



STUDY:

Out-of-MSA moves

How we use the Unacast Migration Patterns dataset to analyze the outflows from US metropolitan areas to locations classified by the distance from its boundary.

Introduction

It is a known fact that when people change their address, more often than less they relocate to a relatively not distant place [[Hendren et al.](#)]. A common trend, especially for certain demographic groups, is to leave busy city centres for a quieter, more family-friendly and cheaper suburban locations [[O'Donnell et al.](#)]. Little is known, however, about how far people move once they decide to relocate beyond the boundary of the metropolitan area where they live.

In this study, we take a closer look at this phenomenon. Specifically, we use the Unacast Migration Patterns dataset to analyze the outflows from US metropolitan areas to locations classified by the distance from its boundary.

unacast.

Research Question

We are interested in answering the following two research questions:

1. *How far do people move once they decide to relocate beyond the boundaries of their home Metropolitan Statistical Area (MSA)?*
2. *Is the distribution of outflow distances similar for all MSAs or is it region-specific?*



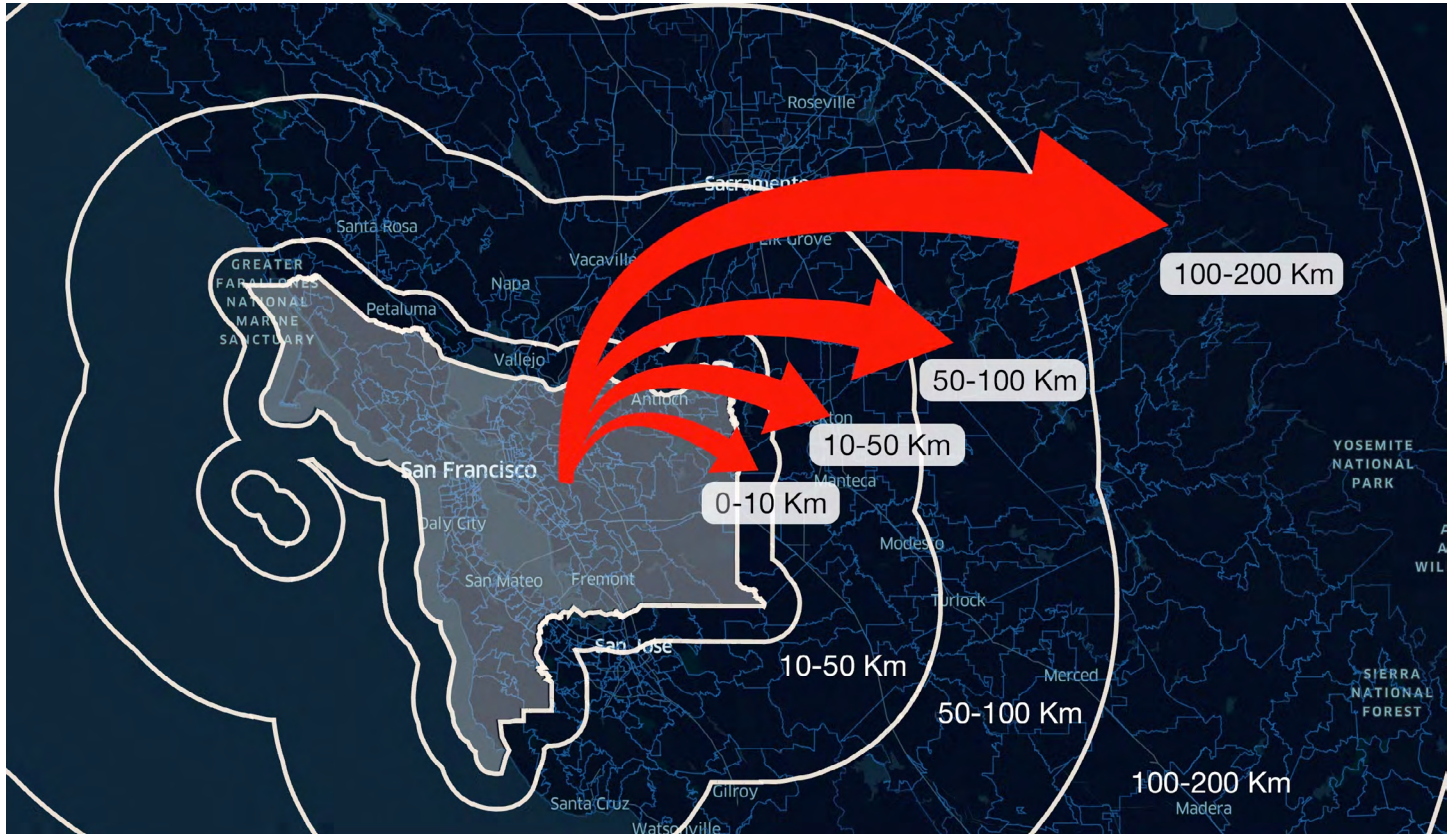
Approach

The origin-destination migration flow in the scope of this study (and related dataset) is defined as the estimated number of people (semi-)permanently changing their home location from one area (origin) to a different area (destination) during a selected period of time. The most granular spatial representation of home location offered by the dataset is ZIP code level or, alternatively, Census Tract level.

We choose to work with the former one and we specifically focus on one aspect of this dataset - outflows from ZIP codes belonging to Metropolitan Statistical Areas (MSA). In other words, we study the number of people relocating from ZIP code areas within a certain MSA to ZIP code areas beyond its boundaries. With 41,683 ZIP codes existing in the US it is highly impractical to analyze flows between individual ZIP code areas, instead, we aggregate the flow originating anywhere in the MSA and introduce bands of destination ZIPs defined by the distance from the boundary of MSA (Fig. 1).

Definition:

Band of outflow is an area that we use for aggregating outflow from an MSA. The area is defined as a buffer polygon outside of the MSA, which is given by the shape of the MSA, the distance from the boundary of the defining MSA polygon, and the width of the buffer. Since our migration is defined on the ZIP level geographies, each band aggregates flows to all ZIP code areas whose centroid falls inside the band's polygon.



Parameters of this study:

We analyze moves during the observation window between 2019-01-01 and 2022-01-01 and we use the following 7 outflow bands:

- 0-10 km
- 10-50 km
- 50-100 km
- 100-200 km
- 200-500 km
- 500-1000 km
- 1000-2000 km

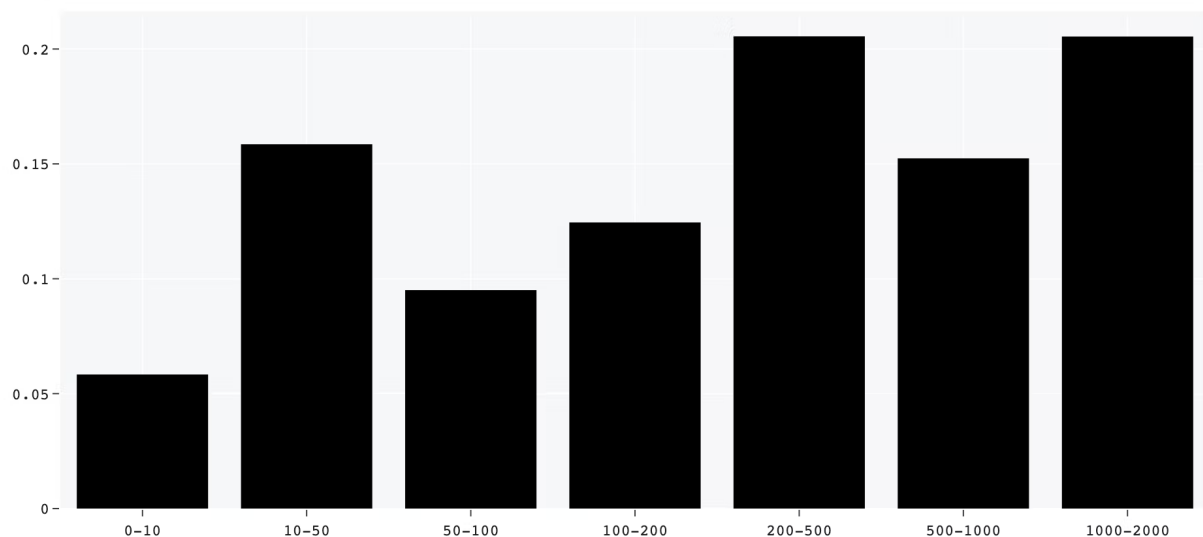
Results

RQ1: How far do people move once they decide to relocate beyond the boundaries of their home Metropolitan Statistical Area (MSA)?

Based on the described approach, we calculate the flows to seven distinct outflow bands for 384 Metropolitan Statistical Areas. To answer our first RQ, we aggregate these results and, as depicted in Fig. 2, we see two trends:

- A) The majority of people find their new home in a distant location (>56% move farther than 200 Km from their home MSA).
- B) A relatively large group of people (15.9%) move to a narrow band which lies only 10-50 Km from the home MSA.

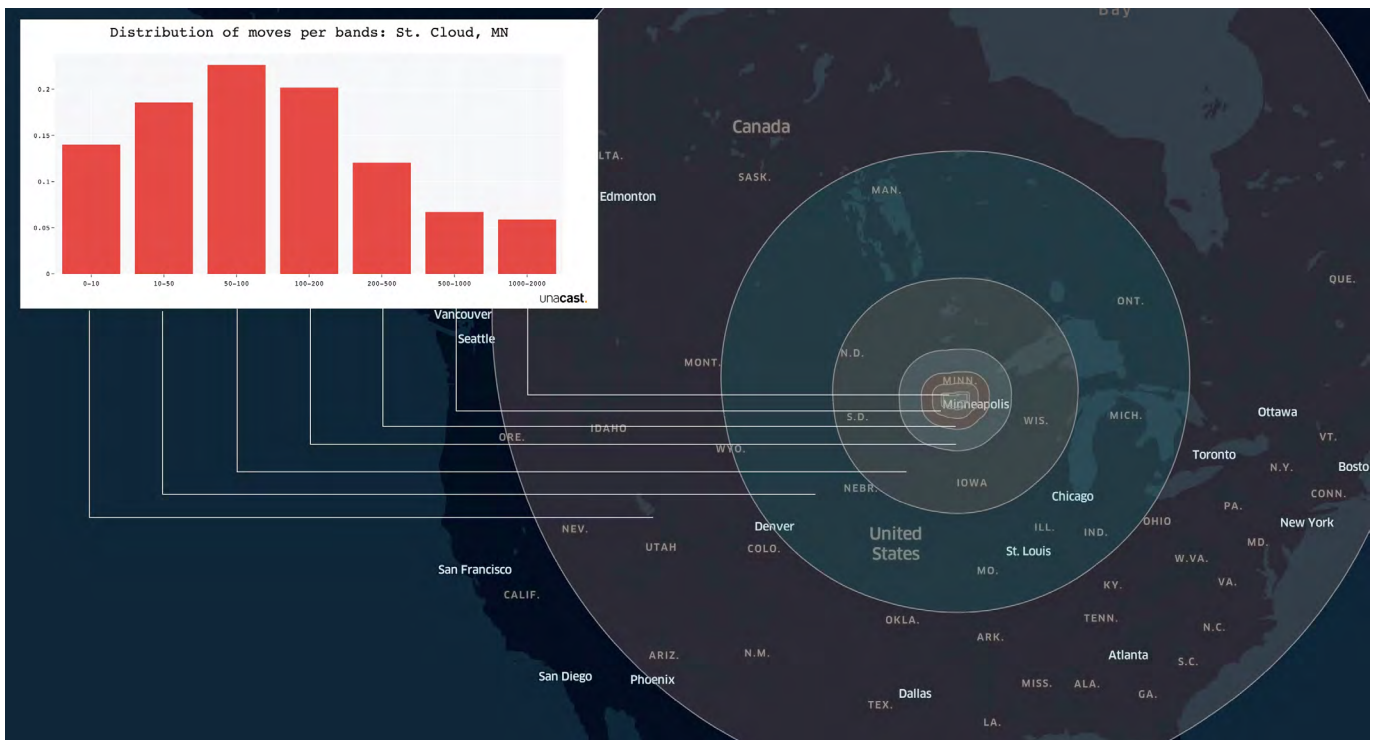
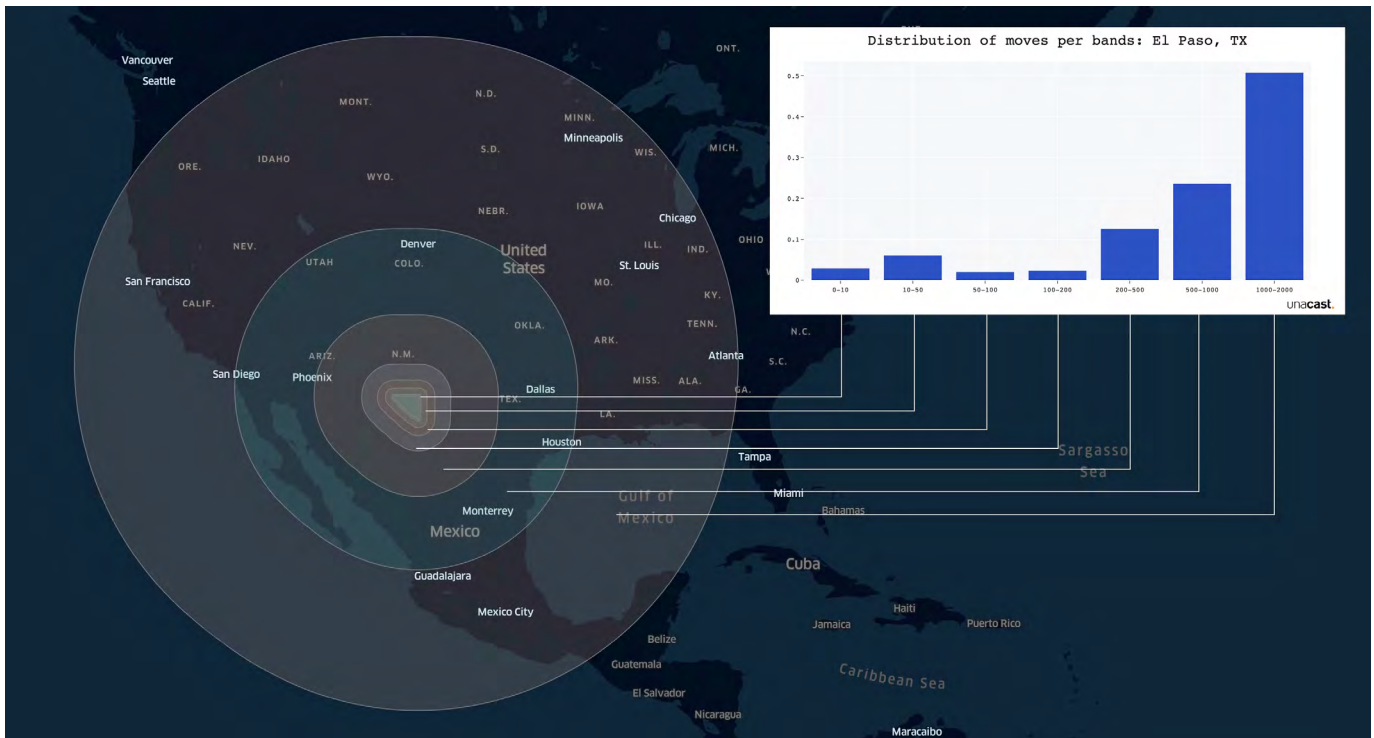
Distribution of moves per bands: All MSAs



The distribution in Fig. 2 aggregates distances from all MSAs in the United States. It would be naive to assume that human migration follows the same pattern in all regions of the country. Therefore, we ask a second question

RQ2: Is the distribution of outflow distances similar for all MSAs or is it region-specific?

In order to answer RQ2, we first calculate the distance distribution for each MSA individually. And, indeed, we see that MSAs differ from the perspective of move distance distribution. Fig. 3 offers a visual comparison between **St. Cloud MSA in Minnesota**, where most of the moves are targeting rather nearby locations, and **El Paso in Texas**, where the most common distance band is the longest one, i.e., 1000-2000 Km.



The difference between the two locations above invites for definition of classes that would describe the nature of MSAs from the outflow distance point of view. Can we classify MSAs into those from which people tend to move far away and those with rather nearby outflow destinations?

Classification of MSAs by distance of moves

To classify MSAs based on the distance of moves we establish a scoring metric. The distance score is a weighted average of relocation distances, where the strength of the flow is used as the weight and the destination band's mean distance represents the value.

The distance score is defined as:

$$\text{score}(\text{msa}) = \frac{1}{n} \sum_{i=1}^n w_i * \text{avg_dist}(b_i)$$

, where w_i is the normalized migration flow of the i -th distance band, $\text{avg_dist}(b_i)$ is the mean relocation distance of the i -th distance band in Km, and n is the total number of distance bands.

Example:

Let's take a look at score calculation for **St. Cloud, MN** MSA. First step is to calculate average distance of moves within each buffer. We get:

- 0-10 km (3390 moves) → 5
- 10-50 km (4500 moves) → 30
- 50-100 km (5485 moves) → 75
- 100-200 km (4890 moves) → 150
- 200-500 km (2920 moves) → 350
- 500-1000 km (1625 moves) → 750
- 1000-2000 km (1425 moves) → 1500

Now, when we know amount of moves to each buffer and approximate distance of these moves, we are ready to make calculation according to the formula described previously.

$(5 * 3390 + 30 * 4500 + 75 * 5485 + 150 * 4890 + 350 * 2920 + 750 * 1625 + 1500 * 1425) / (3390 + 4500 + 5485 + 4890 + 2920 + 1625 + 1425)$

$(16950 + 135000 + 411375 + 733500 + 1022000 + 1218750 + 2137500) / 24235$

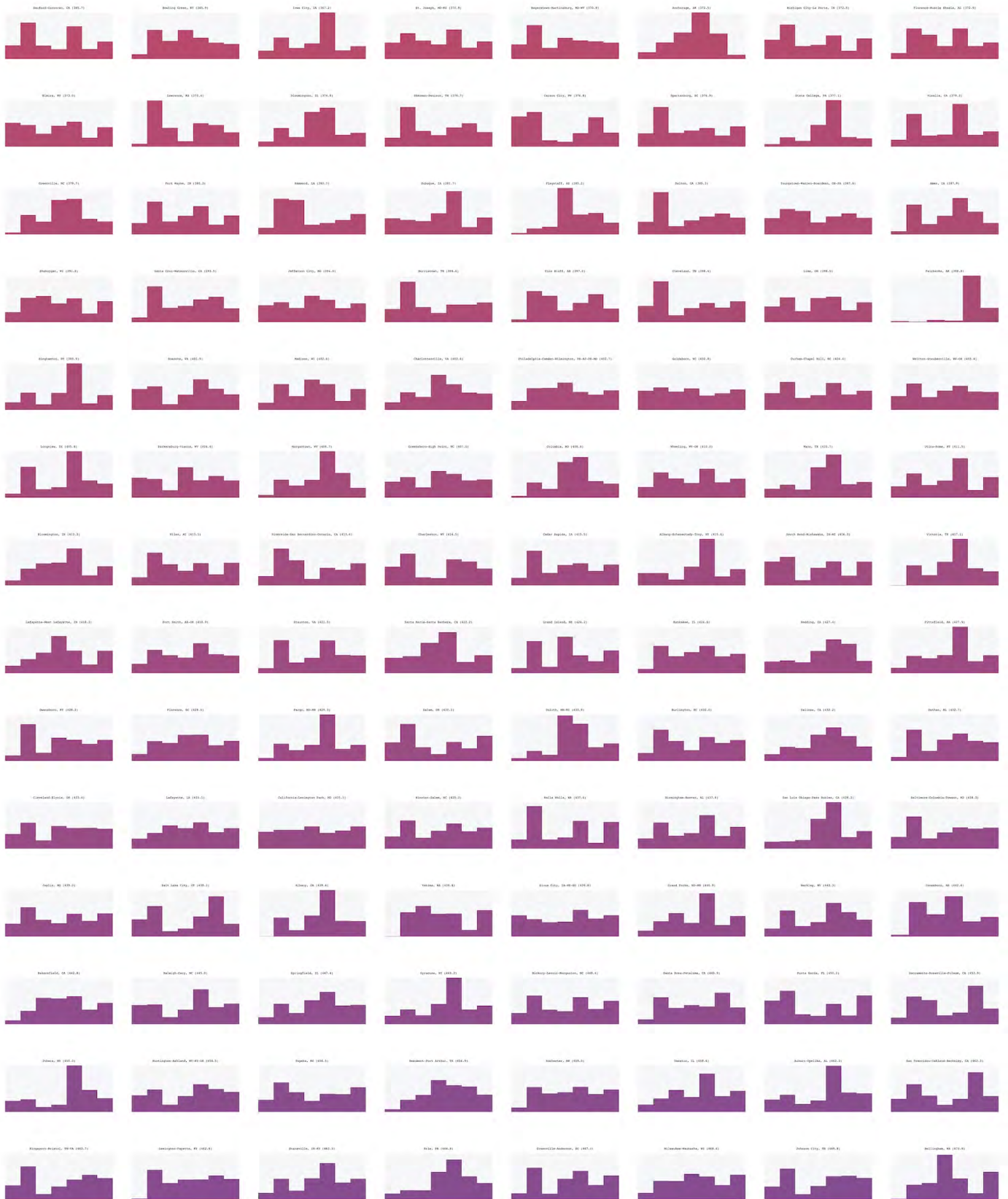
$5675075 / 24235$

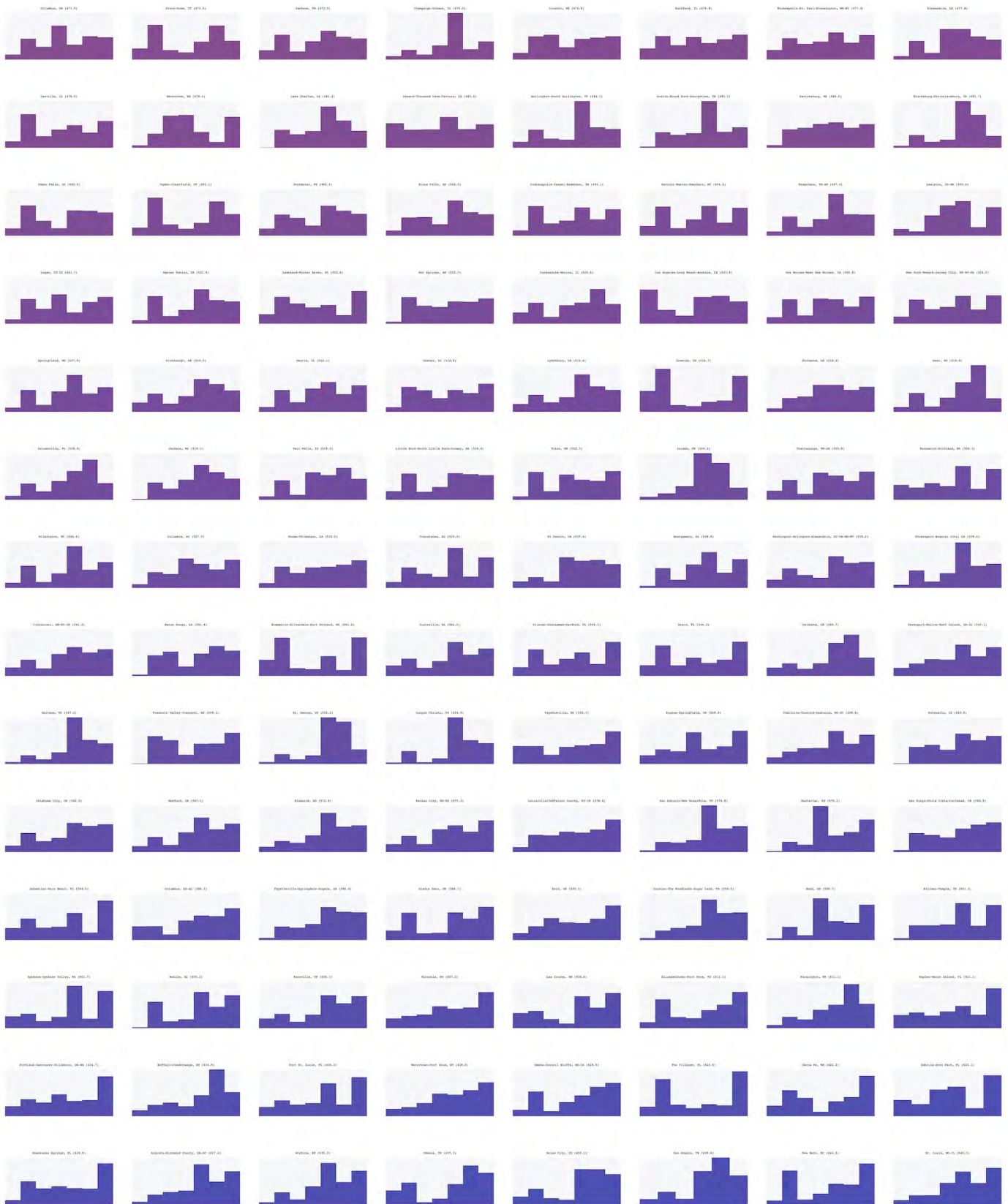
234,168558

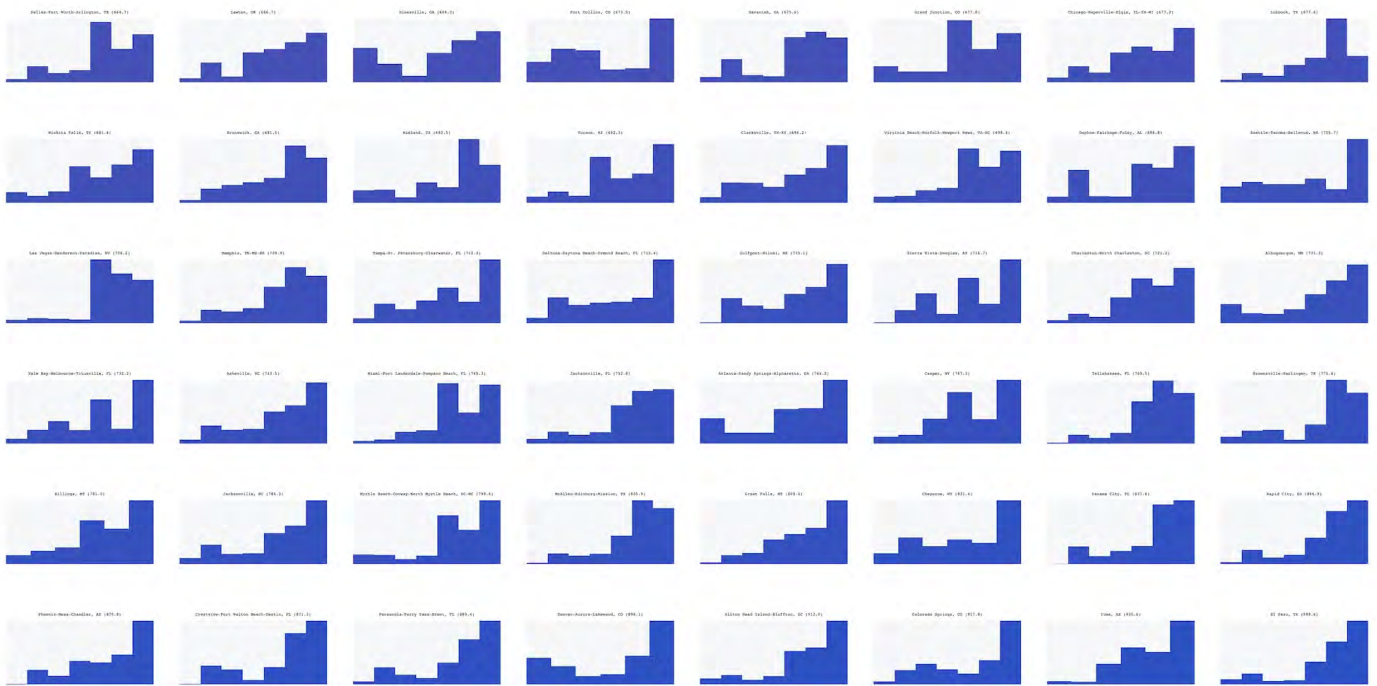
Ranking of MSAs

The single-value score that summarizes the distance distribution is a perfect attribute for sorting MSAs by their distance of relocations. Below we show an overview of MSAs ranked by the score in ascending order, that is, from short outflow distance to long outflow distance MSAs.



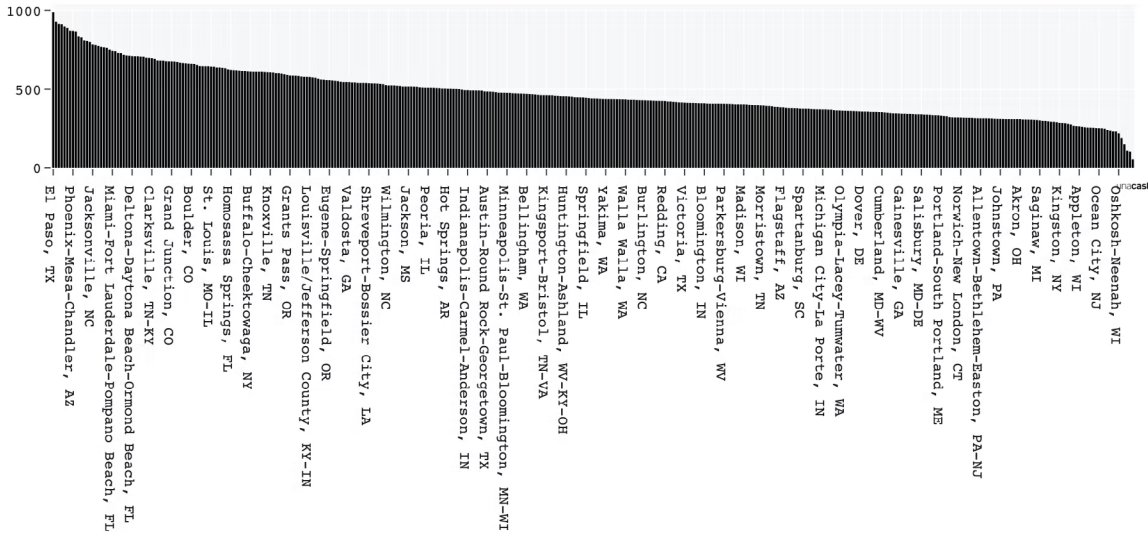






In the visualization above we can see extremes on both ends - majority of moves to the closest bands at the beginning (red plots) versus MSAs with almost all moves towards distant locations at the end (blue plots). A simple way to classify the MSAs by the nature of outflow moves is to define classes based on the score. We label the top 25% of MSAs with the highest score as **long-distance-outflow** MSAs (marked with red colour), bottom 25% as **short-distance-outflow** MSAs (marked with blue colour), and 50% in the middle as **medium-distance-outflow** MSAs (marked with gray colour).

Distribution of distance score

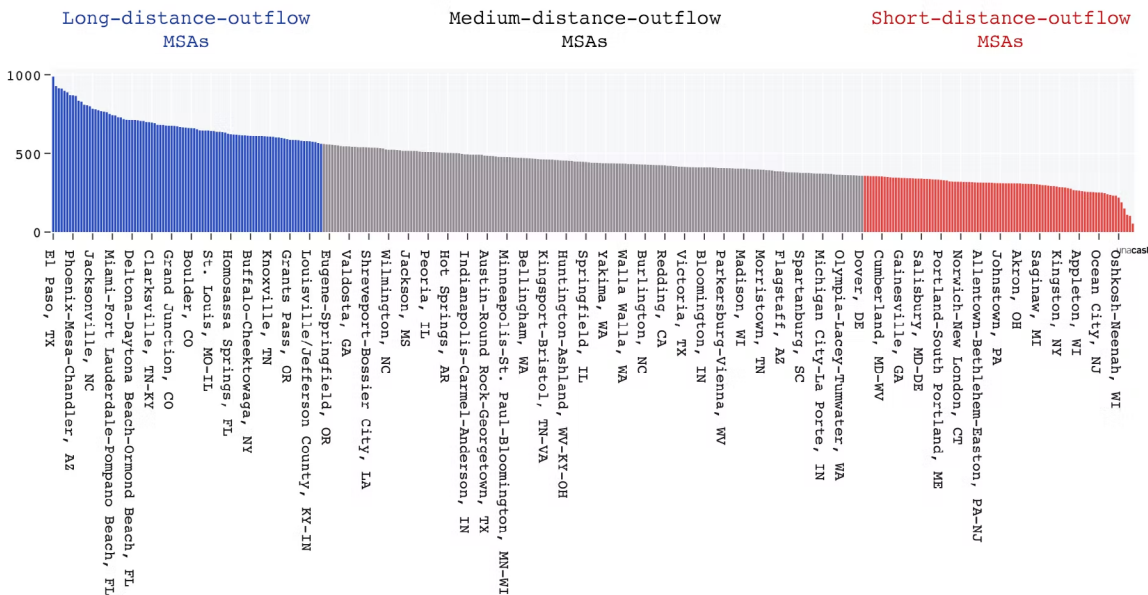


Classification

(25%, 50%, 25%)



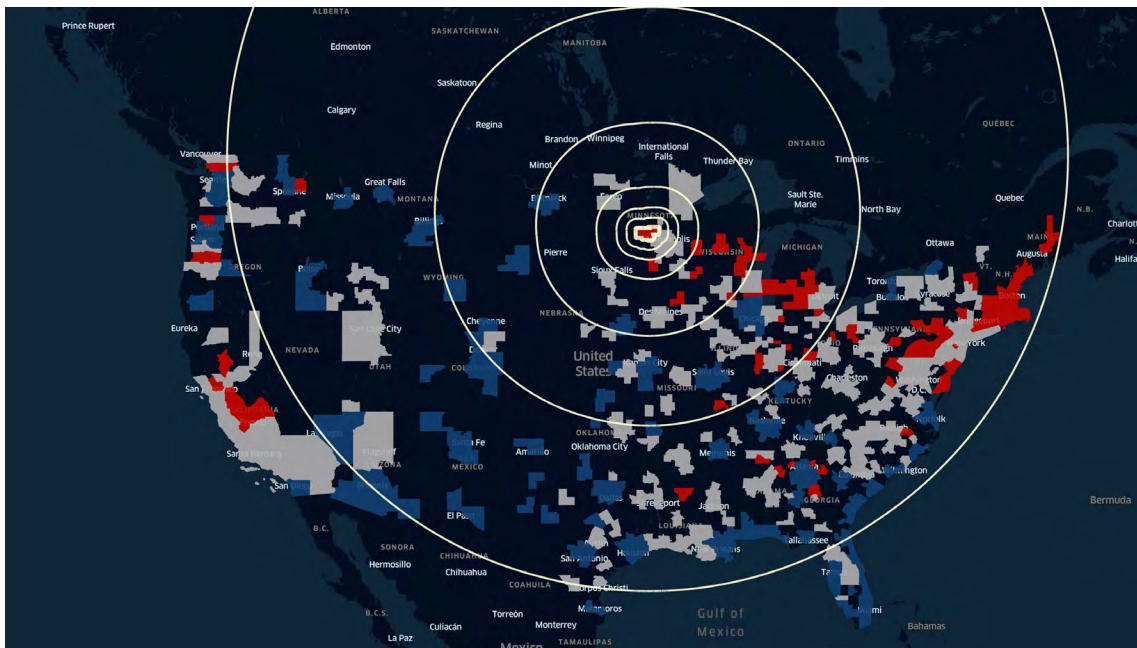
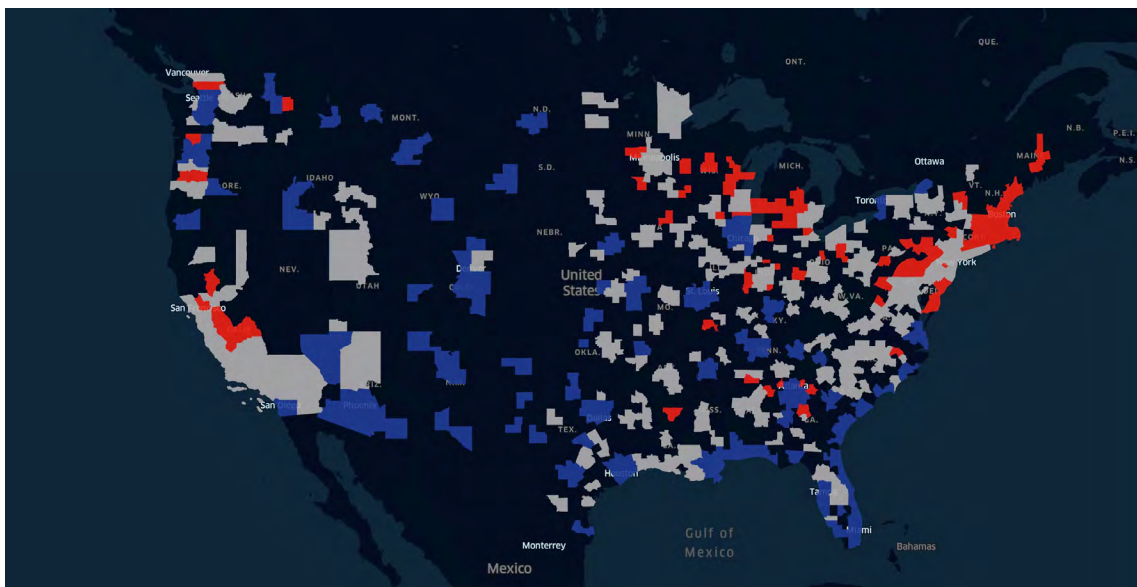
Distribution of distance score



Using the corresponding colours (red=short distance, blue=long distance, gray=medium distance), the map of MSAs in the US reveals an uneven distribution of these classes across the country.

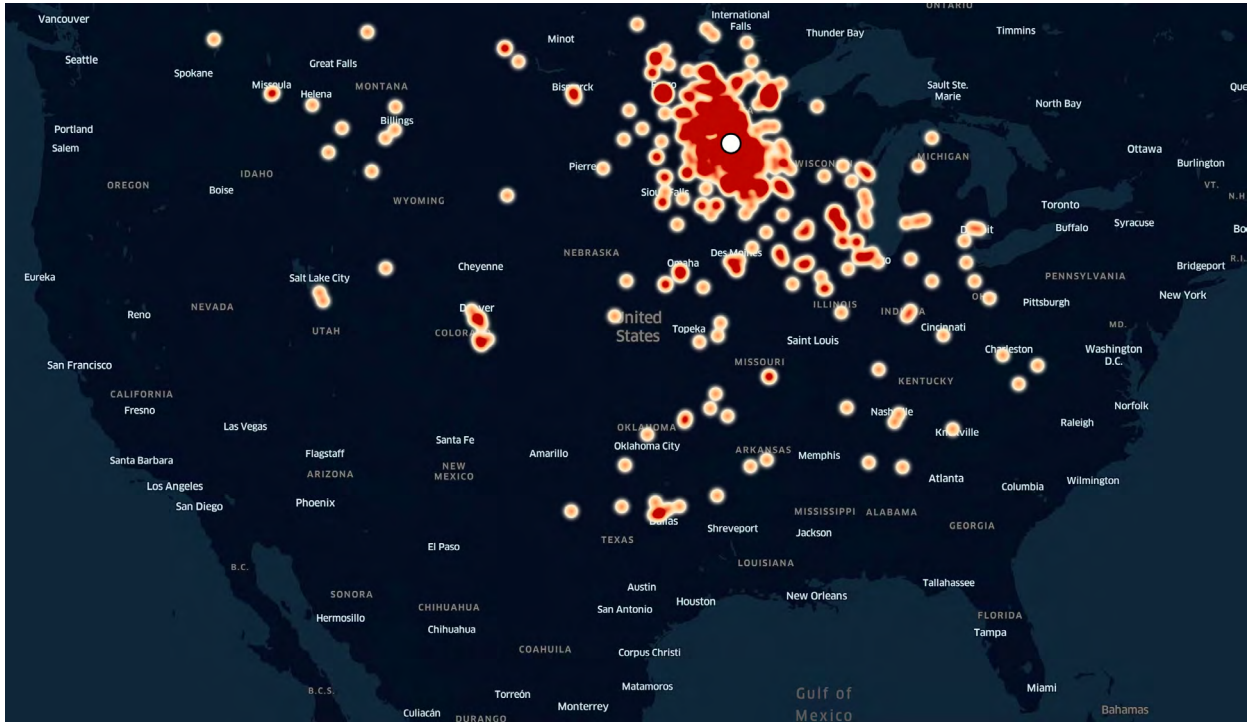
- People from smaller MSAs usually move to big MSAs nearby.
It is a well-known trend that people from smaller cities tend to move to a bigger city nearby. This is mostly caused by more attractive work opportunities.
- The Miami MSA does not attract people nearby but other MSAs from top 10 do.
Miami is considered one of the most expensive cities to live.
- There are some clusters of MSAs where people tend to move within the cluster (New York, San Francisco).

This huge cluster on the northeast is the world’s largest megapolis in terms of economic output.



Revisiting our two examples

One of the key factors that can make differences in terms of outflow range is location. In this example we are comparing St. Cloud - small city close to Minneapolis with population under 70000 people, what shows that most of people are tempted to move to a bigger neighbour nearby; and El Paso what is located right on the border with Mexico in Texas and doesn't have much cities around.



Future work

- Study moves from the most populated “core” of the MSA to the rest of the MSA
- Dynamic view - studying the changing character of the distance over time
- Shifts between counties of various kinds (as in [here](#))
- Follow-up: more detailed study common properties of MSAs within each group (location, population etc.).
- Is there any correlations between range of moves and other metrics (e.g. number of businesses).

Literature

- [Are Millennials leaving town? Reconciling peak Millennials and youthification hypotheses / article](#)



Visit unacast.com to learn more