**Full Project To-Do List:**

**1. Project Overview**

* **Objective:**
  + **Clearly state the objective of the project. For example, "The objective of this project is to identify the key factors that influence house prices using a dataset of property features."**
* **Dataset Description:**
  + **Describe the dataset, including its source, the number of records, and key features (e.g., price, grade, sqft\_living, etc.).**

**2. Initial Data Inspection**

* **Understand the Dataset Structure:**
  + **Data Types: Check the data types of each feature.**
  + **Missing Values: Summarize and handle any missing values in the dataset.**
  + **Basic Statistics: Get basic descriptive statistics for numerical features (mean, median, standard deviation, etc.).**

**3. Correlation Analysis**

* **Calculate Correlation Matrix:**
  + **Identify Strong Correlations: Calculate the correlation matrix focusing on the correlation between price and other features.**
  + **Focus on Key Features: Select features with strong correlations (e.g., grade, sqft\_living, lat, sqft\_living15, sqft\_above, bathrooms) for deeper analysis.**

**4. Targeted Exploration of Key Features**

* **Visual Exploration:**
  + **Histograms and Boxplots: Create histograms and boxplots for the top correlated features to understand their distribution and identify outliers.**
  + **Scatter Plots: Generate scatter plots to visualize the relationship between each key feature and price.**
* **Bivariate and Multivariate Analysis:**
  + **Pair Plots: Use pair plots to visualize interactions between selected features and price.**

**5. Feature Engineering (if needed)**

* **Create New Features:**
  + **Based on your analysis, consider creating new features (e.g., price per square foot).**
* **Transform Existing Features:**
  + **Apply transformations if necessary (e.g., log transformations, binning).**
* **Handle Categorical Variables:**
  + **Encode categorical variables appropriately (e.g., one-hot encoding).**

**6. Modeling**

* **Model Selection:**
  + **Baseline Model: Start with a baseline linear regression model using the key features identified.**
  + **Complex Models: Depending on the baseline model performance, explore more complex models like decision trees or random forests.**
* **Model Evaluation:**
  + **Train-Test Split: Split the data into training and testing sets.**
  + **Metrics: Evaluate model performance using metrics such as R-squared, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).**
  + **Compare Models: Compare the performance of different models to select the best one.**

**7. Results and Interpretation**

* **Summarize Key Findings:**
  + **Discuss which features are most influential in predicting price.**
* **Model Performance:**
  + **Report the performance of the chosen model(s) and discuss any limitations or considerations.**

**8. Conclusion and Future Work**

* **Overall Findings:**
  + **Summarize the overall findings of the project.**
* **Future Work:**
  + **Suggest potential areas for improvement or future research (e.g., collecting more data, exploring additional features, applying different modeling techniques).**

**9-Day Timeline Suggestion**

* **Day 1-2: Project Overview and Initial Data Inspection.**
* **Day 3: Correlation Analysis and Identification of Key Features.**
* **Day 4-5: Targeted Exploration and Feature Engineering.**
* **Day 6: Modeling - Baseline Model Development.**
* **Day 7: Modeling - Complex Models and Evaluation.**
* **Day 8: Results Interpretation and Summarization.**
* **Day 9: Conclusion, Reporting, and Future Work.**