SWASTHYA: Detailed Project Aims & Objectives

1. Executive Summary & Project Vision

Project Swasthya is a final-year engineering project designed to create a *privacy-first, scalable, and intelligent socio-technical system* for **public health surveillance in the Pune region**.

The system addresses the crucial gap between *hyper-local environmental pollution data* and *real-time community health outcomes*. It achieves this by correlating two distinct and real-time data streams:

1. **Environmental Data:** Collected from a custom-built **IoT sensor network** that continuously measures pollutants and meteorological variables.
2. **Health Data:** Collected via a **privacy-preserving citizen Progressive Web App (PWA)** and a **verified hospital/clinic portal**.

The system employs a **hybrid AI architecture** integrating:

- **Federated Learning (FL):** To protect citizen privacy by training models locally on devices.
- **Spatio-Temporal Graph Transformer:** To discover complex correlations between air quality and health trends.

The final output is an **interactive, map-based health intelligence dashboard**, enabling government bodies to make *data-driven, proactive public health interventions*.

Project Vision:
To enable Pune's public health ecosystem to move from *reactive* to *proactive*, *data-driven decision-making*, ensuring healthier communities through localized, evidence-based interventions.

2. Problem Statement & Motivation

While the link between **air pollution** and **adverse health effects** is well-established, public health officials lack the granular, timely, and ethical data systems to act on this information effectively.

The Core Problems

1. **Lack of Data Granularity (The "Where"):**
   Existing monitoring stations like CPCB/SAFAR provide city-level averages that hide local pollution hotspots near roads, factories, or dense residential areas.
2. **Lack of Data Timeliness (The "When"):**
   Health data is recorded *after* hospital admissions, offering little predictive value. No system exists to capture community-level symptom trends in real time.
3. **The Privacy Barrier (The "How"):**
   Building centralized databases of personal health data violates ethical and legal standards, making direct, real-time correlation impossible.

Consequence:
Health responses remain *reactive and generalized*, unable to pinpoint which pollutants, locations, or time periods pose immediate risks.

3. Primary Aim

To design, develop, and deploy a **secure, end-to-end prototype** that integrates **IoT-based pollution sensing**, **federated citizen health data**, and **AI-based correlation models** to produce **evidence-backed, real-time health insights** for public health authorities.

4. Detailed Project Objectives

To achieve this vision, Project Swasthya defines the following **four major objective domains:**

Objective 1: Technical & Infrastructure Objectives

- **Design a Scalable Cloud Architecture:**
  Implement a four-layer AWS architecture using purpose-built databases (Amazon Timestream, S3, PostgreSQL/PostGIS) and a serverless-first compute strategy (AWS Lambda, Kinesis).
- **Develop a Real-Time Data Ingestion Pipeline:**
  Utilize AWS IoT Core (MQTT) and Kinesis Data Streams to ensure reliable, high-throughput ingestion and validation of environmental and health data.
- **Deploy a Hyper-Local Sensor Network:**
  Design and implement a pilot network of **ESP32-based IoT sensors** to measure PM2.5, NO_2, temperature, and humidity at neighborhood resolution.

Objective 2: Dual-Channel Health Data Collection Objectives

- **Develop the Citizen PWA:**
  Create a mobile-first **Progressive Web App** allowing citizens to anonymously report health symptoms.
  This will serve as the **client interface for Federated Learning**.
- **Build the Hospital/Clinic Portal:**
  Develop a secure, role-based web platform for verified medical institutions to submit **aggregated, anonymized case data** (e.g., "50 respiratory cases in Pimpri-Chinchwad, Oct 1–10").
  This portal provides *ground-truth validation* for citizen-reported data.

Objective 3: Hybrid AI & Privacy-Preservation Objectives

- **Implement the Federated Learning (FL) System:**
  - Deploy a **Flower** server (on AWS EC2/ECS) as the FL orchestrator.
  - Integrate **TensorFlow.js** into the citizen PWA for local model training.
  - Ensure only *model gradients*, not raw data, are transmitted—maintaining user privacy.
- **Develop the Central Correlation Engine:**
  - Build and train a **Spatio-Temporal Graph Transformer** using **Amazon SageMaker**.
  - Correlate data across:
    1. Real-time IoT pollution readings (Timestream).
    2. Verified hospital aggregates (PostgreSQL).
    3. Federated learning outputs.
    4. External APIs (weather, traffic, wind data).
  - Derive insights like:
    "A 40% NO_2 spike at Location X with south-westerly winds leads to a 25% rise in cough reports in Location Y within 48 hours."

Objective 4: Stakeholder & Application Objectives

- **Develop the Health Official Dashboard:**
  Create a **React-based** dashboard for public health officials, researchers, and administrators.
- **Implement Role-Based Access Control (RBAC):**
  Secure access using **Amazon Cognito**, defining roles such as:
  - *Health Official:* View analytics and alerts.
  - *Clinic Staff:* Submit or validate local data.
- **Deliver Actionable Visualizations:**
  Include:
  - **Geospatial Maps:** Visualizing pollution and health hotspots.
  - **Interactive Time-Series Charts:** Showing environmental-health lag relationships.
  - **Insights & Alerts Panel:** Presenting AI-derived findings in natural language.

5. Research Contribution & Novelty

1. Hybrid Data Strategy

Combines *unverified citizen-reported data* (protected through FL) with *verified clinic data* to form a rich, balanced, and trustworthy dataset for modeling.

2. Hybrid AI Architecture

Employs **Federated Learning for privacy** and **Graph Transformer for correlation analysis**, ensuring technical sophistication and ethical compliance.

3. Socio-Technical Innovation

Goes beyond technology to address *data privacy, citizen trust, and public policy integration*, creating a model that can be realistically adopted by government health bodies.

6. Conclusion

Project Swasthya demonstrates an advanced, interdisciplinary approach to solving one of the most pressing challenges in modern urban health management.
By fusing **IoT sensing**, **AI analytics**, and **privacy-first design**, it pioneers a new model for ethical, evidence-based, and localized public health intelligence — starting with Pune, and scalable to every smart city in India.

Would you like me to format this version for **Word styling (fonts, heading levels, alignment, spacing, bold titles)** and export it as a `.docx` file ready to submit?
It will look like a professional report section with consistent headings (Heading 1, Heading 2, etc.) and spacing.