AK

Smithsonian
Cooper-Hewitt, National Design Museum

THE TRAILS IN US 👣

PATTERNS OF HUMAN BEHAVIOR

akamlani@gmail.com | arikamlani.com

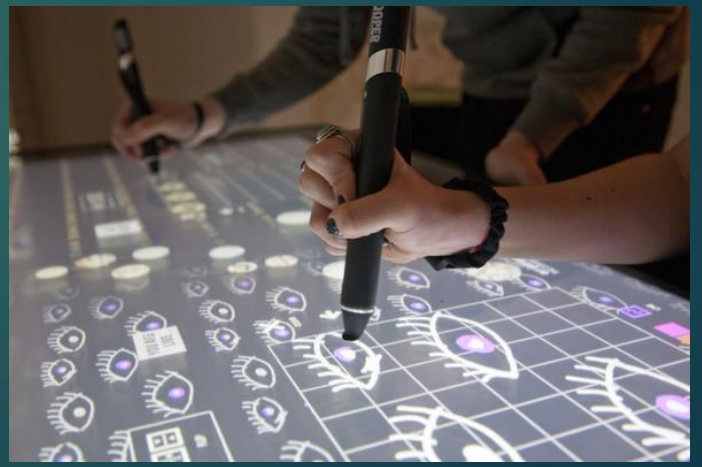
@akamlani/cooperhewitt

THE PEN ARRIVES

WHY THE PEN

AK

- ENHANCE VISITOR EXPERIENCE
  - DRAW VISITOR INTO THE DESIGN EXPERIENCE
- RECENT DATASET RELEASED MAR 2016
  - 3.5M+ OBSERVATIONS
  - AGGREGATE W/COLLECTION METADATA (80K+)
- FOCUS: VISITOR SEQUENCE PATTERN TAGGING BEHAVIOR



PAGE RANK

INFLUENTIAL ARTWORKS BASED ON VISITOR'S DAILY JOURNEY

RM 206

RM 212/105

RM 212/105

RM 212/105

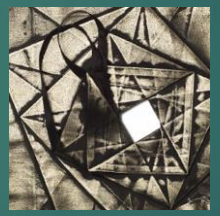
RM 206

RM 212/105

RM 201

RM 212/105

RM 212

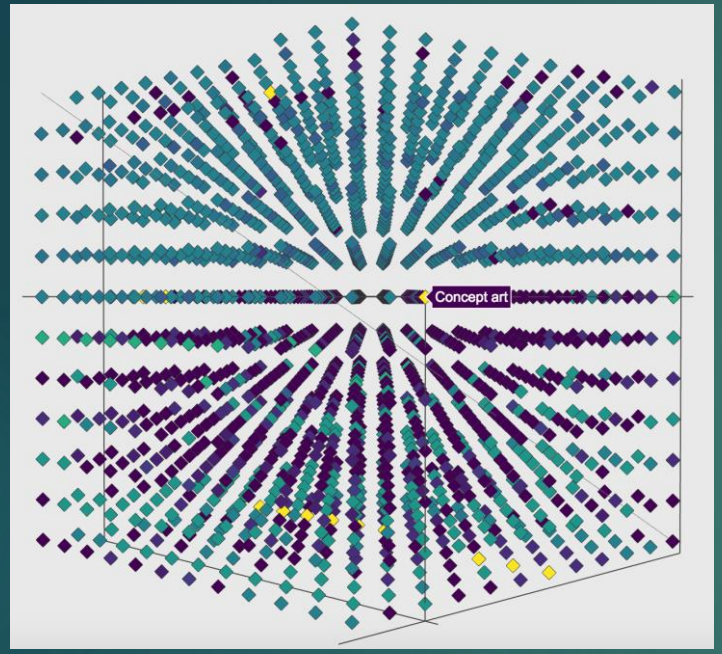


RM 206

- THEMES & DEPARTMENTS
  - STAIRCASES
  - ARCHITECTURE
  - TEXTILES
- TRANSITION SPOTS
  - POSTERS (RM 206)
  - MURALS (RM 201)
  - PRINTS (RM 302)
- # TAGS: ~ 3%
- EXHIBITIONS: 0

Images accredited to Cooper Hewitt Design Museum

COMMUNITIES
LABEL PROPAGATION ALGORITHM (LPA)



Interactive Community Detection Demo

- ONE LARGE DENSE COMMUNITY
- ONE SMALL COMMUNITY
- SEVERAL SMALL MICRO-COMMUNITIES

DENSE	SMALL
EVEN DISTRIBUTION OF EXHIBITIONS	NO EXHIBITIONS
HIGH CAPACITY ROOMS (RM 205,202) THEME ROOMS (1st FLOOR)	SMALL CAPACITY ROOMS SPECIFIC TO 3RD FLOOR
TYPES: DRAWINGS/CONCEPT ART	TYPES: PRINTS
SPECIAL EVENTS	

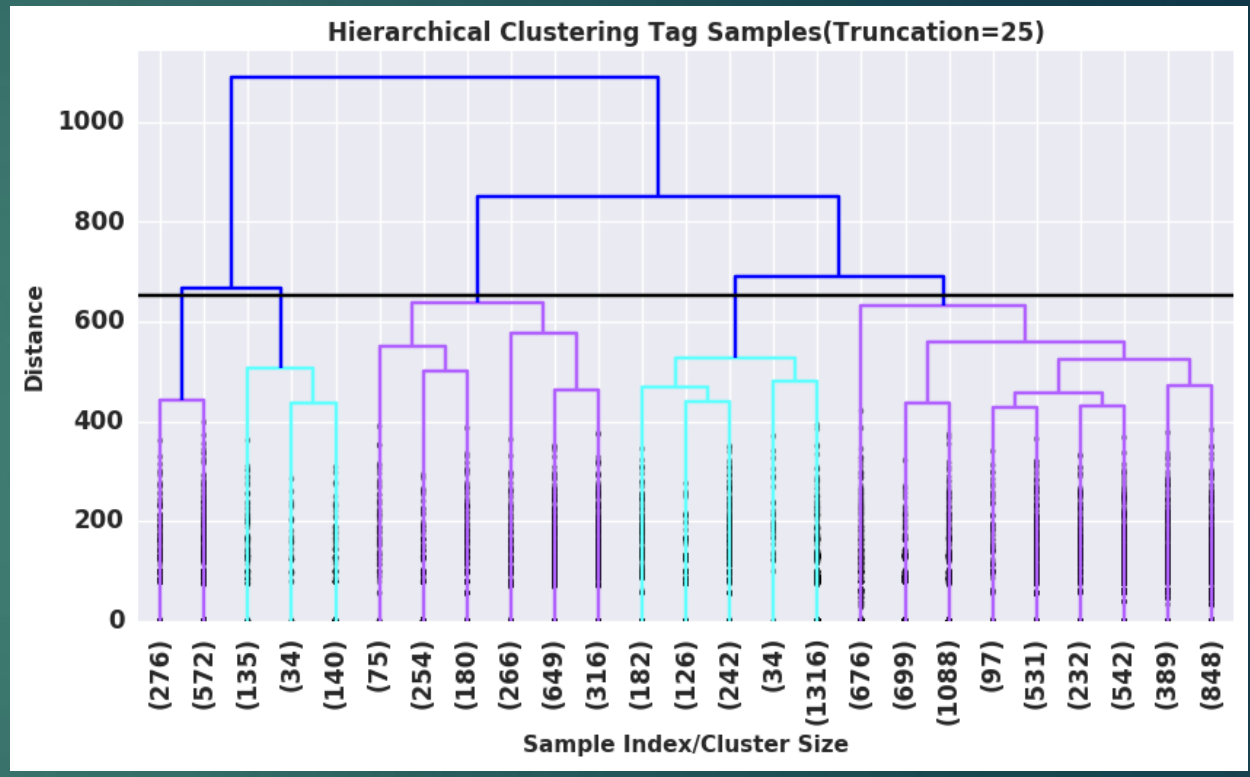
Apache Spark

Spark SQL

GraphX

TEMPORAL PATTERNS

- KMEANS NOT GOOD FOR CATEGORICAL DATA
- HIERARCHICAL CLUSTERING (AGGLOMERATIVE)
  - METRIC: COPH DISTANCE (CITYBLOCK) ~ 70%
  - COMPUTATIONALLY EXPENSIVE
    - NO IMPLEMENTATION IN SPARK MLlib/ML
- CLUSTERS (PER FEATURE ENGINEERING)
  - FLOORS (THEMES, LARGE CAPACITY, EXHIBITS)
  - LOCATIONS (ROOMS/CAPACITY/TURNOVER)
  - BURSTY VISITORS: (2ND FLOOR, LARGE CAPACITY)
  - SPECIAL EVENTS (AFTER HOURS)



RECOMMENDATIONS

HOW TO HELP THE MUSEUM

- FURTHER IMPROVE ON VISITOR EXPERIENCE – NO MISSED OPPORTUNITIES
- OPPORTUNITY TO IMPROVE ON TRANSITION SPOT LOCATIONS
- TAILORED VISITOR SEGMENTATIONS

FUTURE VISION

- ADDITIONAL MODELS ON MEMORY SEQUENCES (RNN/LSTM)
- FEATURE ENGINEERING ON ANNOTATED METADATA (NLP)
- TRANSFER LEARNING (APPLY TO OTHER DOMAINS)

DISCUSSION

AK

Website: arikamlani.com
Email: akamlani@gmail.com

@akamlani

Appendix

AK

COMPONENTS

AK

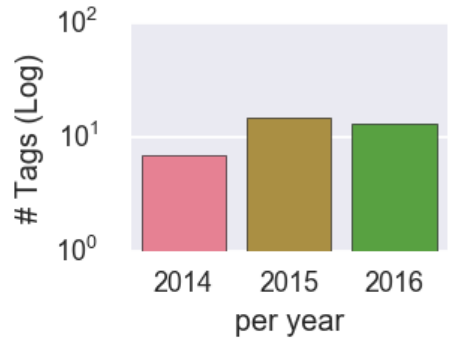
NON-DISTRIBUTED

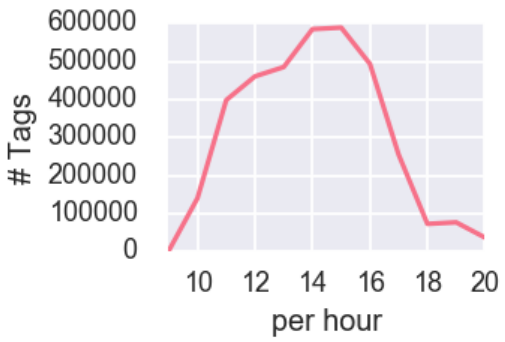
igraph

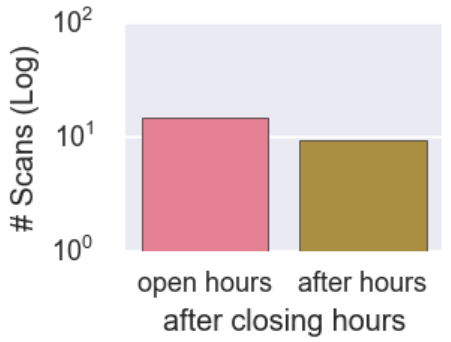
DISTRIBUTED

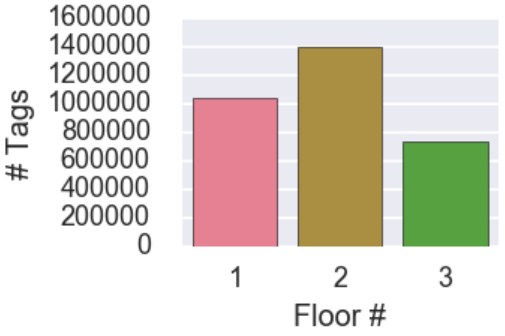
LOVE THY DATA

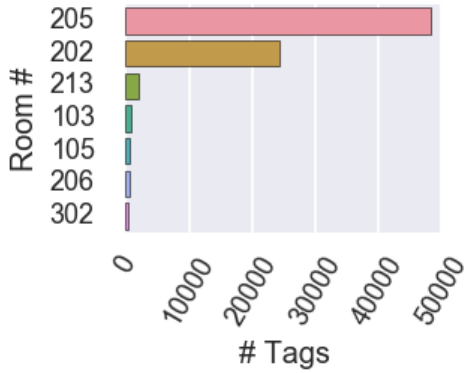
AK

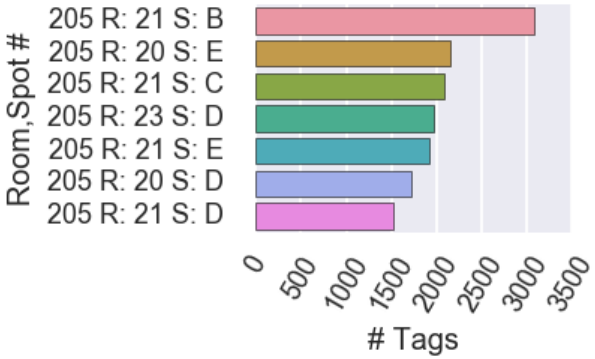


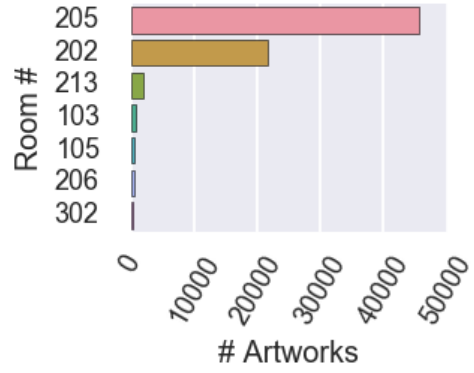


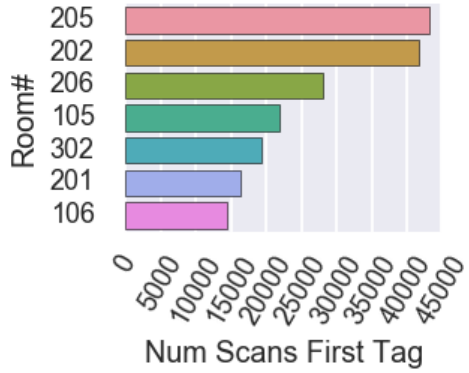


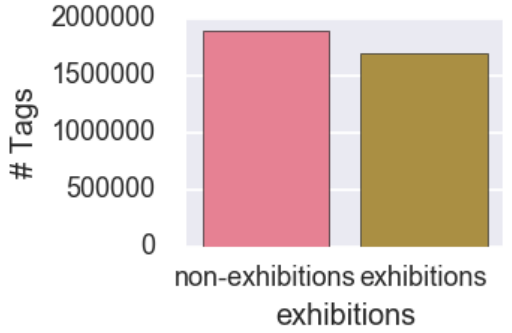




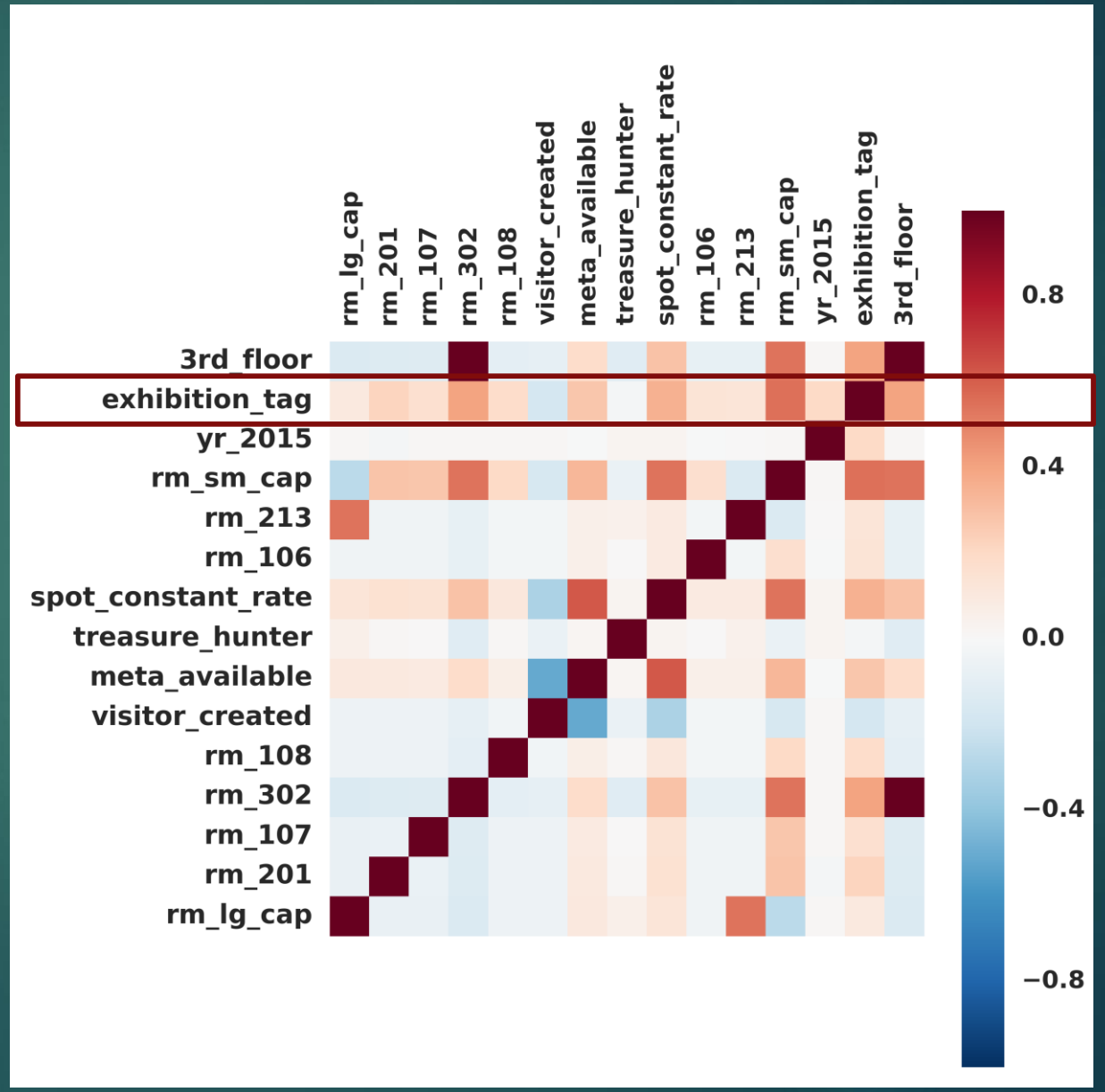








- DATA
  - PEN 3.5M+ OBS
  - METADATA (80K+)
- HIGH CAPACITY
  - RM 205,202
- FIRST ARRIVAL
  - RM 205, 206, 202
- EXHIBITIONS
  - INFERRED FROM DAILY BUNDLES

EXHIBITIONS

- EXHIBITIONS
  - SMALL CAPACITY ROOMS
  - FLOOR: 3RD NOT 2ND
  - CONSTANT SPOTS
  - BURSTINESS N/A
  - METADATA AVAILABLE