Final Project — Milestone 2

Milestone 2

**Output Summary Report**

**Name**: Rupali S. Metkari  
**Date**: 31st March 2021

Output

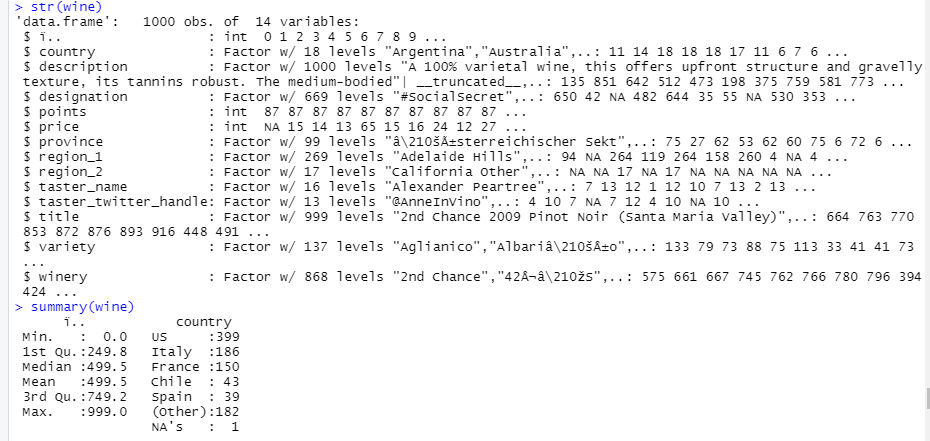


Fig (1)

First, we load Wine\_tasting.csv data. This data contains total 1000 obs. and 14 variables. In this data describe wine taste. Using points and price variable we did hypothesis tasting.

Hypothesis testing is an act in statistics whereby an analyst tests an assumption regarding a population parameter. The methodology employed by the analyst depends on the nature of the data used and the reason for the analysis.

Hypothesis testing is used to assess the plausibility of a hypothesis by using sample data. Such data may come from a larger population, or from a data-generating process. The word "population" will be used for both cases in the following descriptions.

According to given instruction we would to perform two types of t.test . one sample t.test and two sample t.test.

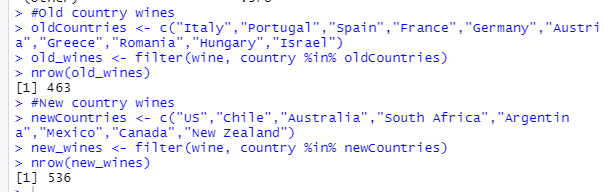


Fig (2)

In Fig (2), we can see create two subsets old countries and new countries from the original dataset. With using three variables country, price, points.

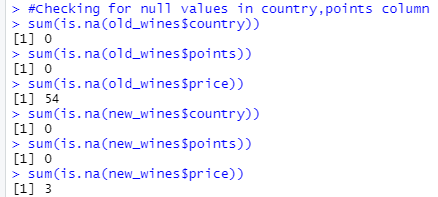


Fig (3)

After that check null values in both subsets. In Fig (3) shows that in price variable null values is exist so next step is these null values replace by median.

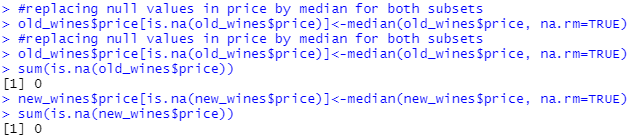


Fig (4)

In Fig (4) shows replace null values by median for both countries. Now all data is clean.

For hypothesis testing and inferential statistics take two questions are:

1. According to points old countries wines are tasty.
2. According to price old countries wines are costly.

**Two sample t.test- based on points:**

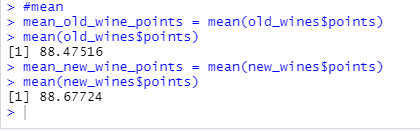


Fig (5)

In Fig (5) we can see that mean of two countries are different. We take significance value is α= 0.05.

Now assume two scenarios:

H0: Old countries wines are tasty.

H1: New countries wines are tasty.

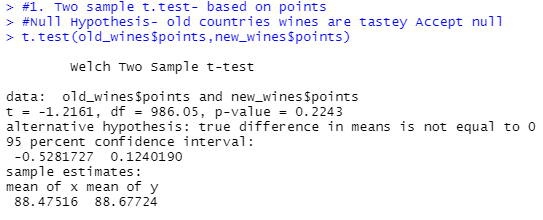


Fig (6)

T- statistics: t = -1.2161, df = 986.05, p-value = 0.2243

for significance value α= 0.05 We accept Null hypothesis as p-value > 0.05 because p value is 0.2243 is greater than significance value. So, it is concluded that “Old countries wines are tasty.”

**Conclusion:** p-value is greater than the significance value so we reject alternative hypothesis and accept null hypothesis.

**One sample t.test- based on price:**

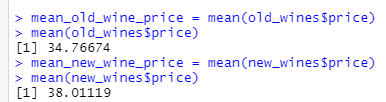


Fig (7)

Old countries mean is 34. 76 and new countries mean is 38.011. Here find two mean. For t.test compare new countries mean to old countries.

Scenario is: Old countries wines are more expensive.

1. H0  is Null hypothesis: Old countries wines are more expensive. (μ old = μ new )
2. H1  is Alternative hypothesis: New countries wines are more expensive. (μ old ≠ μ new )

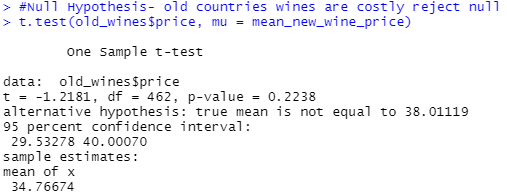


Fig (8)

Here we can see that true mean μ old = 34.76 is not equal to μ new = 38.011, so we reject null hypothesis. And Alternative hypothesis H1 is accepted. Menace old countries wines are not expensive. New countries wines are expensive.

According to me the climates of New country wine regions are often warmer, which tends to result in riper, more alcoholic, full-bodied, and fruit-centered wines. These wines are often made in a more highly extracted and oak-influenced style.

References:

1. Allan G. Bluman Published by McGraw-Hill Education, 2 Penn Plaza, New York, NY 10121. Copyright © 2018 by McGraw-Hill Education. *Elementary Statistics A Step-by-Step Approach, 10th edition*.
2. Hypothesis Testing By CHRISTINA MAJASKI Reviewed By PETER WESTFALL Updated Oct 24, 2020, Investopedia : <https://www.investopedia.com/terms/h/hypothesistesting.asp>
3. Two-Sample-T-test, Chew Jian Chief, 2014, ISIXSIGMA: <https://www.isixsigma.com/tools-templates/hypothesis-testing/making-sense-two-sample-t-test/>
4. The Two-sample t-test, July 2019, jmp: <https://www.jmp.com/en_us/statistics-knowledge-portal/t-test/two-sample-t-test.html>
5. Hypothesis testing and p-values, 2019, Khan Academy: <https://www.khanacademy.org/math/statistics-probability/significance-tests-one-sample/more-significance-testing-videos/v/hypothesis-testing-and-p-values>
6. T Test in R: One Sample and Paired (with Example), May 2018, Guru99: <https://www.guru99.com/r-t-test-one-sample.html>