

BRIDGING DATA SCIENCE AND REAL ESTATE:

PREDICTIVE MODELING FOR PROPERTY
VALUATION





AGENDA

- Background and Business Objectives- Paola Rodriguez
- Terminology, Data Mining Goals & Success Criteria, and Project Plan (Gantt chart) - Tommy Barron
- Data Description, Data Exploration and Data Quality- Austin Mallie

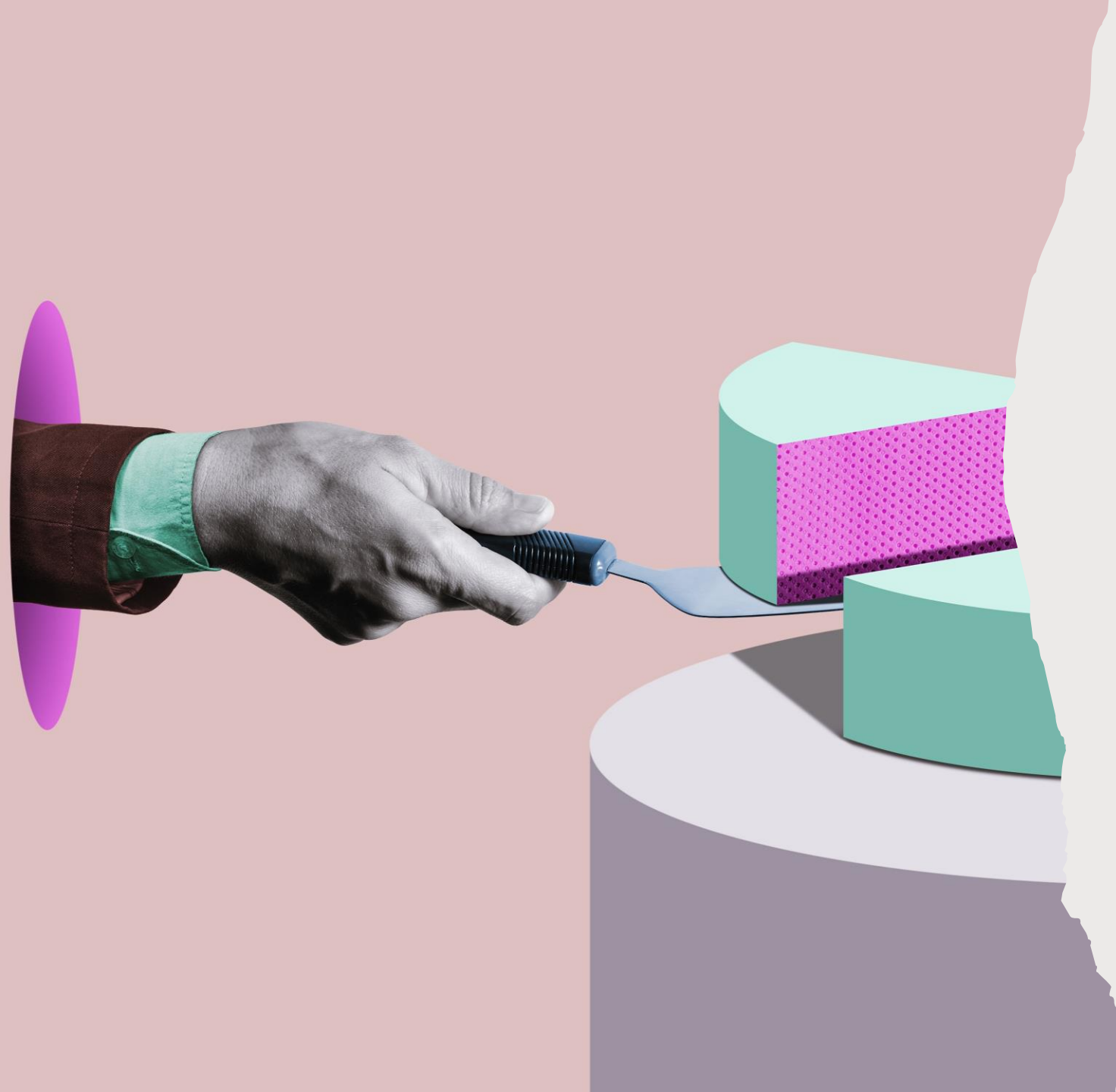
BACKGROUND

- Ames, Iowa housing market
- Leveraging machine learning and data analytics
- Our goal



BUSINESS OBJECTIVES

- Understanding the Data
- Addressing constraints, assumptions and considerations of the project



TERMINOLOGY

- Key terms used throughout the project
- Importance of clear communication





DATA MINING GOALS & SUCCESS CRITERIA

- KPIs
- Business Impact

PROJECT PLAN (GANTT CHART)



Three project phases



Timeline: 12-week
breakdown



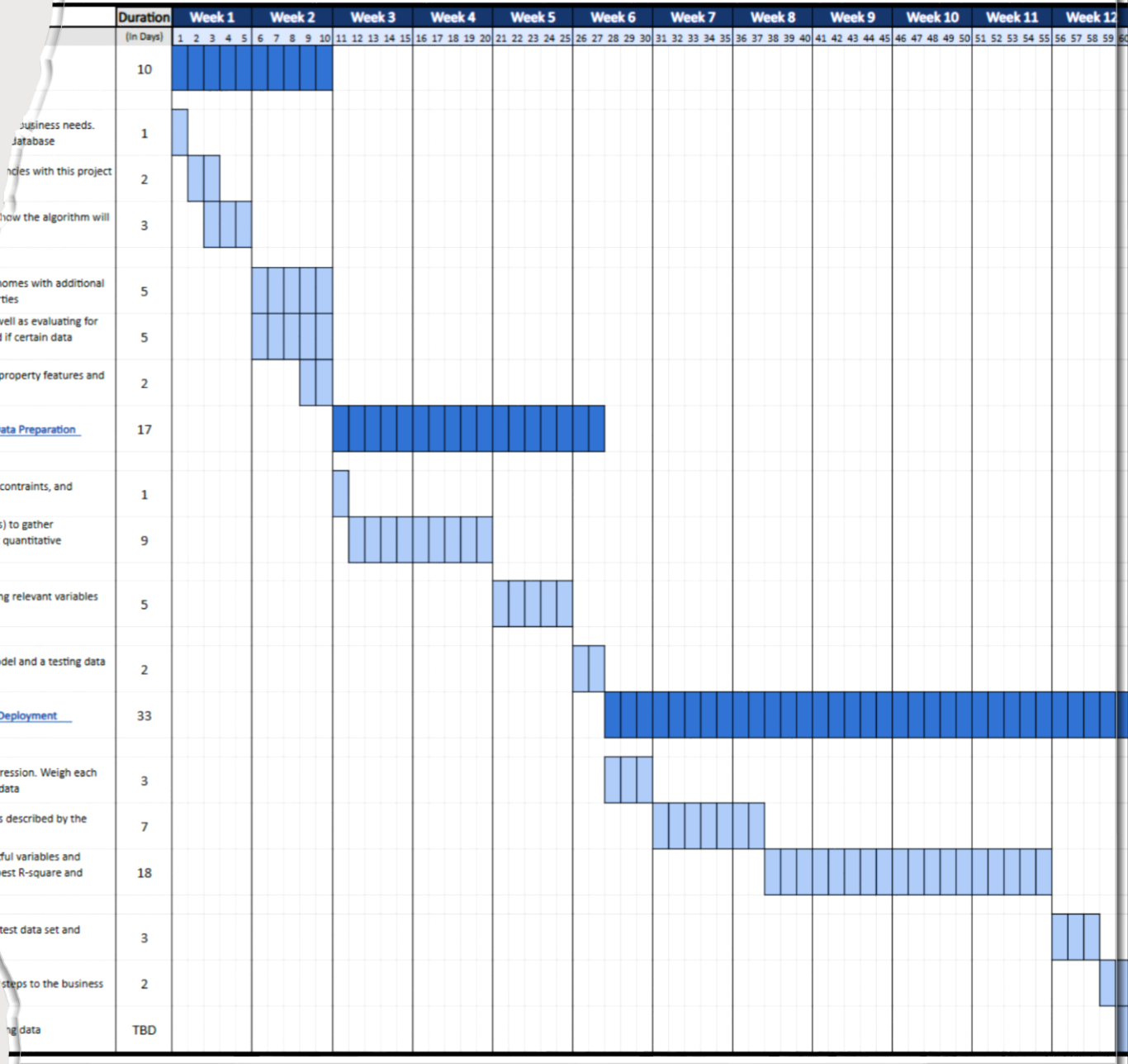
Key tasks: Data understanding, modeling, evaluation



Business and technical milestones



Deliverables and next steps



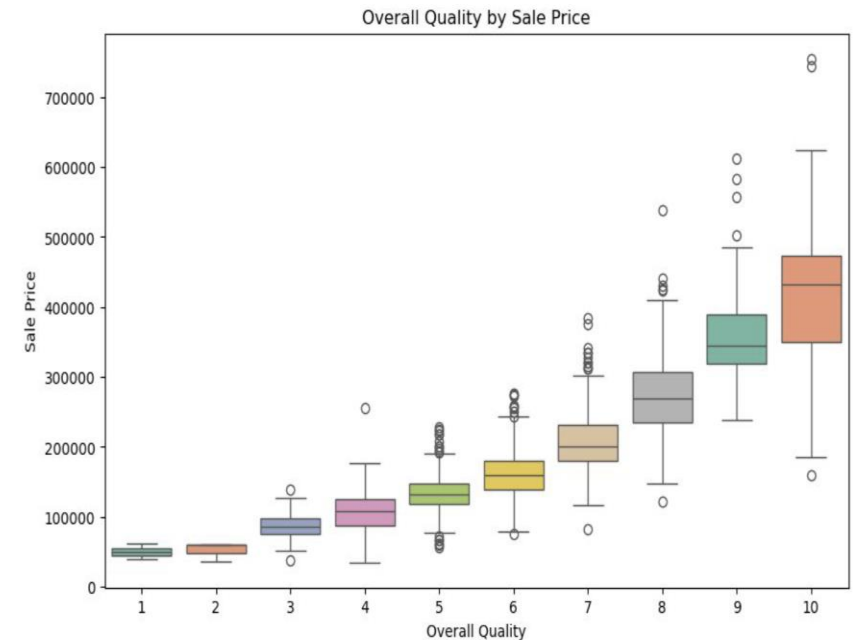


DATA DESCRIPTION

- Source: Ames Housing dataset
- Key attributes: 79 explanatory variables, target variable (SalePrice)
- Categories: Lot, Building, Sale conditions
- Types: Numerical and categorical variables
- Missing values and data quality

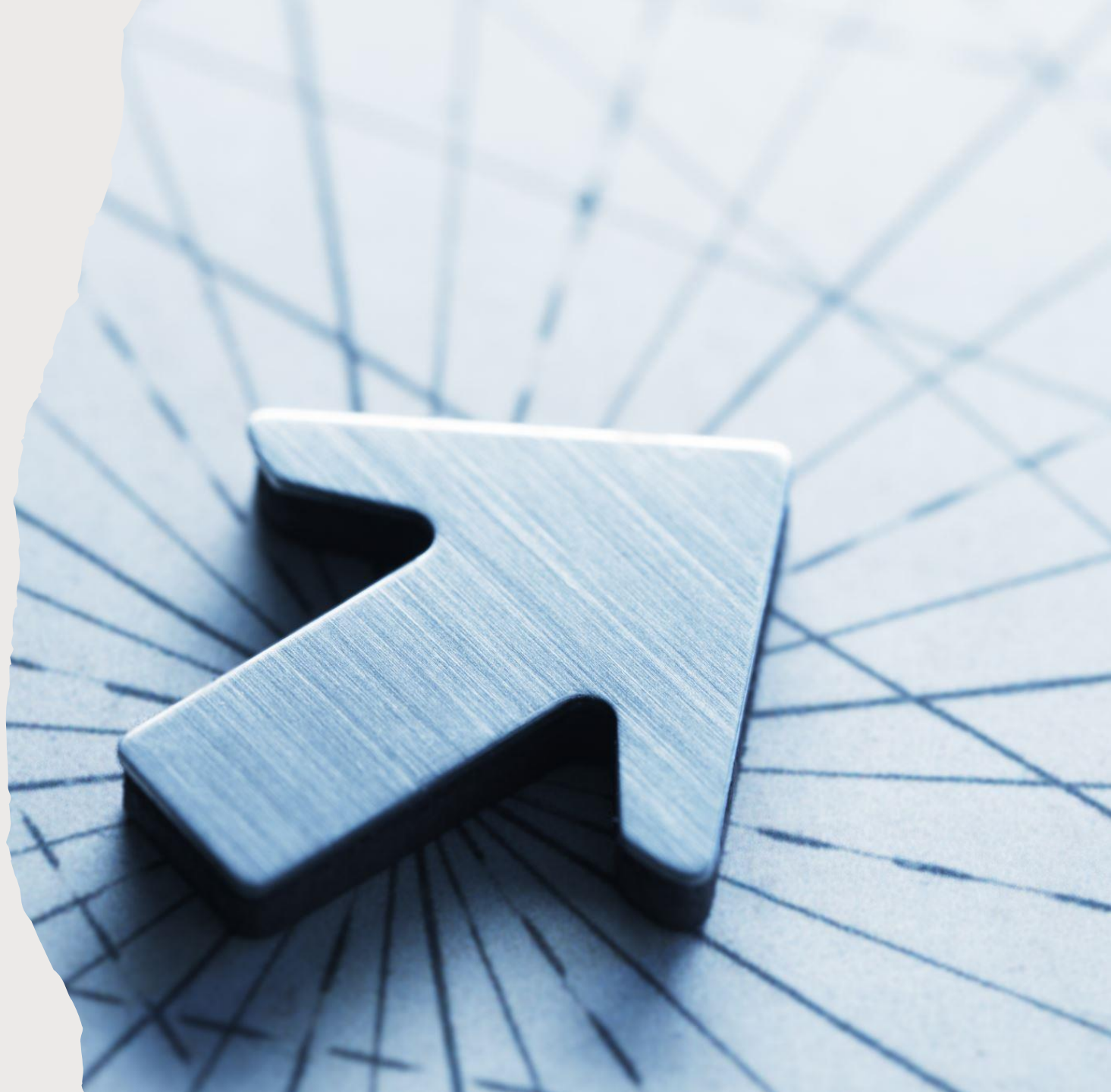
DATA EXPLORATION

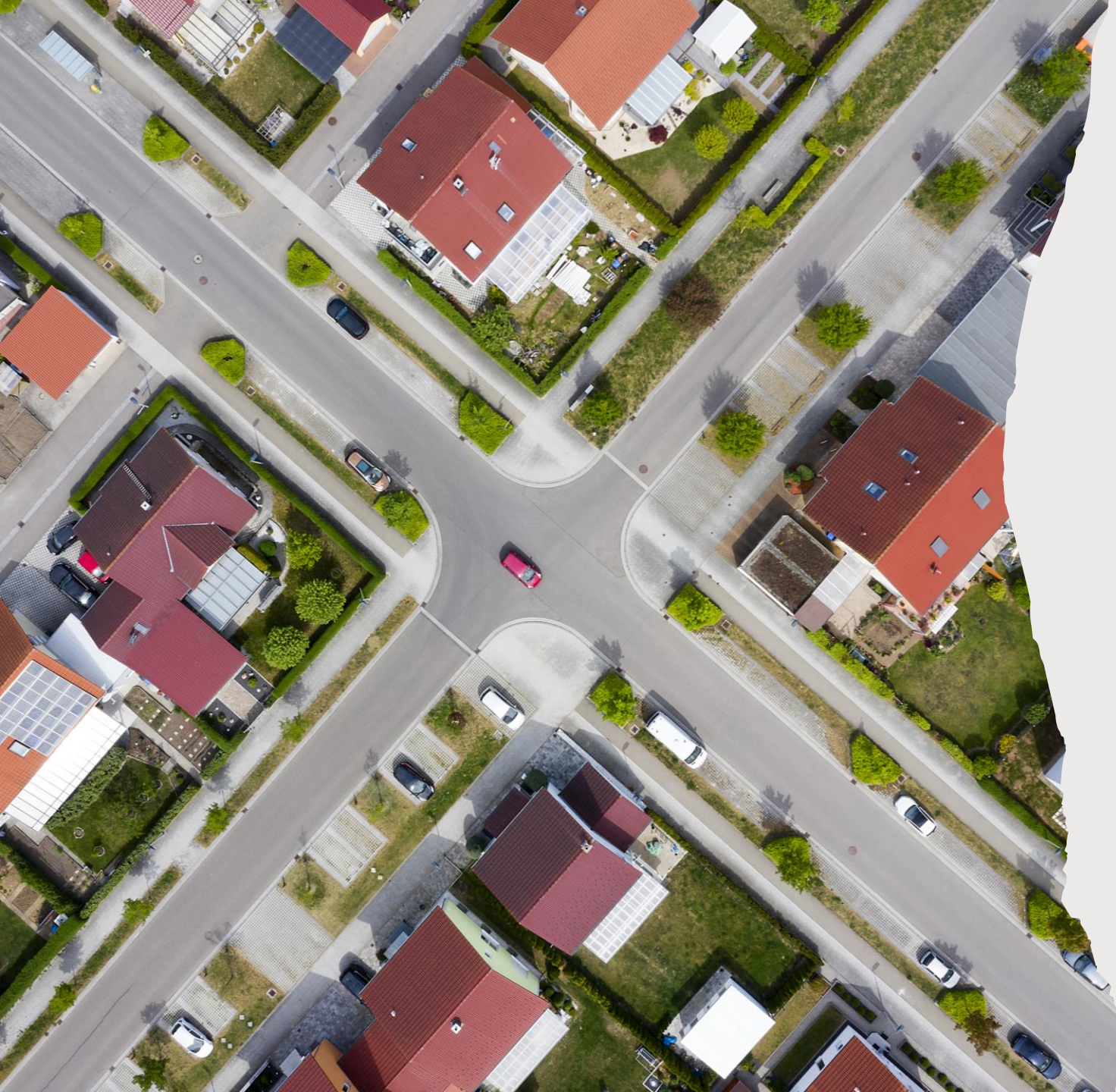
- Examination of key features and variables
- Identifying trends, patterns, and outliers in the data
- Data visualization techniques employed
- Initial hypothesis generation based on exploratory findings
- Insights from exploratory analysis



DATA QUALITY

- Assessing missing and inconsistent data
- Data cleaning and transformation processes
- Ensuring data accuracy, completeness, and reliability
- Impact of data quality on the model
- Steps to maintain data integrity throughout the project





CONCLUSION

- Recap of project goals and progress
- Importance of data quality and model integrity
- Actionable next steps for successful implementation
- Ongoing evaluation and client collaboration