BRIDGING DATA SCIENCE AND REAL ESTATE: search PREDICTIVE MODELING FOR PROPERTY VALUATION Time атан П BUSINESS



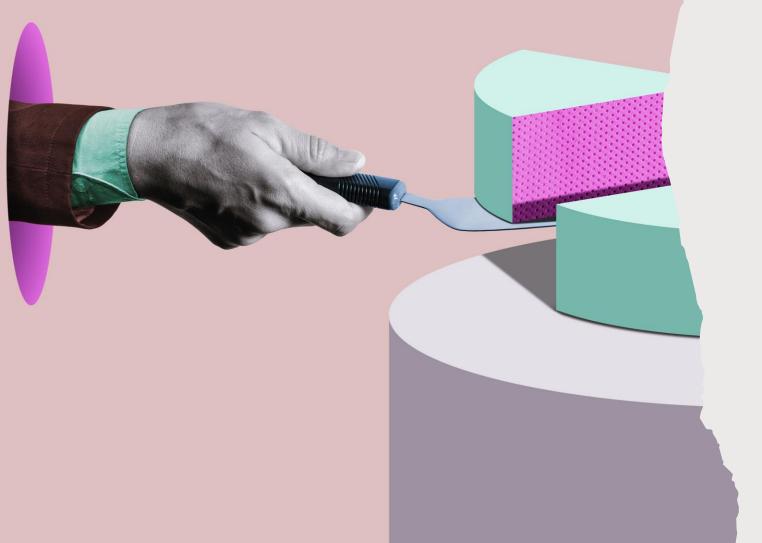
AGENDA

- 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

| + | Duration | - 14 | eek: | | Veek | | Weel | | Wee | | 14/0 | ek 5 | | 14/0 | ek 6 | 146 | eek 7 | , | - 14 | /eek | | Veek | | lant. | 10 | Mank | :11 | Wee |
|---|-----------|------|------|--|------|--|------|--|-----|--|------|------|-------|------|------|-----|-------|---|------|------|--|------|------|-------|-------|------|-----|-----|
| | (In Days) | | | | | | | | | | | | 25 26 | | | | | | | | | | | | 49 50 | | | |
|) | 10 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Jusiness needs. Jatabase | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ncies with this project | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| , how the algorithm will | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| / sold homes with additional . properties | 5 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ness as well as evaluating for understand if certain data | 5 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| nations of property features and | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| rsis, and Data Preparation | 17 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| locument contraints, and | 1 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| tate agents) to gather that aren't quantitative | 9 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| n identifying relevant variables | 5 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| for the model and a testing data | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| nent, and Deployment | 33 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| linear regression. Weigh each ct on the data | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ry variables described by the type | 7 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| nd impactful variables and with the best R-square and | 18 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| with test data set and | 3 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| steps to the business | 2 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ng data | TBD | | | | | | | | | | | | | | | | | | | | | | | | | | | |

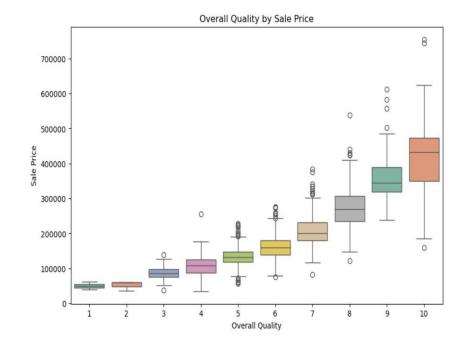


DATA DESCRIPTION

- 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

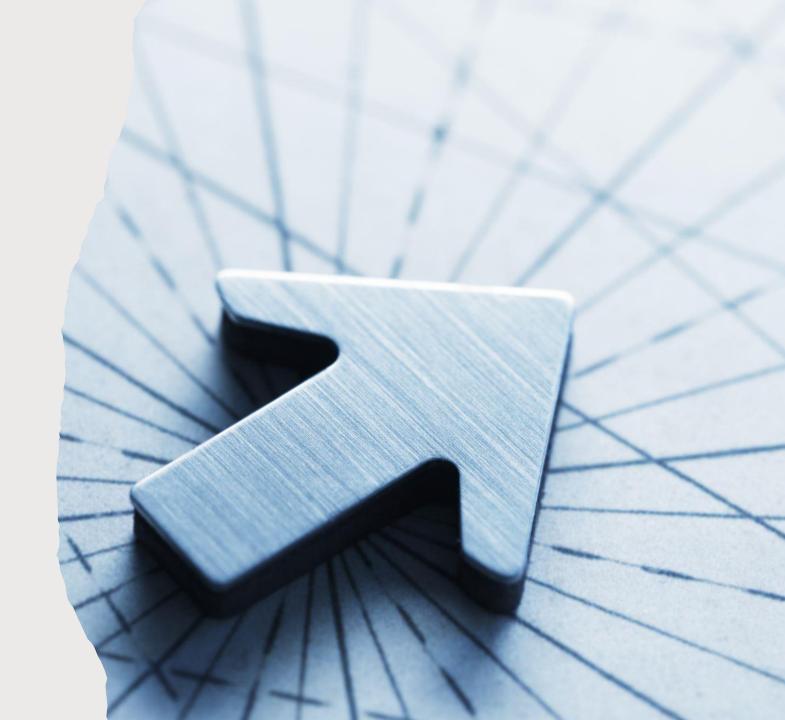
DATA EXPLORATION

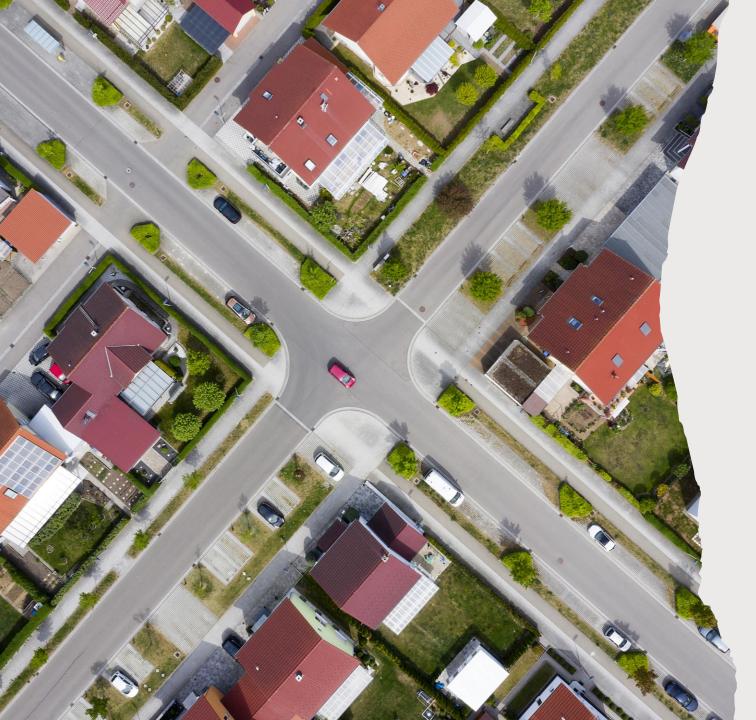
- Examination of key features and variables
- Identifying trends, patterns, and outliers in the data
- Data visualization techniques employed
- Initial hypothesis generation based on exploratory findings
- Insights from exploratory analysis



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