

GSD&M

TEXAS A&M

CAPSTONE PROJECT

PREDICTIVE SEQUENTIAL ASSOCIATION RULE MINING FOR
TRANSACTIONAL CLICKSTREAM DATA

Texas A&M | March 2017



U.S. AIR FORCE

Flow Satellites



ADJUST YOUR PROFILE

We take our role as the world's most preeminent Air Power very seriously. Explore careers that take on this great responsibility.

SEE WEAPONRY & MATERIEL CAREERS

INTERESTED?

CHAT LIVE

Flow Satellites



Users Flow

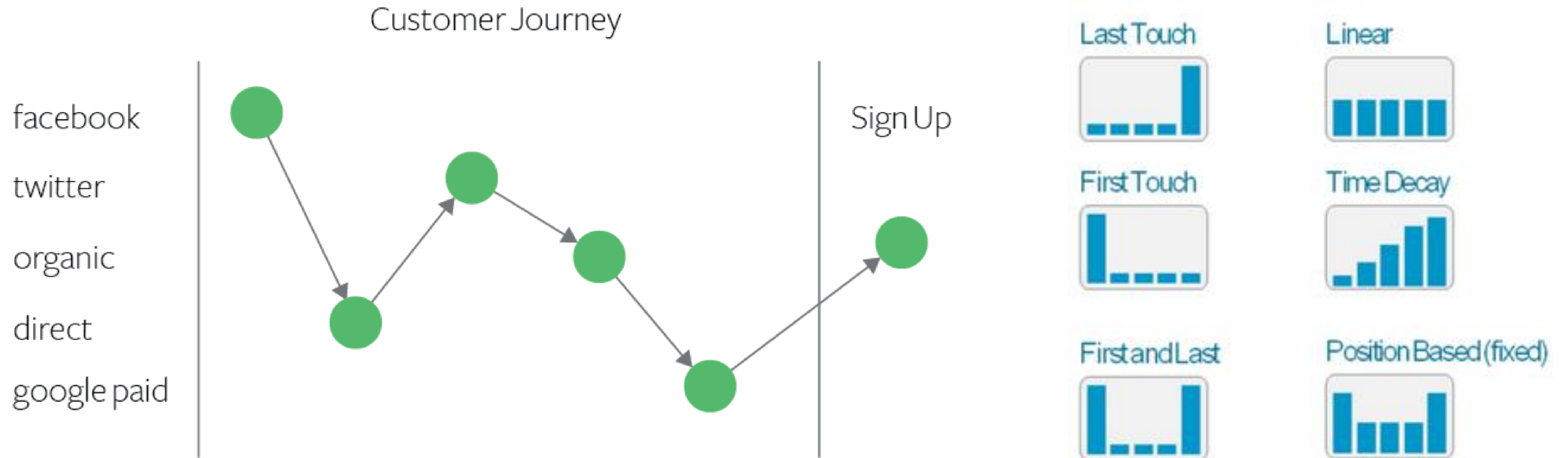
Jan 17, 2017 - Feb 15, 2017

 All Users
100.00% Sessions

This report is based on 97K sessions (71.1% of sessions). [Learn more](#)



MULTI-TOUCH ATTRIBUTION



PRIMARY GOAL:

**OPTIMIZE THE SITE BY DYNAMICALLY SERVING CONTENT THAT
MAXIMIZES THE LIKELIHOOD OF A VISITOR TO APPLY**

SECONDARY GOAL:

**EVOLVE MARKETING MULTI-TOUCH ATTRIBUTION MODELS TO
INCLUDE NON-MEDIA EVENTS**

FRAMEWORK



118,489,114 Obs
50 GB

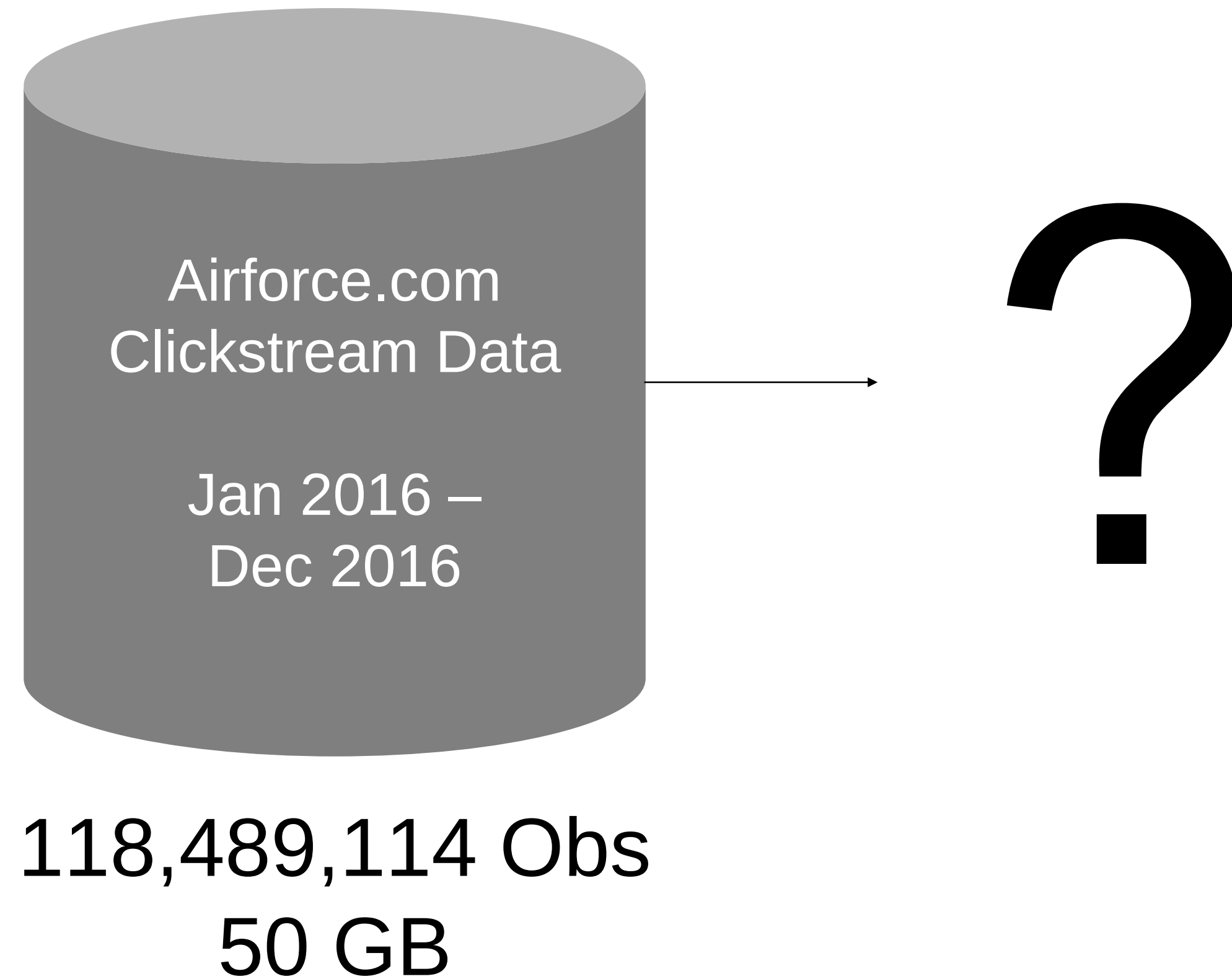
STRUCTURAL CHALLENGES

- TRANSACTIONAL
- 3,841,025 VISITORS
- 52 VARIABLES
- 216 PAGES
- 8 MARKETING CHANNELS

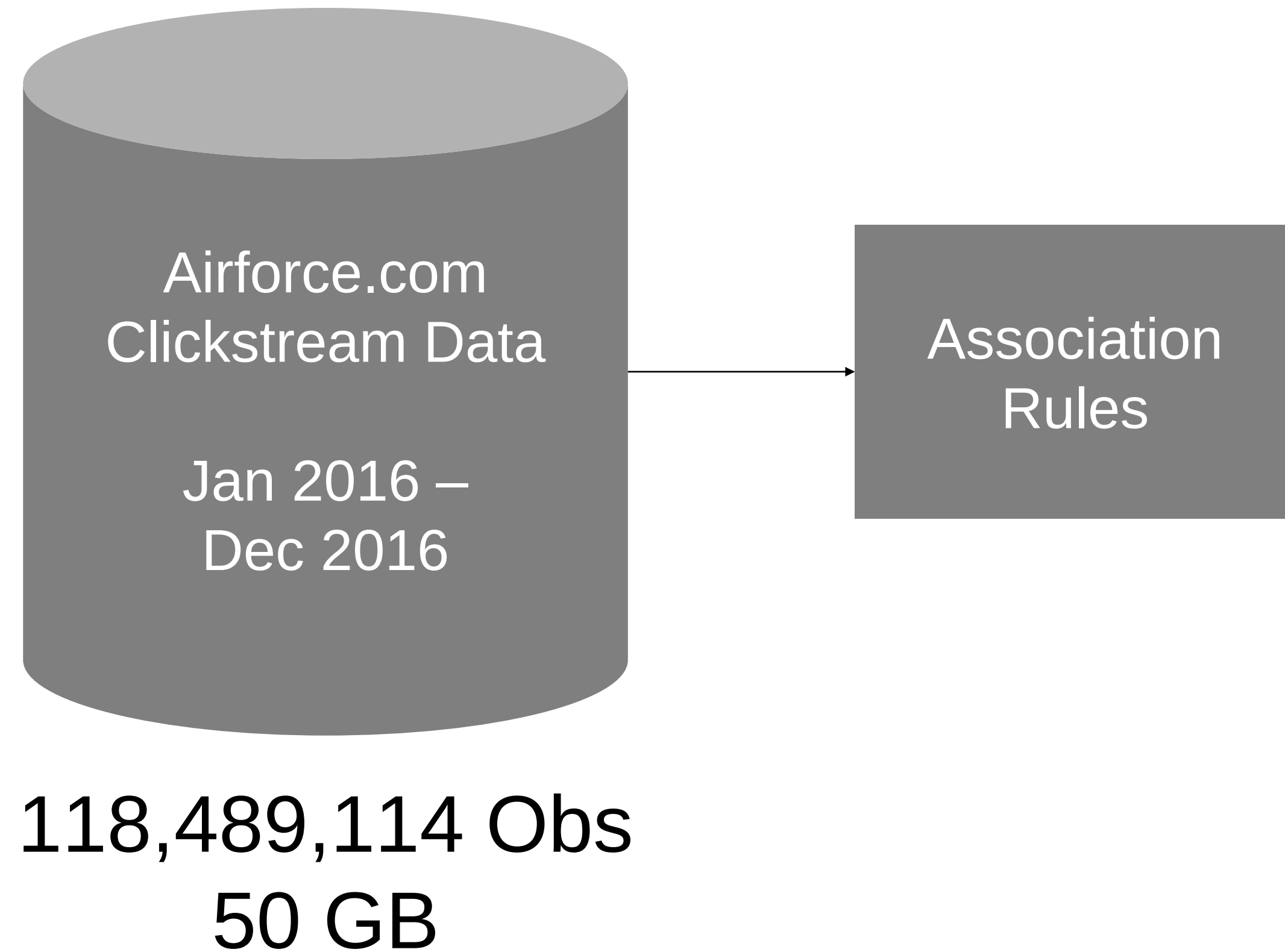
RARE EVENT CHALLENGES

- 3,687,633 VISITORS
- 153,392 APPLICANTS
- 4% APPLICATION RATE

FRAMEWORK



FRAMEWORK



ASSOCIATION RULES

The ASSOC Procedure

[Overview](#)

[Procedure Syntax](#)

[PROC ASSOC Statement](#)

[CUSTOMER Statement](#)

[TARGET Statement](#)

[Details](#)

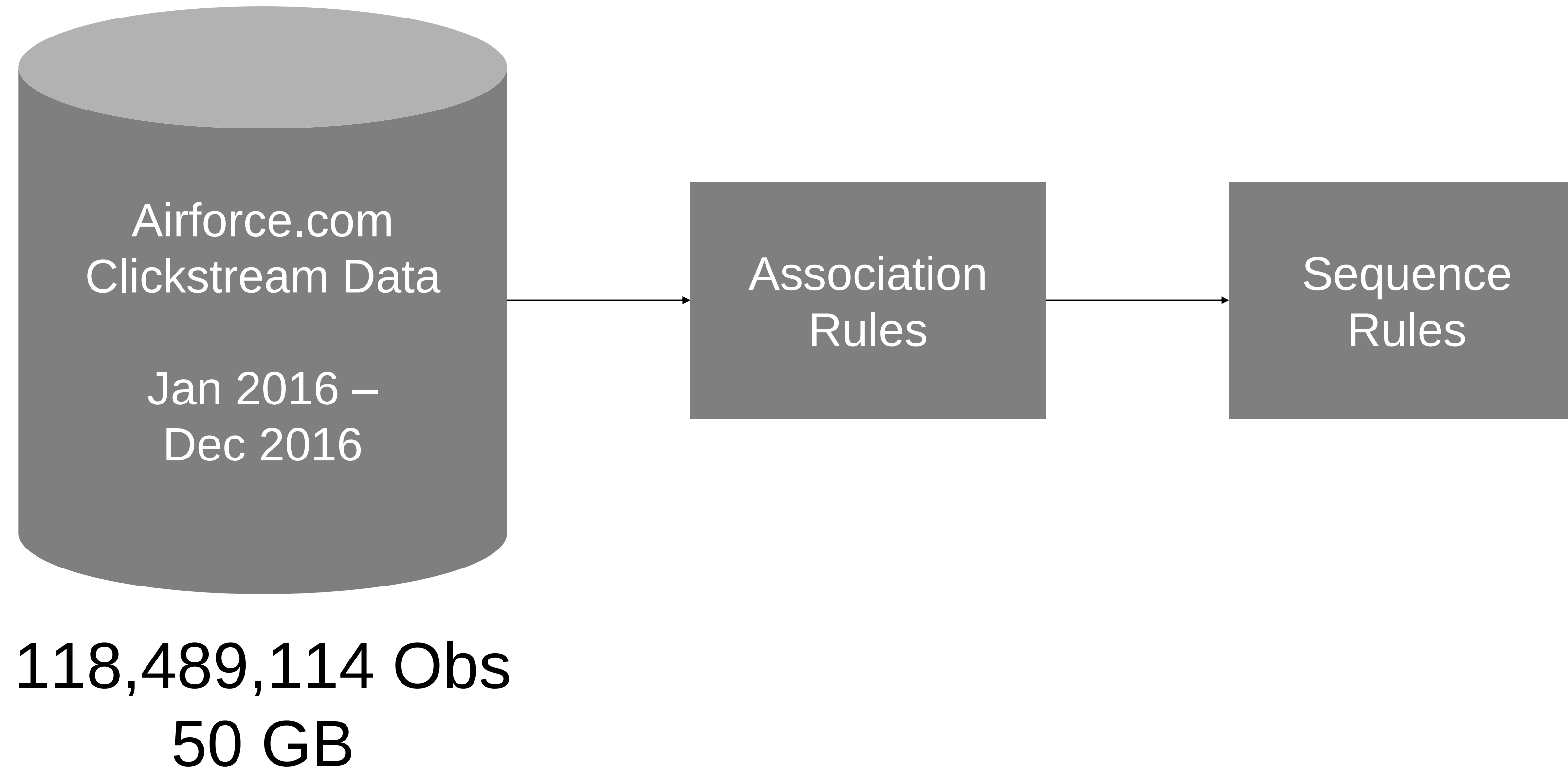
[Example](#)

[References](#)

$$\begin{aligned} \text{Rule: } X \Rightarrow Y & \begin{cases} \text{Support} = \frac{frq(X, Y)}{N} \\ \text{Confidence} = \frac{frq(X, Y)}{frq(X)} \\ \text{Lift} = \frac{\text{Support}}{Supp(X) \times Supp(Y)} \end{cases} \end{aligned}$$



FRAMEWORK



SEQUENCING RULES

The SEQUENCE Procedure

Overview

Procedure Syntax

[PROC SEQUENCE Statement](#)

[CUSTOMER Statement](#)

[TARGET Statement](#)

[VISIT Statement](#)

Details

Examples

[Example 1: Performing a Simple 2-Item Sequence Discovery](#)

[Example 2: Specifying the Maximum Number of Item Events and Setting the Lower Timing Limit](#)

References

The diagram illustrates the relationship between a sequencing rule and its associated metrics. A central rule, $Rule: X \Rightarrow Y$, is shown with three arrows pointing to its metrics: Support, Confidence, and Lift.

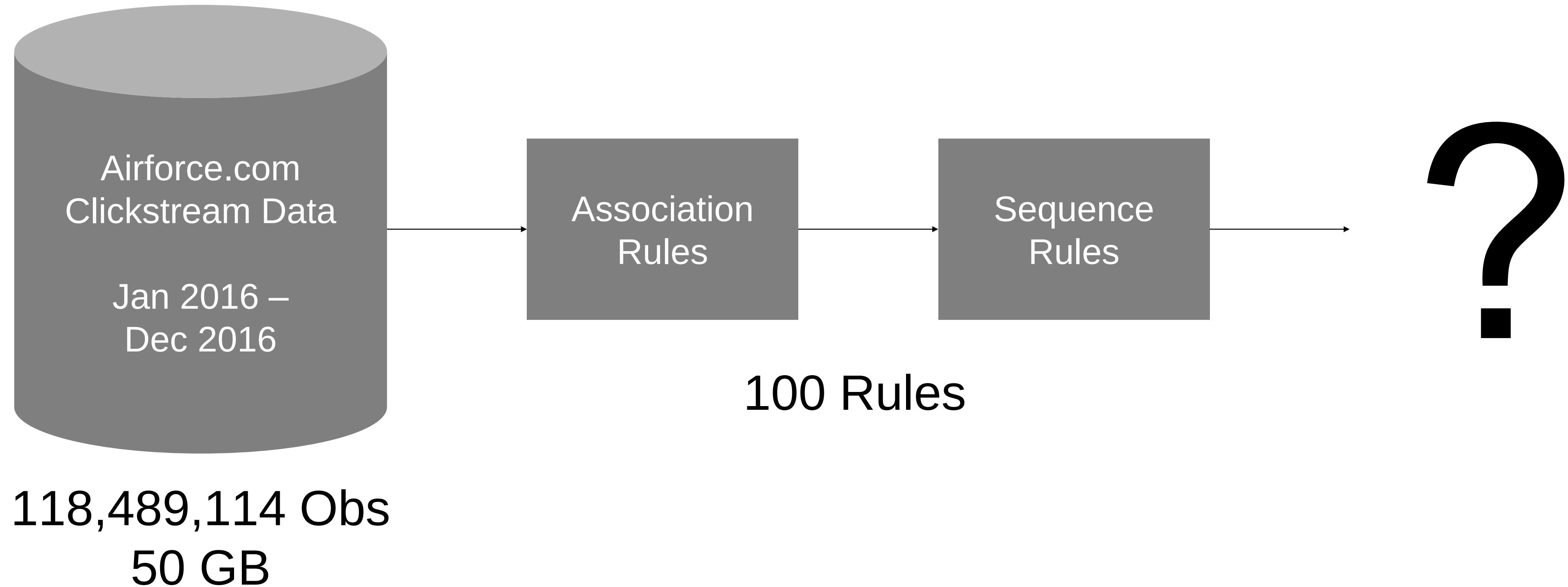
$$Rule: X \Rightarrow Y \begin{cases} \nearrow Support = \frac{freq(X, Y)}{N} \\ \rightarrow Confidence = \frac{freq(X, Y)}{freq(X)} \\ \searrow Lift = \frac{Support}{Supp(X) \times Supp(Y)} \end{cases}$$

RARE EVENT RULES

50 RULES WERE IDENTIFIED INDIVIDUALLY FOR EVENTS AND NON-EVENTS AND SCORED ON THE ORIGINAL DATASET.

Rule Identifier ▲	Rule	Count	Support	Confidence	Chain Size
1	www.airforce.com/apply-now ==> www.airforce.com/apply-now	19610	12.7886	12.9039	2
2	cpc ==> www.airforce.com/careers/pay-and-benefits	12396	8.0840	28.6037	2
3	www.airforce.com/apply-now ==> www.airforce.com/	12029	7.8447	7.9154	2
4	www.airforce.com/apply-now ==> www.airforce.com/find-a-recruiter ==> www.airforce.com/apply-now	11523	7.5147	29.6755	3
5	organic ==> www.airforce.com/ ==> www.airforce.com/how-to-join	10076	6.5710	23.2246	3
6	www.airforce.com/how-to-join/connect ==> www.airforce.com/chat-live	9849	6.4230	24.9930	2
7	www.airforce.com/how-to-join/process ==> www.airforce.com/apply-now	7921	5.1656	41.3392	2
8	www.airforce.com/find-a-recruiter ==> www.airforce.com/find-a-recruiter	7590	4.9498	10.8424	2
9	cpc ==> www.airforce.com/careers	7540	4.9172	17.3985	2
10	www.airforce.com/how-to-join ==> www.airforce.com/how-to-join/prepare-for-success/make-a-decision	7470	4.8715	14.1746	2
11	www.airforce.com/chat-live ==> www.airforce.com/frequently-asked-questions	7446	4.8559	18.6043	2
12	www.airforce.com/apply-now ==> www.airforce.com/find-a-recruiter ==> www.airforce.com/chat-live	7275	4.7444	18.7355	3
13	www.airforce.com/apply-now ==> organic ==> www.airforce.com/apply-now	7184	4.6850	32.0086	3
14	www.airforce.com/apply-now ==> www.airforce.com/careers	7090	4.6237	4.6654	2
15	www.airforce.com/find-a-recruiter ==> www.airforce.com/apply-now ==> www.airforce.com/find-a-recruiter	7081	4.6178	27.2409	3
16	www.airforce.com/careers/pay-and-benefits ==> www.airforce.com/apply-now	6904	4.5024	16.4534	2
17	www.airforce.com/find-a-recruiter ==> www.airforce.com/chat-live ==> www.airforce.com/create-profile-confirmation ==> www.airforce.com/profile	6880	4.4868	99.9564	4
18	www.airforce.com/apply-now ==> www.airforce.com/how-to-join	6784	4.4242	4.4640	2
19	www.airforce.com/how-to-join ==> www.airforce.com/how-to-join/process/enlisted ==> www.airforce.com/apply-now	6757	4.4065	61.7134	3
20	www.airforce.com/chat-live ==> www.airforce.com/profile	6504	4.2416	16.2507	2
21	www.airforce.com/ ==> www.airforce.com/how-to-join ==> www.airforce.com/how-to-join/prepare-for-success/meet-requirements	6450	4.2063	29.6729	3
22	organic ==> www.airforce.com/how-to-join	6365	4.1509	6.7475	2
23	www.airforce.com/profile ==> www.airforce.com/profile	6265	4.0857	16.1832	2
24	referral ==> www.airforce.com/	6171	4.0244	53.8576	2
25	cpc ==> www.airforce.com/find-a-recruiter	5954	3.8829	13.7388	2
26	organic ==> www.airforce.com/careers/pay-and-benefits	5906	3.8516	6.2609	2
27	www.airforce.com/chat-live ==> www.airforce.com/create-profile-confirmation ==> www.airforce.com/profile ==> www.airforce.com/apply-now	5790	3.7759	33.3795	4
28	www.airforce.com/ ==> www.airforce.com/profile	5790	3.7759	7.5079	2
29	www.airforce.com/chat-live ==> www.airforce.com/chat-live	5639	3.6774	14.0894	2
30	www.airforce.com/apply-now ==> organic ==> www.airforce.com/	5619	3.6644	25.0356	3
31	www.airforce.com/how-to-join ==> www.airforce.com/how-to-join/prepare-for-success	5427	3.5392	10.2979	2
32	www.airforce.com/chat-live ==> www.airforce.com/how-to-join/connect	5404	3.5242	13.5022	2
33	www.airforce.com/find-a-recruiter ==> www.airforce.com/	5358	3.4942	7.6540	2
34	www.airforce.com/find-a-recruiter ==> www.airforce.com/chat-live ==> www.airforce.com/find-a-recruiter	5342	3.4838	26.7676	3
35	organic ==> www.airforce.com/ ==> www.airforce.com/careers/browse-careers	5317	3.4675	12.2554	3
36	www.airforce.com/careers/browse-careers ==> www.airforce.com/careers	5284	3.4459	11.6283	2
37	www.airforce.com/profile ==> www.airforce.com/create-profile-confirmation	5275	3.4401	13.6259	2
38	www.airforce.com/how-to-join/process/officer ==> www.airforce.com/apply-now	5259	3.4296	30.7778	2
39	organic ==> www.airforce.com/how-to-join/process/officer	5242	3.4185	5.5570	2
40	www.airforce.com/how-to-join ==> www.airforce.com/how-to-join/process/officer	5237	3.4153	9.9374	2
41	organic ==> www.airforce.com/how-to-join/prepare-for-success/meet-requirements	5236	3.4146	5.5507	2
42	www.airforce.com/how-to-join/prepare-for-success/meet-requirements ==> www.airforce.com/apply-now	5123	3.3409	16.2218	2
43	www.airforce.com/ ==> organic	5023	3.2757	6.5133	2
44	www.airforce.com/careers/browse-careers ==> www.airforce.com/apply-now	4997	3.2588	10.9967	2

FRAMEWORK



MECE TREE

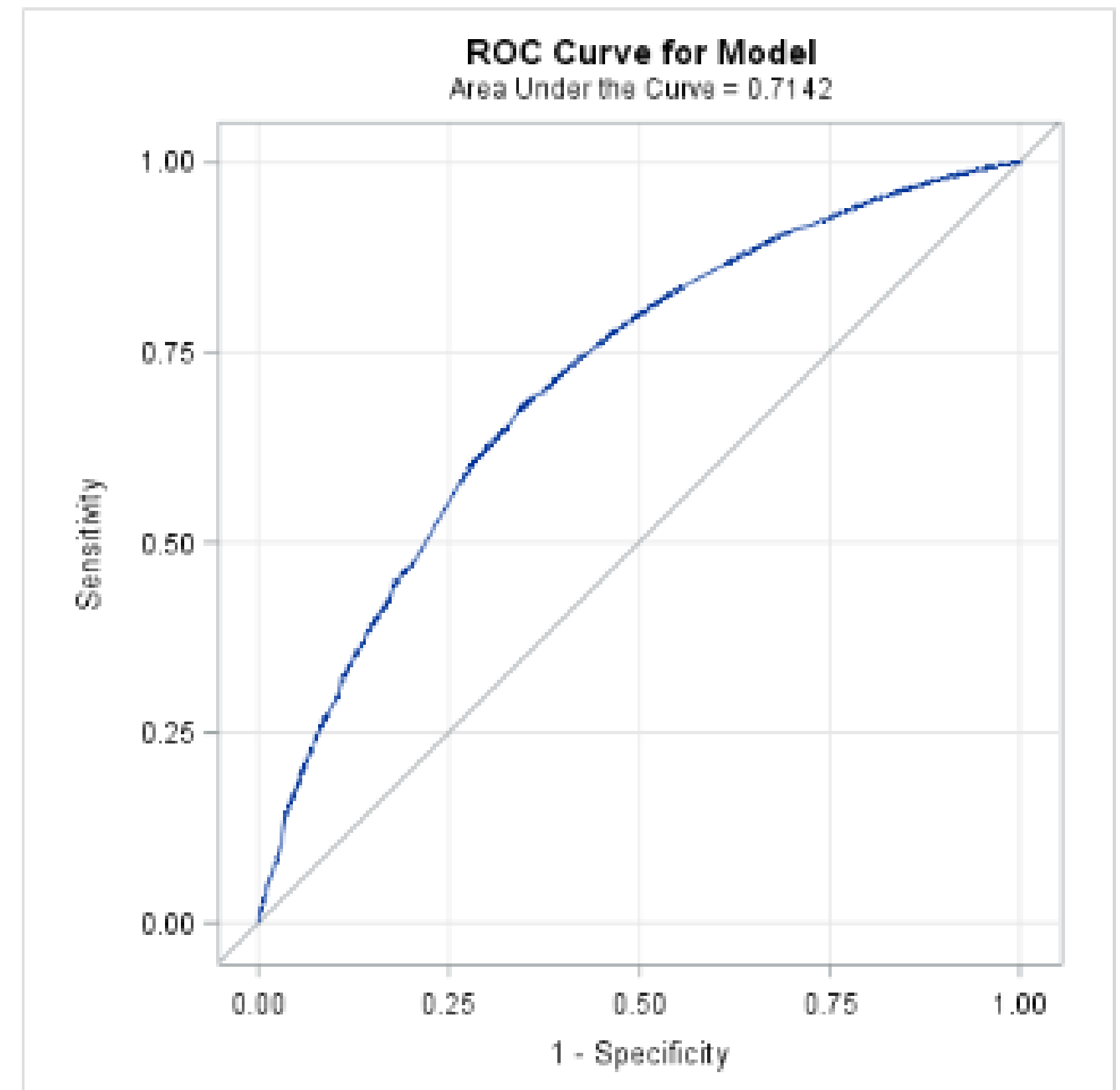
- 1 DEPENDENT (APPLICATION FLAG)
- 100 BINARY (RULES)
- 2 NOMINAL (DEVICE, LOCATION)
- 2 INTERVAL (TIME ON SITE, TIME SINCE 1ST VISIT)
- 1 ORDINAL (NUMBER OF VISITS)

LOGISTIC LASSO

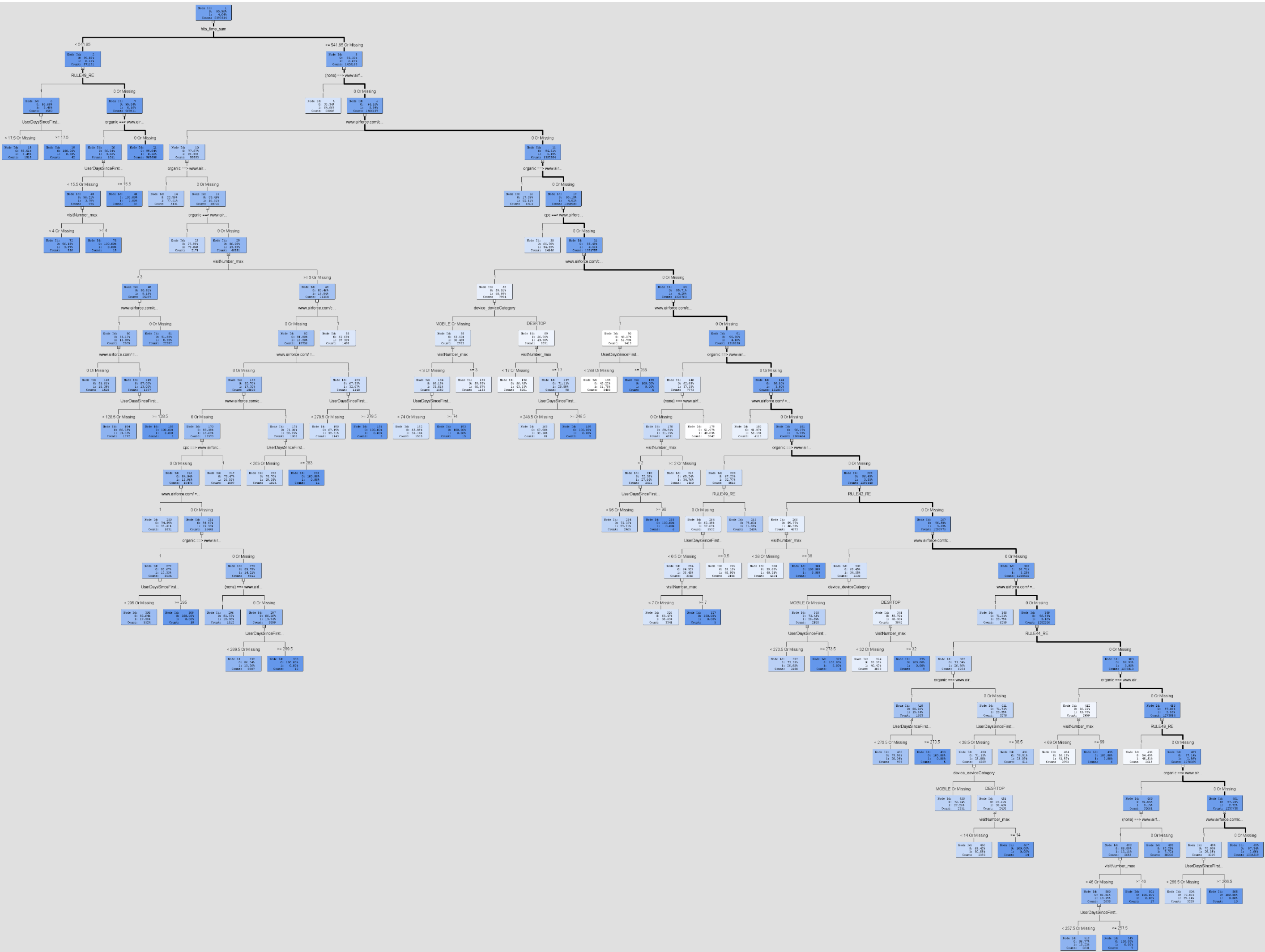
**SAS DOES NOT NATIVELY
SUPPORT LASSO FOR LOGISTIC
MODELS.**

***PROC HPGENSELECT* FOR A
BINOMIAL DEPENDENT WAS
USED.**

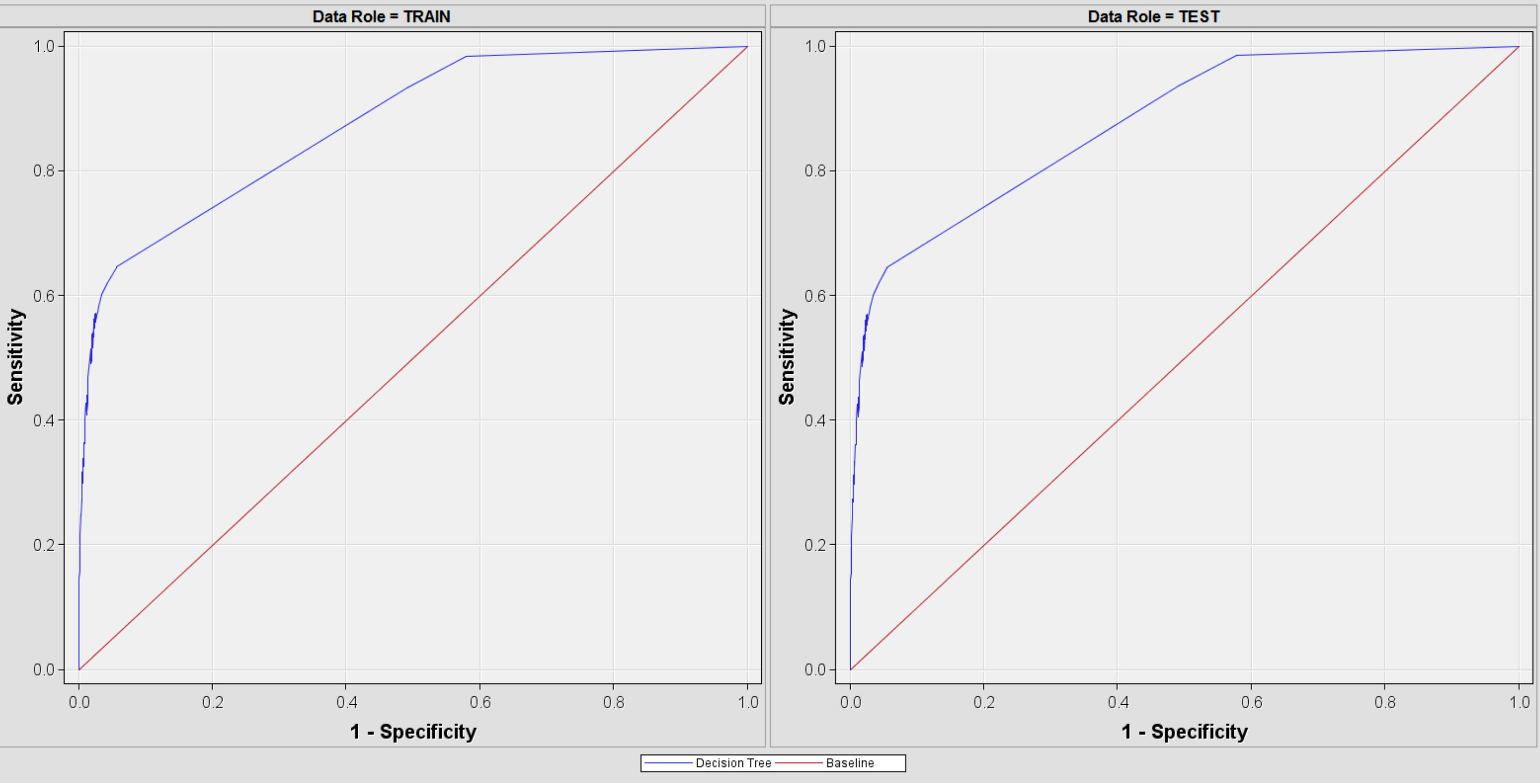
Association of Predicted Probabilities and Observed Responses			
Percent Concordant	70.9	Somers' D	0.428
Percent Discordant	28.0	Gamma	0.433
Percent Tied	1.1	Tau-a	0.194
Pairs	25213418777	c	0.714



DECISION TREE



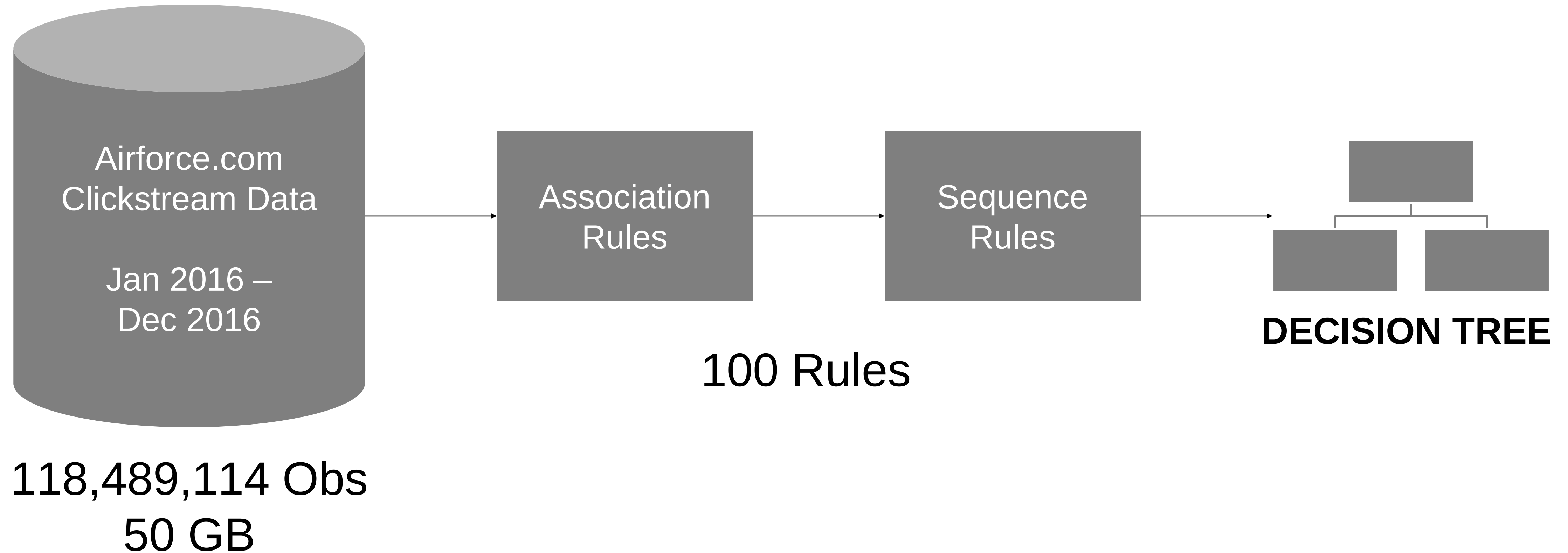
MODEL DIAGNOSTICS



Confusion Matrix		
False Negative	28,929	1%
True Negative	228,6263	95%
False Positive	14,301	1%
True Positive	67,841	3%

30% HOLD-OUT

FRAMEWORK



LEARNINGS

- APPLICANTS SPEND AT LEAST 10M ON THE SITE BEFORE APPLYING
- MORE THAN 50% OF APPLICANTS APPLY ON THE DAY OF THEIR FIRST VISIT
- IMPORTANT PAGE SEQUENCES ARE DEPENDENT ON DEVICE
- GOOD FOR BUSINESS:
 - Recruiter Chat
 - Browsing Careers
 - Searching for Recruiters
 - Requirements
- BAD FOR BUSINESS:
 - Pay and Benefits
 - BMT
 - Education

IMPLEMENTATION

- FLOW SATELLITE RULES ARE UPDATED BASED ON MODEL RESULTS
- MARKETING LANDING PAGES DYNAMICALLY DETERMINED BY MODEL RESULTS
- MARKETING RE-ENGAGEMENT IS SHIFTED TOWARDS DAY ONE VISITORS
- 'APPLY NOW' BUTTON FEATURED ON EVERY PAGE OF THE SITE
- ADJUSTED CONTENT ON:
 - Pay and Benefits
 - BMT
 - Education

RESULTS

- AIRFOCE.COM SITE CONVERSION RATE HAS **DOUBLED**
- MARKETING IS NOW **TWICE** AS EFFICIENT
- ESTIMATED ANNUAL IMPACT OF **\$4.5M** (10% OF BUDGET)



Major General Harencak

