 **Data Cleaning and Preprocessing**:

* **Focus on Relevant Columns**: Remove any non-essential columns (like Version, FP\_Number, TargetWidth, TargetHeight for now, unless they are needed for training), and retain columns like **PlotSize**, **FloorLevel**, and **Room Counts**.
* **Normalize Room Counts**: Standardize the room count data for each plot size (e.g., average number of bedrooms for 5 Marla, 10 Marla, and 20 Marla). This will guide the generation process by ensuring consistency in room distribution.

 **Feature Engineering**:

* **Room Adjacency and Placement**: Create a feature representing typical **room adjacency** and **relationship**. For example:
  + Kitchens typically adjoin dining areas.
  + Bedrooms and bathrooms should be placed near each other, with sufficient privacy considerations.
  + Living areas (lounges, drawing rooms) should be centrally located or close to entrances.
* **Plot Size vs. Room Count Analysis**: Examine how the room counts vary by plot size. For instance, a **5 Marla plot** might typically have 2-3 bedrooms, while a **20 Marla plot** could have 4-5 bedrooms, multiple bathrooms, and larger living spaces. This will help you define ranges for each room type for the model.

 **Conditional Inputs for the Model**:

* Use **room counts** (e.g., number of bedrooms, bathrooms, kitchens) as input conditions for generating floor plans.
* Incorporate **plot size** (5 Marla, 10 Marla, 20 Marla) as a conditioning factor to ensure the floor plan meets the spatial constraints for that plot size.

 **Model Training (CGAN Strategy)**:

* **Quantitative Inputs**: The model should accept **plot size** and **room counts** (e.g., number of bedrooms, bathrooms) as input conditions.
* **Qualitative Inputs**: The model should also incorporate spatial relationships between rooms (room adjacency). This could be represented as a graph or adjacency matrix indicating which rooms are typically adjacent to each other.
* **Training**: Use a **Conditional GAN (CGAN)** approach where the **generator** creates floor plans based on the conditioning factors (plot size, room counts, adjacency), and the **discriminator** evaluates the realism of the generated plans, considering both the quantitative and qualitative constraints.

 **Validation and Refinement**:

* **User Feedback**: After generating floor plans, allow users to review the floor plans to ensure that the room counts and adjacencies match their preferences.
* **Room Quality Checks**: Implement post-processing checks to verify that generated plans respect typical room relationships and functional requirements (e.g., kitchens near dining areas).

 **Integration with Visualization Tools**:

* Ensure that generated floor plans are **color-coded** according to the room types (based on the RGB color mapping provided).
* Allow for easy export of floor plans in **DXF**, **SVG**, and **PNG** formats for further use in CAD software and sharing.

Timeline Suggestion

Here's a suggested timeline for implementing these steps:

1. **Week 1**: Data exploration and visualization; understand patterns in your dataset
2. **Week 2**: Data preprocessing and initial model architecture setup
3. **Week 3-4**: Model training and hyperparameter tuning
4. **Week 5**: Evaluation and refinement of model
5. **Week 6**: Development of the user interface for floor plan generation
6. **Week 7**: Testing with various input configurations and documenting results
7. **Week 8**: Final integration and project documentation

Remember to start with **a simpler model** and gradually increase complexity as you gain confidence in your implementation. Would you like me to elaborate on any specific part of this roadmap?