

National University of Singapore
TCX2002 Introduction to Business Analytics
Tutorial 2

Lesson 4 - Preparing and Comparing Data: t-tests, ANOVA, and Hypothesis Testing

1. One-Sample t-test

Scenario: A coffee chain claims that its average daily sales is \$500, in the past month. You are asked to verify this.

Dataset:

A vector of 30 daily sales figures:

```
sales <- c(595, 447, 477, 423, 570, 547, 541, 490, 593, 557, 533, 482, 516, 590,  
476, 499, 419, 509, 430, 479, 502, 511, 496, 471, 554, 596, 485, 534, 414, 443)
```

1. State the null and alternative hypotheses.
2. Use `t.test(sales, mu = 500)` to test the claim.
3. What is the p-value? Do you **reject or fail to reject** the null hypothesis?

2. Paired Sample t-test

Scenario: A 10-day marketing campaign was launched. You want to test whether customer online visits to the online shop increased during the campaign vs the 10 days before.

Dataset:

```
before <- c(125, 153, 148, 113, 139, 123, 106, 146, 119, 172)  
after <- c(140, 165, 154, 122, 148, 194, 164, 152, 200, 197)
```

1. Run a paired t-test using `t.test(before, after, paired = TRUE)`.
2. Interpret the confidence interval and p-value.
3. What business decisions might be made from this result?

National University of Singapore
TCX2002 Introduction to Business Analytics
Tutorial 2

3. ANOVA: Comparing More Than Two Groups

Scenario: A manager wants to compare average satisfaction scores (about of 5 star rating) across 3 branches: North, South, and Central, during the past 4 weeks of summer holidays.

Dataset:

```
branch <- factor(c(rep("North",4), rep("South",4), rep("Central",4)))  
score <- c(3.2, 3.5, 3.7, 3.2, 4.0, 4.2, 3.9, 3.6, 3.7, 3.6, 3.4, 3.9)
```

1. Run `aov(score ~ branch)` and check `summary()`.
2. What is the F-statistic and p-value?
3. What do the results imply for how customers perceive service across branches?

4. Comparing weight loss in diet plans

Scenario: A nutritionist wants to compare the average weight loss (in kg) for three different diet plans (A, B, and C) over a month. She randomly assigns 5 people to each plan. Their weight losses are:

Dataset:

```
diet <- factor(rep(c("A", "B", "C"), each = 5))  
loss <- c(2.1, 2.5, 2.0, 2.9, 2.4, # Diet A  
          3.0, 3.2, 2.8, 3.5, 3.1, # Diet B  
          1.8, 2.0, 2.2, 1.9, 2.1) # Diet C
```

1. Is there a significant difference in mean weight loss among the three diets?
2. Report the F-statistic and p-value.
3. What is your conclusion

National University of Singapore
TCX2002 Introduction to Business Analytics
Tutorial 2

5. Risk management software in financial trading

Scenario: A financial analyst wants to determine if a new risk management software has reduced the number of trading errors in a brokerage firm. She records the number of errors made by 10 traders in the month **before** and **after** the software was implemented:

Dataset:

```
before <- c(8, 6, 7, 9, 10, 5, 8, 7, 6, 9)
after  <- c(5, 4, 5, 6, 8, 3, 6, 5, 4, 6)
```

1. Has the risk management software significantly reduced trading errors
2. Provide the p-value and interpret the result.

Tutorial 2 Learning Outcomes	
1.	Hypothesis Testing: A structured method for making decisions about population parameters based on sample data, while controlling the risk of incorrect conclusions.
2.	t-Test Purpose: Compare means between groups or against a known value to determine if differences are statistically significant # clarify significance is statistical (not due to random chance).
3.	ANOVA (Analysis of Variance): Statistical method to compare means across three or more groups simultaneously, while controlling for the overall error rate.