

World-POI: Global Point-of-Interest Data Enriched from Foursquare and OpenStreetMap as Tabular and Graph Data

Hossein Amiri^{1,*}, Mohammad Hashemi¹, and Andreas Züfle^{1,*}

¹Department of Computer Science and Informatics, Emory University Atlanta, USA

*Corresponding authors: Hossein Amiri (hossein.amiri@emory.edu) and Andreas Züfle (azufle@emory.edu)

ABSTRACT

Recently, Foursquare released a global dataset with more than 100 million points of interest (POIs), each representing a real-world business on its platform. However, many entries lack complete metadata such as addresses or categories, and some correspond to non-existent or fictional locations. In contrast, OpenStreetMap (OSM) offers a rich, user-contributed POI dataset with detailed and frequently updated metadata, though it does not formally verify whether a POI represents an actual business. In this data paper, we present a methodology that integrates the strengths of both datasets: Foursquare as a comprehensive baseline of commercial POIs and OSM as a source of enriched metadata. The combined dataset totals approximately 1 TB. While this full version is not publicly released, we provide filtered releases with adjustable thresholds that reduce storage needs and make the data practical to download and use across domains. We also provide step-by-step instructions to reproduce the full 631 GB build. Record linkage is achieved by computing name similarity scores and spatial distances between Foursquare and OSM POIs. These measures identify and retain high-confidence matches that correspond to real businesses in Foursquare, have representations in OSM, and show strong name similarity. Finally, we use this filtered dataset to construct a graph-based representation of POIs enriched with attributes from both sources, enabling advanced spatial analyses and a range of downstream applications.

Background & Summary

Points of Interest (POI) datasets are foundational for a wide range of spatial computing applications, including urban analytics^{1,2}, mobility modeling and simulation^{3,4}, place recommendation^{5–8}, anomaly detection^{9–11}, modeling infectious disease spread¹², and enabling location-based services¹³. As these applications continue to evolve, the demand for comprehensive high-quality POI datasets has intensified. However, existing POI datasets often exhibit trade-offs between spatial coverage, semantic richness, and cost, limiting their utility for certain tasks.

There are several publicly available POI datasets, ranging from commercial to open-source options. Prominent paid sources include the Google Places API¹⁴, ADVAN¹⁵, Bing Maps¹⁶, HERE Maps¹⁷, and Precisely¹⁸. These services typically offer extensive global coverage, rich metadata, and verified business listings, but come with licensing fees or usage-based pricing models. For example, the Google Places API provides access to over 200 million POIs but incurs substantial costs at scale. In contrast, open-source and freely available datasets such as Foursquare¹⁹, OpenStreetMap (OSM)²⁰, the USGS Geographic Names Information System (GNIS)²¹, and Overture Maps²² offer cost-free access, though with varying levels of completeness, consistency, and quality. Among these, Foursquare and OSM are the focus of this work, and we validate our integrated dataset using Google Maps as an external reference.

Commercial datasets such as Foursquare generally provide high-precision, verified business entries and up-to-date information. However, they often lack consistent or complete metadata, including standardized category taxonomies, administrative boundaries, or contextual geographic attributes¹⁹. Furthermore, user-generated contributions may introduce inaccuracies or erroneous points of interest (POIs) into the database. In contrast, OSM offers a comprehensive, community-driven platform with rich semantic annotations, flexible tagging schemes, and broad spatial coverage. However, its crowd-sourced nature means that business verification is informal, and metadata can sometimes be inconsistent, incomplete, or outdated²⁰.

In this paper, we address the complementary limitations of these sources by introducing **World-POI**, an enriched and integrated POI dataset that aligns and merges records from Foursquare and OSM. Using a hybrid matching approach based on spatial proximity and name similarity, we construct a high-confidence mapping of POI pairs that combines the precision and reliability of Foursquare with the semantic depth and geographic richness of OSM. The resulting dataset provides detailed metadata, standardized category labels, and accurate spatial annotations. An

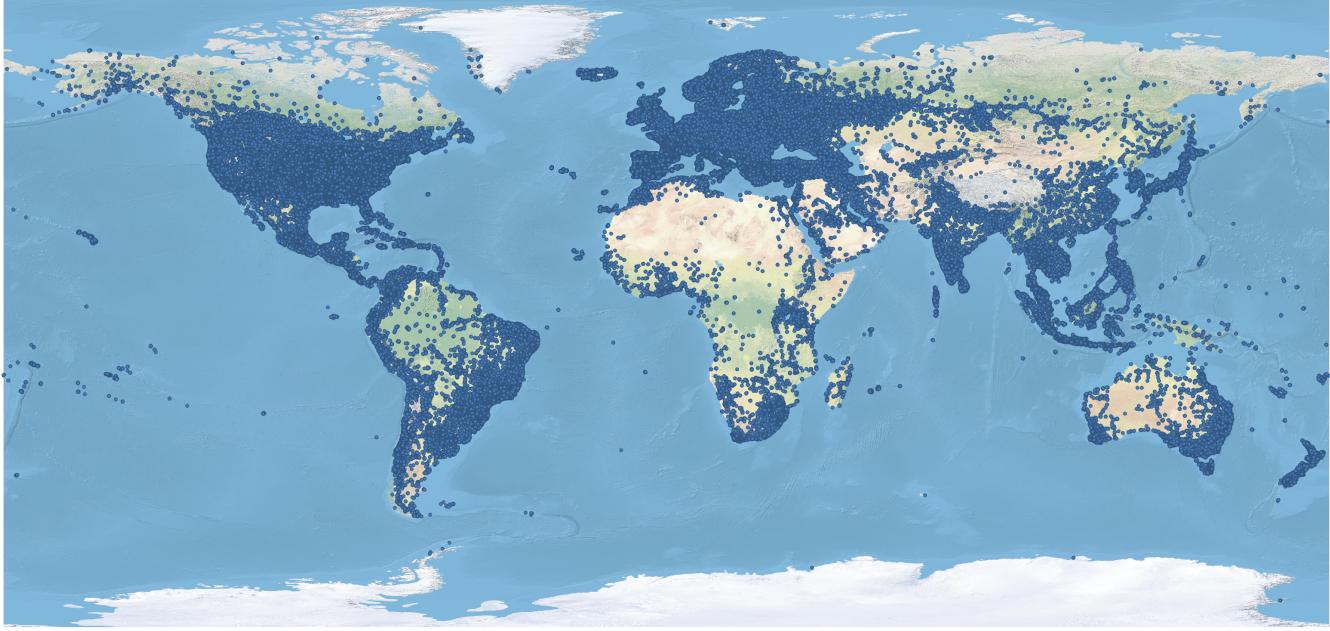


Figure 1. Visualization of a representative sample output from the World-POI database, filtered to include entries with Levenshtein name similarity scores greater than 0.5. The map was generated in QGIS using the ESRI Physical Map basemap layer.

example of a representative sample from the World-POI database is shown in Figure 1, illustrating entries filtered by Levenshtein name similarity scores greater than 0.5. In other words, if a POI appears in both Foursquare and OSM, shares similar spatial attributes, and exhibits high name similarity, it is more likely to represent the same real-world location. Moreover, POIs present in both datasets are more likely to correspond to actual businesses rather than user-generated or fictitious entries added to gain contribution points within the platform.

World-POI is released in both tabular and graph-based formats to support diverse analytical workflows, such as geographic knowledge graph construction, spatial clustering, and simulation-based modeling. By bridging the strengths of two widely used POI sources, World-POI enables more accurate modeling, analysis, and simulation of real-world spatial environments. This dataset serves as a high-utility resource for researchers and practitioners in urban computing, geoinformatics, and spatial data science.

Methods

In this section, we outline the methodology used to construct the World-POI dataset, ensuring both transparency and reproducibility. The complete workflow is illustrated in Figure 2. In summary: (1) collect POI data from Foursquare and OpenStreetMap (OSM); (2) perform data cleaning and initial preprocessing on the Foursquare dataset; (3) import the cleaned Foursquare data and downloaded OSM data into PostgreSQL/PostGIS; (4) conduct additional preprocessing to harmonize attributes and resolve inconsistencies; (5) create spatial indexes; (6) perform spatial joins to identify candidate POI pairs based on geographic proximity; (7) compute name-similarity scores between Foursquare and OSM entries; (8) apply similarity thresholds to retain high-confidence matches; (9) generate tabular outputs; and (10) construct corresponding graph-based representations.

Step (1) — Data Collection

To obtain the Foursquare data, we accessed Foursquare’s publicly available cloud storage (hosted on Amazon S3) and downloaded a comprehensive collection of place-related datasets. These datasets were originally provided in split, compressed Parquet format and organized by category information and place data to facilitate access.

For OpenStreetMap (OSM) data, we downloaded the full dataset directly from the official OSM website. To make the data suitable for analysis and querying, we utilized Nominatim, a geocoding and data conversion tool, to import the raw OSM data into a PostgreSQL-compatible SQL format. This process generated multiple relational tables, among which the places table—containing detailed information about points of interest (POIs)—served as the primary input for integration.

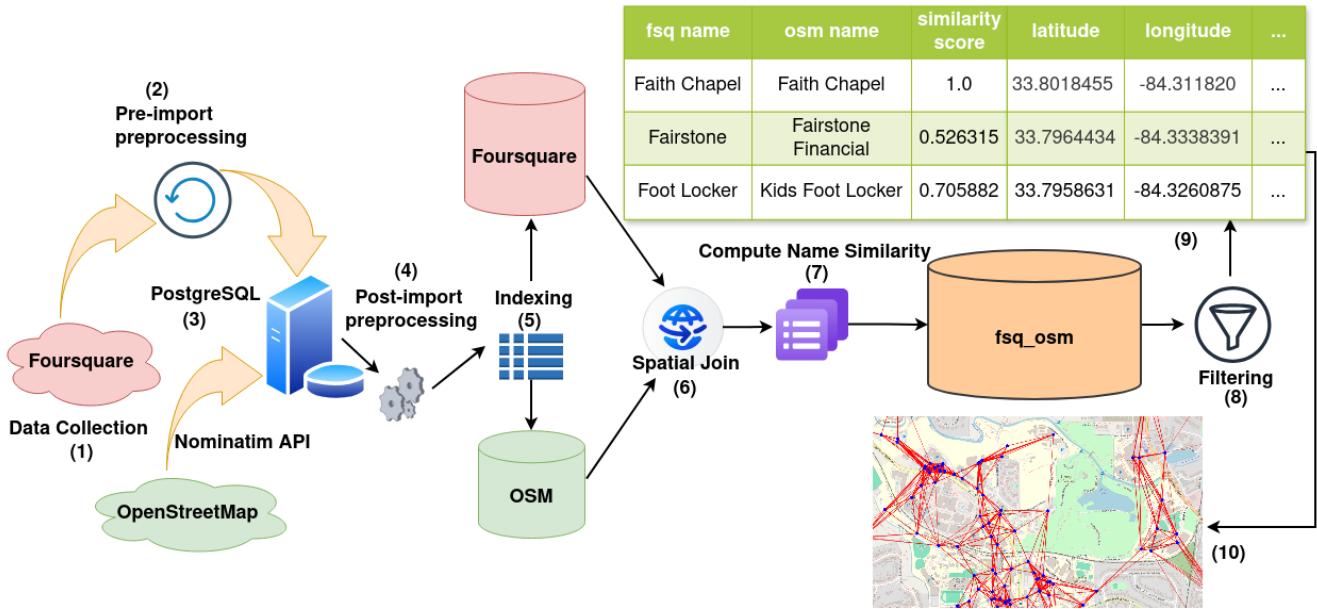


Figure 2. Overview of the pipeline for constructing the World-POI dataset by merging Foursquare and OpenStreetMap POIs through preprocessing, spatial indexing, and similarity-based filtering. The map was generated in Python and the source code is on GitHub

Step (2) — Pre-import Preprocessing

After downloading the Foursquare data, the Parquet files were converted into a more accessible and widely supported format (CSV). All converted files were then aggregated into a single unified dataset. During this process, geometry values were removed, numerical fields were converted to numeric types, and string attributes were enclosed in quotation marks (“”) to ensure proper parsing. This cleaning step ensured consistent formatting and completeness prior to database ingestion. In addition, columns were renamed to maintain a consistent naming convention by prefixing each with “fsq_” if it did not already begin with that prefix. These steps ensured that the dataset was fully standardized and ready for structured ingestion into PostgreSQL/PostGIS.

Step (3) — Data Import into PostgreSQL/PostGIS

The cleaned Foursquare dataset was imported into a PostgreSQL database with the PostGIS extension enabled to support spatial operations. A new database, `fsq-osm`, was created to store both the Foursquare and OpenStreetMap (OSM) datasets in structured form. For the Foursquare data, a dedicated table (`foursquare`) was created following the schema defined in our GitHub repository²³. The data were loaded using the `\copy` command, which efficiently imports large CSV files while preserving headers, delimiters, and quoted strings. The OSM data, downloaded in PBF format, were first processed using the Nominatim import tool to convert the raw OSM file into a PostgreSQL-compatible database.

Step (4) — Post-import Preprocessing

Following the data import, additional preprocessing was conducted to harmonize attributes and resolve inconsistencies between the Foursquare and OSM schemas. This included aligning coordinate systems, validating geometries, and removing duplicate or corrupted records. From the imported OSM database, we used the place table and only features containing valid `name` attributes and not classified as `highways` were extracted and exported as a filtered CSV file. During this step, all column names were prefixed with `osm_` to maintain a consistent naming convention. In addition, the centroid of each original geometry was calculated to obtain representative latitude and longitude coordinates. The filtered data were subsequently exported to a CSV file and imported into the `osm` table within the `fsq-osm` database using the `\copy` command. Then, Geometry columns (`fsq_geom` and `osm_geom`) were generated for both `foursquare` and `osm` tables using PostGIS functions that convert longitude and latitude values into point geometries, standardized under the WGS 84 coordinate reference system (EPSG:4326) to ensure consistency.

Step (5) — Spatial Indexing

To optimize performance for spatial queries and joins, GiST (Generalized Search Tree) indexes were created on the geometry columns of both datasets (`fsq_geom` and `osm_geom`). These spatial indexes substantially improved query efficiency, reducing computation time during spatial matching and distance calculations.

Step (6) — Spatial Join Based on Geographic Proximity

To align and integrate the two data sources, a spatial join was performed between Foursquare POIs (`foursquare` table) and OSM POIs (`osm` table) based on geographic proximity. For each Foursquare record, the nearest OSM feature within a 50-meter radius was identified. This threshold was chosen to balance precision and recall—accurately linking spatially corresponding entries while minimizing false matches and to limit the output size to a manageable volume (approximately 1 TB) for storage and further processing. In this step, the geographic distance between each pair was calculated and stored as a column in degrees. The resulting integrated dataset combined the commercial metadata of Foursquare with the semantic richness of OSM, enabling comprehensive spatial and contextual analysis.

Step (7) — Name Similarity Computation

For each spatially matched POI pair, name similarity between Foursquare and OSM entries was computed using two complementary text-matching methods: trigram-based similarity and the Levenshtein²⁴ distance metric. These measures quantified the textual closeness of place names, complementing spatial proximity as an additional criterion for validating matches. The resulting similarity values were stored in two separate columns—one for trigram-based scores and one for Levenshtein-based scores—within the integrated dataset to facilitate downstream filtering and confidence assessment.

Step (8) — High-Confidence Match Selection

To enhance dataset quality, only high-confidence matches were retained based on a Levenshtein name-similarity threshold of 0.5. Pairs meeting or exceeding this threshold were considered reliable alignments between Foursquare and OSM entries. The resulting dataset, stored as `fsq_osm_filtered_5_lev`, contains harmonized POI records with unique `place_id` identifiers from Foursquare and corresponding `osm_id` values from OSM, forming a consistent and high-quality integrated dataset.

Step (9) — Tabular Dataset Generation

After filtering, the integrated dataset was further refined for tabular export. Each record includes validated spatial coordinates, standardized metadata, name-similarity scores, and computed geographic distances. In addition, we assign a unique `poi_id` as the primary key to the filtered dataset to handle potential duplicates of `fsq_place_id`. This situation can occur when a single Foursquare place ID has multiple high-similarity matches in nearby areas—for example, Emory University, Emory University Hospital, and Emory University Library. The finalized dataset was exported in CSV format to ensure broad compatibility with analytical and visualization tools. This structured output serves as the foundation for subsequent geospatial and semantic analyses. The resulting data are illustrated in Figure 1.

Step (10) — Graph-Based Representation

In addition to the tabular dataset, a graph-based representation of World-POI was generated. Each POI was modeled as a node connected to its N nearest neighbors based on geographic proximity. Edge weights correspond to the spatial distances between nodes, capturing local spatial relationships and neighborhood structure. This graph representation facilitates network-based analyses of urban topology, connectivity, and spatial clustering. An example of the resulting graph for $N = 10$ in the vicinity of Emory University is shown in Figure 3.

Data Records

In this section, we provide a detailed overview of the records contained in the dataset. The tabular format includes all available data fields, whereas the graph representation contains only the place identifiers and the distances between them. To obtain detailed information about a specific place identifier in the graph, the corresponding record from the tabular dataset should be referenced. In addition, we provide multiple versions of the dataset generated using different similarity thresholds, allowing users to select the one most suitable for their specific applications. However, throughout this paper, examples and analyses are based on the dataset filtered using a Levenshtein-based name-similarity threshold greater than 0.5. A complete list of all generated datasets, along with their corresponding thresholds and download links, is available in the project's GitHub repository (<https://github.com/onspatial/world-poi>) and available for download at (<https://osf.io/p96uf>)

Table 1. Descriptions of columns in the World-POI dataset.

Column Name	Description	Example Value
fsq_place_id	Unique identifier of the place (Foursquare)	4a4a0fd6f964a52087ab1fe3
fsq_name	Name of the place (Foursquare)	Starbucks
fsq_latitude	Latitude coordinate (Foursquare)	33.768107
fsq_longitude	Longitude coordinate (Foursquare)	-84.34941113
fsq_address	Street address (Foursquare)	506 Moreland Ave NE
fsq_locality	Neighborhood or locality (Foursquare)	Birmingham
fsq_region	Region, state, or province (Foursquare)	GA
fsq_postcode	Postal or ZIP code (Foursquare)	30307
fsq_admin_region	Higher-level administrative division (Foursquare)	England
fsq_post_town	Post town (Foursquare)	Birmingham
fsq_po_box	PO Box number (Foursquare)	P.O. Box 41404
fsq_country	Country name (Foursquare)	GB
fsq_date_created	Date the record was created (Foursquare)	2009-06-30
fsq_date_refreshed	Date the record was last updated (Foursquare)	2025-06-22
fsq_date_closed	Date the place was closed, if applicable (Foursquare)	2025-01-01
fsq_tel	Telephone number (Foursquare)	(404) 230-9085
fsq_website	Website URL (Foursquare)	http://starbucks.com
fsq_email	Email address (Foursquare)	info@starbucks.com
fsq_facebook_id	Facebook page identifier or URL (Foursquare)	22092443056.0
fsq_instagram	Instagram handle or URL (Foursquare)	starbucks
fsq_twitter	Twitter handle or URL (Foursquare)	starbucks
fsq_category_ids	List of category IDs (Foursquare)	['4bf58dd8d48988d1e0931735']
fsq_category_labels	Human-readable category names (Foursquare)	['Dining and Drinking > Cafe, Coffee, and Tea House > Coffee Shop']
fsq_placemaker_url	URL to Placemaker or API resource (Foursquare)	https://foursquare.com/placemakers/review-place/4a4a0fd6f964a52087ab1fe3
fsq_unresolved_flags	Quality issues reported for a POI (Foursquare)	['duplicate']
fsq_bbox	Bounding box of the place (Foursquare)	{'xmin': -84.34941113, 'ymin': 33.768107, 'xmax': -84.34941113, 'ymax': 33.768107}
fsq_geom	Point geometry (latitude/longitude) in WGS84 (Foursquare)	01010..1E24040
osm_id	Unique identifier of the OSM object	1237615380
osm_class	Main category or class (OSM)	amenity
osm_type	More specific type (OSM)	cafe
osm_name	Name of the object (OSM)	Starbucks
osm_address	Address or location description (OSM)	"city"=>"Little Five Points Village", "state"=>"GA", "street"=>"Moreland Avenue", "postcode"=>"30307", "housenumber"=>"506"
osm_extratags	Additional tags or metadata (OSM)	"phone"=>"+1 404-230-9085", "branch"=>"Little Five Points", "cuisine"=>"coffee_shop", "takeaway"=>"yes",...
osm_geometry	Geometry of the object (OSM)	01010..1E24040
osm_latitude	Latitude in decimal degrees derived from osm_geometry	33.7680684
osm_longitude	Longitude in decimal degrees derived from osm_geometry	-84.3494503
osm_geom	Point geometry derived from osm_latitude and osm_longitude	01010..1E24040
fsq_osm_name_similarity_score_trg	Trigram-based name similarity	0.99
fsq_osm_name_similarity_score_lev	Levenshtein-based name similarity	0.99
fsq_osm_distance	Distance between Foursquare and OSM locations in degrees	5.499317140236702e-05

Table 2. Example subset of k -nearest-neighbor connections and distances between Foursquare–OSM POIs.

fsq_place_id_source	fsq_place_id_destination	distance_m
4ec132eb7ee54e4cd348d18b	4c48c4076594be9a5fba2e24	313.03
4ec132eb7ee54e4cd348d18b	5097f4fe4b9033ce2157b74d	347.70
4ec132eb7ee54e4cd348d18b	4c1e90a4eac020a18aec49c2	377.12
4ec132eb7ee54e4cd348d18b	0842aab84b994f9990b21a4d	385.50
4ec132eb7ee54e4cd348d18b	4be835d4d837c9b60d05a506	378.52
4ec132eb7ee54e4cd348d18b	b6184fc4efce46d3f37e756b	393.30
...

Tabular Data Structure

The `fsq_osm` table constitutes the final integrated dataset that merges detailed place information from Foursquare and OpenStreetMap (OSM). Each row in the table corresponds to a point-of-interest (POI), enriched with spatial, semantic, and contextual metadata from both sources.

Table 1 lists the columns of the provided dataset along with their descriptions. Columns originating from Foursquare are prefixed with `fsq_`, while those from OpenStreetMap are prefixed with `osm_`. Foursquare-derived fields include unique identifiers, place names, latitude and longitude coordinates, street address components, city, state, country, postal code, contact details (e.g., phone number, website), and associated social media handles (e.g., Twitter, Facebook). In addition, each record contains one or more categorical tags describing the place type or function. The spatial location is stored using a `POINT` geometry in WGS 84 (EPSG:4326) format, which enables efficient geospatial queries using PostGIS or other spatial databases. OSM-derived fields enrich these records with community-curated metadata, including OSM identifiers, POI names, classification types (e.g., amenity, tourism, leisure), administrative levels, and freeform tag-value pairs contributed by mappers. These fields offer rich geographic and semantic context, such as whether the place is within a park, located on a university campus, or tagged with specific amenities. We provide two geometries for each OSM record: the original geometry from the OSM dataset and a point geometry calculated from its centroid. The centroid-based geometry ensures consistency with the Foursquare point geometries, aligned with the same spatial reference system, enabling uniform spatial operations between the two datasets. To facilitate transparent integration and downstream filtering, we include three computed fields:

- **`fsq_osm_distance`:** the great-circle distance (in degrees) between the Foursquare and OSM POI coordinates.
- **`name_similarity_score_trg`:** a normalized trigram-based string similarity metric that quantifies the semantic resemblance between the names of the matched POIs.
- **`name_similarity_score_lev`:** a normalized Levenshtein-based string similarity metric that quantifies the semantic resemblance between the names of the matched POIs.

We note that using these fields, filtering can be performed based on different thresholds and criteria. For instance, candidate pairs can be selected using a minimum name-similarity score (e.g., `name_similarity_score_lev` ≥ 0.5) and a maximum spatial distance (e.g., `fsq_osm_distance` ≤ 0.001 degrees). Additionally, other similarity metrics or attributes can be computed and appended to the join table to enhance matching accuracy. However, the calculation and integration of these additional metrics are beyond the scope of this paper and are left for users or future work.

Graph Data Structure

Figure 3 illustrates the spatial point-graph generated from the integrated tabular dataset. Each blue node represents a POI, and each red line denotes an edge connecting a POI to its ten geographically nearest neighbors, as determined using the k -nearest-neighbor (k NN) algorithm. As shown, some clusters are disconnected from the main graph, reflecting a natural outcome of the k NN-based connectivity, where isolated regions may form independent subgraphs. In contrast, denser urban areas such as downtown and midtown exhibit a higher edge density, while predominantly residential or low-business regions contain fewer or no nodes in the graph.

Table 2 presents a sample subset of pairwise distances between POIs, where each node is connected to its N nearest neighbors. As shown, the graph data include only three fields: the source place identifier, the destination place identifier, and the distance between them (in meters). The distances were computed using a PostGIS function that calculates great-circle distances in meters rather than degrees, providing a consistent and meaningful global scale. Using meters ensures that edge weights are directly interpretable and avoids distortions that can occur

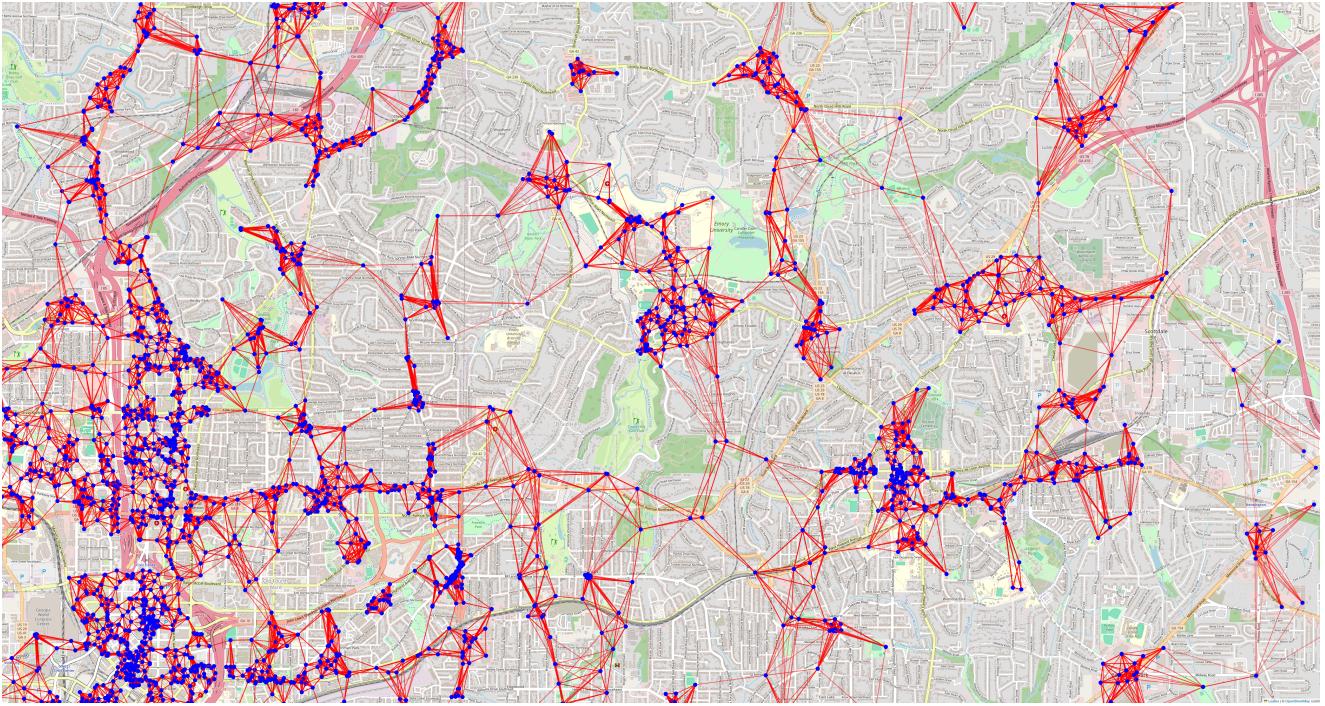


Figure 3. Visualization of the k -nearest-neighbor (kNN) spatial graph generated from tbular dataset around the Emory area. Blue dots represent POIs and red lines indicate edges connecting each POI to its ten nearest neighbors based on geodesic distance ($k=10$). The map was generated in Python and the source code is on GitHub

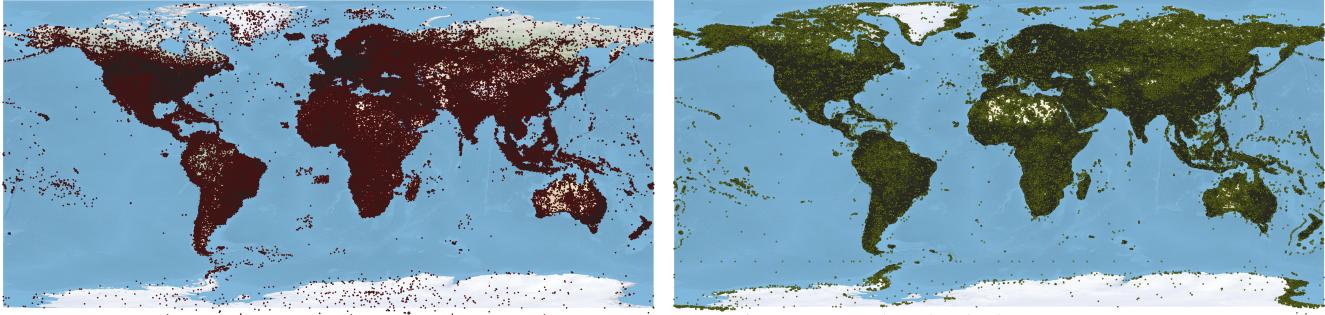
when using degrees, which vary with latitude (e.g., between the equator and the poles). We note that a single `fsq_place_id_source` may correspond to multiple rows in the tabular dataset. This occurs when a single Foursquare place is matched to more than one OSM location with sufficiently high name-similarity scores. As a result, a single node in the graph can represent multiple candidate matches in the tabular data. To maintain transparency and flexibility, we include all such records in the dataset and retain their corresponding place identifiers. This design allows users to apply their own filtering criteria or tie-breaking rules—such as selecting the record with the highest similarity score or the smallest spatial distance—based on the requirements of their specific analyses. An alternative solution would be to assign a new unique identifier to each record and use it as the primary key. However, this approach would create multiple distinct identifiers for what is essentially the same location, potentially leading to ambiguity and misinterpretation in the graph representation. To avoid this issue, we preserved the original `fsq_place_id` as the node identifier, ensuring that all records referring to the same place remain linked, while still allowing users to manage duplicate matches according to their analysis needs.

Technical Validation

To ensure data quality and correctness of the integration process, multiple verification checks were applied at each stage of the pipeline described in the Methods section. In this section, we present an in-depth validation and comparison against external sources to verify the accuracy, consistency, and reliability of the integrated dataset.

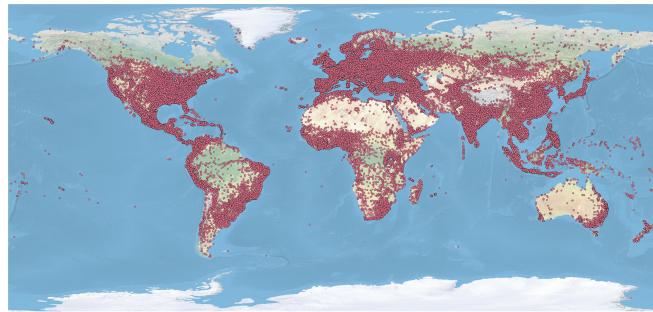
Comparison to City Population

Figure 4 provides a visual comparison of World-POI with other data sources. As shown in Figure 4a, the Foursquare data visualization indicates that POIs are broadly distributed across all regions, including areas that are likely uninhabited. In contrast, Figure 4b illustrates the dense clustering of OSM POIs, which generally correspond to areas that users have physically visited and mapped. The peer-reviewed and community-validated nature of OSM further enhances the accuracy and reliability of these locations. To further validate our results, we visualize city populations in Figure 4c, highlighting cities with populations exceeding 1,000—regions where POI presence is expected to be higher. Finally, Figure 4d presents the output of our integration approach, where results were filtered using a Levenshtein name-similarity threshold greater than 0.5. The strong spatial correlation between World-POI and the population map demonstrates that World-POI provides a realistic representation of human

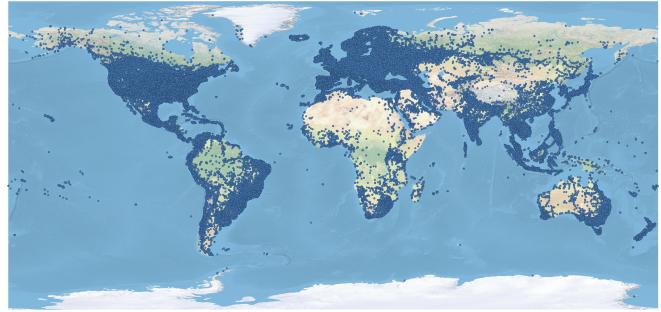


(a) Foursquare POIs show broad spatial coverage, including entries in sparsely populated or uninhabited regions.

(b) OSM POIs exhibit denser clustering, primarily in areas where contributors have recorded detailed map data.



(c) City population distribution (population > 1,000) illustrates expected regions of human settlement and activity



(d) World-POI, filtered by a Levenshtein name-similarity threshold greater than 0.5

Figure 4. Visualization of Foursquare POIs, OSM points, city population distribution, and the integrated World-POI dataset. The World-POI dataset aligns closely with populated regions, demonstrating enhanced spatial accuracy and semantic consistency. The map was generated in QGIS using the ESRI Physical Map basemap layer.

activity patterns across populated regions. In other words, our approach effectively removes noise by retaining only POIs that exist in both Foursquare and OSM datasets and meet the similarity criteria, resulting in a more accurate and valid representation of real-world locations.

Greenland Case Study

Manually examining every point in the dataset is not feasible; therefore, we conducted a focused case study to assess data quality. We selected Greenland as the study region because it is sparsely populated yet contains several notable points of human activity and businesses. Moreover, its POI distributions in Foursquare, OSM, population data, and World-POI showed significant variation, making it a strong candidate for validation.

Figure 5 presents this case study, highlighting how our integration approach refines noisy POI datasets. In Figure 5a, Foursquare POIs appear sparsely scattered across the island, including in uninhabited regions. In contrast, Figure 5b shows dense clustering along the coasts but also includes numerous points in areas without permanent populations, likely resulting from automated or crowd-sourced mapping artifacts. Figure 5c visualizes cities with populations exceeding 1,000. The World-POI subset for Greenland, shown in Figure 5d, includes only POIs with a Levenshtein name-similarity score greater than 0.5. This filtering criterion effectively minimizes false positives and results in a spatial distribution that closely aligns with populated coastal regions. The strong correspondence between World-POI and known inhabited areas demonstrates that the proposed filtering and integration methodology yields a realistic, accurate, and reliable representation of true human settlements.

The visualization provided in Figure 5 demonstrates the overall validity of World-POI. To further evaluate the dataset, we performed a manual validation of 150 POIs: 50 from Foursquare, 50 from OSM, and 50 from World-POI. Each location was cross-verified using Google Maps, and when the site could not be identified, we consulted additional sources such as Wikipedia (<https://wikipedia.com>), Google Search (<https://google.com>), and Mapcarta (<https://mapcarta.com/>) to obtain contextual or supporting evidence.

The results of this manual validation are summarized in Table 3, Table 4, and Table 5. Overall, the majority of Foursquare POIs lacked verifiable real-world locations; despite extensive searches, many entries appeared to

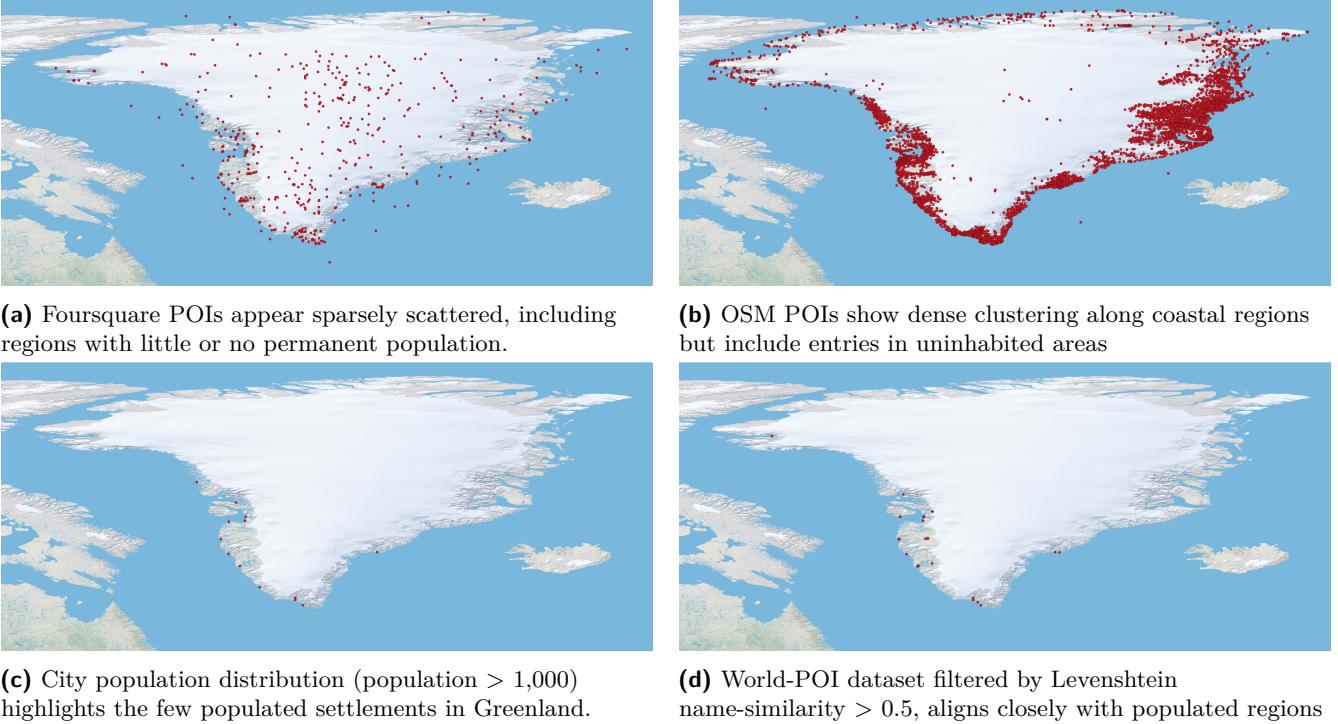


Figure 5. Comparison of POI distributions in Greenland across different data sources.

be synthetic or mislabeled. In contrast, OSM contained more physically realistic coordinates, yet many points corresponded to generic or unverified map features such as islands, bays, or geographic formations rather than functional POIs. World-POI, however, produced a curated subset of POIs that corresponded to real, identifiable locations with accurate names and coordinates. Most entries in World-POI were confirmed through Google Maps or other sources, indicating high spatial and semantic validity. This demonstrates that our integration and filtering pipeline effectively eliminates non-existent or ambiguous POIs while retaining accurate and meaningful locations suitable for spatial and behavioral analysis.

Foursquare sample (Table 3). The Foursquare subset contains a mix of verifiable venues (e.g., *Restaurant Nasaasaaq*, *Sisimiut Tandklinik*, *Brugseni*) and many entries that are unverifiable, out of region, or semantically non-POI tokens. Notable issues include abstract or cultural terms (*Christmas 2012*, *Purgatory*), planetary or foreign toponyms (*Mars*, *Amsterdam*, *Rypefjord* in Norway), vague abbreviations (*KNR*, *HTX*), and generic areas rather than discrete POIs (*South Greenland*, *Qaasuitsup*). Several records could not be verified through any external or corroborating source, even after targeted manual searches. Overall, the sample reveals substantial noise in the Foursquare data, characterized by name inflation, redundant entries, and semantic drift away from true ground-truth venues.

Specifically, among the 50 Foursquare records examined, 11 correspond to real locations, 1 represents a permanently closed site, and 38 do not correspond to any verifiable real-world location. Of the verified real locations, 10 qualify as actual points of interest (POIs) such as businesses or facilities. Notably, the entry “Northern International Crypto Kingdom’S” appears twice in the dataset—this duplication reflects distinct Foursquare place IDs (5f5b74765148947f3da6f71b and 5f5b8389fe774e63aabdfbce) assigned to identically named records.

OSM sample (Table 4). The OSM subset is geographically coherent but skewed toward natural and administrative features rather than consumer POIs. Many entries are islands, bays, capes, lakes, straits, glaciers, or municipalities (e.g., *Innaarsulik*, *Nordfjord*, *Kakivfaiit*, *Avannaata Kommunia*), which are appropriate cartographic features but not business destinations. There is a healthy presence of genuine facilities and businesses (*Spar*, *Restaurant Ulo*, *Hotel Arctic*, *Royal Greenland*), and occasional duplicates or multi-tag representations for the same landmark (*Knud Rasmussen sculpture* as both *viewpoint* and *memorial*). Ambiguous points such as *N 90* drinking water and parking laybys illustrate minor tagging inconsistencies. In sum, OSM provides realistic coordinates but often for non-venue entities.

For the OSM sample, 44 entries correspond to real locations, 3 represent real-world geographic areas (such as bays,

islands, or glaciers), and 3 could not be verified as real-world locations. Among the verified records, 11 correspond to identifiable points of interest (POIs) such as businesses or public facilities. Notably, Avannaata Kommunia appears twice in the dataset, with OSM IDs 405944737 and 405944738, both categorized as amenity/townhall. Similarly, Ikerasaarsuk occurs twice, with OSM IDs 413832375 and 413833463, both labeled as natural/strait. These repetitions represent legitimate multiple mappings within OSM rather than data duplication errors.

World-POI sample (Table 5). The matched table yields high-confidence, real venues with small, interpretable positional offsets rather than semantic noise. Most pairs are legitimate places verified in maps or imagery (*Nuuk Kunstmuseum*, *Café Inuk*, *Ilisimatusarfik*, *STARK Sisimiut*, *Hotel Kulusuk*). A common pattern is sub-building to street-side displacement where the FSQ point lies just outside the structure and OSM lies on the building footprint (*Fitness GL*, *Hotel Sisimiut*, *Blue Café*). The table also captures operational status and context (*SparQuick* permanently closed; *Nuugaatsiaq* now abandoned), and highlights occasional name variants or multiple OSM IDs for the same venue (*Hotel Disko Island*). One clear geocoding error (*Ilulissat Vandrehjem* matched to a house due to a misplaced FSQ point) underscores the value of dual-source matching. Overall, the matched set preserves genuine, mappable venues with building-level offsets typical of heterogeneous geocoders, supporting its suitability for spatial analysis.

For the World-POI sample, 48 entries correspond to real locations, 2 represent permanently closed sites, and 2 correspond to broader areas or villages. Among these, 45 out of 48 verified entries are valid points of interest (POIs). The entry Sulorraq illustrates a unique case—its Foursquare record corresponds to an apartment complex, while the OSM data include both Sulorraq 11 (a building) and Sulorraq (a nearby statue). Because the primary match represents residential housing rather than a commercial or functional POI, Sulorraq was excluded from the POI count. Similarly, Ilulissat Vandrehjem corresponds to a private house rather than a POI. Additionally, duplicate mappings were identified where Brugseni was matched to two OSM records (180978036 and 1913382879) and Hotel Disko was mapped to two OSM records (715178079 and 4399888527), both representing legitimate but redundant OSM entries of the same establishment.

Usage Note

The World-POI dataset offers a unified, high-quality view of global points of interest (POIs) by integrating Foursquare and OpenStreetMap (OSM) data through spatial and semantic alignment. This integration enables a wide range of downstream research and analytical applications across geospatial, data science, and urban informatics domains. Below, we describe key potential use cases.

1. Geographic Knowledge Graph Construction. The dataset provides a foundation for building geographic knowledge graphs that link semantically and spatially related POIs across heterogeneous data sources. Each record includes identifiers, metadata, and similarity metrics that facilitate relationship extraction and entity resolution. This structure enables machine-readable representations of urban spaces, supporting tasks such as graph-based reasoning, spatial relationship modeling, and semantic query expansion.

2. Location-Based Clustering and Classification. The spatial coordinates and category labels of POIs can be used for unsupervised clustering and classification analyses. Researchers can identify spatial clusters of economic activity, cultural landmarks, or service facilities and classify regions based on POI density and diversity. The combination of spatial and semantic similarity fields allows advanced methods such as DBSCAN, k-means, or hierarchical clustering to reveal patterns of urban structure and human activity.

3. Semantic Enrichment of Commercial POI Data. World-POI enables the semantic enrichment of commercial or proprietary POI datasets using OSM's open and extensible tagging framework. Users can cross-link business listings with OSM attributes such as amenities, building types, or accessibility indicators, improving data completeness and interoperability. This feature is especially useful for data fusion in urban analytics, retail mapping, and smart city applications.

4. Urban and Regional Planning Analysis. The integrated dataset can be employed to assess spatial accessibility, infrastructure distribution, and land-use balance across cities or regions. Planners can identify underserved neighborhoods, evaluate commercial diversity, and monitor urban growth trends over time by leveraging the dataset's standardized geographic and categorical attributes.

5. Mobility and Human Behavior Modeling. Because the dataset accurately reflects real-world activity centers, it is suitable for mobility modeling, travel demand forecasting, and agent-based simulations. Each verified POI represents

a potential anchor for human movement, allowing integration into population mobility models or epidemic simulations where realistic behavioral and spatial dynamics are required.

6. Data Quality Assessment and Benchmarking. World-POI can serve as a benchmark for evaluating geocoding accuracy, record linkage algorithms, or POI data quality assessment methods. Its verified matches between Foursquare and OSM records, coupled with name similarity metrics and distance fields, allow researchers to systematically test and calibrate entity-matching pipelines or evaluate spatial accuracy in noisy or incomplete datasets.

7. Network and Graph-Based Analyses. The provided graph representation of POIs, where nodes represent locations and edges denote k-nearest spatial relationships, enables graph-theoretic analyses such as centrality, clustering coefficients, and community detection. These analyses can reveal patterns of urban connectivity, identify local hubs of activity, and support transportation or accessibility modeling.

Together, these use cases highlight World-POI as a robust and versatile resource for spatial data integration, geoinformatics, and computational urban studies. Its reproducible design and global coverage make it suitable for both methodological research and applied spatial analysis at multiple geographic scales.

References

1. Kandt, J. & Batty, M. Smart cities, big data and urban policy: Towards urban analytics for the long run. *Cities* **109**, 102992 (2021).
2. Ommi, S. & Hashemi, M. Machine learning technique in the north zagros earthquake prediction. *Appl. Comput. Geosci.* **22**, 100163 (2024).
3. Amiri, H. *et al.* Massive trajectory data based on patterns of life. In *SIGSPATIAL'23*, 1–4, <https://doi.org/10.1145/3589132.3625592> (2023).
4. Amiri, H., Yang, R. & Züfle, A. Geolife+: Large-scale simulated trajectory datasets calibrated to the geolife dataset. In *SIGSPATIAL GeoSim'24 Workshop*, 25–28, <https://doi.org/10.1145/3681770.3698573> (2024).
5. Saiph Savage, N., Baranski, M., Elva Chavez, N. & Höllerer, T. I'm feeling loco: A location based context aware recommendation system. In *Advances in Location-Based Services: 8th International Symposium on Location-Based Services, Vienna 2011*, 37–54 (Springer, 2012).
6. Takeuchi, Y. & Sugimoto, M. Cityvoyager: An outdoor recommendation system based on user location history. In *International Conference on Ubiquitous Intelligence and Computing*, 625–636 (Springer, 2006).
7. Hashemi, M., Amiri, H. & Zufle, A. Placefm: A training-free geospatial foundation model of places using large-scale point of interest data. *arXiv e-prints* arXiv-2507 (2025).
8. Hashemi, M. & Zufle, A. From points to places: Towards human mobility-driven spatiotemporal foundation models via understanding places. *arXiv e-prints* arXiv-2506 (2025).
9. Zhang, Z., Amiri, H., Liu, Z., Zhao, L. & Züfle, A. Large language models for spatial trajectory patterns mining. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Geospatial Anomaly Detection*, 52–55 (2024).
10. Zhang, Z. *et al.* Transferable unsupervised outlier detection framework for human semantic trajectories. In *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*, 350–360 (2024).
11. Amiri, H., Kong, R. & Züfle, A. d. Urban anomalies: A simulated human mobility dataset with injected anomalies. In *SIGSPATIAL GeoAnomalies'24 Workshop*, 1–11, <https://dl.acm.org/doi/10.1145/3681765.3698459> (2024).
12. Kohn, W., Amiri, H. & Züfle, A. Epipol: An epidemiological patterns of life simulation (demonstration paper). In *Proceedings of the 4th ACM SIGSPATIAL International Workshop on Spatial Computing for Epidemiology*, 13–16 (2023).
13. Amiri, H. *et al.* The patterns of life human mobility simulation. In *Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems*, 653–656 (2024).
14. Google. Google places api. <https://www.google.com/maps> (2025). [Accessed 10-16-2025].
15. AdvanResearch. Location and sales data for finance, retail & commercial real estate. <https://www.advanresearch.com> (2025). [Accessed 10-16-2025].

16. Microsoft. Bing maps api. <https://www.bing.com/maps> (2025). [Accessed 10-16-2025].
17. Technologies, H. Here location services. <https://www.here.com> (2025). [Accessed 10-16-2025].
18. Precisely. Precisely points of interest data. <https://www.precisely.com/> (2025). [Accessed 10-16-2025].
19. Foursquare. The future of geospatial technology. <https://www.foursquare.com/> (2025). [Accessed 10-16-2025].
20. OpenStreetMap. Openstreetmap. <https://www.openstreetmap.org> (2025). [Accessed 10-16-2025].
21. Survey, U. G. Geographic names information system (gnis). <https://www.usgs.gov/> (2025). [Accessed 10-16-2025].
22. Foundation, O. M. Overture maps: Open map data. <https://www.overturemaps.org> (2025). [Accessed 10-16-2025].
23. Amiri, H. Foursquare-openstreetmap integration repository. <https://github.com/onspatial/fsq-osm> (2025). [Accessed 10-16-2025].
24. Yu Jian, L. & Bo, L. A normalized levenshtein distance metric. *IEEE transactions on pattern analysis machine intelligence* **29**, 1091–1095 (2007).

Code Availability

The complete codebase used for data collection, preprocessing, spatial integration, and export is publicly available at: <https://github.com/onspatial/world-poi>

The repository includes:

Scripts for downloading and parsing Foursquare and OSM data.

SQL and Python code for spatial processing and matching.

Tools for generating graph representations and exporting structured CSVs.

Configuration files and reproducibility documentation.

Acknowledgments

This research was supported by the National Science Foundation (NSF) under Award No. 2109647, *Data-Driven Modeling to Improve Understanding of Human Behavior, Mobility, and Disease Spread*.

Author contributions statement

HA collected and processed the data and prepared the draft manuscript. MH generated the graph data and also contributed to the draft manuscript. AZ designed and conducted the analysis and participated in writing the draft manuscript.

Competing interests

The authors declare no competing interests.

Table 3. Manual validation of a random sample of 50 **Foursquare** records from Greenland. Validation was performed using Google Maps and other online sources to determine whether each listed place exists and, if so, whether it represents an actual business.

fsq_name	fsq_category_labels	Business	Observation
Akiki	['Retail > Food and Beverage Retail > Grocery Store']	YES	Real Location
Brugseni	['Retail > Food and Beverage Retail > Supermarket']	YES	Real Location
STARK Sisimiut	['Retail > Hardware Store']	YES	Real Location
Restaurant Nasaasaaq	['Dining and Drinking > Restaurant']	YES	Real Location
Nuuk Kunstmuseum	['Arts and Entertainment > Museum > Art Museum']	YES	Real Location
Sisimiut Tandklinik	['Health and Medicine > Dentist']	YES	Real Location
Sisimiut Harbour	['Landmarks and Outdoors > Harbor or Marina']	YES	Real Location
The Culture Centre Taseralik	['Arts and Entertainment > Performing Arts Venue > Concert Hall']	YES	Real Location
Nana's Thai Takeaway	['Dining and Drinking > Restaurant > Asian Restaurant > Thai Restaurant']	YES	Real Location
Paamaap Kuua 13-307	['Community and Government > Assisted Living']	YES	Real Location
sermeq kujataleq	['Landmarks and Outdoors > Scenic Lookout']	NO	Real Location
Spar Quick	['Retail > Miscellaneous Store']	N/A	Real Location (Closed)
Qaasuitsup	['Landmarks and Outdoors > States and Municipalities > Town']	N/A	NO Real Location
South Greenland	['Landmarks and Outdoors > Island']	N/A	NO Real Location
Nuuk	['Travel and Transportation > Transport Hub > Airport']	N/A	NO Real Location
Groelandia	['Landmarks and Outdoors > Field']	N/A	NO Real Location
Biblioteket - The Library	['Community and Government > Library']	N/A	NO Real Location
Christmas 2012	['Arts and Entertainment']	N/A	NO Real Location
Purgatory	['Landmarks and Outdoors > Campground']	N/A	NO Real Location
48 GU 556	['Travel and Transportation > Road']	N/A	NO Real Location
Mars	['Travel and Transportation > Rest Area']	N/A	NO Real Location
Amsterdam	['Landmarks and Outdoors > Mountain']	N/A	NO Real Location
KNR	['Business and Professional Services > TV Station']	N/A	NO Real Location
Fit Life Fitness	['Sports and Recreation']	N/A	NO Real Location
Antartika	['Landmarks and Outdoors > Mountain']	N/A	NO Real Location
Marlin Shadow Master Lloyd Waters	['Landmarks and Outdoors > Park > National Park']	N/A	NO Real Location
Tall Junk 45 Kite	['Landmarks and Outdoors > Park > National Park']	N/A	NO Real Location
Big Aster Dong Water Over Lloyd Forest 67	['Landmarks and Outdoors > Park > National Park']	N/A	NO Real Location
Davis Strait	['Landmarks and Outdoors > Other Great Outdoors']	N/A	NO Real Location
Last Intermediate National Kartel	['Landmarks and Outdoors > Park > National Park']	N/A	NO Real Location
Mine Of New Youth Making Erosion	['Landmarks and Outdoors > Park > National Park']	N/A	NO Real Location
Party Hard On the Boulevard	['Community and Government > Spiritual Center > Temple']	N/A	NO Real Location
Northern International Crypto Kingdom'S	['Landmarks and Outdoors > Park > National Park']	N/A	NO Real Location
Northern International Crypto Kingdom'S	['Landmarks and Outdoors > Park > National Park']	N/A	NO Real Location
Eslant Sentinel Castle	['Landmarks and Outdoors > Castle']	N/A	NO Real Location
Afsane	['Dining and Drinking > Restaurant > Diner']	N/A	NO Real Location
Bayramyeri Sok	['Retail > Big Box Store']	N/A	NO Real Location
Rypefjord	['Landmarks and Outdoors > Bay']	N/A	NO Real Location
Polo Norte	['Landmarks and Outdoors > Mountain']	N/A	NO Real Location
Igloo Mountain	['Sports and Recreation > Snow Sports > Ski Lodge']	N/A	NO Real Location
Griffith South Center	['Landmarks and Outdoors > Structure']	N/A	NO Real Location
Tacan	['Landmarks and Outdoors > Field']	N/A	NO Real Location
baghimiz	['Landmarks and Outdoors > Botanical Garden']	N/A	NO Real Location
Haluk Kece	['Event > Conference']	N/A	NO Real Location
Afrika	['Travel and Transportation']	N/A	NO Real Location
Ekvator	['Arts and Entertainment > Aquarium']	N/A	NO Real Location
Alpler	['Landmarks and Outdoors > Mountain']	N/A	NO Real Location
HTX	['Community and Government > Education > College and University > College Cafeteria']	N/A	NO Real Location
Bar Aaveq	['Dining and Drinking > Bar']	N/A	NO Real Location
The Hut	['Landmarks and Outdoors > Scenic Lookout']	N/A	NO Real Location

Table 4. Manual validation of a random sample of 50 **OpenStreetMap (OSM)** records from Greenland. Validation was performed using Google Maps and other online sources to determine whether each listed place exists and, if so, whether it represents an actual business.

osm_name	osm_class	osm_type	Business	Observation
Akiki	shop	supermarket	YES	Real Location
Hotel Arctic	building	hotel	YES	Real Location
Spar	shop	supermarket	YES	Real Location
Kangerluarsuk	natural	bay	NO	Real Location
Innaarsulik	place	islet	NO	Real Location
Angisunnguaq	place	island	NO	Real Location
Kap Udkiggen	natural	cape	NO	Real Location
Nordfjord	natural	bay	NO	Real Location
Krumodden	natural	cape	NO	Real Location
Pooqattaaq	place	island	NO	Real Location
Arctic Umiaq Line	amenity	ferry_terminal	YES	Real Location
Aqqutarfik	tourism	museum	YES	Real Location
Royal Greenland	building	industrial	YES	Real Location
Royal Arctic Line	building	warehouse	YES	Real Location
Polar Oil	building	yes	YES	Real Location
Sunset Boulevard	amenity	fast_food	YES	Real Location
Knud Rasmussen sculpture	tourism	viewpoint	NO	Real Location
Knud Rasmussen sculpture	historic	memorial	NO	Real Location
Restaurant Ulo	amenity	restaurant	YES	Real Location
Orpissup Tasia	water	lake	NO	Real Location
Butik Sara	shop	gift	YES	Real Location
Eqalunnguit Nunaat	place	island	NO	Real Location
Tasersuaq	water	lake	NO	Real Location
Kangilat	place	locality	NO	Real Location
Apuliliip Apusiaa	natural	glacier	NO	Real Location
Qeqertarsuatsiaq	place	island	NO	Real Location
Sermeq (2012)	natural	glacier	NO	Real Location
Iggiaanut - Bro 1	amenity	parking	NO	Real Location
Inalugartuut Iluat	natural	bay	NO	Real Location
Kangiussap Paava	natural	bay	NO	Real Location
Aarfit Timaat	place	island	NO	Real Location
Kangeq	place	island	NO	Real Location
Amitsuarssuk	natural	bay	NO	Real Location
Aasiaat	place	island	NO	Real Location
Ungoorsivik	natural	bay	NO	Real Location
Eqalunnguit	natural	bay	NO	Real Location
Majooq	place	island	NO	Real Location
Niaqornaq	place	islet	NO	Real Location
Kakivfaat	natural	strait	NO	Real Location
Anarusuk	place	island	NO	Real Location
Eqalugaarsuit	natural	bay	NO	Real Location
Ikerasaarsuk	natural	strait	N/A	Real Location
Ikerasaarsuk	natural	strait	N/A	Real Location
Nuuk Imeq A/S	building	industrial	N/A	Real Location (Closed)
Avannaata Kommunia	amenity	townhall	N/A	AREA
Avannaata Kommunia	amenity	townhall	N/A	AREA
Pingorsuaq	place	neighbourhood	N/A	AREA
Fiskefabrik	shop	seafood	N/A	NO Real Location
N 90 E	amenity	drinking_water	N/A	NO Real Location
N 90	amenity	drinking_water	N/A	NO Real Location

Table 5. Manual validation of a random sample of 50 **World-POI** records from Greenland. Validation was performed using Google Maps and other online sources to determine whether each listed place exists and, if so, whether it represents an actual business.

fsq_name	osm_name	osm_type	Business	Observation
Nuuk Kunstmuseum	Nuuk Kunstmuseum	museum	YES	Real Location
Fitnessgl	Fitness GL	fitness_centre	YES	Real Location
Salon Nauja	Salon Nauja	hairdresser	YES	Real Location
Arktisk Kommando	Arktisk Kommando	government	YES	Real Location
Comby A/S	Comby	it	YES	Real Location
Nukissiorfítt	Nukissiorfítt	government	YES	Real Location
Café Inuk	Café Inuk	cafe	YES	Real Location
Ilisimatusarfik	Ilisimatusarfik	university	YES	Real Location
Grønlands Naturinstitut	Grønlands Naturinstitut	parking	YES	Real Location
Nuuk Golfklub	Nuuk Golf Club	yes	YES	Real Location
Suloraq	Suloraq 11	apartments	NO	Real Location
Suloraq	Suloraq	artwork	NO	Real Location
Pisiffik Qinngorput	Pisiffik Qinngorput	supermarket	YES	Real Location
Maniitsoq Harbour	Maniitsoq	ferry_terminal	YES	Real Location
Hotel Søma Sisimiut	Hotel SØMA	hotel	YES	Real Location
Hotel Sisimiut	Hotel Sisimiut	hotel	YES	Real Location
STARK Sisimiut	STARK Sisimiut	hardware	YES	Real Location
Brugseni	Brugseni	supermarket	YES	Real Location
Brugseni	Brugseni	supermarket	YES	Real Location
Nana's Thai Takeaway	N/Anas Thai Take away	fast_food	YES	Real Location
Pub Raaja	Pub Raaja	pub	YES	Real Location
Panorama	Panorama Hostel	hotel	YES	Real Location
Hotel Disko	Hotel Disko Island	hotel	YES	Real Location
Hotel Disko	Hotel Disko Island	hotel	YES	Real Location
Cafe Blue	Blue Café	cafe	YES	Real Location
Skansen - Dit hjem	Skansen Dit Hjem	chalet	YES	Real Location
Hotel Kulusuk	Hotel Kulusuk	hotel	YES	Real Location
ammassalik museum	Ammassalik Museum	museum	YES	Real Location
Hotel Angmassalik	Angmagssalik Hotel	hotel	YES	Real Location
Efterskole Villads Villadsen	Efterskole Villads Villadsen	school	YES	Real Location
Cafe Nuka	Cafe Nuka	cafe	YES	Real Location
Hangout Bistro	Hangout Bistro	fast_food	YES	Real Location
Best Western Plus Hotel Ilulissat	Best Western Plus Hotel Ilulissat	hotel	YES	Real Location
Hotel Icefiord	Hotel Icefiord	hotel	YES	Real Location
Akiki	Akiki	supermarket	YES	Real Location
The Glacier Shop	Glacier Shop	gift	YES	Real Location
Cafennguaq	Cafennguaq	cafe	YES	Real Location
Hotel Arctic Igloos	Hotel Arctic	hotel	YES	Real Location
Ilulissat Vandrehjem	Vandrehjem	house	NO	Real Location
Ilulissat Airport	Ilulissat	terminal	YES	Real Location
Eqip Sermia / Eqip Glacier	Eqip Sermia	viewpoint	YES	Real Location
Eqi Glacier Lodge	Glacier Lodge Eqi	camp_site	YES	Real Location
Kangerlussuaq International Science Support Center	KISS (Kangerlussuaq Intn'l Science Support)	hotel	YES	Real Location
Kang Mini Marked	Kang mini marked	supermarket	YES	Real Location
Kangerlussuaq Youth Hostel / Vandrehjem	Kangerlussuaq Vandrehjem	hostel	YES	Real Location
Polar Lodge	Polar Lodge	hotel	YES	Real Location
Spar Quick	SparQuick	convenience	N/A	Real Location (Closed)
Cafe Naapiffik	Cafe Naapiffik	cafe	N/A	Real Location (Closed)
Nuugaatsiaq (city)	Nuugaatsiaq	village	N/A	AREA (Abandoned)
Qeqertat	Qeqertat	village	N/A	AREA