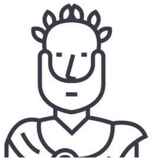
Data analysis

- Qualitative
  - Interview analysis
  - Hypothesis generation
  - Deductive analysis
- Quantitative
  - Exploratory Data Analysis
  - Integration
  - Survey analysis
  - Modeling and validation
- Best practices

Qualitative Data Analysis

Deductive vs. Inductive Reasoning



DEDUCTIVE

Theory

↓

Hypothesis

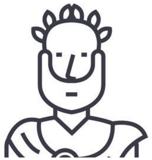
↓

Observation

↓

Confirmation

Deductive vs. Inductive Reasoning



DEDUCTIVE

Theory

↓

Hypothesis

↓

Observation

↓

Confirmation

INDUCTIVE

Theory

↑

Hypothesis

↑

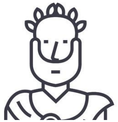
Pattern

↑

Observation

Qualitative analysis is *usually* inductive

INDUCTIVE



DEDUCTIVE

Punch, K. F., (2014) Introduction to social research: Quantitative and qualitative approaches (3rd ed.). Thousand Oaks, CA: Sage..

How qualitative analysis fits into mixed methods

Exploratory Sequential

Deep, foundational study first

- Understand people's relationships with what you are studying
- Develop hypotheses about user satisfaction within this domain to explore quantitatively
- Tends to involve induction exclusively

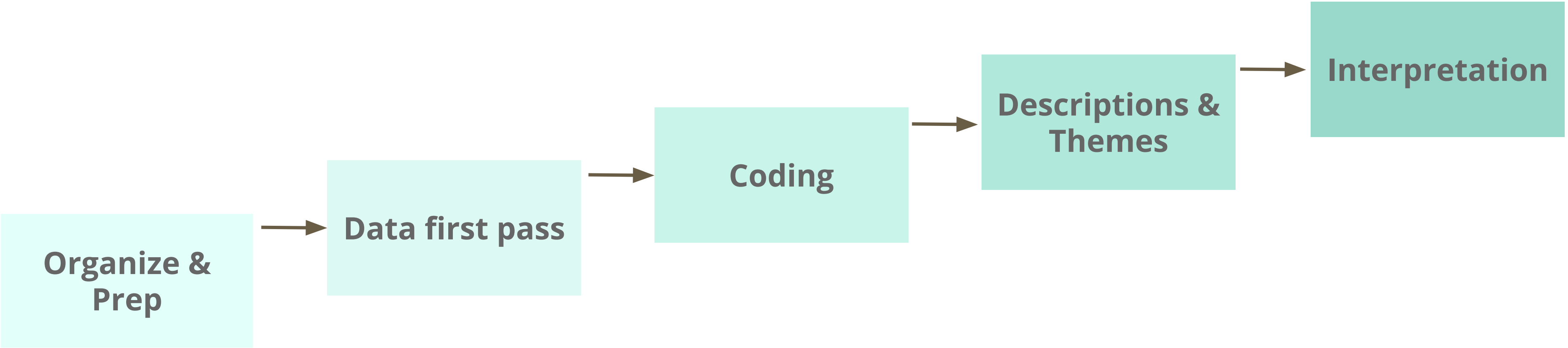
Explanatory Sequential

Explain why we are seeing quantitatively documented behaviors

- Quant narrows the scope of the qualitative investigation
- Points to specific signals or patterns to explore from the users perspective
- Tends to mix induction and deduction

Creswell, J.W. and Clark, V.L.P. (2017). 'The Nature of Mixed Methods Research.'In Designing and conducting mixed methods research. Sage publications.

Overview of interview analysis process

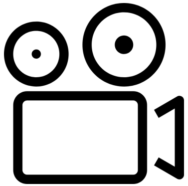


Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Organize and prepare

Transcriptions

Videos



Images

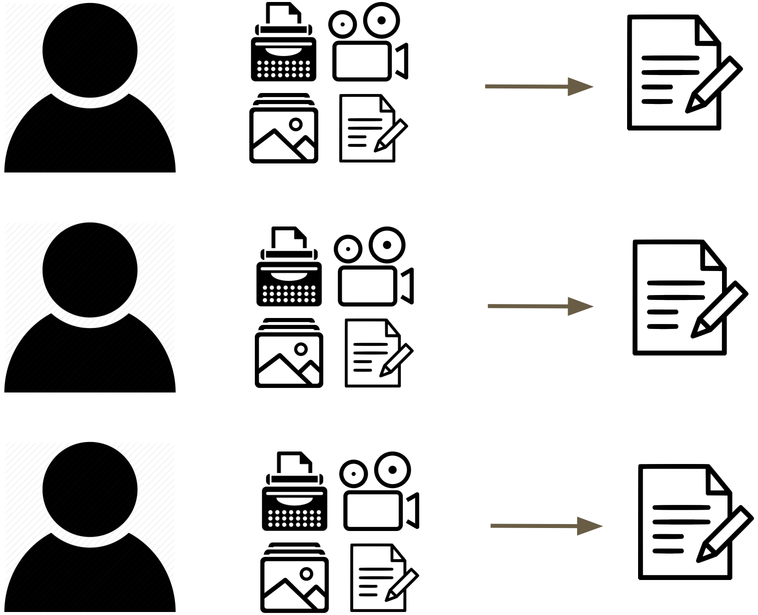
Notes

Gather all assets and tag by participant + participant metadata, date, time.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

First pass through data

Go through all assets by each participant individually.



Take notes on each to document high level ideas capturing the individual's experience as they relate to one another.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Code the data

"Coding means naming segments of data with a label that simultaneously categorizes, summarizes, and accounts for each piece of data."

-Charmaz, 2006

Open coding
Implement coding that allows for a flexible structure; codes can be added and re-worked as more information comes through.

→

Memos
Systematically record ongoing ideas about meaning of codes, their relationships to one another, open questions, etc.

→

Code review
Review coding system for consistency in code usage and identifying/resolving redundancies.

Charmaz, K. (2006). Constructing grounded theory. Thousand Oaks, CA: Sage.; Strauss, A., & Corbin, J. (1998) Basics of qualitative research: Grounded theory procedures and techniques (2nd ed.) Thousand Oaks, CA, Sage

Code the data

"Coding means naming segments of data with a label that simultaneously categorizes, summarizes, and accounts for each piece of data."

-Charmaz, 2006

Open coding
Implement coding that allows for a flexible structure; codes can be added and re-worked as more information comes through.

→

Memos
Systematically record ongoing ideas about meaning of codes, their relationships to one another, open questions, etc.

→

Code review
Review coding system for consistency in code usage and identifying/resolving redundancies.

Charmaz, K. (2006). Constructing grounded theory. Thousand Oaks, CA: Sage.; Strauss, A., & Corbin, J. (1998) Basics of qualitative research: Grounded theory procedures and techniques (2nd ed.) Thousand Oaks, CA, Sage

Code the data

"Coding means naming segments of data with a label that simultaneously categorizes, summarizes, and accounts for each piece of data."

-Charmaz, 2006

Open coding
Implement coding that allows for a flexible structure; codes can be added and re-worked as more information comes through.

→

Memos
Systematically record ongoing ideas about meaning of codes, their relationships to one another, open questions, etc.

→

Code review
Review coding system for consistency in code usage and identifying/resolving redundancies.

Charmaz, K. (2006). Constructing grounded theory. Thousand Oaks, CA: Sage.; Strauss, A., & Corbin, J. (1998) Basics of qualitative research: Grounded theory procedures and techniques (2nd ed.) Thousand Oaks, CA, Sage

Code the data

"Coding means naming segments of data with a label that simultaneously categorizes, summarizes, and accounts for each piece of data."

-Charmaz, 2006

Open coding		Memos		Code review
Implement coding that allows for a flexible structure; codes can be added and re-worked as more information comes through.	→	Systematically record ongoing ideas about meaning of codes, their relationships to one another, open questions, etc.	→	Review coding system for consistency in code usage and identifying/resolving redundancies.

Charmaz, K. (2006). Constructing grounded theory. Thousand Oaks, CA: Sage.; Strauss, A., & Corbin, J. (1998) Basics of qualitative research: Grounded theory procedures and techniques (2nd ed.) Thousand Oaks, CA, Sage

Coding types (meta codes)

Expected codes

Codes you would expect to see based on the interview questions, existing knowledge of the domain, and logic.

Surprising codes

Codes that you did not expect to emerge from the data.

Codes of conceptual interest

Codes that seems particularly interesting or relevant to developing overarching theory.

Tesch, R. (1990). Qualitative research: Analysis types and software tools. New York: Falmer.

Coding types (meta codes)

Expected codes

Codes you would expect to see based on the interview questions, existing knowledge of the domain, and logic.

Surprising codes

Codes that you did not expect to emerge from the data.

Codes of conceptual interest

Codes that seems particularly interesting or relevant to developing overarching theory.

Tesch, R. (1990). Qualitative research: Analysis types and software tools. New York: Falmer.

Coding types (meta codes)

Expected codes

Codes you would expect to see based on the interview questions, existing knowledge of the domain, and logic.

Surprising codes

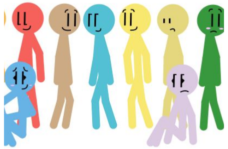
Codes that you did not expect to emerge from the data.

Codes of conceptual interest

Codes that seems particularly interesting or relevant to developing overarching theory.

Tesch, R. (1990). Qualitative research: Analysis types and software tools. New York: Falmer.

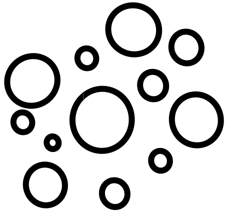
Generate descriptions and themes



Descriptions

Capture the context and experiences of each individual.

Summarize at the individual level.



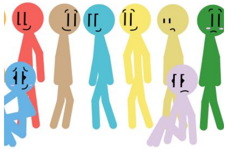
Themes

Cluster and tie together codes and categories that emerge within and across participant.

Should display multiple perspectives and be well supported by diverse set of quotes.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

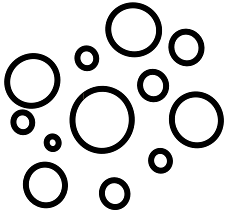
Generate descriptions and themes



Descriptions

Capture the context and experiences of each individual.

Summarize at the individual level.



Themes

Cluster and tie together codes and categories that emerge within and across participant.

Should display multiple perspectives and be well supported by diverse set of quotes.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Interpretation

What lessons did you learn?

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.; Lincoln, Y. S., & Guba, E.G. (1985) Naturalistic inquiry. Beverly Hills, CA: Sage

Interpretation

What lessons did you learn?

Theory

Themes and how they relate to one another.

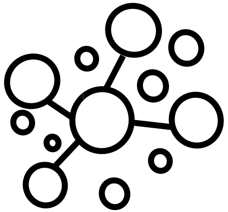
Hypotheses

Pieces of the theory to explore further.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.; Lincoln, Y. S., & Guba, E.G. (1985) Naturalistic inquiry. Beverly Hills, CA: Sage

Interpretation

What lessons did you learn?



Theory

Themes and how they relate to one another.

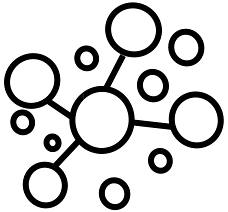
Hypotheses

Pieces of the theory to explore further.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.; Lincoln, Y. S., & Guba, E.G. (1985) Naturalistic inquiry. Beverly Hills, CA: Sage

Interpretation

What lessons did you learn?



Theory

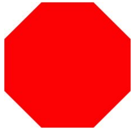
Themes and how they relate to one another.

Hypotheses

Pieces of the theory to explore further.

Contextualize

Explain how this fits or is at odds with existing knowledge.



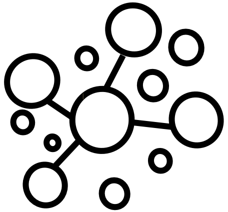
Acknowledge limits

Describe potential biases and shortcomings.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.; Lincoln, Y. S., & Guba, E.G. (1985) Naturalistic inquiry. Beverly Hills, CA: Sage

Interpretation

What lessons did you learn?



Theory

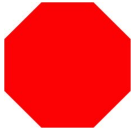
Themes and how they relate to one another.

Hypotheses

Pieces of the theory to explore further.

Contextualize

Explain how this fits or is at odds with existing knowledge.



Acknowledge limits

Describe potential biases and shortcomings.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.; Lincoln, Y. S., & Guba, E.G. (1985) Naturalistic inquiry. Beverly Hills, CA: Sage

Interpreting for satisfaction metrics

Generate hypotheses around:

Interpreting for satisfaction metrics

Generate hypotheses around:

Definitions of good and bad experiences

If individuals differ in definition, how can we understand those differences?

What behavioral signals mean

What is a positive, negative, neutral or ambiguous signal?

What can help us interpret ambiguous signals?

How to operationalize those signals

How can we best represent signals in the data?

Aggregate? Temporally? Normalized? Average? Min? Max?

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Interpreting for satisfaction metrics

Generate hypotheses around:

Definitions of good and bad experiences

If individuals differ in definition, how can we understand those differences?

What behavioral signals mean

What is a positive, negative, neutral or ambiguous signal?

What can help us interpret ambiguous signals?

How to operationalize those signals

How can we best represent signals in the data?

Aggregate? Temporally? Normalized? Average? Min? Max?

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Interpreting for satisfaction metrics

Generate hypotheses around:

Definitions of good and bad experiences

If individuals differ in definition, how can we understand those differences?

What behavioral signals mean

What is a positive, negative, neutral or ambiguous signal?

What can help us interpret ambiguous signals?

How to operationalize those signals

How can we best represent signals in the data?

Aggregate? Temporally? Normalized? Average? Min? Max?

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Validity and reliability

Validity

+ Triangulate with different data sources to gain corroboration and confidence in findings
+ Generate rich descriptions to contextualize findings
+ Present discrepant information

Reliability

+ Check for accurate representation of raw data (e.g, transcriptions accurate)
+ Confirm code consistency
+ Use independent coders to check for interrater reliability

Gibbs, G.R. (2007). Analyzing qualitative data. In U. Flick (Ed.), The Sage qualitative research kit. Thousand Oaks, CA: Sage.

Validity and reliability

Validity

+ Triangulate with different data sources to gain corroboration and confidence in findings
+ Generate rich descriptions to contextualize findings
+ Present discrepant information

Reliability

+ Check for accurate representation of raw data (e.g, transcriptions accurate)
+ Confirm code consistency
+ Use independent coders to check for interrater reliability

Gibbs, G.R. (2007). Analyzing qualitative data. In U. Flick (Ed.), The Sage qualitative research kit. Thousand Oaks, CA: Sage.

Example: Discover Weekly coding

Codes

Find a new
favorite song

Songs fit
together

No emotional
reactions

Background
while working

Good for
running

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Example: Discover Weekly coding

Codes

Find a new favorite song

Songs fit together

No emotional reactions

Background while working

Good for running

Categories

Good experience

Bad experience

Background listening

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Example: Discover Weekly coding

Codes

Find a new favorite song

Songs fit together

No emotional reactions

Background while working

Good for running

Categories

Good experience

Bad experience

Background listening

Themes

Experiences

Goals

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Example: Discover Weekly coding

Codes

Find a new favorite song

Songs fit together

No emotional reactions

Background while working

Good for running

Categories

Good experience

Bad experience

Background listening

Themes

Experiences

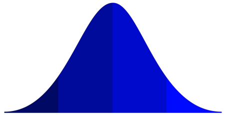
Goals

Hypotheses

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Example: Discover Weekly hypotheses

Behaviors provide clearer signals within the context of users' goals.



Metrics should be normalized relative to each user's typical behavior.

Just one great recommendation can have a large positive effect on satisfaction.

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Example: Discover Weekly hypotheses

There were four overarching goals, and the definition of satisfaction looked different for each.

Play new background music

- No skipping
- Saves or adds
- Listening time
- Sessions per week

Listen to new music now and later

- Saves or adds
- % tracks heard by EOW
- Streams over half the song
- Downstream listening

Find new music for later

- Saves or adds
- Streams
- Downstream listening

Engage with new music

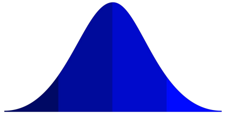
- Artist page views
- Album page views
- Downstream listening

unambiguously positive signal

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Example: Discover Weekly hypotheses

Behaviors provide clearer signals within the context of users' goals.

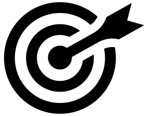


Metrics should be normalized relative to each user's typical behavior.

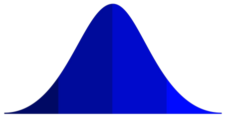
Just one great recommendation can have a large positive effect on satisfaction.

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Example: Discover Weekly hypotheses



Behaviors provide clearer signals within the context of users' goals.

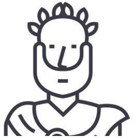


Metrics should be normalized relative to each user's typical behavior.

Just one great recommendation can have a large positive effect on satisfaction.

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

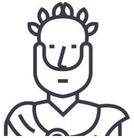
When to use deductive analysis?



DEDUCTIVE

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

When to use deductive analysis?



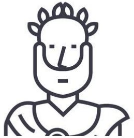
DEDUCTIVE

Literature review

There is existing literature that you believe generalizes to your area of interest, and can analyze qualitative data through the lens of existing theory.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

When to use deductive analysis?



DEDUCTIVE

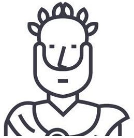
Literature review

There is existing literature that you believe generalizes to your area of interest, and can analyze qualitative data through the lens of existing theory.

You/your colleagues have been studying this already…

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

When to use deductive analysis?

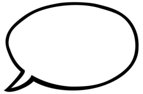


DEDUCTIVE

Literature review

There is existing literature that you believe generalizes to your area of interest, and can analyze qualitative data through the lens of existing theory.

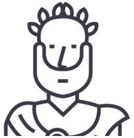
You/your colleagues have been studying this already…



Qualitatively:
you have an existing codebook to work from

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

When to use deductive analysis?

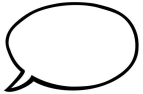


DEDUCTIVE

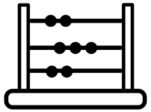
Literature review

There is existing literature that you believe generalizes to your area of interest, and can analyze qualitative data through the lens of existing theory.

You/your colleagues have been studying this already...



Qualitatively: you have an existing codebook to work from



Quantitatively: you can follow an explanatory sequential design

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Qualitative studies in explanatory sequential design

Previous quantitative analysis uncovered what is happening but we don't know why it's happening.

Creswell, J. W., & Clark, V. L. P. (2017). Designing and conducting mixed methods research. Sage publications.

Qualitative studies in explanatory sequential design

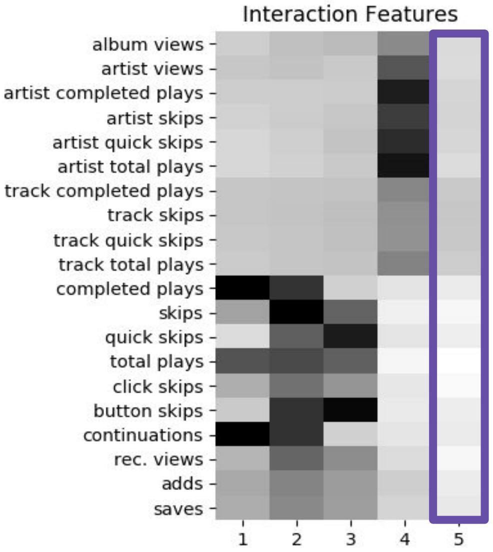
Previous quantitative analysis uncovered what is happening but we don't know why it's happening.

- What participants to recruit
- What questions to ask
- What signals to look at
- What stimuli to present

Creswell, J. W., & Clark, V. L. P. (2017). Designing and conducting mixed methods research. Sage publications.

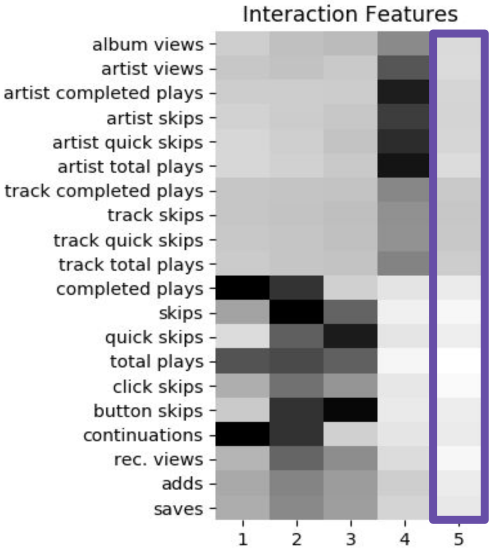
Example: Discover Weekly Neverminds

We clustered Discover Weekly behaviors to identify patterns.



Example: Discover Weekly Neverminds

We clustered Discover Weekly behaviors to identify patterns.

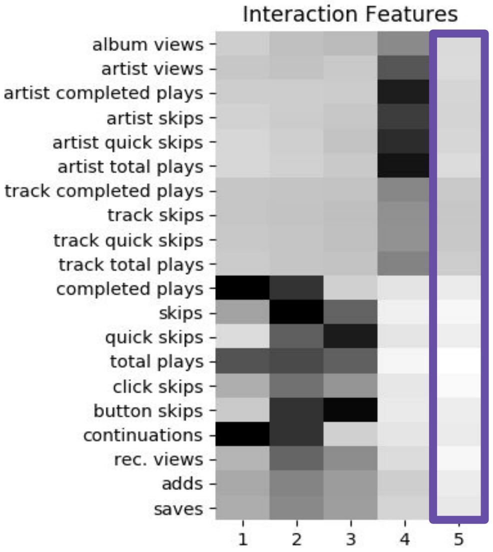


4 clusters mapped onto goals from interviews.

But the largest cluster did not! Neverminds!

Example: Discover Weekly Neverminds

We clustered Discover Weekly behaviors to identify patterns.



4 clusters mapped onto goals from interviews.

But the largest cluster did not! Neverminds!

- Recruit these users
- Discover Weekly habits
- Discover Weekly evaluation
- Probe around lack of key signals
- Show example nevermind behavior and walk through

Quantitative Data Analysis

Quantitative Analysis

- **Quantitative Analysis role in Mixed Methods Research**
- Exploratory Data Analysis
  - Normalization and maintaining interpretability
  - Identifying structure of user behavior through visualization
  - Selecting key behaviors
- Integration
  - Semantic vs Observed behaviors
- Satisfaction Survey
- Modeling and Validation

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- **Exploratory Data Analysis**
  - **Normalization and maintaining interpretability**
  - **Identifying structure of user behavior through visualization**
  - **Sizing impact of user behavior through modeling**
- Integration
  - Semantic vs Observed behaviors
- Satisfaction Survey
- Modeling and Validation

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- Exploratory Data Analysis
  - Normalization and maintaining interpretability
  - Identifying structure of user behavior through visualization
  - Sizing impact of user behavior through modeling
- **Integration**
  - **Semantic vs Observed behaviors**
- Satisfaction Survey
- Modeling and Validation

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- Exploratory Data Analysis
    - Normalization and maintaining interpretability
    - Identifying structure of user behavior through visualization
    - Sizing impact of user behavior through modeling
- Integration
    - Semantic vs Observed behaviors
- **Satisfaction Survey**
- Modeling and Validation

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- Exploratory Data Analysis
  - Normalization and maintaining interpretability
  - Identifying structure of user behavior through visualization
  - Sizing impact of user behavior through modeling
- Integration
  - Semantic vs Observed behaviors
- Satisfaction Survey
- **Modeling and Validation**

How Does Quantitative Analysis Fit In...

Exploratory Sequential

Presence of Qualitative behaviors

- Correct logging
- Match observed patterns to described patterns
- Detect described goals

Differences from Qualitative behaviors

- Missing Qualitative patterns
- Significant patterns not described by in Qualitative study
- Differences by segment or goal

Explanatory Sequential

Discover behaviors we can describe

- Correct logging
- Give qualitative descriptions of observed patterns

What will be hard to study qualitatively

- Patterns exclusive to low engagement
- Difficult to reach clusters or groups
- Stay aware of assumptions

Creswell, J.W. and Clark, V.L.P. (2017). 'The Nature of Mixed Methods Research.' In *Designing and conducting mixed methods research*. Sage publications.

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- **Exploratory Data Analysis**
  - **Normalization and maintaining interpretability**
  - Identifying structure of user behavior through visualization
  - Sizing impact of user behavior through modeling
- Integration
  - Semantic vs Observed behaviors
- Satisfaction Survey
- Modeling and Validation

Importance of Interpretability

- Need to compare qualitative and quantitative results, so need to retain interpretability of metrics

Importance of Interpretability

- Need to compare qualitative and quantitative results, so need to retain interpretability of metrics

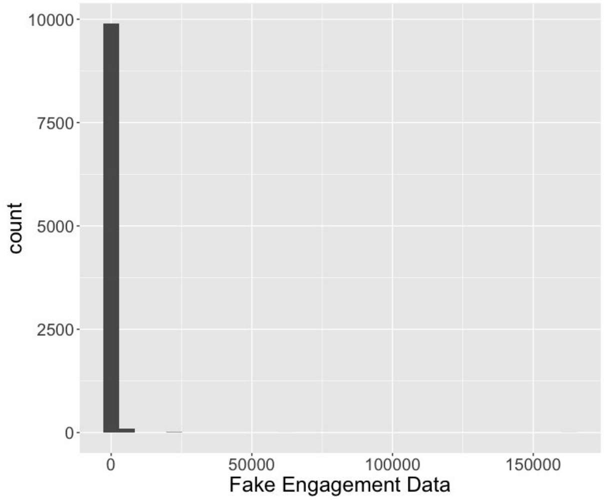
- … but ALSO need to make data transformations so that we can do data analysis

Importance of Interpretability

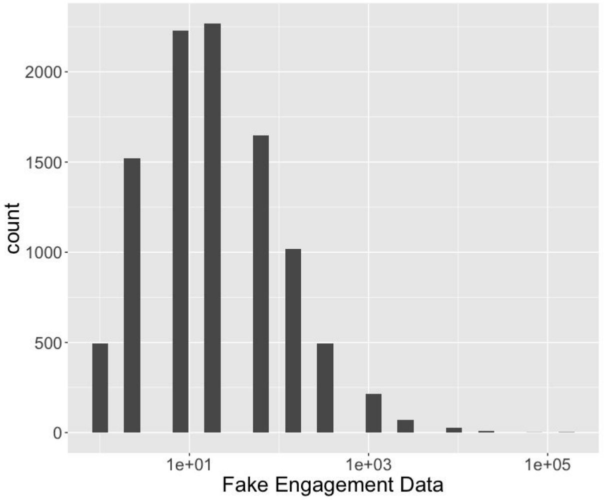
- Need to compare qualitative and quantitative results, so need to retain interpretability of metrics
- ... but ALSO need to make data transformations so that we can do data analysis
- Tension between feature engineering, data processing, and interpretability

Normalization and Standardization

Most engagement will be heavily positively skewed



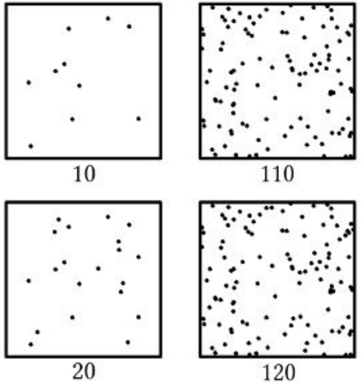
Log
Transform



Fechner's Law Applied to Numerical Cognition

"Subjective sensation is proportional to the Log of the stimulus intensity"

Applies to how we interpret data and how users generate data

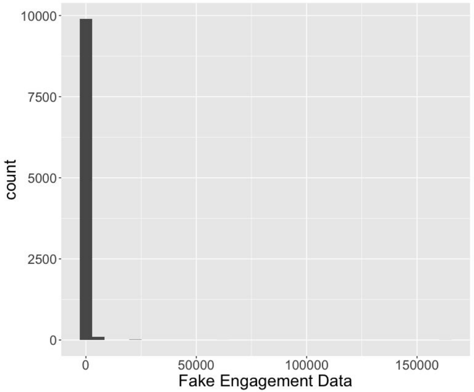


Fechner, Gustav Theodor (1966) [First published .1860]. Howes, D H; Boring, E G, eds. *Elements of psychophysics* [*Elemente der Psychophysik*]. volume 1. Translated by Adler, H E. United States of America: Holt, Rinehart and Winston.

Moyer R.S., Landauer T.K. (September 1967). "Time required for judgements of numerical inequality". *Nature*. **215** (5109): 1519–20. doi:10.1038/2151519a0

Dealing with Zero

Good quant and qual reasons to use log-normalization, but most engagement data is zero inflated

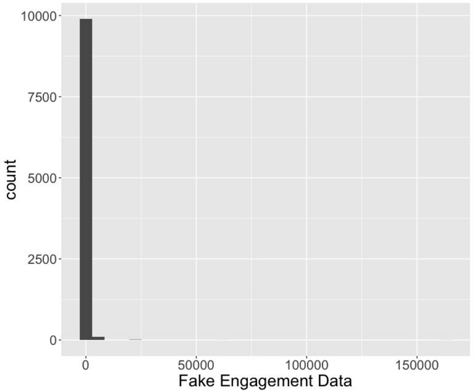


Log(0) = 

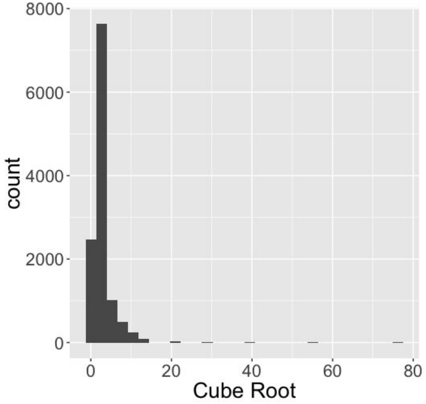
MacKinnon, James G & Magee, Lonnie, 1990. "Transforming the Dependent Variable in Regression Models,"International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 31(2), pages 315-339, May.

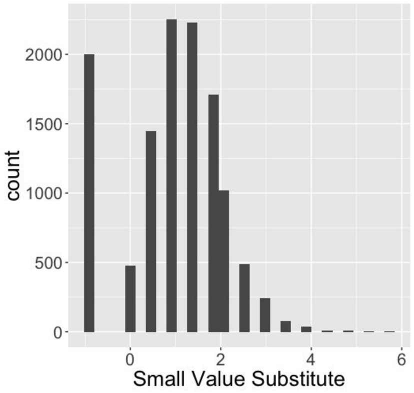
Dealing with Zero

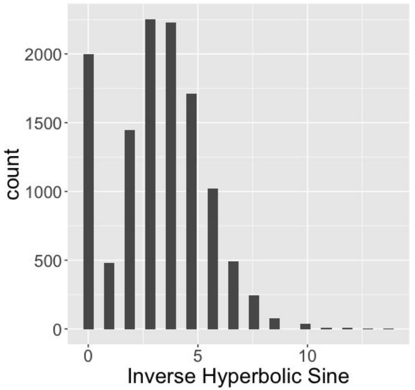
Good quant and qual reasons to use log-normalization, but most engagement data is zero inflated



Log(0) = 😡



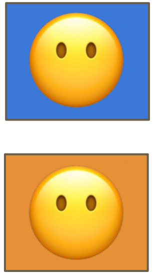


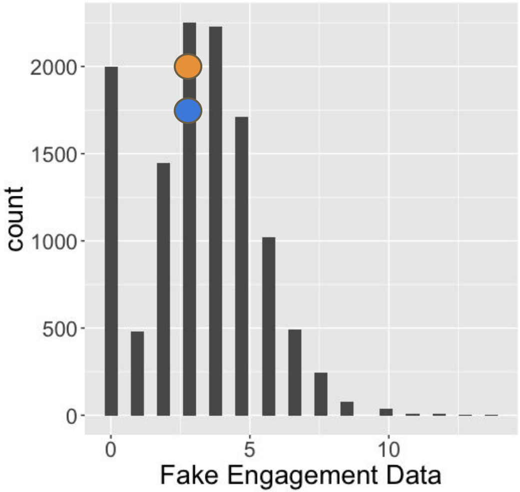


MacKinnon, James G & Magee, Lonnie, 1990. "Transforming the Dependent Variable in Regression Models,"International Economic Review, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, vol. 31(2), pages 315-339, May.

Normalizing Data For Habits

Users tend to compare current experiences in context of past experiences



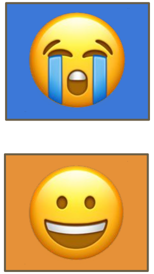


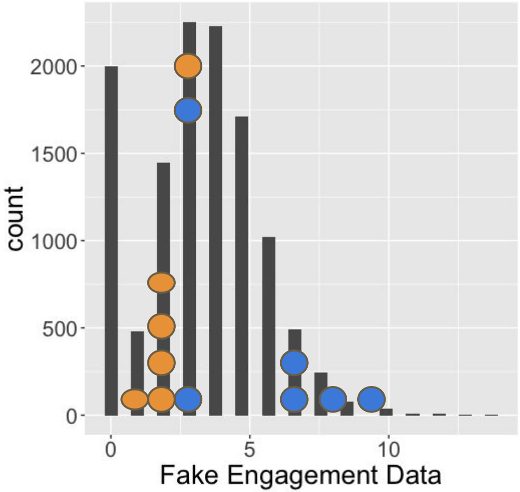
Kahneman, Daniel & Tversky, Amos, 1979. "Prospect Theory: An Analysis of Decision under Risk,"Econometrica, Econometric Society, vol. 47(2), pages 263-291, March.

Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Normalizing Data For Habits

Users tend to compare current experiences in context of past experiences





Kahneman, Daniel & Tversky, Amos, 1979. "Prospect Theory: An Analysis of Decision under Risk,"Econometrica, Econometric Society, vol. 47(2), pages 263-291, March.

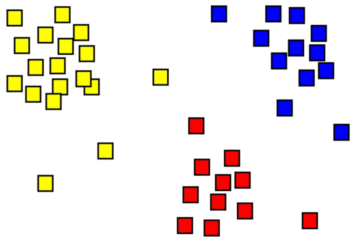
Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- **Exploratory Data Analysis**
  - Normalization and maintaining interpretability
  - **Identifying structure of user behavior through visualization**
  - Sizing impact of user behavior through modeling
- Integration
  - Semantic vs Observed behaviors
- Satisfaction Survey
- Modeling and Validation

Goals and Segments

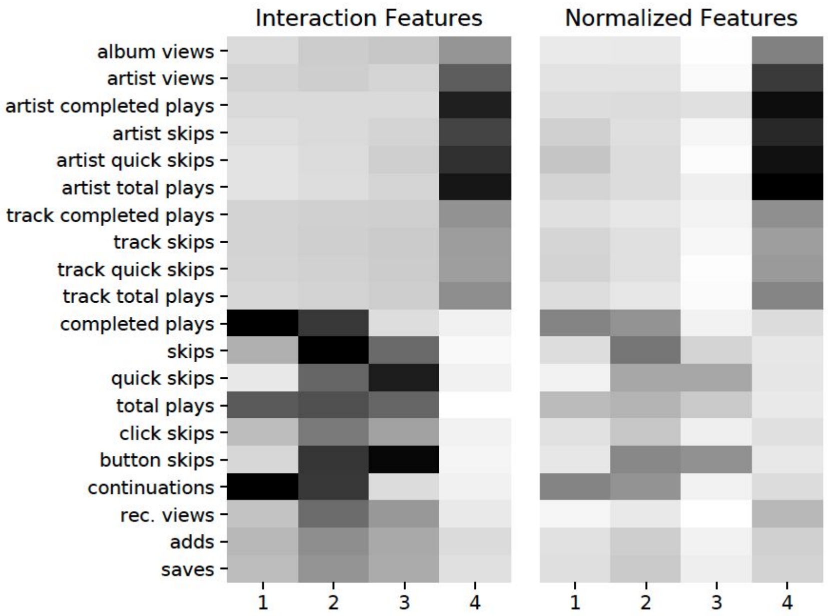
- Find clusters with enough support in the data to be qualitatively relevant
- Perform clustering on meaningful and complete sets of variables
- “Clustering is only difficult when it does not matter”
    - Results should be robust to perturbations in data



Daniely, Amit; Linial, Nati; Saks, Michael. 05/2012. Clustering is difficult only when it does not matter. arXiv:1205.4891v1

Example: Discover Weekly User Goals

- Example: Music Recommendation, [Garcia-Gathright et al, 2018]
  - Clustering on engagement behaviors corresponded to different stated user goals
  - Clusters consistent across days of week
  - Clusters consistent over small changes to the underlying data
- Identifying and understanding goals gives context for user satisfaction



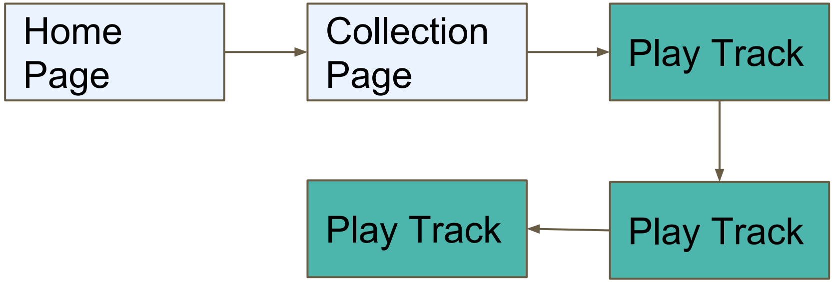
Jean Garcia-Gathright, et al. 2018. Understanding and Evaluating User Satisfaction with Music Discovery. In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18). ACM, New York, NY, USA, 55-64. DOI: https://doi.org/10.1145/3209978.3210049

Correlation of Engagement Actions

- Most platforms require multiple actions to achieve any goal, meaning platform engagement will always look correlated

Semantic objective:
Listen to music

Observed Behavior

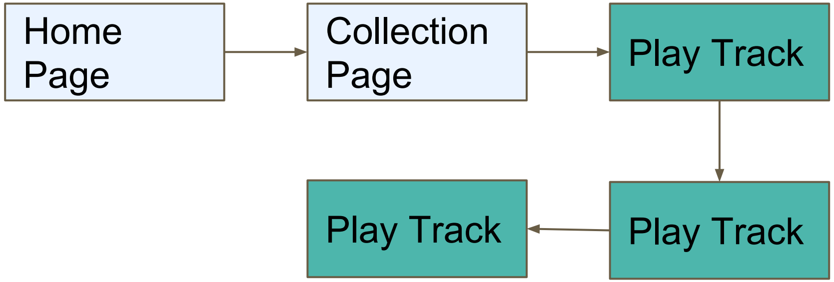


Correlation of Engagement Actions

- Most platforms require multiple actions to achieve any goal, meaning platform engagement will always look correlated
- But actions may not be unique to just one semantic objective

Semantic objective:
Listen to music

Observed Behavior

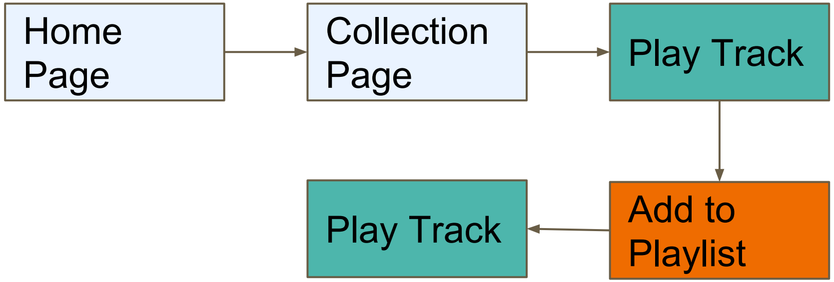


Correlation of Engagement Actions

- Most platforms require multiple actions to achieve any goal, meaning platform engagement will always look correlated
- But actions may not be unique to just one semantic objective

Semantic objective:
Make a playlist

Observed Behavior

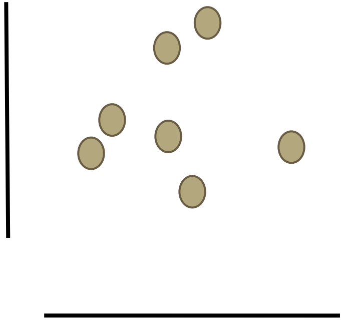


Correlation of Engagement Actions

- Most platforms require multiple actions to achieve any goal, meaning platform engagement will always look correlated
- But actions may not be unique to just one semantic objective

Semantic objectives

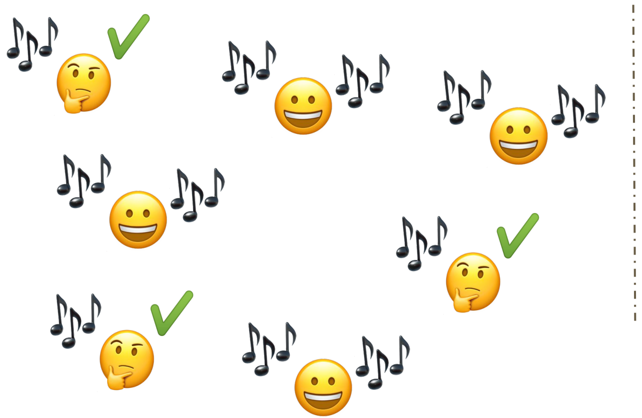
Observed Behavior



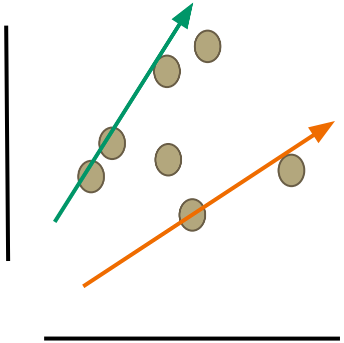
Correlation of Engagement Actions

- Most platforms require multiple actions to achieve any goal, meaning platform engagement will always look correlated
- But actions may not be unique to just one semantic objective

Semantic objectives



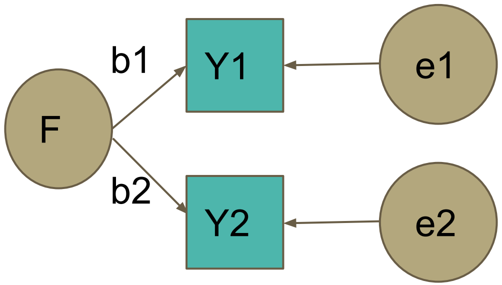
Observed Behavior



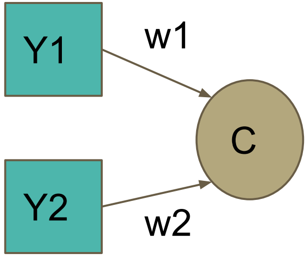
Decompositions of the data set help identify types of actions

Exploratory Factor Analysis or Principal Component Analysis?

Exploratory Factor Analysis



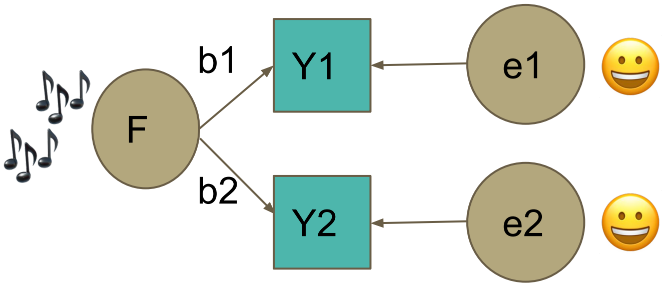
Principal Component Analysis

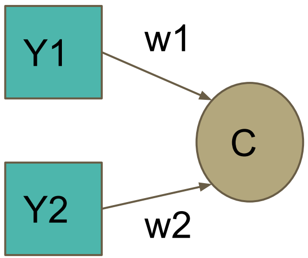


Brian Feldstein. “Principal Component Analysis vs Exploratory Factor Analysis”. February, 2016.
http://activisiongamescience.github.io/2016/02/09/Principal-Component-Analysis-vs-Exploratory-Factor-Analysis/

Exploratory Factor Analysis or Principal Component Analysis?

EFA structure matches our assumptions ...

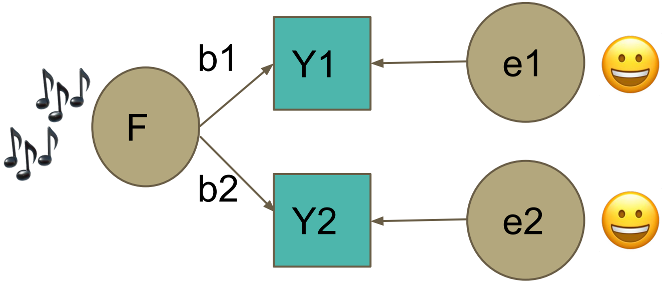


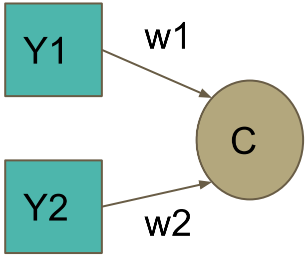


Brian Feldstein. “Principal Component Analysis vs Exploratory Factor Analysis”. February, 2016.
http://activisiongamescience.github.io/2016/02/09/Principal-Component-Analysis-vs-Exploratory-Factor-Analysis/

Exploratory Factor Analysis or Principal Component Analysis?

EFA structure matches our assumptions … but with lots of data F and C converge as measurement error shrinks

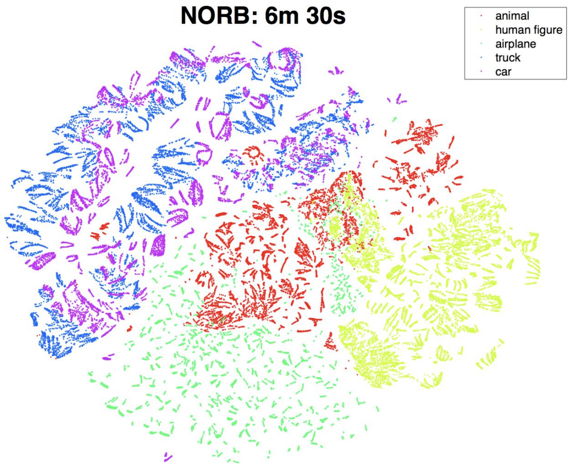


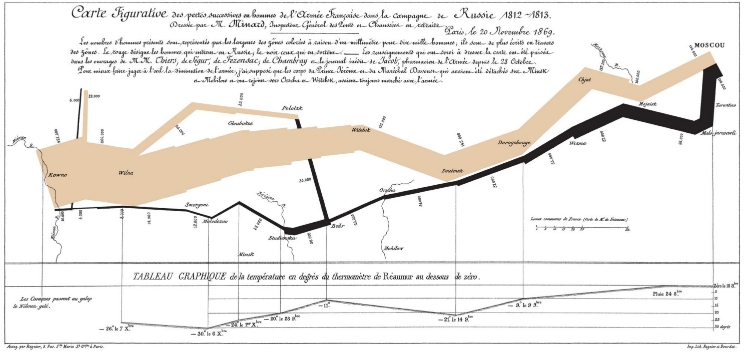


Brian Feldstein. “Principal Component Analysis vs Exploratory Factor Analysis”. February, 2016.
http://activisiongamescience.github.io/2016/02/09/Principal-Component-Analysis-vs-Exploratory-Factor-Analysis/

Visualizing Differences Between Groups

- It’s helpful to understand how different segments can change search behavior holistically to improve our ability to successfully run qualitative follow ups
- Helpful plots are
  - Original data
  - PCA/EFA
  - t-SNE
  - Sankey





van der Maaten, L. & Hinton, G. (2008). Visualizing Data using t-SNE . *Journal of Machine Learning Research*, 9, 2579--2605.
Jesse H. Krijthe (2015). Rtsne: T-Distributed Stochastic Neighbor Embedding using a Barnes-Hut Implementation, URL: https://github.com/jkrijthe/Rtsne
L.J.P. van der Maaten. Accelerating t-SNE using Tree-Based Algorithms. Journal of Machine Learning Research 15(Oct):3221-3245, 2014.

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- **Exploratory Data Analysis**
  - Normalization and maintaining interpretability
  - Identifying structure of user behavior through visualization
  - **Sizing impact of user behavior through modeling**
- Integration
  - Semantic vs Observed behaviors
- Satisfaction Survey
- Modeling and Validation

Sizing Impact of Behavior on Differences

- May have demographic data, cluster, segment, or a satisfaction proxy we want to understand the relationship to remaining engagement variables

Segment	PC1 (probably listening)	PC2 (probably saving)	PC3 (probably …)	sin(Streams 1am-2am)
	++	--	+	+
	+	++	-	-
	-	--	-	+
	+	++	+	-

Sizing Impact of Behavior on Differences

- May have demographic data, cluster, segment, or a satisfaction proxy we want to understand the relationship to remaining engagement variables

Segment	PC1 (probably listening)	PC2 (probably saving)	PC3 (probably …)	sin(Streams 1am-2am)
	++	--	+	+
	+	++	-	-
	-	--	-	+
	+	++	+	-

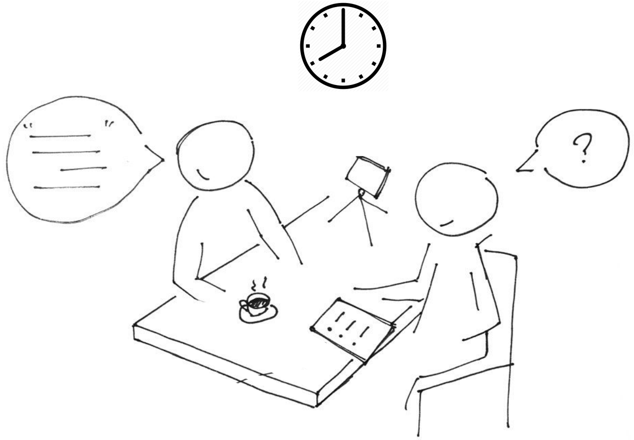
Many, Many features →

Sizing Impact of Behavior on Differences

- May have demographic data, cluster, segment, or a satisfaction proxy we want to understand the relationship to remaining engagement variables
- End goal is to make qualitative comparisons or plan a study

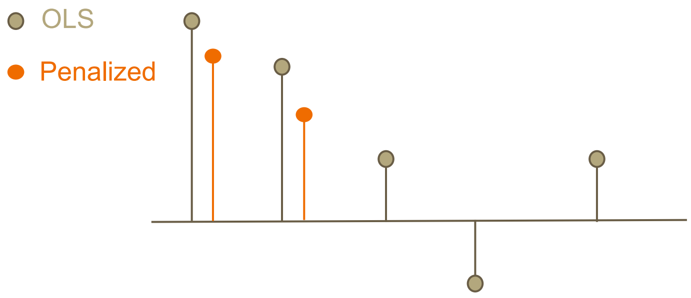
Segment	PC1 (probably listening)	PC2 (probably saving)	PC3 (probably …)	sin(Streams 1am-2am)
	++	--	+	+
	+	++	-	-
	-	--	-	+
	+	++	+	-

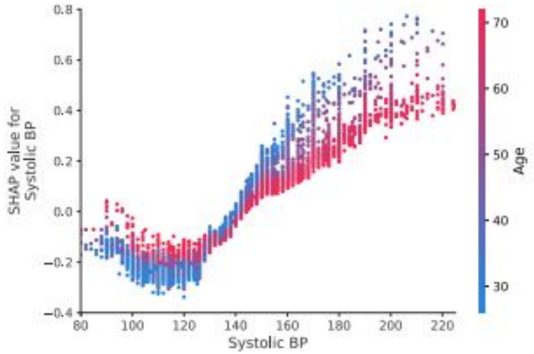
Many, Many features



Sizing Impact of Behavior on Differences

- May have demographic data, cluster, segment, or a satisfaction proxy we want to understand the relationship to remaining engagement variables
- End goal is to make qualitative comparisons or plan a study
- Choose sparse, penalized regressions to help identify key variables
- Understand directionality through coefficients, SHAP values



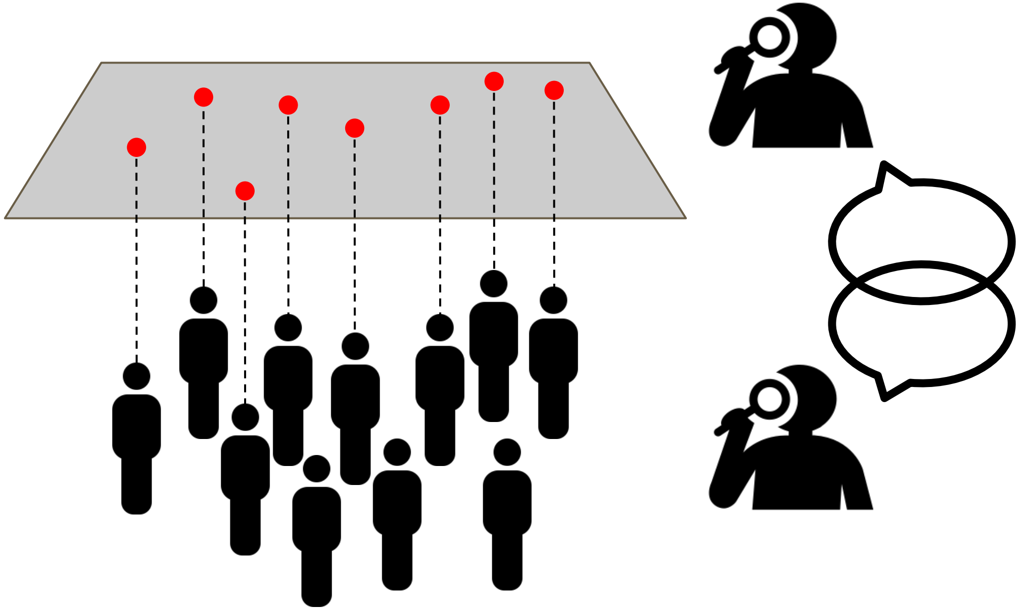


Lundberg, Scott M., Erion, Gabriel G., Lee, Su-In. "Consistent Individualized Feature Attribution for Tree Ensembles." 2017. arXiv:1802.03888v2

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- Exploratory Data Analysis
  - Normalization and maintaining interpretability
  - Identifying structure of user behavior through visualization
  - Sizing impact of user behavior through modeling
- **Integration**
  - **Semantic vs Observed behaviors**
- Satisfaction Survey
- Modeling and Validation

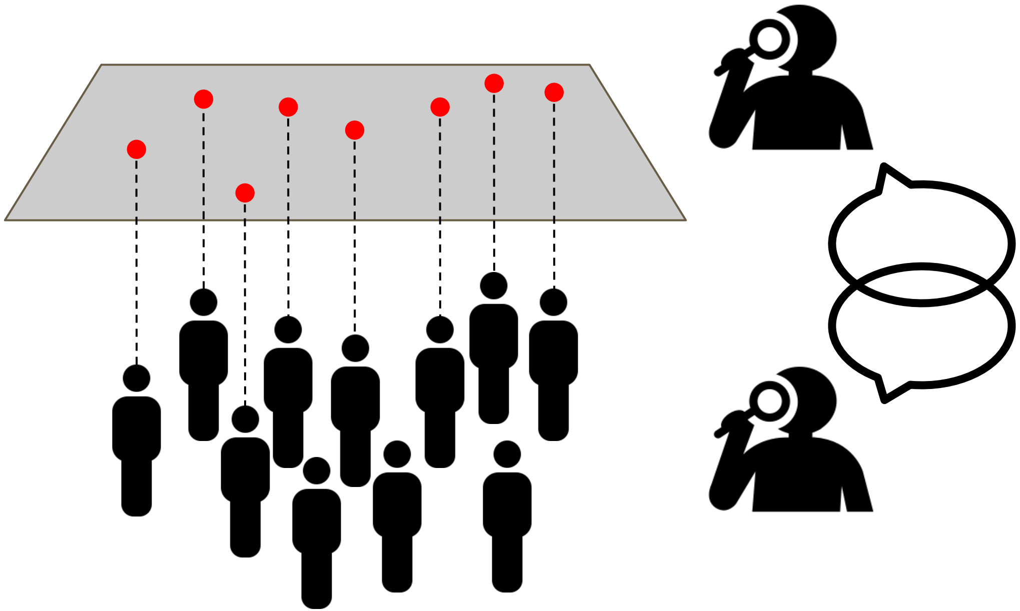
Integration



What patterns did we discover that might be related to satisfaction?

What segments lend themselves to sampling? Do behaviors persist?

Integration



What patterns did we discover that might be related to satisfaction?

What segments lend themselves to sampling? Do behaviors persist?

What assumptions are we making about user mental models?

“Users engage in actions _, when they are in state _”

“Users do not engage in actions _, when they are in state _”

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- Exploratory Data Analysis
  - Normalization and maintaining interpretability
  - Identifying structure of user behavior through visualization
  - Sizing impact of user behavior through modeling
- Integration
  - Semantic vs Observed behaviors
- **Satisfaction Survey**
- Modeling and Validation

Satisfaction Survey

Business Metrics		Behavioral Metrics		
$			++	-
$			--	+
	$		++	+
	$		+	++

Satisfaction	Drivers of satisfaction			
😀	5		1	✔️
😒	1		0	
😀	3		0	✔️
😒	4		1	

Hendrik Müller and Aaron Sedley. 2014. HaTS: large-scale in-product measurement of user attitudes & experiences with happiness tracking surveys. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design (OzCHI '14). ACM, New York, NY, USA, 308-315. DOI=http://dx.doi.org/10.1145/2686612.2686656

Jiepu Jiang, Ahmed Hassan Awadallah, Xiaolin Shi, and Ryen W. White. 2015. Understanding and Predicting Graded Search Satisfaction. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15). ACM, New York, NY, USA, 57-66. DOI: https://doi.org/10.1145/2684822.2685319

Satisfaction Survey

What self reported actions are drivers of satisfaction?

Business Metrics		Behavioral Metrics		
$			++	-
$			--	+
	$		++	+
	$		+	++

Satisfaction	Drivers of satisfaction			
😀	5	🦨	1	✔️
😒	1	🐵	0	
😀	3	🐱	0	✔️
😒	4	🫎	1	

Hendrik Müller and Aaron Sedley. 2014. HaTS: large-scale in-product measurement of user attitudes & experiences with happiness tracking surveys. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design (OzCHI '14). ACM, New York, NY, USA, 308-315. DOI=http://dx.doi.org/10.1145/2686612.2686656

Jiepu Jiang, Ahmed Hassan Awadallah, Xiaolin Shi, and Ryen W. White. 2015. Understanding and Predicting Graded Search Satisfaction. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15). ACM, New York, NY, USA, 57-66. DOI: https://doi.org/10.1145/2684822.2685319

Example: Discover Weekly

Range of goals confirmed earlier hypotheses

Self-reported achievement of goals were most correlated with satisfaction

Found track love was important

Table 2: Pearson correlations with survey responses and satisfaction this week and overall.

Survey question	This week	Overall
Overall Satisfaction and Drivers		
Satisfaction	0.465	1
Ease of use	0.116	0.167
Meets needs	0.550	0.607
Fits taste in general	0.558	0.567
User Goals		
Has goal: add for later	0.016	0.069
Has goal: artist exploration	0.040	0.086
Has goal: background	0.046	0.032
Has goal: genre exploration	0.044	0.061
Has goal: new music right now	0.073	0.086
Has goal: specific activity	0.038	0.034
This Week Satisfaction and Drivers		
Satisfied this week	1	0.465
Achieved this week: add for later	0.608	0.356
Achieved this week: artist exploration	0.569	0.341
Achieved this week: background	0.354	0.187
Achieved this week: genre exploration	0.456	0.296
Achieved this week: new music right now	0.489	0.305
Achieved this week: specific activity	0.571	0.242
At least one song annoyed	-0.206	-0.152
At least one song loved	0.403	0.245
Songs fit music taste	0.632	0.407
Understands why songs are chosen	0.255	0.188

Satisfaction Survey

Are any business or behavioral metrics related to satisfaction that can serve as a proxy?

Business Metrics		Behavioral Metrics		
$			++	-
$			--	+
	$		++	+
	$		+	++

Satisfaction	Drivers of satisfaction			
😀	5		1	✔
😒	1		0	
😀	3		0	✔
😒	4		1	

Hendrik Müller and Aaron Sedley. 2014. HaTS: large-scale in-product measurement of user attitudes & experiences with happiness tracking surveys. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design (OzCHI '14). ACM, New York, NY, USA, 308-315. DOI=http://dx.doi.org/10.1145/2686612.2686656

Jiepu Jiang, Ahmed Hassan Awadallah, Xiaolin Shi, and Ryen W. White. 2015. Understanding and Predicting Graded Search Satisfaction. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15). ACM, New York, NY, USA, 57-66. DOI: https://doi.org/10.1145/2684822.2685319

Satisfaction Survey

In lieu of simple explanations, we can build a model to learn relationships between behavior and satisfaction

Business Metrics		Behavioral Metrics		
$			++	-
$			--	+
	$		++	+
	$		+	++

Satisfaction	Drivers of satisfaction			
😀	5		1	✔
😒	1		0	
😀	3		0	✔
😒	4		1	

Hendrik Müller and Aaron Sedley. 2014. HaTS: large-scale in-product measurement of user attitudes & experiences with happiness tracking surveys. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design (OzCHI '14). ACM, New York, NY, USA, 308-315. DOI=http://dx.doi.org/10.1145/2686612.2686656

Jiepu Jiang, Ahmed Hassan Awadallah, Xiaolin Shi, and Ryen W. White. 2015. Understanding and Predicting Graded Search Satisfaction. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15). ACM, New York, NY, USA, 57-66. DOI: https://doi.org/10.1145/2684822.2685319

Quantitative Analysis

- Quantitative Analysis role in Mixed Methods Research
- Exploratory Data Analysis
    - Normalization and maintaining interpretability
    - Identifying structure of user behavior through visualization
    - Sizing impact of user behavior through modeling
- Integration
    - Semantic vs Observed behaviors
- Satisfaction Survey
- **Modeling and Validation**

Modeling Survey Data

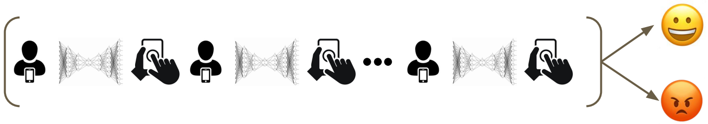
- Functional form of the model is not important
  - Target: Satisfaction or highly correlated proxy
  - Covariates: Data from Logs only
  - Goal: Learn a model that generalizes to the full population
- Survey responses can be limited, so choose meaningful features
  - Helpful to collaborate with qualitative research on feature engineering

Behavioral Metrics		
	++	-
	--	+
	++	+
	+	++

Satisfaction
😀
😡
😀
😡

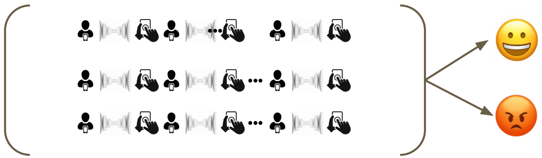
Modeling Other Targets

Task Success



can treat as a sequence labeling task

Retention

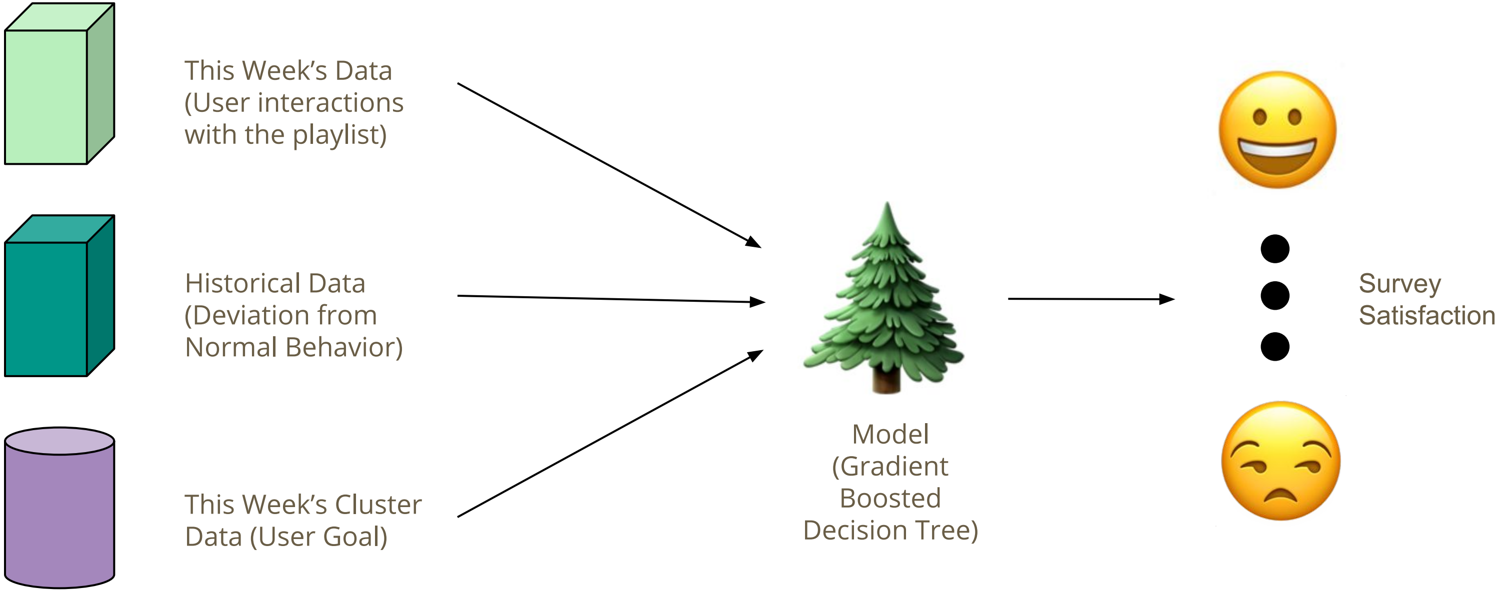


can treat as a classification task

R. Mehrotra, A. Hassan Awadallah, M. Shokouhi, E. Yilmaz, I. Zitouni, A. El Kholy, M. Khabsa. Deep sequential models for task satisfaction prediction.CIKM 2017.
Ryen W. White, Ashish Kapoor, and Susan T. Dumais. Modeling long-term search engine usage. In Paul De Bra, Alfred Kobsa, and David Chin, editors, User modeling, adaptation, and personalization, 28--39, Berlin, Heidelberg, 2010.
Qingyun Wu, Hongning Wang, Liangjie Hong, and Yue Shi. Returning is believing: optimizing long-term user engagement in recommender systems. CIKM 2017.

Example: Discover Weekly

- Features informed by hypotheses from user interviews



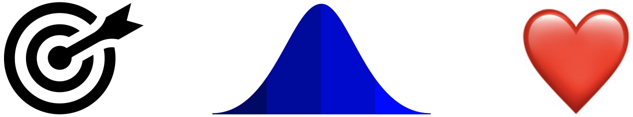
Example: Discover Weekly

Consistency in results between Survey ~ Log Data and Survey ~ Survey make us confident in results

Top predictor is fulfillment of user goals

Normalized features give context for users experiencing success in their goals

Peak experiences (weekly max and sum) inform us about loving a track



	Cluster	Normalized	Max	Sum
Gain (%)	70.6	15.4	5.6	2.3
Weight (%)	54.1	30.7	7.5	1.0

Validating Metrics

- How does the metric behave when applied to the out-of-sample population
- Looking forward to Hypothesis Testing
    - Can algorithms trained with the metric be used to make better recommendations?
    - Does the metric move in the expected direction when applied to A/B test data where one system is known to be better?

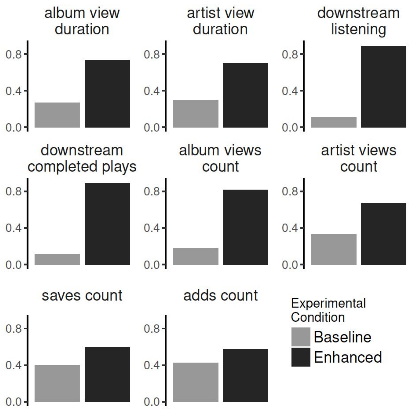


Figure 5: Density for positive interaction signals between baseline and enhanced conditions. Density used to protect user data.

Best practices for mixed methods teams

Levels of collaboration (Rosenfield 1992, Tress 2005)

Loose collaboration between disciplines for exchange of knowledge, producing the sum of each part

Working together to produce new knowledge by sharing different perspectives, but in a disciplinary-specific way

Using a shared conceptual framework that brings together disciplinary-specific theories and approaches

Rosenfield, P.L., 1992. The potential of transdisciplinary research for sustaining and extending linkages between the health and social sciences. Social science & medicine, 35(11), pp.1343-1357.

Tress, B., Tress, G. and Fry, G., 2005. Defining concepts and the process of knowledge production in integrative research. From landscape research to landscape planning: Aspects of integration, education and application, 12, pp.13-26.

Characteristics of successful teams (O'Cathian, 2008)
Not specific to mixed methods

- Successful teams...
  - Are adaptable in their communication styles
  - Develop mutual trust with their team members
  - Are often co-located

O'Cathain, A., Murphy, E. and Nicholl, J., 2008. Multidisciplinary, interdisciplinary, or dysfunctional? Team working in mixed-methods research. Qualitative health research, 18(11), pp.1574-1585.

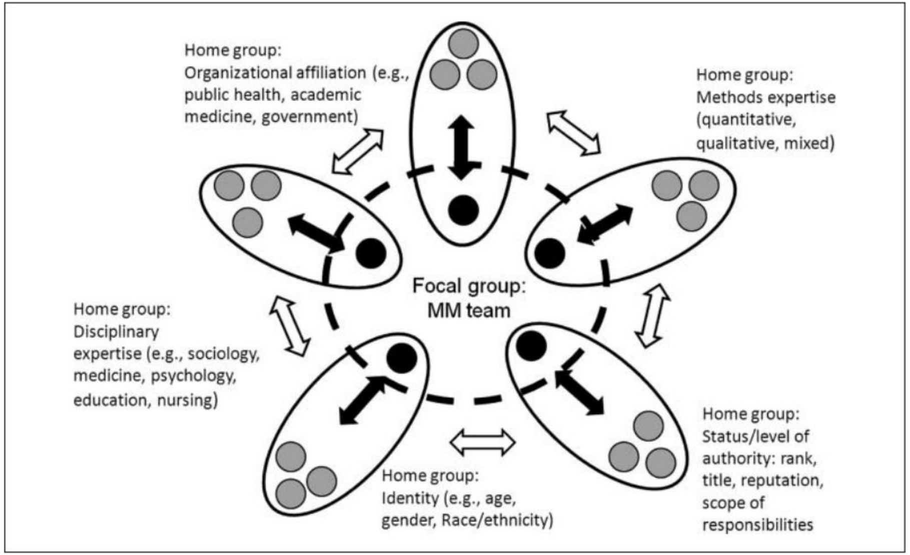
Characteristics of successful teams (O'Cathian, 2008)
Specific to mixed methods

- Successful mixed methods teams...
  - Respect methodological differences
  - Meet frequently to communicate about emerging findings
  - Have a principal investigator who enables each component to make a full contributions

O'Cathain, A., Murphy, E. and Nicholl, J., 2008. Multidisciplinary, interdisciplinary, or dysfunctional? Team working in mixed-methods research. Qualitative health research, 18(11), pp.1574-1585.

Guiding principles for mixed methods teams (Curry, 2012)

- Authors examined group dynamics from their own experiences as mixed methods researchers
- Representational group theory: team members participate as individuals and as representatives of various groups
- Study offers guidelines for managing team dynamics from the group-intergroup view



Curry, L.A., O'Cathain, A., Clark, V.L.P., Aroni, R., Fetters, M. and Berg, D., 2012. The role of group dynamics in mixed methods health sciences research teams. Journal of mixed methods research, 6(1), pp.5-20.

Guiding principles for mixed methods teams (Curry, 2012)

Embrace differences

- Honor group memberships
- Foster and sustain respect

Trust each other

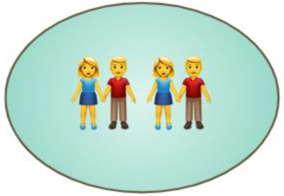
- Make a safe space for discussing both competencies and limitations
- Encourage and support candor

Manage conflict

- Normalize the essential tensions
- Recognize the temptation to avoid conflict
- Establish mechanisms for conflict resolution

Curry, L.A., O'Cathain, A., Clark, V.L.P., Aroni, R., Fetters, M. and Berg, D., 2012. The role of group dynamics in mixed methods health sciences research teams. Journal of mixed methods research, 6(1), pp.5-20.

Guiding principles for mixed methods teams (Curry, 2012)



Create a meaningful group

- Establish a shared commitment to the project's overall goal
- Enable team members to speak freely
- Develop a common language
- Establish time and processes for information exchange

Enact effective leadership

- Value and promote integration
- Balance issues of relationship and task
- Leader doesn't have to be the most "senior" member of the team

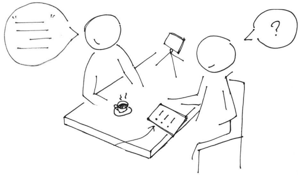
Mixed methods is a way to robustly evaluate user satisfaction

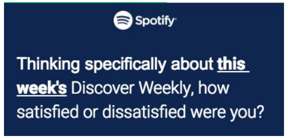
- Qualitative research can help us understand the user perspective on satisfaction

- Combine with quantitative behavioral data at scale to get a more complete understanding

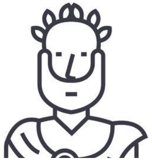
What we have learned so far

Mixed methods study designs

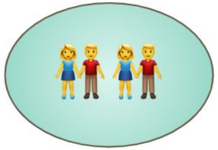




Data collection:
interviews, surveys, logs

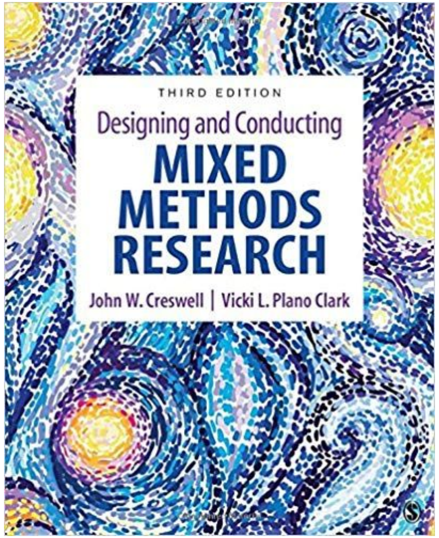


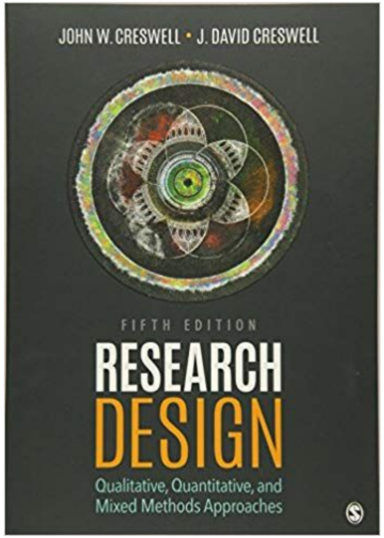
Data analysis:
Inductive, deductive, exploration, satisfaction modeling

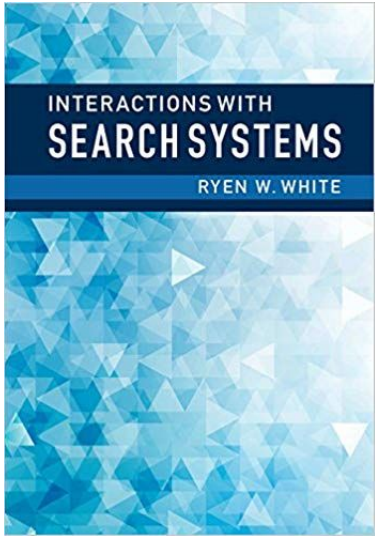


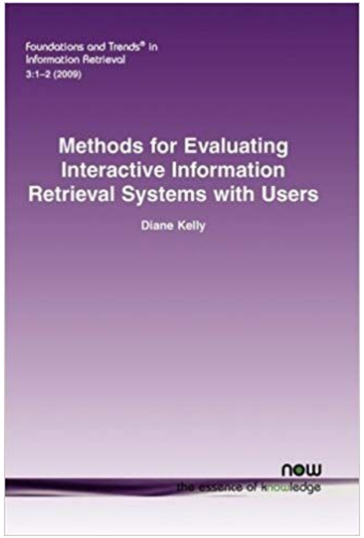
Teamwork

References


THIRD EDITION
Designing and Conducting
MIXED METHODS RESEARCH
John W. Creswell | Vicki L. Plano Clark


JOHN W. CRESWELL • J. DAVID CRESWELL
FIFTH EDITION
RESEARCH DESIGN
Qualitative, Quantitative, and Mixed Methods Approaches


INTERACTIONS WITH
SEARCH SYSTEMS
RYEN W. WHITE


Foundations and Trends® in Information Retrieval
3:1–2 (2009)
Methods for Evaluating Interactive Information Retrieval Systems with Users
Diane Kelly
now
the essence of knowledge

References

Rishabh Mehrotra. Learning from user interactions. Russian Summer School in Information Retrieval, 2018.

Rishabh Mehrotra, Emine Yilmaz, and Ahmed Hassan. Understanding and inferring user tasks and needs. The Web Conference, 2018.

Susan Dumais, Robin Jeffries, Daniel M. Russell, Diane Tang, and Jaime Teevan. Design of Large Scale Log Analysis Studies: A Short Tutorial. Human Computer Interaction Consortium, 2010.

Slides for this tutorial

https://github.com/jeanigarcia/recsys2018-evaluation-tutorial

Break

Coming up at 2pm:

- Hypothesis testing