|  |  |
| --- | --- |
| **Task** | **Member** |
| Ingest and merge 19 air quality sensor CSV files into a unified DataFrame | Humaid & Omar |
| Handle file I/O errors, encoding issues, and column alignment for all sensor files | Humaid & Omar |
| Load and validate the Chicago community‐area shapefile (ensure correct CRS and geometry) | Humaid & Omar |
| Clean raw sensor records: remove nulls, negative or impossible PM₂.₅ values | Humaid & Omar |
| Clean raw taxi records: filter out invalid coordinates and timestamps | Humaid & Omar |
| Parse and standardize all timestamp fields to a common datetime and timezone | Mahmoud |
| Filter both datasets to points strictly within Chicago city limits | Moaaz |
| Encode sensor and taxi coordinates into geohash cells at precision level 6 | Abdullah & Mahmoud |
| Aggregate sensor readings into hourly mean PM₂.₅ per geohash cell | Moaaz |
| Aggregate taxi pickup events into hourly counts per geohash cell | Moaaz |
| Perform a spatial join on (geohash, hour) to merge PM₂.₅ and taxi counts | Mahmoud |
| Develop sampling logic to isolate high‐pollution events (PM₂.₅ > 20 µg/m³) | Humaid |
| Cap per‐cell sample counts at 200 to balance spatial coverage | Humaid |
| Engineer features: normalized pickup rates, pollution‐to‐traffic ratios, community‐area IDs | Abdullah & Mahmoud & Moaaz |
| Normalize/scale feature matrix [latitude, longitude, PM₂.₅] for clustering | Abdullah & Moaaz |
| Run baseline DBSCAN on the normalized feature set | Mahmoud |
| Inspect and validate baseline cluster labels via mapping and silhouette scores | Mahmoud |
| Design a custom weighted distance metric combining geographic distance and pollution difference | Abdullah & Mahmoud |
| Implement the weighted distance function as a Python callable | Abdullah & Mahmoud |
| Compute the full pairwise distance matrix for the hotspot subset using the custom metric | Abdullah & Mahmoud |
| Optimize the distance computation with NumPy vectorization | Abdullah & Mahmoud |
| Integrate the precomputed distance matrix into scikit‐learn’s DBSCAN (metric=’precomputed’) | Abdullah & Mahmoud |
| Automate a parameter sweep over ε (neighborhood radius), min\_samples, and α (spatial vs. pollution weight) | Abdullah & Mahmoud |
| Record silhouette scores for each parameter configuration | Abdullah & Mahmoud |
| Identify the optimal DBSCAN parameters (ε, min\_samples, α) that maximize cluster cohesion | Abdullah & Mahmoud |
| Compute Pearson correlation coefficients between taxi counts and PM₂.₅ at the community‐area level | Abdullah & Mahmoud |
| Test statistical significance of correlations | Abdullah & Mahmoud & Moaaz |
| Generate scatter plots with regression lines and confidence intervals for traffic vs. pollution | Abdullah & Mahmoud & Moaaz |
| Create a daily‐average PM₂.₅ time series plot over the full study period | Abdullah & Mahmoud & Moaaz |
| Highlight major pollution events (e.g., heat waves, regional haze) on the time series | Abdullah & Mahmoud & Moaaz |
| Produce weekday vs. weekend PM₂.₅ comparison charts | Abdullah & Mahmoud & Moaaz |
| Plot hourly PM₂.₅ cycles for rush‐hour analysis | Abdullah & Mahmoud & Moaaz |
| Annotate temporal plots with peak percentages and time‐window labels | Abdullah & Mahmoud & Moaaz |
| Craft choropleth maps of average PM₂.₅ by community area | Abdullah |
| Produce taxi‐density heatmaps using geohash grid counts | Abdullah & Mahmoud & Moaaz |
| Map DBSCAN clusters on Chicago’s basemap with color‐coded labels | Abdullah & Mahmoud |
| Adjust figure styles: color scales, legends, annotations for publication quality | Omar & Humaid |
| Cross‐validate spatial visualizations against traffic zones | Omar & Humaid |
| Draft each section of the manuscript (Abstract, Introduction, Related Work, etc.) | Omar & Humaid |
| Prepare and format all figures, tables, and equations for the paper | Omar & Humaid |
| Conduct literature review and summarize key related works | All Members |
| Draft Introduction, Related Work, Methodology, Results, Discussion, Summary, and Future Work sections | All Members |
| Compose and refine the Abstract to align with paper content | All Members |
| Format the manuscript using the IEEEtran LaTeX template | All Members |
| Manage citations and compile the References section in IEEE style | All Members |
| Integrate and position figures, tables, and equations within the LaTeX document | All Members |
| Proofread the manuscript for grammar, clarity, and consistency | All Members |
| Coordinate internal peer reviews and incorporate feedback into revisions | All Members |

A screenshot of a computer

AI-generated content may be incorrect.  
  
A screenshot of a computer

AI-generated content may be incorrect.

**Abdullah Soubhi Abdulkarim U24107870**

A black text on a white background

AI-generated content may be incorrect.

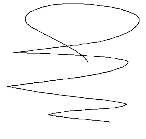
A black signature on a white background

AI-generated content may be incorrect.**Mahmoud Madi U23103334**

**Moaaz Saed Alshehadat U24102854**

A black and white logo

AI-generated content may be incorrect.

**Omar Alhamadi U23200847**

**Humaid AlHadidi U23103346**

A black line on a white background

AI-generated content may be incorrect.