

Statistics for Decision Making

Coding Questions

Instructions:

- These problems will test your ability to manipulate real-world data and answer statistical questions.
 - Use the property.csv dataset to arrive at the solutions.
 - The data is on real estate prices in Australia. Use this data to test the following hypothesis:
1. For the suburb of Altona, it is postulated that a typical property sells for \$800,000.
 - Use the data at hand to test this assumption.
 - Is the typical property price really \$800,000 or has it increased?
 - Use a significance level of 5%.
 2. For the year 2016, is there any difference in the prices of properties sold in the summer months vs winter months?
 - Consider months from October till March as winter months and rest as summer months.
 - Use a significance level of 5%.
 3. For the suburb of Abbotsford, what is the probability that out of 10 properties sold, 3 will not have a car parking space?
 - Use the column car in the dataset.
 - Round off your answer to 3 decimal places.
 4. In the suburb of Abbotsford, what are the chances of finding a property with 3 rooms? Round your answer to 3 decimal places.
 5. In the suburb of Abbotsford, what are the chances of finding a property with 2 bathrooms? Round your answer to 3 decimal places.
 6. One-Sample Hypothesis Test (Industry Pricing) A real estate firm claims that the average property price in Richmond is \$1,000,000. Using the dataset, test whether the actual average price is significantly different from this claim at a 5% significance level. Clearly state:
 - Null and alternative hypotheses
 - Test statistic
 - p-value
 - Final business conclusion

7. Independent Two-Sample T-Test (Feature Impact)

Do properties with car parking sell at a higher average price than properties without car parking, across the entire dataset? Use a 5% significance level and justify:

- Choice of test
- Interpretation of p-value
- Business implications for developers

8. Two-Way ANOVA (Location & Property Type)

Investigate whether property prices are influenced by:

- Suburb
- Type of property
- Interaction between suburb and property type

Use Two-Way ANOVA and explain which factors significantly affect price.

9. p-Value Interpretation (Decision Making)

A hypothesis test comparing prices across two suburbs results in a p-value of 0.032.

Answer:

- What does this p-value indicate?
- Should the null hypothesis be rejected at $\alpha = 0.05$?
- How should a business stakeholder interpret this result?

10. Industry-Style Hypothesis Validation (Policy Decision)

A housing policy group believes that properties with more than 2 bathrooms command a premium price.

Design and execute a statistical test to validate this claim:

- Identify the correct test
- State hypotheses
- Report p-value
- Give a clear recommendation to policymakers

Submission Instructions:

- Jupyter Python file(.ipnyb)
- Artifacts generated need to be submitted in vLearn on or before the deadline.
- File Name: File name: firstname_lastname_CPDA_Batch.ipnyb
E.g., Kartik_Mudaliar_CPDA_B1.ipynb