Technical Task

Aim

Convert raw GPS data into meaningful metrics and insights that a coach and athlete can easily understand.

Instructions & Details

Leaderboard Generation

Create a leaderboard ranking athletes based on the three key metrics:

1. **Total Distance** – The total distance covered by each athlete.
2. **Distance at Speed Zone 5** – The cumulative distance an athlete runs while moving at a speed between 19.8 km/h and 25.1 km/h.
3. **Top Speed** – The highest speed recorded for each athlete.

Data Processing Steps

- **Noise Reduction**: Apply appropriate filtering to remove noise from speed data. Be specific about which filtering method you use. Consider whether there is anything you could do with both the speed and position data.
- **Pitch Boundaries**: Filter the data to include only movements within the valid pitch area:
    - `pitch_x`: -52.5 to +52.5 (105m)
    - `pitch_y`: -34 to +34 (68m)

Visuals and ball interactions.

Explore additional insights and visualisations that may be valuable to the coach. Create some examples of:

- **Heatmap Analysis**: Generate a heatmap to visualise where the team spent most of their time on the pitch.
- **Ball Data Integration**: Consider how ball tracking data could be incorporated to provide further insights. What interesting metrics could be created using this data?

Dataset Overview

The dataset consists of the following columns:

- `participation_id` – A unique identifier for each athlete.
- `time` (s) – Timestamp of the recorded data point.
- `pitch_x` and `pitch_y` – Positional data on the pitch.
- `speed` (m/s) – The athlete's speed at the given time.

Deliverables

1. **GitHub Repository**
   - Include all scripts used for data analysis, including all visualisation tools.
   - Provide a **README** with instructions on how to run the scripts and any relevant documentation.

2. **Presentation (≤10 minutes)**
   - Explain your approach to solving the task.
   - Present the **Leaderboard** for the three key metrics.
   - Show your **Heatmap(s)** and **Ball analysis**.
   - Summarise key findings and additional insights.

Important:
If you reach the final interview, you will get a chance to talk over your presentation. For the deliverable above, please just send over your slides in a suitable format.