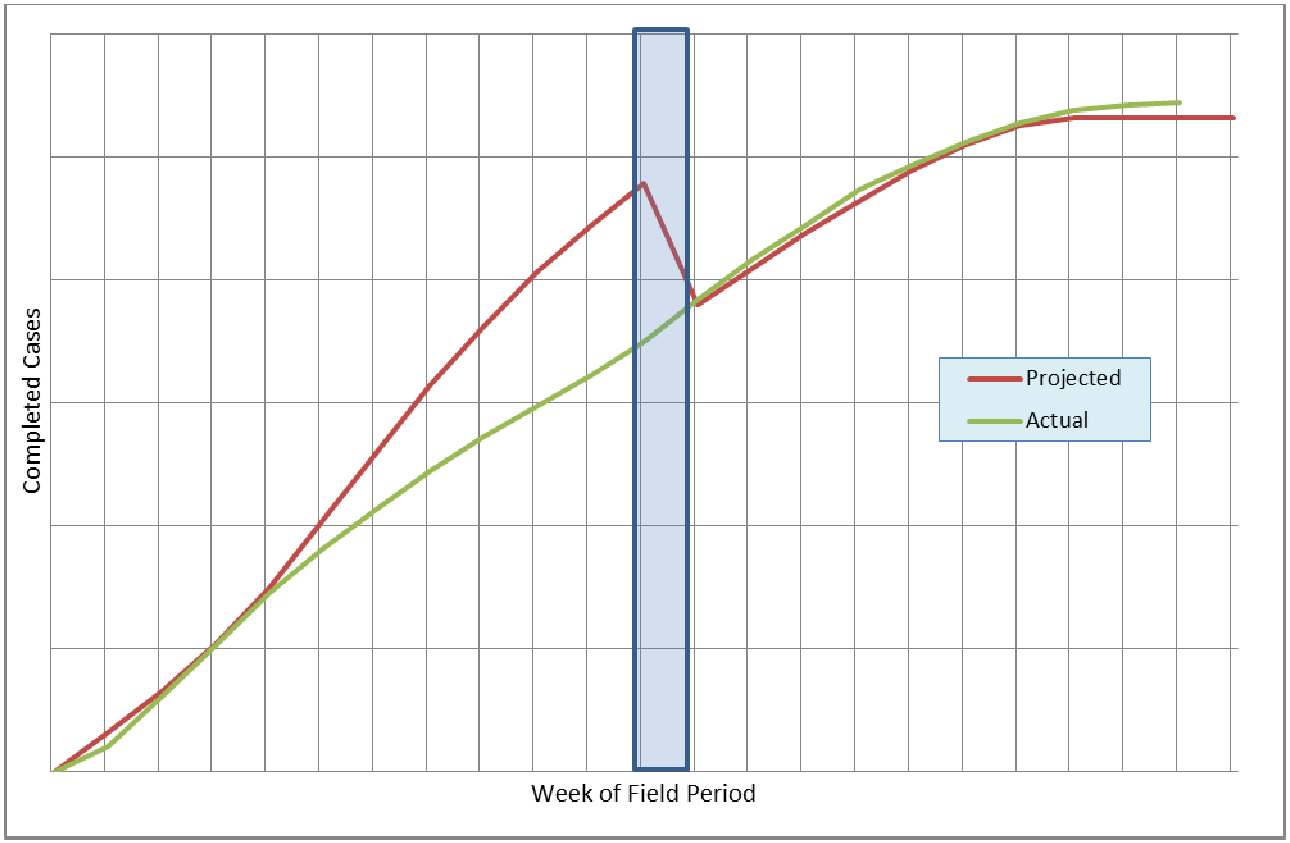
Example: Sample Status

	Week 1	Week 2
1. SAMPLE STATUS		
Assigned Lines	7725	7725
(+) Missed Housing Units	37	46
(-) Out of Scope	249	323
NET SAMPLE	7513	7448
2. PRE-HEF PENDING CODES		
No Action (00)	1642	1001
FI Locating (01)	168	177
FM Locating Needed (02)	0	69
CO Locating Needed (03)	0	0
Case Returned WITH New Info (04)	0	0
Case Returned WITHOUT New Info (05)	0	0
Not Accessible HU (06)	21	151
No One Home (07)	10	1644
Spanish Needed (09)	324	56
Callback (General) (10)	49	782
HEF Appointment (11)	0	0
HEF Broken Appointment (12)	0	0
Temporary Refusal for HEF (13)	245	368
TOTAL PENDING PRE-HEF	4892	4251
3. HEF PENDING CODES		
Partial/HEF (14)	0	53
Partial/HEF Done/Quex Pnd (15)	0	272
HEF Other (16)	85	4
TOTAL PENDING HEF	433	452
4. QUEX PENDING CODES		
R not Home (18)	0	254
R Permanently Incapacitated (19)	0	0

	Week 1	Week 2
3. HEF PENDING CODES		
Partial/HEF (14)	0	53
Partial/HEF Done/Quex Pnd (15)	0	272
HEF Other (16)	85	4
TOTAL PENDING HEF	433	452
4. QUEX PENDING CODES		
R not Home (18)	0	254
R Permanently Incapacitated (19)	0	0
Quex Refused by HU Contact (29)	196	104
Quex Refused by R (31)	0	267
Letter Sent (33)	0	0
Letter Request (32)	0	0
Quex Appt by HU Contact (34)	15	16
Quex Appointment by R (36)	0	327
Quex Broken Appointment (37)	356	115
Partial/Quex (38)	104	18
Quex Other (39)	17	14
Interm Unlocatable (50)	0	0
Interm Not Accessible HU (51)	0	0
Intm R Absent All Field Prd (53)	0	0
Interm Entire HU Unavailbl (54)	0	0
Interm Refusal for HEF (55)	0	0
Interm Refusal for Quex (56)	0	0
Intm R Permntly Incapacitd (57)	0	0
Interm Other (58)	0	0
5. FM SPECIAL CODES		
FM Hold (59)	0	3
TOTAL PENDING QUEX	1069	1186

Lesson: Paradata can support responsive designs.

Example: Actual vs. Projected Production



Lesson: Use paradata to adjust expectations early on.

Example: NIR/OOS Reasons

NORC at the UNIVERSITY of CHICAGO

Final NIR/OOS Summary						
		Region				
Description	Total	1	2	3	4	5
All Final NIR/OOS	**416**	**83**	**80**	**73**	**88**	**92**
Other NIR reason	0	0	0	0	0	0
Inaccessible - prison/other	16	5	3	2	5	1
Respondent too ill/handicapped	6	1	0	0	0	5
Inaccessible - military	15	6	2	4	2	1
Unlocatable	46	12	5	12	8	9
Very hostile refusal	20	8	3	5	2	2
Hostile refusal	19	1	6	1	10	1
Refusal	122	5	12	25	30	50
Gatekeeper refusal	27	8	4	8	5	2
Deceased in current round	4	2	0	1	1	0
Blocked Cases	110	25	40	12	16	17
Deceased in prior round	31	10	5	3	9	4

Lesson: Sometimes detail is useful, but be prepared for noise

Example: Contact Times

NORC at the UNIVERSITY of CHICAGO

Contact attempts by day of week and time of day								
Contact window	**Sun**	**Mon**	**Tues**	**Wed**	**Thurs**	**Fri**	**Sat**	**All Days**
Slot 1: Before 9 AM	1		1		1	2	2	7
Slot 2: 9 AM-12 PM	13	29	5	18	11	21	45	142
Slot 3: 12 - 3 PM	33	31	16	19	34	44	38	215
Slot 4: 3 - 6 PM	24	20	20	12	24	24	45	169
Slot 5 : 6 - 9 PM	8	19	9	21	13	14	4	88

Lesson: Be ready to take action if data show divergence from protocols

Example: Benchmarks

Benchmark Update				
Benchmark	Target	Actual	% of target	Progress
Interviews during the week	100	59	59	
Interviews yesterday	17	21	124	
Cases worked yesterday	285	283	99	
Interviewers working yesterday	47	49	104	

- The weekly target is calculated by subtracting the actual interview total as of the end of the prior week from the cumulative production goal for the selected week.
- Because weekend production has been slower than weekday production, daily production targets for Saturday and Sunday are 50% of the weekday target. The daily production target for weekdays during the selected week is 17.
- During the last 7 days, the project has averaged one completed interview for each 17.12 cases worked by an FI. The targeted number of cases to work equals the daily interview target multiplied by 17.12.
- During the last 7 days, interviewers have worked an average of 6.02 cases on the days they worked. The targeted number of FIs working equals the goal for cases worked divided by 6.02.

Lesson: Be careful about sharing data if some stakeholders do not want performance to be public

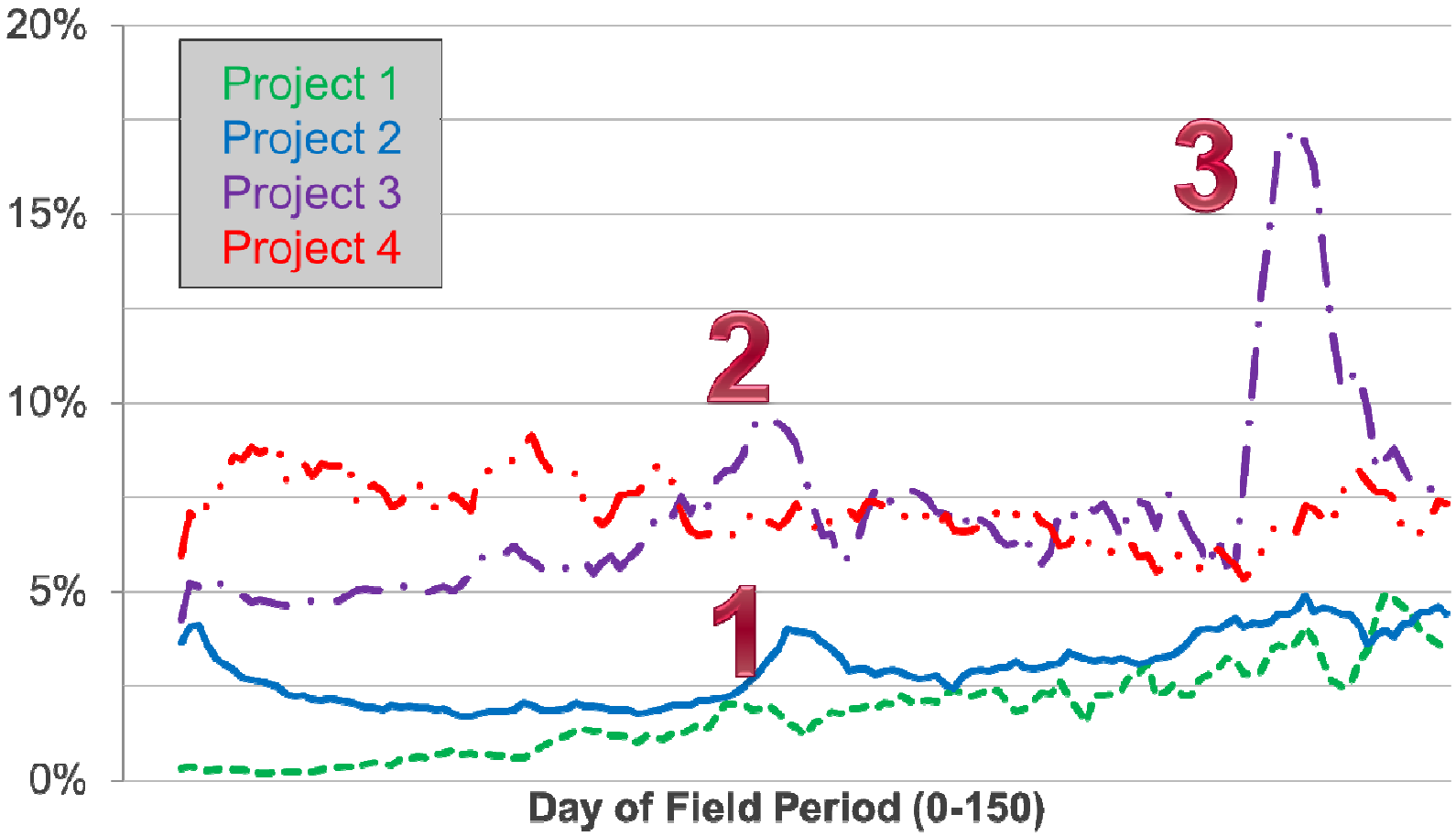
Example: Monitoring Labor

NORC at the UNIVERSITY of CHICAGO

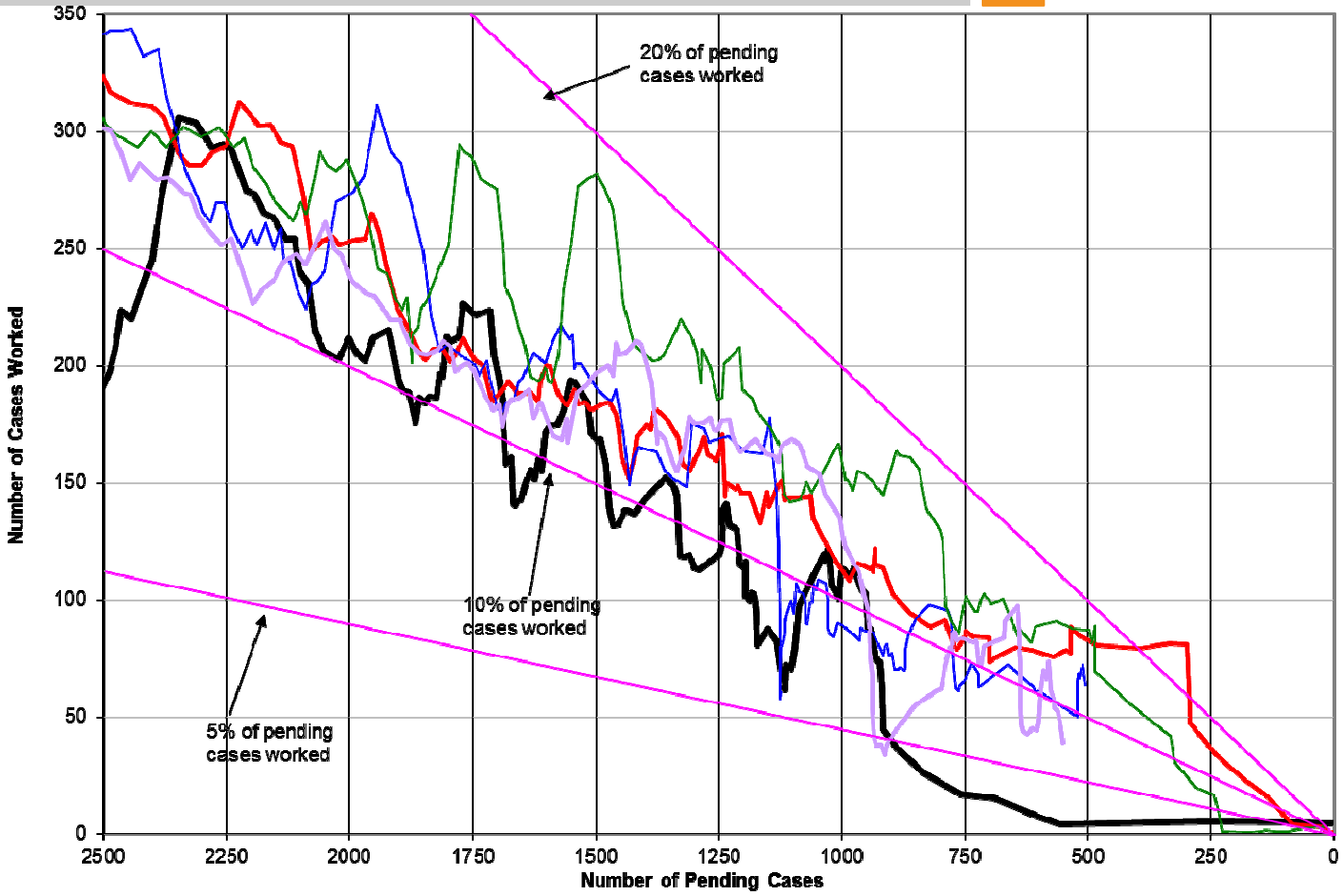
	Interviewers			Labor		Productivity	
Week	Active	Working	% Working	Hours	Hours/FI	FIs Completing a Case	% Completing a Case
Week 1	90	85	94%	2,771	32.60	81	95%
Week 2	140	130	93%	2,559	19.68	130	100%
Week 3	200	175	88%	2,500	14.29	149	85%
Week 4	250	200	80%	4,500	22.50	180	90%
Week 5	250	225	90%	4,250	18.89	169	75%

Lesson: Production and efficiency issues are often first evident in labor reports. Pay close attention to these data.

Example: Refusals (7 day running average)

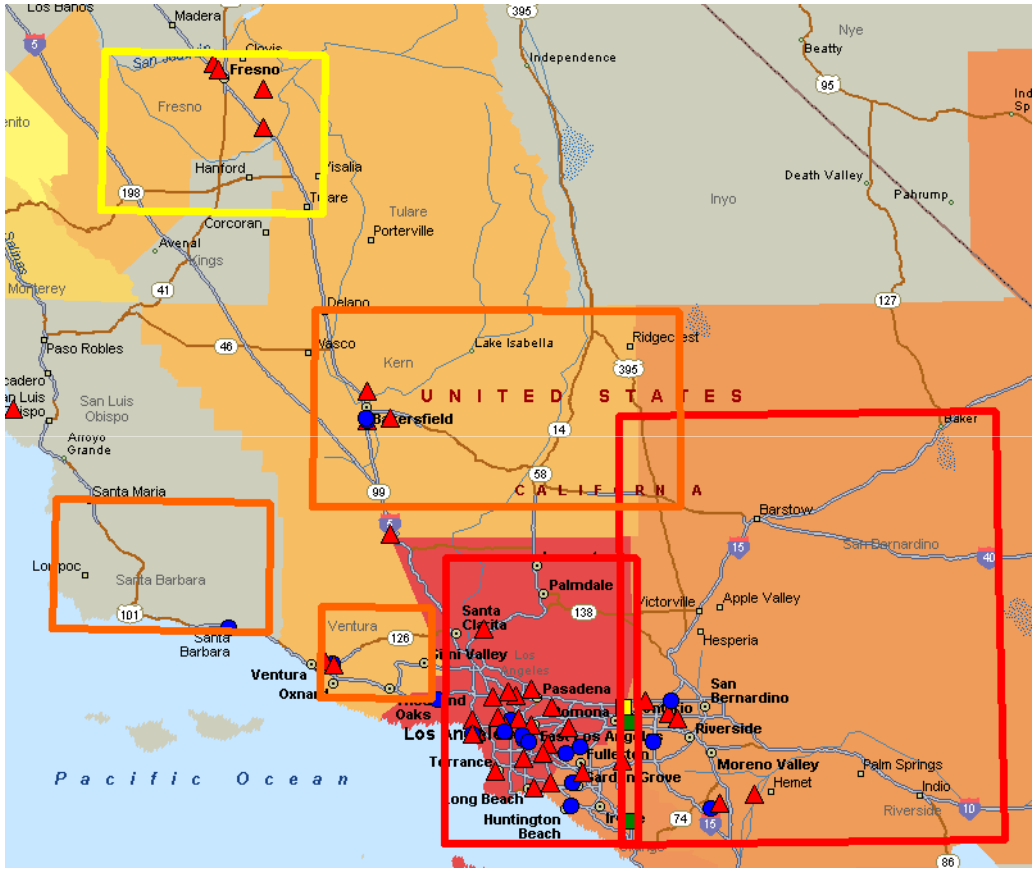


Example: Monitoring Level of Effort



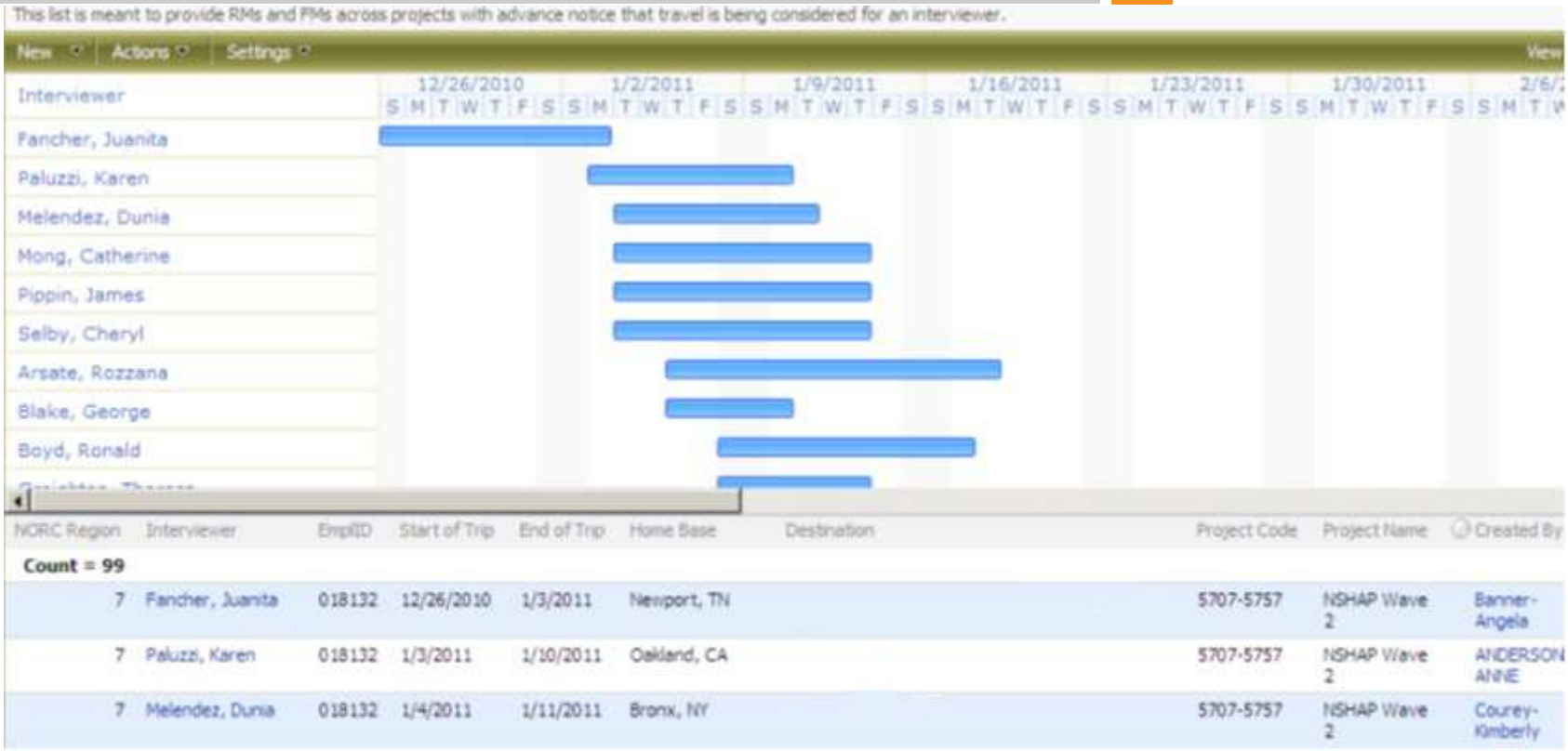
Lesson: Watch for spikes and dips in level of outreach. These can be early indicators of struggles in the field.

Example: Matching Staff to Work Locations



Lesson: Mapping and use of GIS can help clarify patterns in your data. They can also distract.

Example: Coordinating Travel



Lesson: Sometimes it is best to share raw data rather than a report.

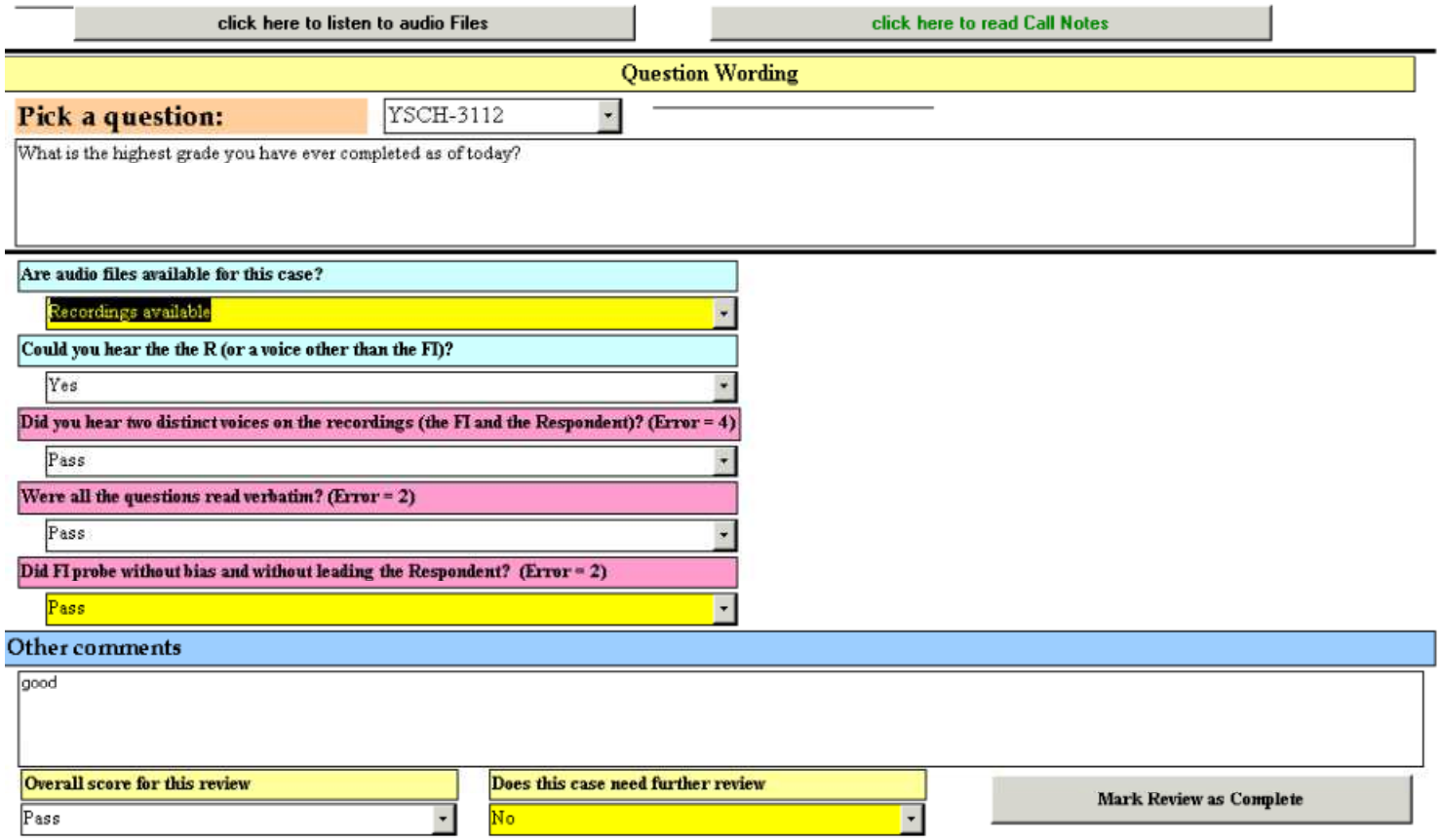
Example: Quality Metrics

NORC at the UNIVERSITY of CHICAGO

Report Week	Region	Inteviewer	Overall Status	# Send /Receives	# Of ROCs	Item Non-response	CARI consent
Week 12	C	Aaron	Attention suggested	12	94	5.71%	100%
Week 12	C	Beverly	Doing Fine	25	63	3.39%	100%
Week 12	C	Candace	Review performance	3	15	8.83%	100%
Week 12	C	Doug	Review performance	2	8	4.18%	100%
Week 12	C	Ethel	Review performance	15	20	7.69%	89%
Week 12	C	Fran	Review performance	3	39	4.95%	89%
Week 12	C	Grace	Doing Fine	34	190	4.42%	90%
Week 12	C	Howard	Attention suggested	9	42	6.09%	100%
Week 12	C	Inez	Attention suggested	5	65	3.90%	95%

Lesson: Use text and color to help guide interpretation of data (but test your strategy with users first).

Example: CARI Review



Lesson: Be wary of plans which require extensive manual coding of paradata before they can be used.

Topic 6

Framework for using paradata

NORC at the UNIVERSITY of CHICAGO

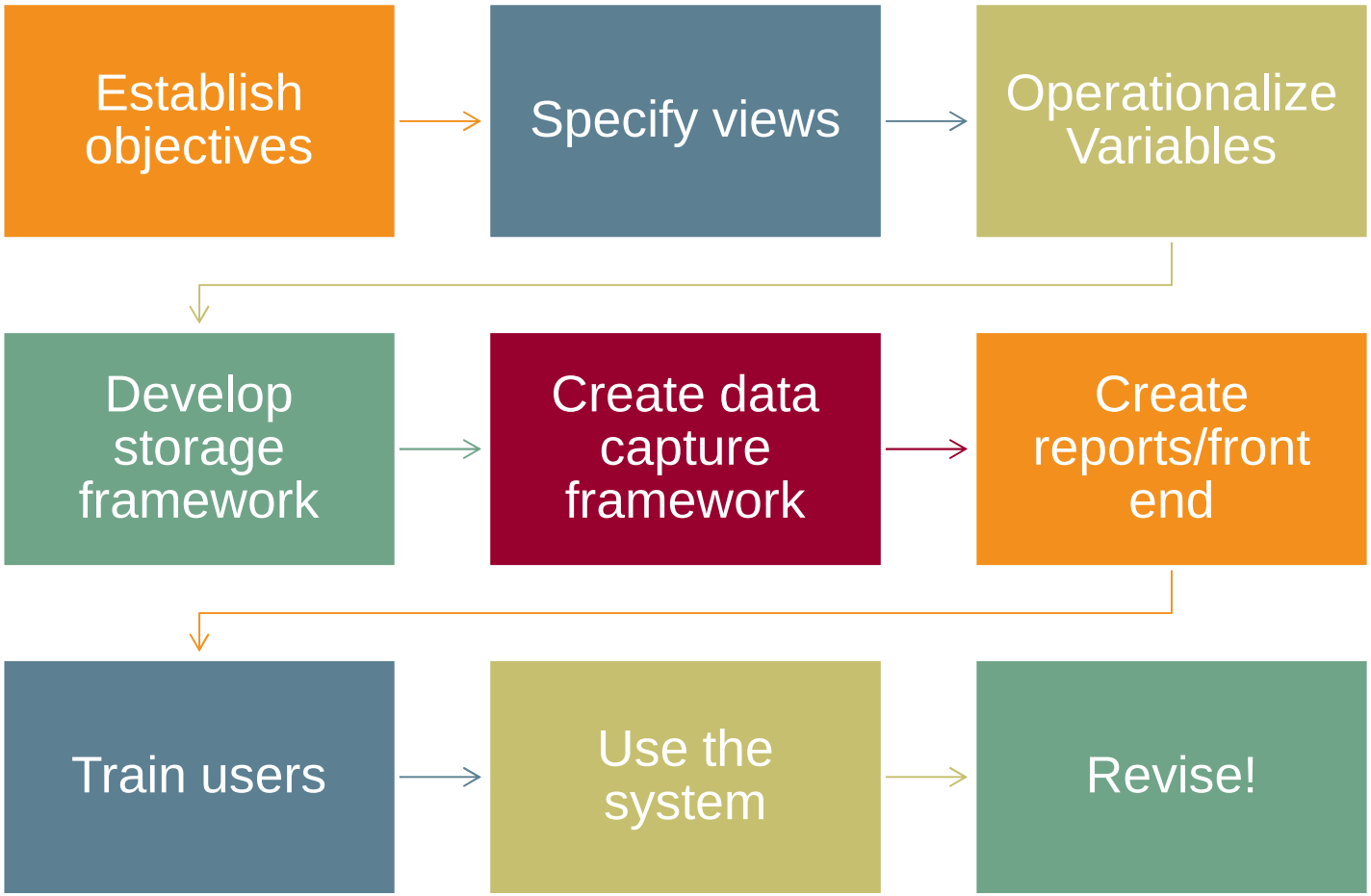
Identify your Priorities

NORC at the UNIVERSITY of CHICAGO

Empower, understand, or control?

Raw data, reports, or dashboards?

Suggested framework



Topic 7

NORC at the UNIVERSITY of CHICAGO

Discussion

Kyle Fennell
NORC
55 East Monroe St, 30th Floor
Chicago, IL 60618
312-759-4055
Fennell-kyle@norc.org

Thank You!

NORC at the UNIVERSITY of CHICAGO

insight for informed decisions™