RE: Data

David Lanham

Wed 8/22/2018 2:57 PM

To: David Harshbarger <DHarshbarger@brookings.edu>;

Cc: Alec Friedhoff <AFRIEDHOFF@brookings.edu>;

Thanks for checking in here, David. I spoke with Andre today as well on a few project scenarios, so I might be up to date on release dates for these three pieces we've discussed. I have in my notes approximate release dates of September 18 for how majority-black cities became majority-black cities, October 9 for the housing piece, and early/mid November for the businesses piece. For filling in the handoff date, that's helpful to include whatever you think is an accurate date where you'll have full material ready for C&E (text, data). As for time needed for publication material, it's a bit of a case by case scenario. A blog without data is doable in 2 days, a blog with data I'd say 2-3 days approx., a research brief with data (piece in the 1,500-3,000 word range) I'd say 3 days. Full research reports (PDF, landing page, etc) are in the 4-6 week range. All of which are variable depending on what else is on the publication calendar.

The variable here on these is how much time Alec will need to create the dataviz for these three pieces, so Alec will want to chime in as well. But the sooner we have everything, the easier we can move on everything.

David Lanham | 202 238 3594 |dlanham@brookings.edu | @BrookingsMetro

---

From: David Harshbarger
Sent: Wednesday, August 22, 2018 11:40 AM
To: David Lanham <DLanham@brookings.edu>
Cc: Alec Friedhoff <AFRIEDHOFF@brookings.edu>
Subject: RE: Data

I talked with Andre and figured out some dates to aim for. As far as uploading to Salesforce, these would be Release Dates. However, there's also a field for Handoff Dates which I'd like to fill in—is there a difference in time needed for blogs vs. reports, and how long before the release date would that be for each?

David

From: David Lanham
Sent: Monday, August 13, 2018 5:15 PM
To: David Harshbarger <DHarshbarger@brookings.edu>
Cc: Alec Friedhoff <AFRIEDHOFF@brookings.edu>
Subject: RE: Data

Gotcha, thanks. Let's get Salesforce updated with prospective release dates for each once Andre's back.

David Lanham | 202 238 3594 |dlanham@brookings.edu | @BrookingsMetro

---

From: David Harshbarger
Sent: Monday, August 13, 2018 5:13 PM
To: David Lanham <DLanham@brookings.edu>
Cc: Alec Friedhoff <AFRIEDHOFF@brookings.edu>
Subject: RE: Data

No exact dates as far as I know, but I imagine he will want to move quickly on both. This shorter piece with the historical Census data has been in the works for a while, so I think he'll be eager to get it out within a couple weeks of his return. The whitepaper report was slated for early October, but since we've made good progress there he may want to aim for late September. I'll be sure to speak with him about it next week, however.

From: David Lanham
Sent: Monday, August 13, 2018 5:07 PM
To: David Harshbarger <DHarshbarger@brookings.edu>
Cc: Alec Friedhoff <AFRIEDHOFF@brookings.edu>
Subject: RE: Data

David, do you have a sense from Andre when he might want to get these pieces published by? Just want to be on top of our editorial calendar, if any timing intel you have while Andre's on vacation.

David Lanham | 202 238 3594 |dlanham@brookings.edu | @BrookingsMetro

---

From: David Harshbarger
Sent: Friday, August 10, 2018 2:46 PM
To: Alec Friedhoff <AFRIEDHOFF@brookings.edu>
Cc: David Lanham <DLanham@brookings.edu>
Subject: RE: Data

Great – and yes, I'm using RStudio so that is perfect. In that case then here's a .RData file which might be better than the csv. I know sometimes weird things can happen with NAs for example when I export to csv. Also, the .RData file includes a list-column of city names that I couldn't put in the csv.

From: Alec Friedhoff
Sent: Friday, August 10, 2018 2:38 PM
To: David Harshbarger <DHarshbarger@brookings.edu>
Cc: David Lanham <DLanham@brookings.edu>
Subject: Re: Data

Great, I'll have a look. Looks like you might be using R? If so (and assuming you're using RStudio) I can share an R notebook with my findings. For now, I think it's fine to go back and forth via email.

-Alec

Alec Friedhoff
Communications Officer and Associate Fellow
Metropolitan Policy Program
BROOKINGS

From: David Harshbarger
Sent: Friday, August 10, 2018 2:32:40 PM
To: Alec Friedhoff
Cc: David Lanham
Subject: RE: Data

Ok, I've attached a .csv file of the data. This represents all the Places for which we have data in each of the census years from 1970-2010, and for which the median of the total population at each of those 5 collection points is at least 2500. Some of the column names may need some clarification:

- cityclass is the metro status. 1 = Principal city of metro and 2 = ring city of metro. 3 = Principal city of micro, 4 = ring city of micro, 5 = not in any cbsa
- distPrincipal is distance in miles to the Principal city of metro or micro if in one at all
- slopeW is slope of the best-fit line between the five records for the white population in a place
- slopeB is slope of the best-fit line between the five records for the black population in a place
    - I was messing around with a way to test the edge cases that are borderline between quadrants, so I made some columns for R^2 and p-values of the best-fit lines which are still included here

- slopePCTB is the trend between each of the five records for the pct black in a place (the steepest lines are those inner-ring quadrant 2 cities)
- mylog10.slopeW is the "trend score" for the white population of a place. Essentially just the logged slopeW number to be able to plot them all at once. Likewise for mylog10.slopeB

For reference I've also attached a couple other rough draft images that I was exploring (but in the future would you rather send data/images on Salesforce instead? wasn't sure if this file might be too big). Eager to see what you come up with as well.

David

From: Alec Friedhoff
Sent: Friday, August 10, 2018 1:42 PM
To: David Harshbarger <DHarshbarger@brookings.edu>
Cc: David Lanham <DLanham@brookings.edu>
Subject: Re: Data

David, this is a great start and I like the way you're thinking about the data. If you have the data, I'd love to take a look at few other plots and maps. I'm happy reading it in any format you currently have it in.

Alec Friedhoff
Communications Officer and Associate Fellow
Metropolitan Policy Program
BROOKINGS

From: David Harshbarger
Sent: Friday, August 10, 2018 1:26:50 PM
To: Alec Friedhoff
Cc: David Lanham
Subject: RE: Data

Hi Alec,

Thanks for reaching out. We're making good progress on both projects (both on Salesforce, although we haven't uploaded any data or text to them yet). Here are some details on the first one:

Andre is currently out on vacation (back on the 20th) but before he left we discussed some dataviz ideas. For the first, shorter blog-type piece, we are of the opinion that a zoomable, national map might be overwhelming for such a small piece. However, there is one chart for the piece which may still be a good candidate for some interactive work, and which I can share data for. Here are the ideas for the findings and graphic(s):

- Motivation: Andre wants to study majority-black cities (simply put, >50.0% black population), so he wants to know when/where they emerged. They often (though not always) emerged when the black population increased and the white population decreased, so when/where did that happen?
- Data: Decennial Census data for each Census from 1970 to 2010, by "Place" (best approximation of cities/towns), with population estimates by race
  - Lat/Long coordinates added for each Place
  - Flagged in a column as principal city of a metro area, in metro but not the principal city, or out of metro area
  - From the two above, the straight-line distance from each Place to the principal city in its metro (if it is in one)
  - Flagged in a column if in a top 100 metro
  - Category for one of 4 Census Regions (https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf)
  - A numeric column for each of the "trend" of the white population and "trend" of the black population since 1970 (essentially just the slope of a best-fit line between the 5 data points from 1970-2010)
- Graphic:
  - "Quadrant" chart – we think it's interesting to plot each of these Places with the white population trend as the x axis and the black population trend as the y-axis. The quadrant of the chart in which

each city falls speaks to the trajectory the city is on. Attached is a draft png of what that could potentially look like, with regions emphasized. I've been using the convention of referring to quadrants as Quadrant 1 – top right, Quadrant 2 – top left, Quadrant 3 – bottom left, Quadrant 4 – bottom right.

The interactive idea would be to provide that chart, but rather than the still image attached to be able to highlight/filter based on region (maybe even particular states?), metro status, %black in 2010, and distance from principal city in metro. When I've done this by hand a couple things have stood out:

- There's a regional divide. Northeastern cities are largely found in the top-left quadrant Q2, where white population has trended down and black population has trended up. Many southern cities have seen upward trends in black population as well. Western cities are growing in both black and white populations in Q1.
- Older industrial cities such as Pittsburgh and St. Louis fall in Q3, where they've just seen blanket loss of population. But their suburbs have seen growing black populations (Q1 or Q2).
- Which leads us to a final point: Q2 Places in top 100 metros tend to be clustered at the edge of the principal cities, followed by Q1 Places in the outer ring. This was what we were originally envisioning as being the purpose of the map, but we think it might be simpler and easier to just provide a few snapshots of certain metro areas as examples.

I briefly touched base with Andre about this before he left, but he still would like to discuss it further after he returns. I can go ahead and send some of the data to you if you'd like to work with it, although I suspect Andre will still want to add/remove certain elements of it.

David

--
David Harshbarger
Research Assistant | Metropolitan Policy Program | The Brookings Institution
1775 Massachusetts Avenue NW, Washington, DC 20036
202.238.3116 | dharshbarger@brookings.edu

From: Alec Friedhoff
Sent: Friday, August 10, 2018 11:42 AM
To: David Harshbarger <DHarshbarger@brookings.edu>
Cc: David Lanham <DLanham@brookings.edu>
Subject: Data

Hey David,

I wanted to touch base with you on any data you might have related to Andre's upcoming of work. I thought it might be helpful to focus in on one paper first. It sounded like the demographic change piece was further along? Any data you can share for that? Of course, any high level thoughts on findings that you guys have put together would be helpful too.

Also, are these projects in Salesforce? I've been finding it helpful to use salesforce as a way to share notes and (small) data files...

Cheers,
Alec