Teaching Case
California Dreamin' Housing Market Visualizations in Tableau

Aimee Jacobs, Jacquelin J. Curry, Concetta A. DePaolo, and Fernando Parra

Find archived papers, submission instructions, terms of use, and much more at the JISE website:
https://jise.org

Teaching Case
California Dreamin' Housing Market Visualizations in Tableau

Aimee Jacobs
Jacquelin J. Curry
Craig School of Business
California State University Fresno
Fresno, CA 93740, USA
ajacobs@csufresno.edu, jacquelinc@csufresno.edu

Concetta A. DePaolo
Scott College of Business
Indiana State University
Terre Haute, IN 47809, USA
cdepaolo@indstate.edu

Fernando Parra
Craig School of Business
California State University Fresno
Fresno, CA 93740, USA
parra@csufresno.edu

ABSTRACT

This manuscript describes the use of real data applied to a fictional real-estate firm for teaching data visualization to university students. In the case study, students employ data analytic techniques in Tableau to clean, organize, and analyze real estate data. By creating visualizations, students address several questions about how selling prices of homes are affected by various factors such as location, industry trends, and property characteristics. The case has been used in business analytics courses; students reported finding the case study relevant and useful, and they were found to largely meet the learning goals of the case, including proficiency with cleaning and filtering data, and creating clear and useful visualizations to convey meaning in Tableau.

Keywords: Data analytics, Data visualization, Tableau, Data mining, Data exploration

1. CASE SUMMARY

Orchard Grove Realty, a real estate firm in the San Joaquin Valley, California, is committed to providing exceptional customer service and assisting clients in finding their ideal properties. Orchard Grove aims to analyze real estate market trends and key indicators to better serve its clients. In pursuit of this goal, Orchard Grove has access to a large, updated, and complex real estate market data set spanning thirteen years, covering various property attributes and transaction details. Orchard Grove is particularly interested in leveraging data analysis and visualization techniques using Tableau (2023) to gain insights into market trends and assist clients in making informed decisions. Students, acting as data analysts, will need to apply all stages of the process from data cleansing to

reporting, including exploration, mining, visualization, and model building. After becoming familiar with the case background and data set, students are asked to cleanse the data (see Section 3.2) and to create visualizations that address questions about the status and trends in the real estate market (see Section 4).

2. CASE BACKGROUND

Orchard Grove Realty is a full-service real estate firm located in the San Joaquin Valley in California. Orchard Grove prides itself on providing excellent customer service and a desire to assist customers in finding the ideal property. Orchard Grove Realty offers a range of comprehensive services to meet the diverse needs of its clients in its service area. The firm is

interested in analyzing trends regarding real estate market valuations and other key relevant indicators to better serve its clients more effectively. For example, Orchard Grove can assist sellers in determining the optimal selling price for their property. This ensures that sellers receive the best value for their investments while attracting numerous potential buyers. Orchard Grove Realty's market value estimations also provide buyers with a clear understanding of the current market conditions, helping them identify properties that align with their budget and expectations. Data analysts and scientists use visualizations to tell the story of the data. Orchard Grove Realty also understands the importance of technology in the real estate industry. The firm leverages advanced tools and technologies to enhance their market analysis capabilities. By utilizing data-driven approaches and innovative software, Orchard Grove can provide even more accurate estimations and insights into market trends, ensuring that clients receive the most up-to-date and reliable information.

Orchard Grove maintains a large data set covering 13 years of real-world measures that relate to real estate market value including days on market, list price, and selling price among others. Orchard Grove has asked you, as one of their experienced data analysts, to help generate real estate metrics and visualizations in Tableau to provide more market insight. As such, the overall purpose of your analysis will be to describe the valuation of real estate property trends by developing a Tableau Storyboard. You will report your findings to your supervisor, the real estate broker, in the form of a Tableau storyboard with visuals outlining your analysis and recommendations. You may be asked to write a report or present your findings to your boss.

3. CASE DATA

3.1 Data Description
The data set consists of 96,625 residential properties sold during 2008-2021. The variables and definitions that will be used in the case can be found in Table 1. *Note:* Your data file may not have all of the variables so you may have to create calculated fields for any missing variables.

3.2 Data Cleansing
Before beginning the data analysis process, it is important to undertake data cleansing techniques to improve data quality and remove errors or inconsistencies. The data should be normalized by converting different units of measurement into a standardized unit. For example, some homeowner association dues are expressed as monthly costs while others are yearly.

All data errors should be corrected through a process of identifying and rectifying erroneous data, in particular, addressing extreme or invalid values. Any inconsistencies or conflicts within text variables should be resolved. For example, the homeowner association frequency variable (HOA Freq) might say dues are collected "Monthly" or "Per Month" and so those text entries should be edited for uniformity. Furthermore, it is recommended to review the variable names to ensure consistent naming conventions.

4. ASSIGNMENT

As one of Orchard Grove's experienced data analysts, you are tasked with creating visualizations to address the following real estate questions in order to provide a market analysis of your data set. Students could also be directed to various real estate YouTube instructional videos or specific websites in order to better understand real estate terminology, current issues and concerns. Some website examples include but are certainly not limited to the following: https://www.opendoor.com/articles/real-estate-terms-you-should-know (Opendoor Team, 2019) or https://www.investopedia.com/mortgage/real-estate-investing-guide (The Investopedia Team, 2022).

Database Variable	Description
Selling Price	Price of the property when sold
APN	Assessor's Parcel Number - identification number assigned to parcels of property by the tax assessor of a particular jurisdiction
ML_Number	Multiple Listing Service number
# of Fireplaces	Number of fireplaces in the property
Address - ZIP	Postal code used to provide a location of a property to the United States Postal Service (USPS)
Total Bathrooms	Number of bathrooms on the property
Bedrooms	Number of bedrooms on the property
Financing Desc	Type of financing used by the buyer to purchase the property
Garage Spaces	Number of spaces for cars to fit in the garage
HOA Dues	Homeowner Association fee charged at a given frequency
HOA Freq	Frequency of payments to the Homeowner Association
Listing Price	The initial suggested selling price of the property
Listing Date	Date property is listed on the market for sale
Selling Date	Date property is sold
Lot Size	Size of the land according to boundary lines
Lot Type	The type of lot size, i.e., acres or square feet
Pool	Whether the property has a pool or not, Y=Yes/N=No
Roofing Desc	Type of roof on the property
School District	School district where the property is located
Solar Display	Whether the property has a solar or not, Y=Yes/N=No
Square Footage	Size of the property in square feet
Year Built	Year the property was first built
DOM	Number of days the property was on the market before being sold
SP%LP	Ratio of selling price to listing price

Table 1. Variables and Definitions

4.1 Case Questions

1. What is the median selling price for your data set? Why do we use median and not mean selling price?
2. What characteristics do homes with higher selling prices have in common?
3. What proportion of homes were sold based on number of bathrooms?
4. How does a pool affect a home's days on market and selling price?
5. What is the average percent of selling price over listing price?
6. What is the average price per square foot?
7. How did market growth change across the years? Explain any drastic changes.
8. Is there a difference in selling price based on financing, e.g., cash, mortgage, etc.?
9. How do selling prices vary by the school district? Would you recommend clients move into a particular school district?
10. What periods of the year do attributes have an impact on home prices? (for example, pool, # of fireplaces, solar, etc.)?
11. How many days are comparable houses (possibly sqft., bathrooms, bedrooms, etc.) staying on the market?
12. What are the median selling price trends? Add a forecast.
13. When is the best time of the year to sell your house? When is the best time to buy?
14. How has COVID-19 impacted the local market?
15. Did you find anything interesting or unusual in your data set?

4.2 Visualization Requirements

Your dashboards, charts, and storyboard should:

- Utilize the data from your data source appropriately;
- Show evidence of thoughtful and effective design within worksheets and between worksheet elements of dashboards and stories;
- Convey a meaningful message to the viewer;
- Provide the user with an opportunity to explore the visuals and make their own discoveries; and
- Organize your workbook so it follows the appropriate hierarchy considering the story sequence.

4.3 Deliverables

Submit your visuals in a single packaged Tableau workbook (.twbx) file, so the file contains both the workbook properties and an extract of the data used in the workbook.

Within your workbook, sequence the tabs so the story you want the viewer to explore comes first. Place any supporting worksheets or dashboards after these presentation tabs. Do not hide the supporting worksheets, just place their tabs to the right of your dashboards and storyboard tabs. Ensure that each of the tabs has descriptive names.

If you are asked to prepare a report, then provide the following information:

1. Create a professionally written report that discusses each of the tabbed items in your workbook by name.
2. Describe the main message(s) of your analysis and final recommendation that you intend the viewer to receive from your storyboard.

3. Report how you incorporated the elements of data visualization in your work.
4. Detail any opportunities for the viewer to use selection and filtering to explore the visuals.
5. Add any additional aspects of your workbook that you'd like to mention.

If you are asked to prepare a Tableau presentation, then follow these guidelines: All data analyst team members should understand how the Tableau workbook views were created and how the workbook functions. You may be asked to recreate an element in the submitted workbook based on your client's feedback.

5. REFERENCES

Opendoor Team. (2019, October 17). *52 Essential Real Estate Terms You Should Know.* Opendoor. https://www.opendoor.com/articles/real-estate-terms-you-should-know

Tableau Software. (2023). https://www.tableau.com/

The Investopedia Team. (2022, June 28). *A Beginner's Guide to Real Estate Investing.* Investopedia. https://www.investopedia.com/mortgage/real-estate-investing-guide/

AUTHOR BIOGRAPHIES

Aimee Jacobs is an associate professor in data analytics. Dr. Jacobs holds a BSc in Management Information Systems from Indiana State University, MBA from Indiana State University, and PhD in Informatics from the University of Reading – UK. Since joining California State University, Dr. Jacobs has been involved with studies related to digital and social technologies, statistics education, and teaching with technology, and data analytics.

Jacquelin J. Curry earned a Bachelor of Science degree in business administration with an option in real estate and land economics from Fresno State. She subsequently earned a Juris Doctor from San Joaquin College of Law in Clovis, California. Dr. Curry's scholarly pursuits encompass a spectrum of research interests including areas of real estate valuation; environmental and sustainability issues in real estate; ethics, inclusion, and equity, as well as business pedagogy and education including data analytics.

Concetta A. DePaolo is a professor of operations & supply chain management in the Scott College of Business at Indiana State University. She earned her master's and Ph.D. in Operations Research from Rutgers University, and an undergraduate degree in Mathematical Sciences from Worcester Polytechnic Institute. She became a Certified Analytics Professional (CAP®) in 2015. Her research interests include optimization, statistical methods, statistics education, and teaching with technology. She has taught statistics, business analytics and management science at the Scott College since 2000.

Fernando Parra, an associate professor of accountancy and Director of the Institute for Family Business at Fresno State, prioritizes experiential learning and aligns his teaching with current professional standards. His research centers on pedagogy, digital transformation's impact on businesses, and socio-economic factors.

INFORMATION SYSTEMS & COMPUTING ACADEMIC PROFESSIONALS
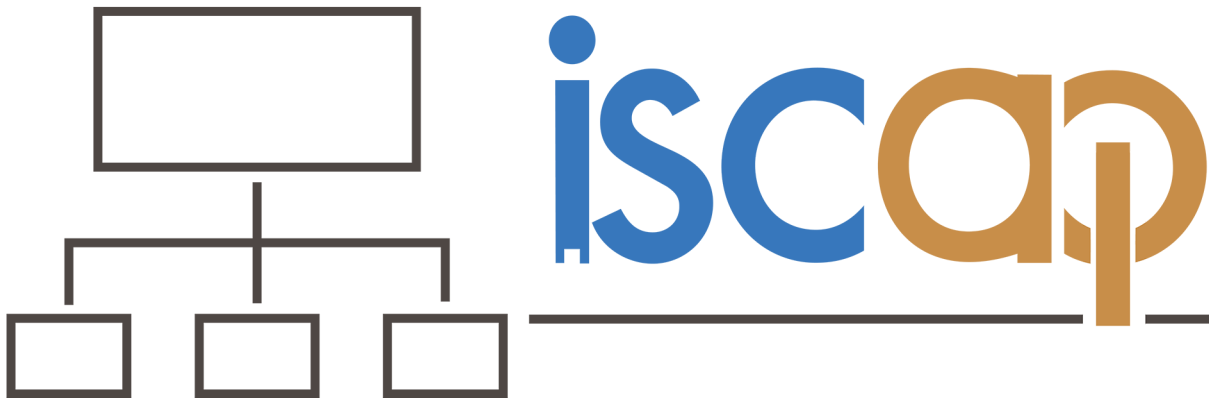
STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the *Journal of Information Systems Education* have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.