



NBRC

Timothy Slaper
tslaper@Indiana.edu

dstillery

&

AUDIENCE SOLUTIONS ▾

AUDIENCE STUDIO ▾

CONTENT HUB ▾

DSTILLERY



YOUR CUSTOM AUDIENCE PARTNER

BUILDING YOUR BEST AUDIENCES WITH CUSTOM AI MODELS

DIGITAL VAPOR TRAILS

Using Website Behavior to Nowcast
Entrepreneurial Activity

- Three goals for official socio-economic statistics:
 - **Accuracy – Timeliness – Relevance**
- Entrepreneurship is **relevant**:
 - business formation rate has been in decline
 - small businesses have created 62 percent of all net new private-sector jobs (BLS/DOL)
- E-shop data on business formation currently **not very timely**

DIGITAL VAPOR TRAILS

Using Website Behavior to Nowcast Entrepreneurial Activity

- E-shop data is often **geographically unspecific** → **reduces relevancy**
 - lack of regional specificity to compare program outcomes, policies and cultural influences
 - “everything happens in space”
 - Several efforts to use unconventional data to “nowcast” economic performance – Gleaser et al. Yelp!
 - Met **dstillery** 2 years ago at the NABE tech conference and...
 - ...we started to collaborate
- Shared work with Peter Lenz and Alyssa Bianca of dstillery

Nowcasting Entrepreneurship

BEHAVIORAL AUDIENCES

Research prompt:

“I want to analyze an audience of people who are starting a business.”



MODEL FEATURES

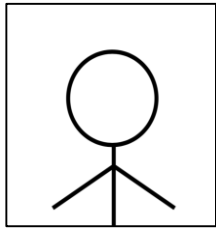
**Intended audience: “People
looking for entrepreneurial
resources”**

Entrepreneurship crafted audience, shareable features 07/01/2018

FEATURES	WEIGHTS
www.paloalto.com	3.40535736
www.mplans.com	3.24590421
www.business-plans-guide.com	3.0856905
www.fastbusinessplans.com	2.73994374
www.startupdaddy.com	2.43793988
www.planware.org	2.35142303
www.futurpreneur.ca	2.24276352
www.score.org	2.22950125
www.accion.org	2.04976988
www.partnersinleadership.com	1.92100525

INDIVIDUAL ANONYMOUS USER PROFILE

Faceless
anonymous user



Anonymous user ID:
3M20D8SR60210

Device type: 32-3B
First seen: 2/15/18
Total actions: 35

DSTILLERY

Every day we see...

- 160+ billion data points
- 450+ MM devices
- 300+ MM individuals

Dynamically scored audience
assignments refresh daily to reflect the
latest user behavior

Dynamic Audiences:

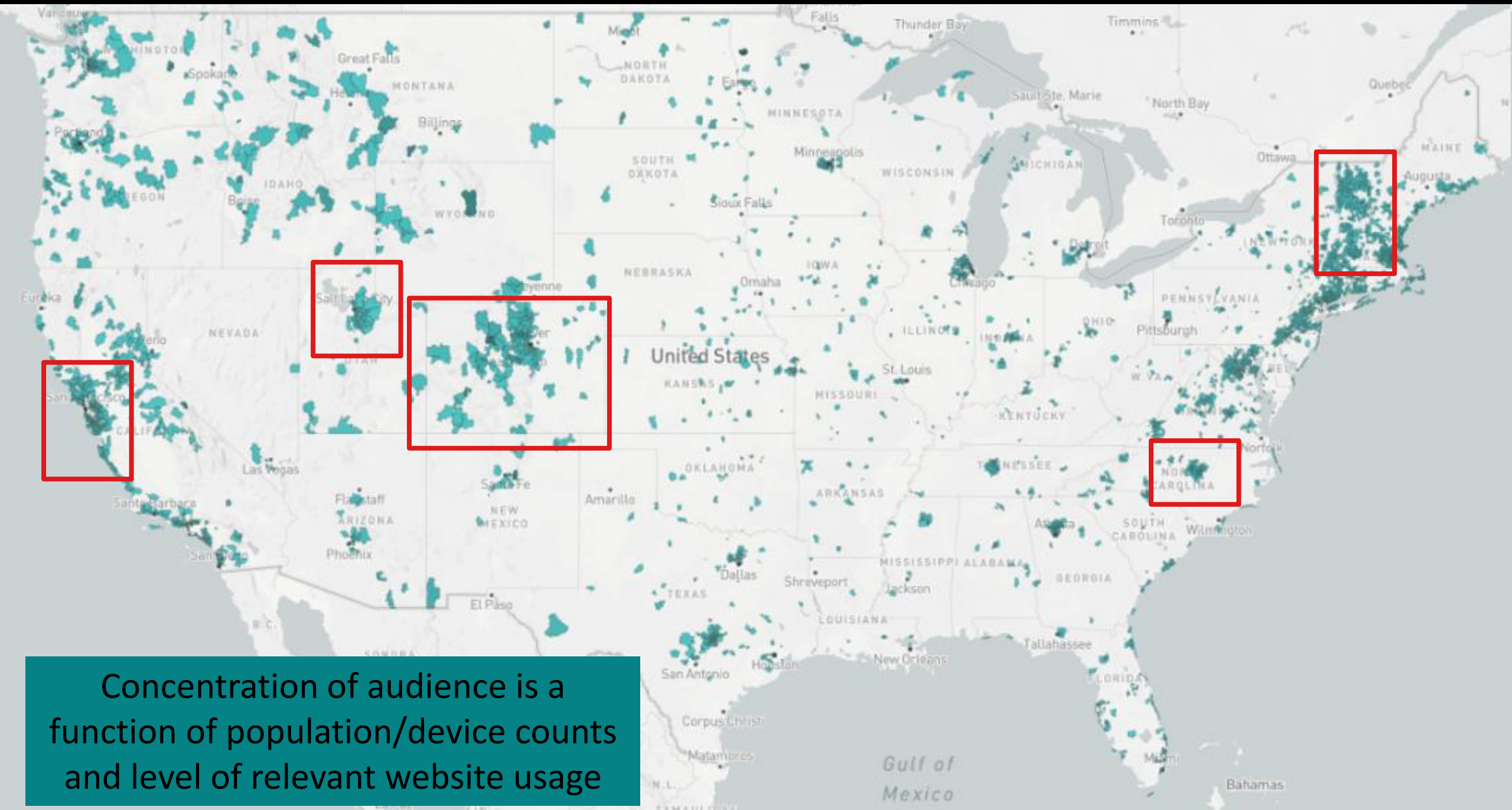
- Business News Readers
- **Entrepreneurship**
- Executives and C-Suite
- Men's Clothing Shoppers

Act-alike Audiences:

- Client C High Propensity
- Client D High Propensity

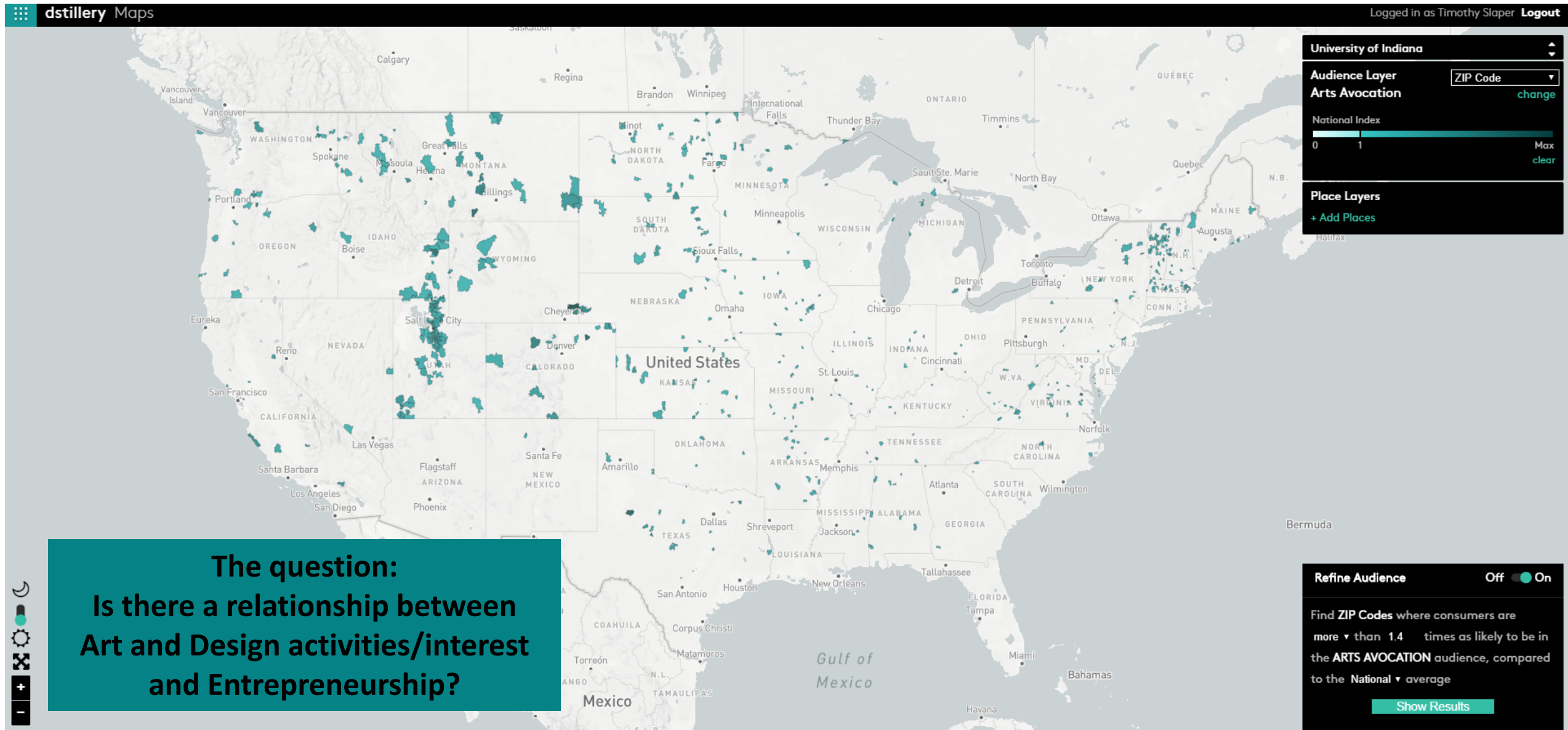
Intended audience

sources"



Concentration of audience is a function of population/device counts and level of relevant website usage

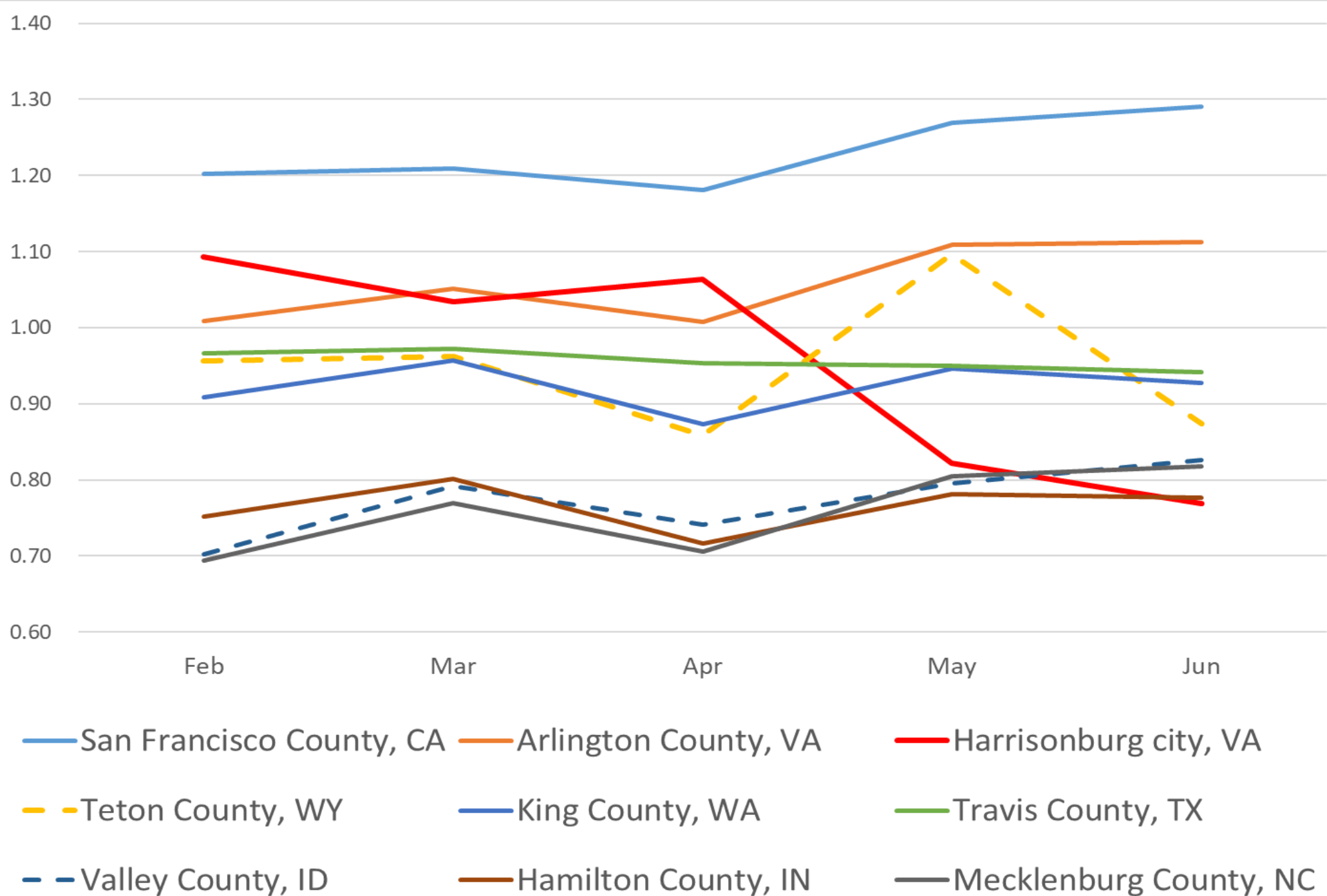
Concentration of Arts Avocations



**Do the
data
make
sense?**

Generally
speaking: Yes

**How does
it align with
Ground
Truth?**



Aligning the data to other E-ship measures

Kauffman Foundation publishes an index on Entrepreneurship

- The index differentiates **start-up** formation activity from **early business growth**
 - See Kauffman Foundation reports (2017a, 2017b) for more detail on what the rate, “opportunity” E-ship share and density measures capture
- Latest iteration of the Kauffman Index data from 2017
- The Kauffman index are available for the top U.S. 40 metro areas (But not any more. State only.)

Aligning the data to other E-ship measures

Dstillery Entrepreneurship Audience

- Dstillery started capturing data for E-ship behavior in January 2018
- Aggregated ZIP code concentration to MSA
- Critical assumption: region's website behavior in current year is consistent with that region's experience in previous years
- Similar assumption made for regional personality profiles to explain entrepreneurial activity, Obschonka et al. (2015).

Compare data for 40 metro areas

Find correlations between Kauffman indexes and Dstillery indexes across three sub-measures defined in the Kauffman data set

Start-up Formation	
Rate	0.32
Share	0.00
Density	0.19

Entrepreneurial Growth	
Rate	0.49
Share	0.27
Density	0.42

SourceLink staff commented that the different formation versus growth results make sense.
They have different resource needs!

SourceLink is an E-Ship ecosystem gardener - <https://www.joinsourcelink.com/>

Yelp! data had weak to moderate predictability for forecasting new establishments
Glaeser, Kim & Luca. (2017) "Using Yelp Data..."

Why is an unconventional data *user* on this panel?

- “Crowd source” experimenting with unconventional data
- State and Federal statistical agencies are resource-constrained and can’t be expected to test every wild hair idea
- We researchers can test if unconventional data is good for estimating “best change” (rather than level) for geo-graphically specific socio-economic data
 - If accurate (enough), these can be adopted to produce economic activity estimates, monitor program/policy effectiveness, assess real-time changes in economic conditions...
- Big question: How do we keep the data flowing as firms change, get purchased, CEO succession, etc.?