**Workshop Problem Statement: Creative Exploration with Ames Housing Dataset**

The objective of this workshop is to uncover unique insights from the Ames housing dataset using exploratory data analysis (EDA) techniques. Participants will approach the data with curiosity and creativity, aiming to present findings that go beyond standard analyses. This workshop emphasizes discovering relationships, identifying data-driven trends, and thinking critically about feature interactions.

**Problem Statement**

The Ames housing dataset includes various characteristics of houses sold in Ames, Iowa, making it an ideal resource for EDA due to its diverse features. Your goal is to analyze this dataset creatively to uncover meaningful insights about the housing market in Ames. You are encouraged to go beyond typical EDA by considering non-linear relationships, interaction effects, and multi-dimensional comparisons.

**Dataset**

The dataset can be found here and includes features such as lot size, house type, year built, neighborhood, and sale price, among others.

**Tasks and Subtasks**

1. **Data Understanding and Cleaning**
   * **Task 1.1**: Familiarize yourself with the dataset's features. Identify continuous, categorical, and ordinal variables. Give descriptions for each features.
   * **Task 1.2**: Identify and handle missing values in a way that retains the dataset's integrity. Consider different approaches for different feature types.
   * **Task 1.3**: Assess and address outliers, especially in sale price and other numerical features.
2. **Feature Analysis**
   * **Task 2.1**: Examine the distribution of sale prices across various categories (e.g., neighborhoods, house types, and overall quality). Look for price patterns that indicate market trends.
   * **Task 2.2**: Investigate correlations between sale price and other numerical variables, but go beyond basic correlations. Visualize and analyze non-linear relationships or interaction effects that could impact sale price.
   * **Task 2.3**: Generate new features that might have a predictive value, such as age at the time of sale, or combinations of existing features that capture unique insights.
3. **Uncovering Relationships and Patterns**
   * **Task 3.1**: Compare houses with similar sizes but different prices. What distinguishes them? Consider exploring renovation years, neighborhood quality, or specific amenities.
   * **Task 3.2**: Analyze the effect of categorical variables like HouseStyle, Neighborhood, and SaleCondition on sale price. Visualize these relationships to find any patterns that could be useful for predictive models.
   * **Task 3.3**: Investigate seasonal trends. Explore if certain times of the year result in higher prices or more listings. This could be related to school schedules, weather, or other factors.
4. **Creative Visualization and Storytelling**
   * **Task 4.1**: Develop creative and informative visualizations that communicate your findings. Go beyond simple charts; use color, layout, and design to enhance understanding.
   * **Task 4.2**: Create a story or narrative based on your analysis. Imagine that you are presenting these findings to potential buyers, real estate agents, or city planners. Focus on insights that are actionable or reveal underlying trends in the Ames housing market.
     1. **Price Differences by Neighborhood**: Higher median sale prices are found in neighborhoods with better overall quality ratings. This may suggest that buyers pay a premium for quality.
     2. **Impact of Sale Condition**: Houses sold under ‘Partial’ (new constructions) conditions generally have higher sale prices, likely reflecting new builds.
     3. **Seasonality**: A noticeable increase in sale prices appears during spring and summer months, likely tied to moving seasons and weather conditions in Iowa.
5. **Final Presentation**
   * **Task 5.1**: Summarize your findings in a short presentation (5-10 minutes), focusing on the most insightful or surprising discoveries.
   * **Task 5.2**: Discuss any challenges faced during analysis and how you overcame them. Highlight areas where you exercised creativity in handling or exploring the data.