### **Slide 1: Title Slide**

* **Title: Data Analysis and Feature Engineering on Property Dataset**
* **Subtitle: An Exploratory Data Analysis and Machine Learning Approach**
* **Your Name/Team Name**
* **Date**

### **Slide 2: Introduction**

* **Objective:**
  + **Explore and analyze a property dataset.**
  + **Handle missing data and outliers.**
  + **Perform feature engineering and encoding.**
  + **Build a predictive model for property price estimation.**
* **Dataset Overview:**
  + **Number of rows and columns: dataset.shape**
  + **Key features: Mention a few key columns (e.g., PropPrice, PropertySize, SaleYr).**

### **Slide 3: Data Inspection**

* **Key Insights:**
  + **Data types of features: Numerical and Categorical.**
  + **Presence of missing values and duplicates.**
* **Visual: Include a table or snippet from dataset.info() or dataset.head().**

### **Slide 4: Handling Missing Values**

* **Numerical Columns:**
  + **Filled using mean values (e.g., PropertyFrontage, ExteriorCladdingArea).**
* **Categorical Columns:**
  + **Filled using mode values.**
* **Outcome:**
  + **Null values reduced significantly.**
* **Visual: Bar chart or table showing the null values before and after handling.**

### **Slide 5: Outliers Detection and Removal**

* **Methodology: IQR Method**
  + **Calculated lower and upper bounds.**
  + **Removed rows containing outliers in numerical columns.**
* **Outcome: Cleaner dataset for analysis.**
* **Visual: Box plots before and after outlier removal.**

### **Slide 6: Feature Engineering**

* **New Features Created:**
  + **property\_age: SaleYr - YearBuilt**
  + **TotalSF: Combined square footage of floors and basement.**
  + **price\_per\_sqft: PropPrice / PropertySize**
* **Impact: Enhanced dataset for better predictive insights.**
* **Visual: Highlight new features in a table or chart.**

### **Slide 7: Encoding Categorical Data**

* **Techniques Used:**
  + **One-Hot Encoding for Nominal Columns (e.g., Neighborhood, SaleType).**
  + **Ordinal Encoding for Ordered Categories (e.g., ExterQual, Functional).**
* **Outcome: All categorical columns transformed into numerical formats.**
* **Visual: Example of encoded columns with charts or tables.**

### **Slide 8: Data Visualization**

* **Key Insights:**
  + **Distribution of PropPrice: Histogram with KDE.**
  + **Correlation analysis: Heatmap for PropPrice vs other features.**
  + **Relationship between PropertySize and PropPrice: Line plot.**
* **Visuals:**
  + **Histogram, heatmap, and line plot.**

### **Slide 9: Machine Learning Model**

* **Model: Random Forest**
  + **Trained using one-hot-encoded features and target PropPrice.**
* **Steps:**
  + **Data split into train (70%) and test (30%).**
  + **Model trained on x\_train and tested on x\_test.**
* **Visual: Model pipeline diagram or process flow.**

### **Slide 10: Model Evaluation**

* **Metrics:**
  + **R² Score: Display value.**
  + **Accuracy: Display percentage.**
  + **Mean Absolute Error (MAE): Display value.**
* **Outcome: Summarize performance of the model.**
* **Visual: Table or bar chart of metrics.**

### **Slide 11: Key Findings**

* **Handled missing values and outliers effectively.**
* **Created new features to improve model performance.**
* **Encoded categorical data for ML readiness.**
* **Built a Random Forest model with satisfactory performance.**

### **Slide 12: Conclusion and Next Steps**

* **Conclusion:**
  + **Dataset cleaned and enhanced for predictive analysis.**
  + **Random Forest provides a solid baseline for property price estimation.**
* **Next Steps:**
  + **Experiment with other models (e.g., Gradient Boosting, XGBoost).**
  + **Perform hyperparameter tuning for better results.**
  + **Deploy model in a real-world application.**

### **Slide 13: Q&A**

* **Title: Questions & Discussion**
* **Invite queries from the audience.**

### **Tips for Enhancing the Presentation:**

1. **Use visuals like graphs, tables, and flowcharts to represent data and processes.**
2. **Keep text concise and avoid overcrowding slides.**
3. **Add consistent branding (color scheme, font) for professional aesthetics.**

**Would you like help with creating visual elements or refining specific slides?**