Exploratory Data Analysis and Contract

Section 1: Contract

Group number: 003E02

Name: Nicole Agnel Paikkatt

Github Link: https://github.sydney.edu.au/rden7271/003E02.git

I agree to:

- Abide by the terms of this contract in relation to the group assessment for DATA2002/2902.
- Store all my written and code contributions to the assessment in the GitHub repository.
- Keep a record of my other contributions to the assessment (e.g. discussions, emails, meetings attended). A copy of this may be requested by the coordinator.
- General Conduct:
    a. Respectful Treatment: Group members are expected to treat each other with respect and dignity at all times, regardless of gender, religion, race, sexuality, or disability, in accordance with the University's commitment to fostering an inclusive and respectful community.
    b. Honesty and Ethical Behavior: Honesty and ethical conduct are paramount in all interactions within the group and with the University. Members must act honestly and ethically in all dealings with the University and fellow group members, promoting a culture of integrity.
    c. Bullying, Harassment, and Discrimination: Members are prohibited from engaging in bullying, harassment, or discriminatory behavior, including on social media platforms. This commitment aligns with the University's policies on bullying, harassment, and discrimination prevention.
    d. Cooperation and Collaboration: Maintaining a cooperative and collaborative approach to relationships with group members, teaching staff, and University affiliates is essential for the group's success and the promotion of a positive learning environment.
    e. Equitable Use of Resources: Group members should use University resources equitably, responsibly, and with consideration for the needs of others, ensuring that access and use are not impeded. This includes adherence to relevant University rules and policies governing resource use.
    f. Academic Integrity: Members must commit to a culture of academic integrity, adhering to all applicable policies, procedures, and unit of study outlines. Academic honesty, ethical conduct in academic matters, and the promotion of high academic standards are fundamental expectations.
    g. Collaboration and Equitable Contribution: In group work, projects, and other learning experiences, members should work collaboratively and contribute equitably. This reflects the University's emphasis on cooperative learning and the advancement of knowledge.
    h. Conflict resolution: In the case of conflict with group member opinions, group members are expected to conduct themselves in a mature manner and issues will be discussed in an open discussion. If the conflict cannot be resolved internally, external support from tutors and unit coordinator may be utilized.
- Communication Protocol: Group communication will primarily take place through the following channels:
    a. Facebook Messenger: For quick updates, notifications, and informal discussions.
    b. Google Drive: For collaborative document editing and sharing of project-related files.
    c. Personal and University Email Accounts: For official announcements and additional communication, when necessary.
    d. Face-to-face communication is facilitated during class tutorials and scheduled meetings to enhance effective discussions. Additional group meetings will be arranged on an ad hoc basis.
    e. Any proposed changes to this communication process must be initiated through a formal email request addressed to all group members, outlining the proposed replacement method. These proposed changes will be further discussed and negotiated during weekly meetings to ensure consensus and efficiency in our communication strategies.

- Communication Frequency: These expectations are established to ensure that group members can effectively collaborate and stay informed while allowing for flexibility to accommodate individual schedules and preferences.
  a. Facebook Messenger: Each group member is expected to check Facebook Messenger at least once every 24-48 hours. Additionally, there is a group responsibility that ensures at least one member responds to another member's communication within 24 hours.
  b. Email Communication (Official Announcements): Official announcements sent via email will allow group members up to a maximum of 48 hours to check and respond if necessary.
  c. Google Drive Communication (Updates and Edits): Group members are expected to check Google Drive communication, including updates and edits, at least weekly. A preference is given for more frequent checks, ideally semi-weekly, to facilitate collaborative work and timely responses.
  d. In-person communication (catchup on other members work): Group members are expected to attend an in person meeting minimum once a week. The time for the meeting will be scheduled at the beginning of the week open to negotiation and schedule of each member. In case of conflict, an online Zoom meeting will be conducted instead.

I understand that:

- My agreement to these terms is indicated through the act of submitting this document in Canvas.
- If I fail to meet my obligations as detailed in this group contract then I have failed to meet the assessment requirements for DATA2002/2902 and may be awarded a mark of zero for some or all of the project components.

Section 2: Exploratory Data Analysis:

Data set: Housing Prices
Dependent variable: Land Value
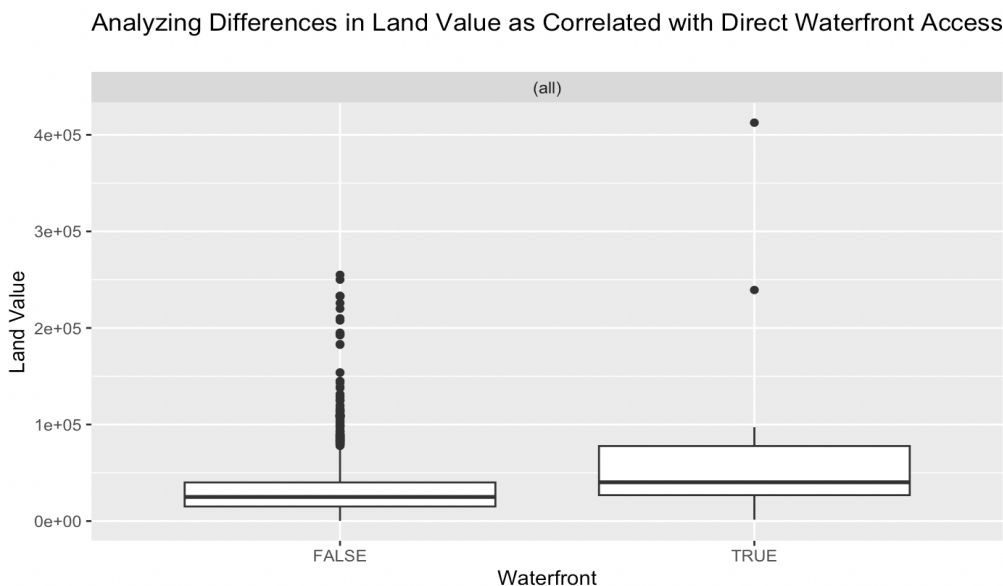
Section 2.1: Initial Visualizations

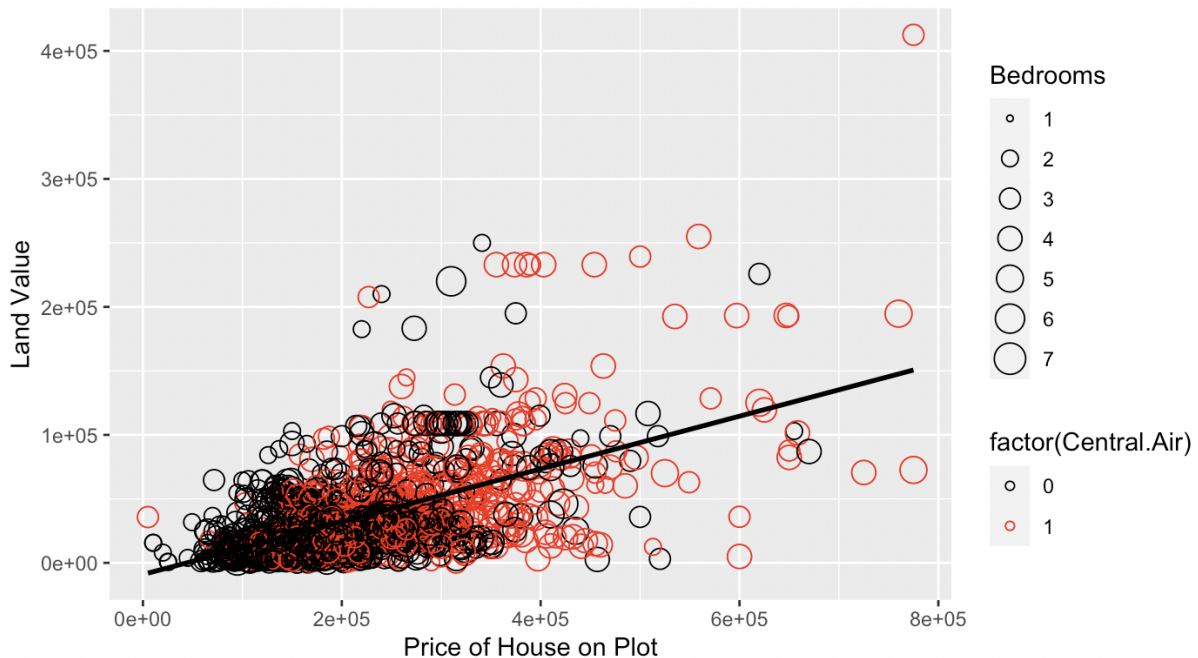*Fig 1: Side by Side Boxplot for variables "Waterfront" and " Land.Value"*

Fig: Scatter Plot of Land Value vs Price, Colour: Central.Air, Size: Bedrooms

Section 2.2: Summary

Fig 1: The median of the value of land with direct waterfront access is clearly higher than the median of the value of land without direct waterfront access, which means it is a suitable factor to predict land value.

Fig 2: The land value appears to be higher for those plots with houses that are more expensive. From the visualization alone, it seems like the houses with central air conditioning are more likely to be built on land with higher value. Bedroom size, from a naive approach, should have been an indicator of area. But clearly, that alone isn't an indication of land value because we can see that some of the pieces of land with the highest values do not necessarily have a large number of bedrooms.

Conclusion: Some of the best predictors for Land Value from initial Analysis are:
1) Waterfront Access
2) Price of the house which is built on the land in question
3) Centralized Air Conditioning