Leveraging Geolocation Data to Analyze and Optimize Distributed Workforce Strategies

Sidney T. Thomas, Ph.D. – Senior People Analytics Consultant, Allstate

Ajay V. Somaraju, Ph.D. – Lead People Analytics Consultant, Allstate

Lauren M. DeYoung – Workplace Futurist, Allstate

Allstate

People Analytics

DATA THAT EMPOWERS AND DRIVES CHANGE

SIOP 2025 Annual Conference

Denver, Colorado

Date Presented:

April 3, 2025

Why Geolocation Data for Workforce Analytics?

Geolocation data refers to *location-based information* that identifies where employees work, how they move between locations, and how they interact with physical and virtual workspaces.

How can geolocation data inform organizational strategy?

Workforce Distribution – Identify where employees are working and optimize talent allocation.

Collaboration Insights – Analyze movement and interaction patterns to improve team connectivity.

Space Utilization – Determine which offices, hubs, or coworking spaces are in high demand.

Employee Experience – Align workspace strategies with employee preferences and behaviors.

Analytical Process

Identification of workforce problem

E.g., Where is everyone?? Who is going into the office? Are they more engaged?

01

Data collection

- *Workplace Entry Logs*
- *Employee Surveys*
- *Device-Based Tracking (Opt-In Only)*
- *Office & Workspace Reservations*

02

Data Cleaning

- *Remove missing or inaccurate location points*
- *Standardize location formats*

03

Data integration

- *HR employee metrics (Job family, salary)*
- *Engagement*
- *Performance*
- *Collaboration (Teams, Outlook, Recognition)*

04

Data analysis

- *Descriptives (e.g., counts by location)*
- *Predictive (E.g., proximity to an office X engagement)*

05

Data visualization

- *Heatmaps*
- *Cluster maps*
- *ONA*
- *Supplemental graphs/charts (e.g., bar, line, Sankey)*

06

Data insights

- *Identify relevant stuff*
- *Communication of results*

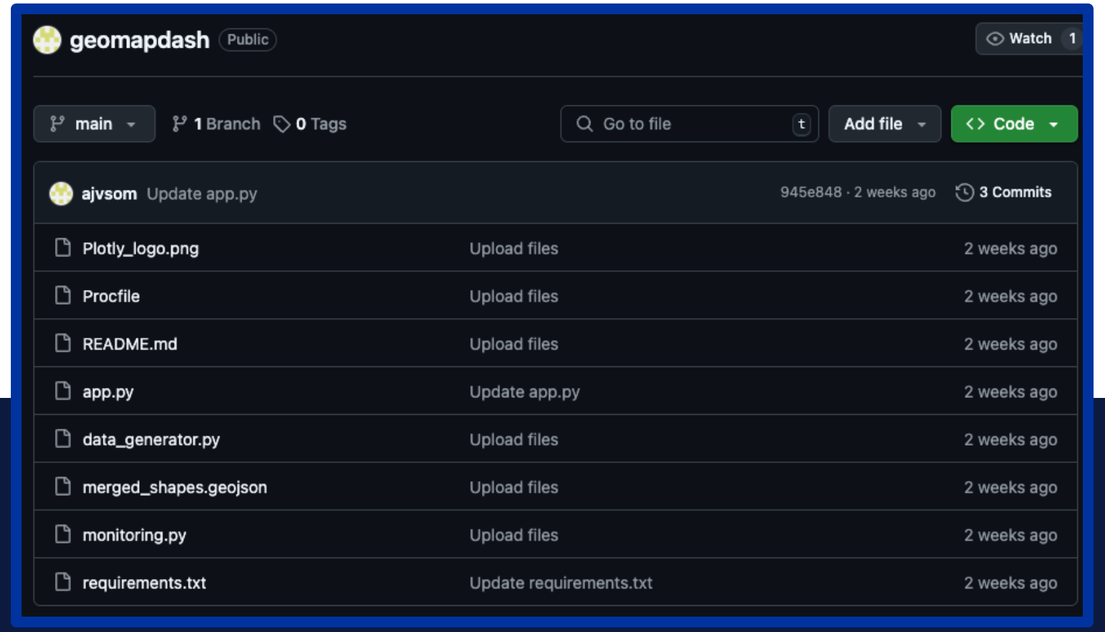
Application of insights to workforce strategy

Everyone is in XYZ metro area and they're actually not near any office, so let's put office options there.

Python Tutorial

Geolocation Dash App

Repo: https://github.com/ajvsom/geomapdash.git

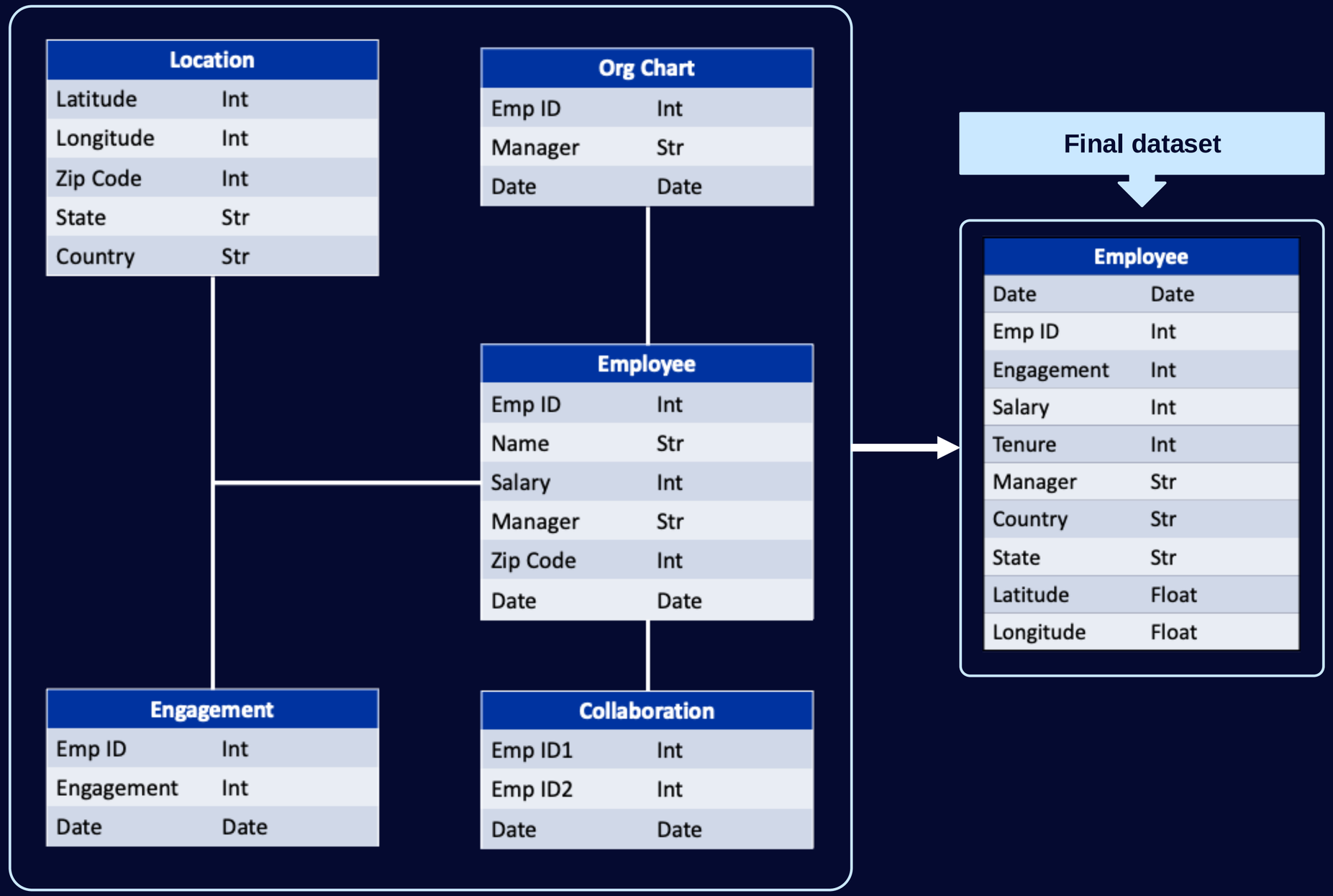


Data collection & structure

Simulated dataset

- Columns:
  - Date
  - Employee ID
  - Engagement Score
  - Salary
  - Reporting Manager
  - Tenure Range
  - Country
  - State
  - Region
  - Latitude
  - Longitude

Example Python Script:
data_generator.py



Data cleaning & integration

Cleaning Data

- Remove missing or inaccurate location points
- Standardize formats (e.g., longitude/latitude, city names)
  - For heatmaps, you will need a *shapefile* (see the Github repo)
  - For pin dropped locations, you can use latitude and longitude using census data (these will become outdated quickly)
  - Using lat/long for center of zip code as to not give away employee home locations

Recommended packages:
Python: *geopy*
R: *tidygeocoder*

Examples of common employee location data errors:

Employee ID	Name	Zip Code	City	State	*Notes*
12345678	John Doe	90210	Bevrly Hils	CA	*Misspelled city name*
13579246	Alicia White	1001	Hampden	MA	*ZIP code too short*
24681357	Omar Patel	10001	NYC	NY	*Abbreviated city name*
11223344	Maria Lopez	77001	Houston	TXS	*Invalid state abbreviation*
99887766	Sam Lee		Austin	TX	*Missing ZIP code*
44332211	Angela Wu	60601	Chicago	IL	*Accurate entry*
33221100	Leah Brooks	30303-9999	Atlanta	GA	*Unusual ZIP extension*
10101010	Emily Johnson	ABCDE	Los Angeles	CA	*Non-numeric ZIP code*
44086000	Carlos Ramos				*Missing address details*

Example of integration with latitude/longitude data

Employee ID	Name	Zip Code
12345678	John Doe	90210
13579246	Alicia White	01001

Latitude	Longitude
34.073620	-118.400352
42.06262	-72.62521

Data Analysis & Visuals

Dash from Python

- Contains three main parts
  1. Initialize the app
  2. Define the Layout
  3. Connect the visual components in the user interface to the input components

Advantages over Enterprise Tools

- Easy version control
- More native analytics capabilities (e.g., showing regression or ML models)
- Integration w/ AI tools like Cursor
  - Direct and detailed prompts
  - Positive feedback for accurate representation
  - Restarting when running into loops

Example Python Script:

app.py

```
# Import packages
from dash import Dash, html, dash_table, dcc, callback, Output, Input
import pandas as pd
import plotly.express as px
import dash_design_kit as ddk

# Incorporate data <-- can be direct csv upload or connected to server tables
```

```
# Initialize the app
app = Dash()

# App layout <--- similar to UI in R Shiny
app.layout = ddk.App([ # <--- this defines the different buckets
    ddk.Header(ddk.Title('My First App with Data, Graph, and Controls')),
    dcc.RadioItems(options=['pop', 'lifeExp', 'gdpPercap'],
                   value='lifeExp',
                   inline=True,
                   id='my-ddk-radio-items-final'),
    ddk.Row([
        ddk.Card([
            dash_table.DataTable(data=df.to_dict('records'), page_size=12,
                  style_table={'overflowX': 'auto'})
        ], width=50),
        ddk.Card([
            ddk.Graph(figure={}, id='graph-placeholder-ddk-final')
        ], width=50),
    ]),

])
```

```
# Add controls to build the interaction <-- similar to Server in R Shiny
@callback( # <-- contains the "id" of the target visual and its property and connects it to the input
component
    Output(component_id='graph-placeholder-ddk-final', component_property='figure'),
    Input(component_id='my-ddk-radio-items-final', component_property='value')
)
def update_graph(col_chosen): # <--- function that updates the visual based on the user action
    fig = px.histogram(df, x='continent', y=col_chosen, histfunc='avg')
    return fig
```

```
# Run the app
if __name__ == '__main__':
    app.run(debug=True)
```

Data Insights & Comms

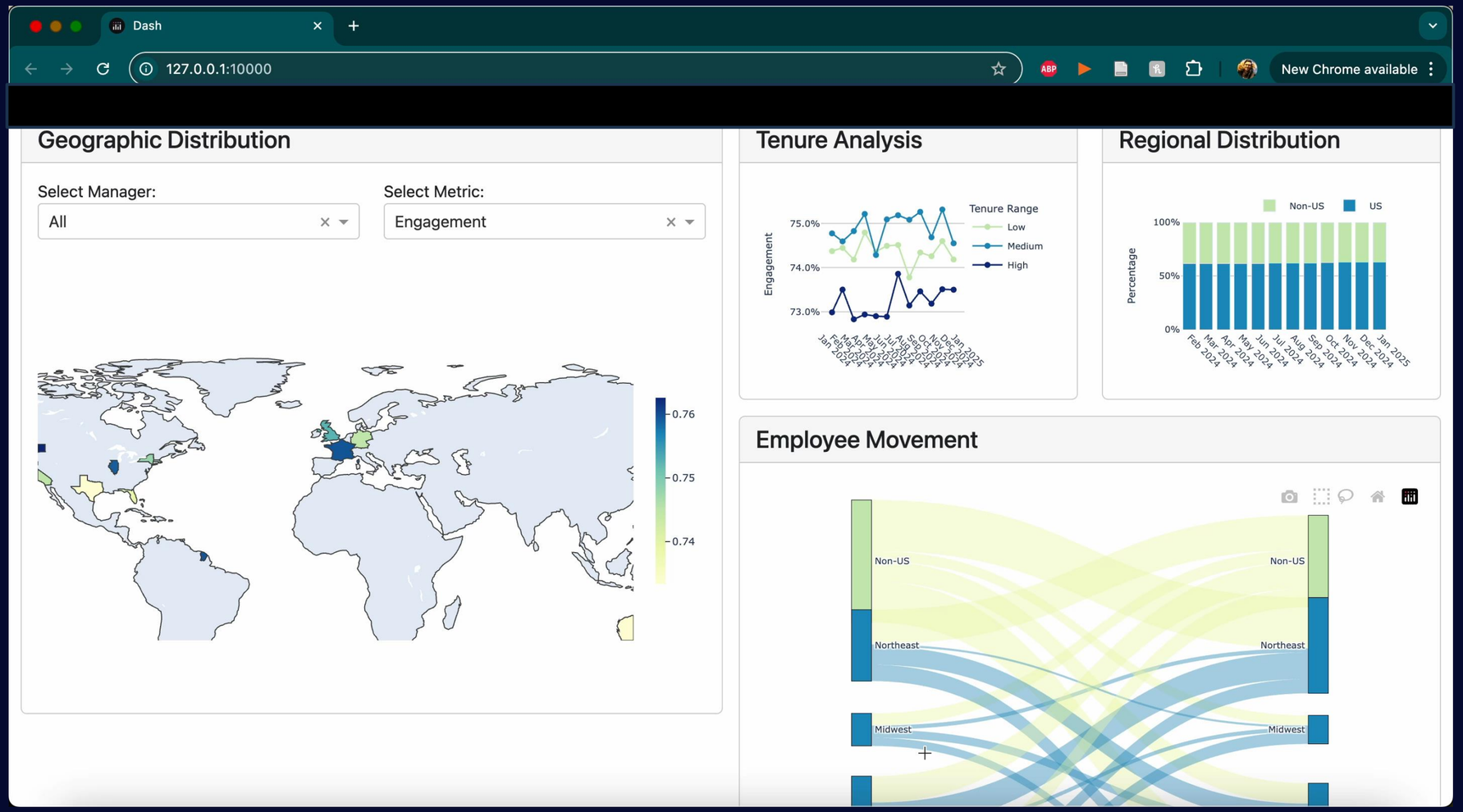
Geographical Maps

- Heat map
- Cluster map

Supplemental graphs/charts

- Bar graph
- Line graph
- Sankey

Example Dash App:



Application of Insights to Workforce Strategy

Geolocation data has been essential in helping us understand where our workforce actually is— not just at the individual level, but in terms of how teams are distributed across functions and their proximity to an office.

For example, we found that most teams were not fully co-located near an office. So rather than trying to bring teams together in person, we shifted to meeting people where they are. This insight led to the introduction of flex pods in high-density areas—flexible workspace options informed directly by location data.

Dashboards built from these insights have supported leadership strategy throughout the journey— from early decision-making to now using the data in execution of our distributed work strategy.

Importantly, this data is not used to monitor individual employees, but to better understand workplace preferences and patterns at scale. We use these insights in conjunction with outcome measures to evaluate what's working and where opportunities remain.

Lauren DeYoung, Workplace Futurist @ Allstate

Lauren leads our employee experience team and spearheads our distributed work strategy at Allstate. She's been a close partner with our People Analytics team to design flexible, location-informed workplace solutions.

Ethical Considerations & Limitations

Ethical Considerations in Using Geolocation Data

- **Employee Privacy & Consent**
  - Employees should be **informed** about how their location data is collected and used.
  - Transparent policies help build **trust and compliance**.
- **Data Anonymization & Security**
  - Use **aggregated or de-identified data** to prevent tracking individual employees.
  - Implement **secure storage** and restrict access to sensitive workforce data.
- **Legal Compliance & Organizational Policies**
  - Align with **U.S. privacy laws (e.g., CCPA)** and internal corporate policies.
  - Ensure that **geolocation tracking is voluntary and opt-in**.

Limitations of Geolocation Data

- **Incomplete or Misleading Data**
  - Not all employees opt into sharing location data, creating potential biases.
  - Device-based tracking may **lack precision or be affected by VPNs/IP masking**.
- **Correlation ≠ Causation**
  - Just because employees **work from a certain location** doesn’t mean it **drives productivity or engagement**.
  - Use geolocation insights alongside **qualitative data** to make informed decisions.
- **Potential Employee Resistance**
  - Misuse or lack of transparency can **erode trust** and impact morale.
  - Organizations should focus on **how data benefits employees**, not just leadership.

Questions?

NAME	EMAIL
Sidney Thomas	sidtaytho@gmail.com
Ajay Somaraju	ajay.somaraju@gmail.com

People Analytics

DATA THAT EMPOWERS AND DRIVES CHANGE