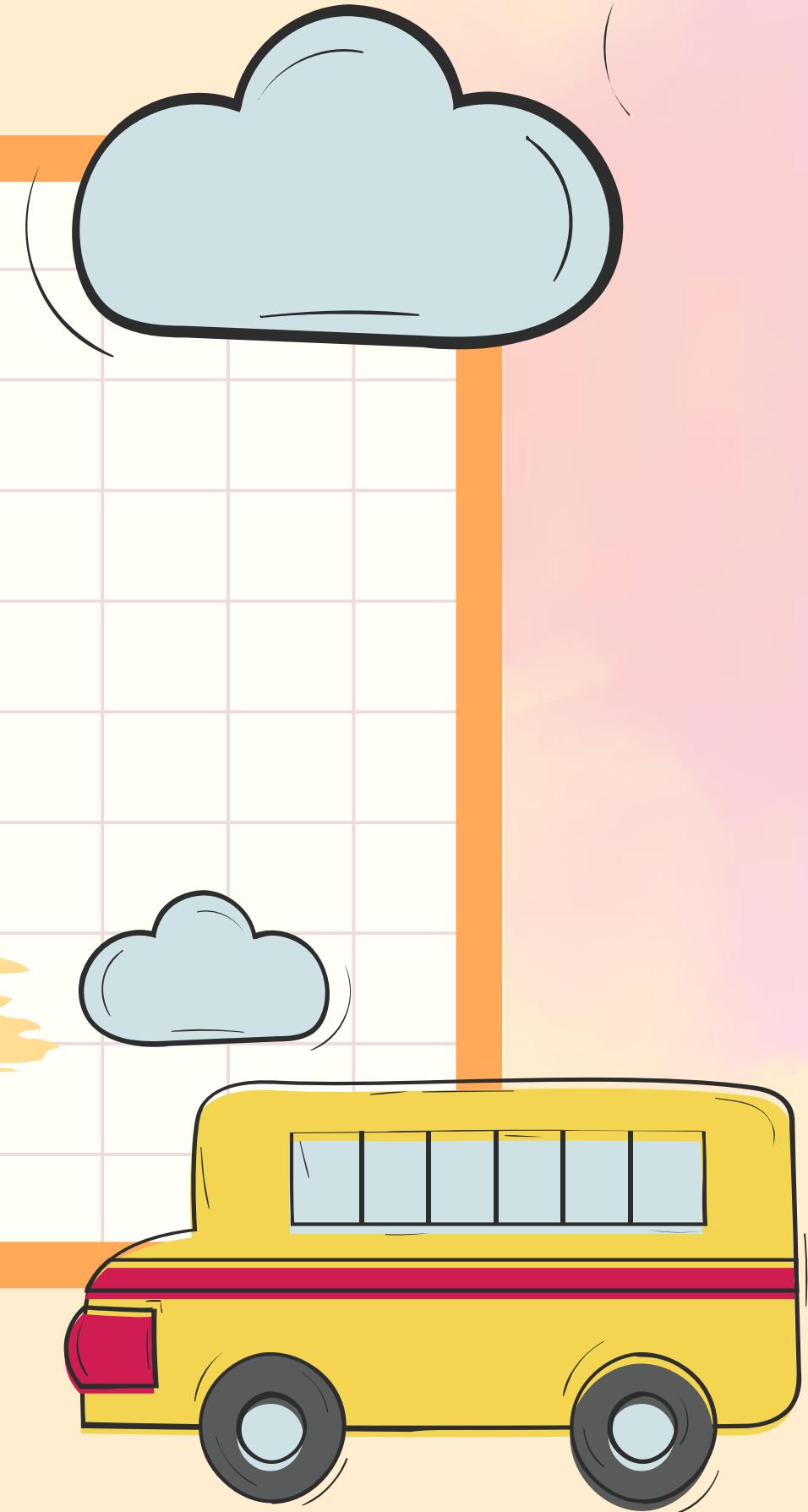


# Group Project

Presented By kon na ta dee Group



# Group Member



Thanyanan Khampooal  
65102010160



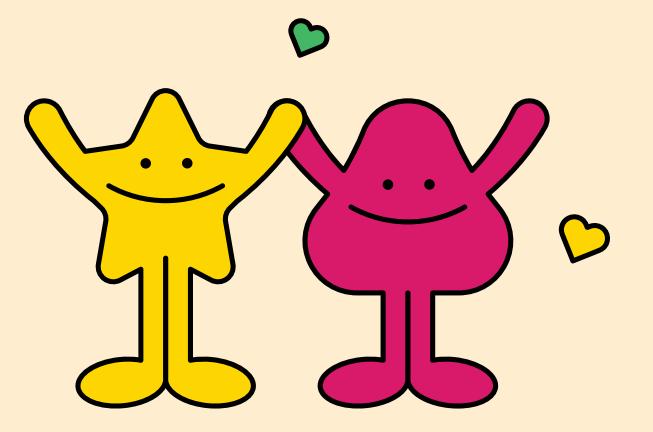
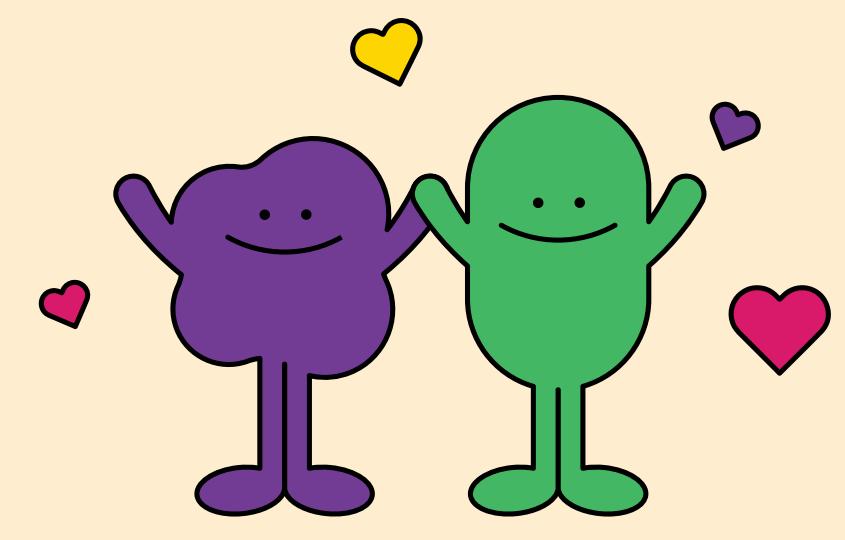
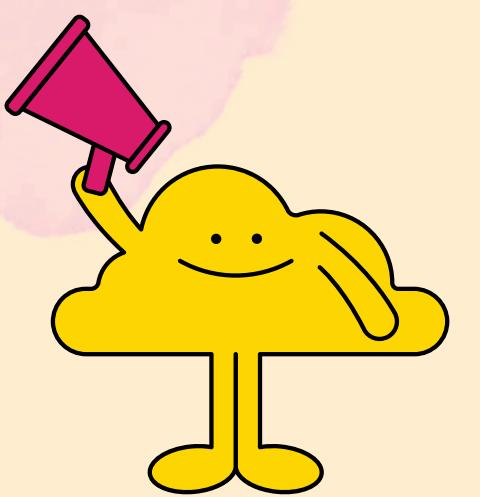
Phantipha Sripadupwong  
65102010208



Werapat Phasit  
65102010210



Thanis Kleebrakson  
65102010688



# Contents

- 
- 1 Preface
  - 2 Introduction
  - 3 Data Collection
  - 4 Data Exploration
  - 5 Normality Test
  - 6 Anova
  - 7 Conclusion
  - 8 Reference

# Preface

This report is part of a course focusing on hotel statistics. In learning statistics for hotel management, we will become familiar with applying statistics to analyze data on room reservations, pricing adjustments, and customer satisfaction. Understanding statistical concepts used in the hotel and tourism industry will empower us as managers to make effective decisions in complex situations.

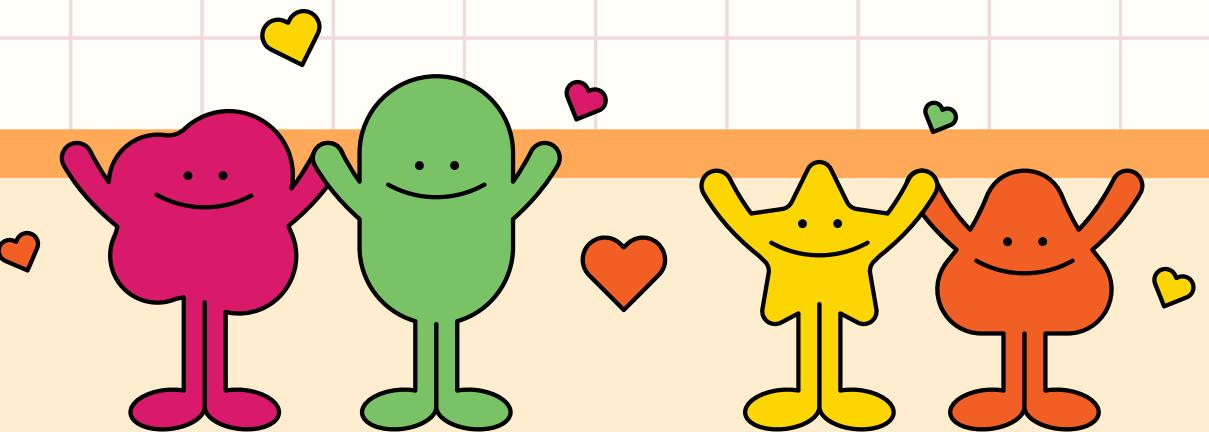
I hope that this report will be particularly beneficial to those studying. If there are any errors in this report, I apologize for them here.



# Introduction

Online Travel Agencies (OTAs) such as Agoda, Traveloka, and Booking aim to provide users with a convenient and diverse platform for comparing hotel prices, making informed decisions, and efficiently booking accommodations. These platforms prioritize user experience, offer competitive pricing, and ensure the security of user data, fostering a seamless and satisfying hotel booking process for travelers.

Our objective is to determine whether there is a population-level difference in the average nightly price of beach pool villas among three Online Travel Agencies (OTAs) – Agoda, Traveloka, and Booking. This involves exploring the data, plotting histograms to assess normal distribution, conducting Q-Q plots for distribution insight, and examining confidence intervals to estimate population parameters. The process includes formulating null hypotheses ( $H_0$ ) and alternative hypotheses ( $H_a$ ). Additionally, the utilization of ANOVA and chi-square charts aids in comprehending the data collection methodology, helping us understand how data was gathered and facilitating the formulation of hypotheses for subsequent statistical testing.



# DATA Collection

I

The data collected comprises information from three websites: Agoda, Traveloka, and Booking compile prices for each hotel in an excel file. The data collection period is from 23 August 2023 to 22 October 2023.

Date	Pool Villa name	Agoda	Traveloka	Booking.com
8/23/2023	Marrakesh Hua Hin Resort&spa	3,897	4,431	4,587
8/24/2023	Avila Resort Pattaya	4,722	5,200	5,525
8/24/2023	Sea Two Pool Villa	8,602	8,935	9,934
8/24/2023	Manhattan Pattaya Hotel	6,800	7,407.72	4,231
8/25/2023	Turtle's Cove	8,331	8,742	9,000
8/26/2023	Royal Cliff Beach Hotel Pattaya	6,800	25,051.27	16,242
8/27/2023	The Gems Mining Pool Villas Pattaya SHA Extra Plus	4,827	25,221.54	25,919
9/1/2023	KW pool villa pattaya	13,925	15,800.13	16,900
9/1/2023	Exquisite Pool Villa K	20,561	22,006.24	22,000
9/1/2023	Baba Beach Club Hua Hin ChaAm Luxury	23,444	32,766.87	38,535
9/1/2023	The X10 Nordic Tent and Glamping Pool Villa	5,509	7,500	7,237
9/1/2023	Paradise Pool Villa Pattaya in Tropicana Village	12,310	12,052.57	13,172
9/2/2023	La Miniera Pool Villas Pattaya	22,590	24,345.33	36,400

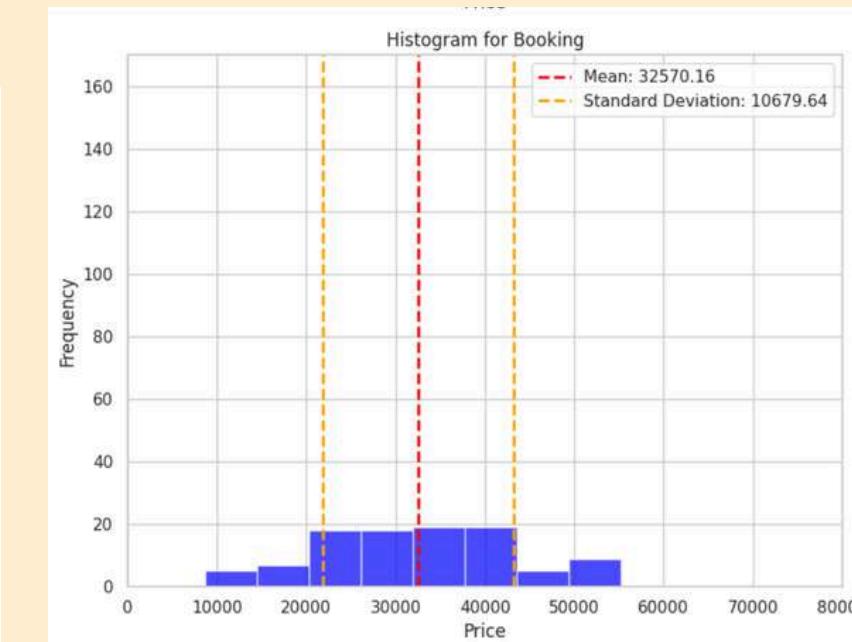
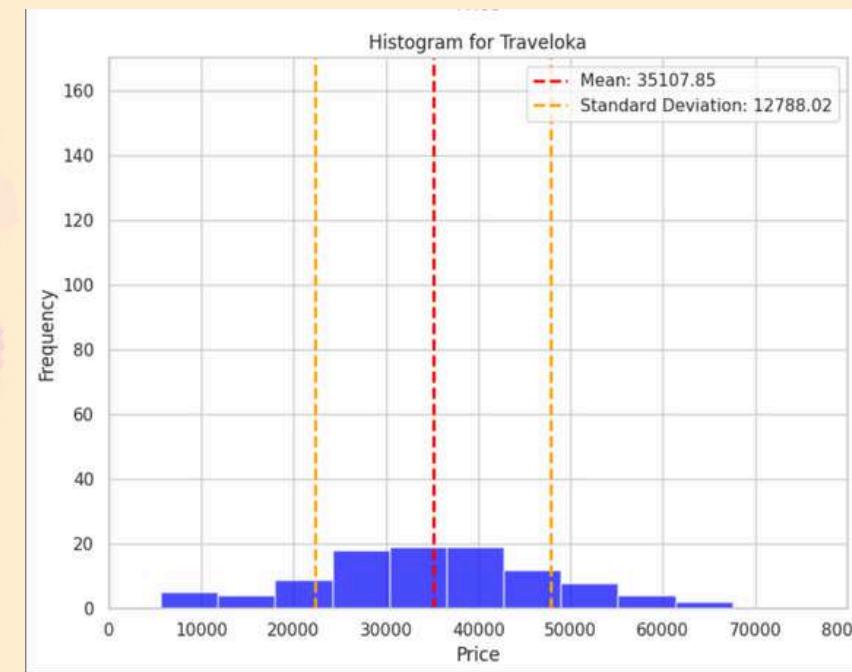
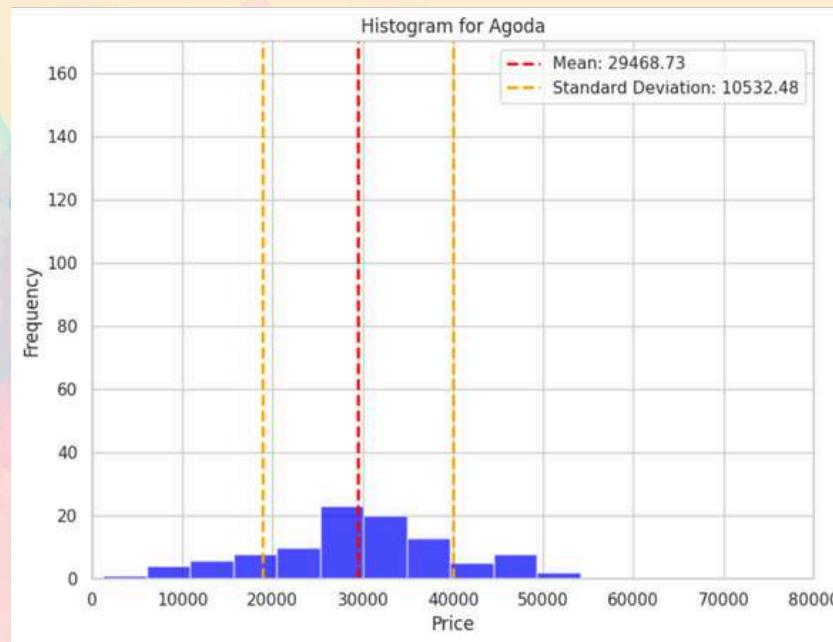
[Project282\(3\).xlsx \(sharepoint.com\)](#)



# Data Exploration

To create a histogram depicting price distributions for each OTA using 5 bins and provide statistical values like mean and standard deviation

- Agoda has a mean value of 29468.73 and standard deviation of 10532.48
- Traveloka has a mean value of 35107.85 and standard deviation of 12788.02
- Booking has a mean value of 32570.16 and standard deviation of 10679.64



(quantile-quantile plot)

# Normality Test

This graph is used to visually represent the distribution of data. It facilitates a comparison between the actual data and a specific distribution of interest, often exemplified by a normal distribution. If the data aligns along a 45-degree diagonal line, it indicates a correspondence with the targeted distribution. Other patterns or deviations in the graph may suggest a divergence from the expected distribution.



The Q-Q plot analysis suggests that our data doesn't fully meet the criteria for a normal distribution. While some data points align with the expected curve, there are notable deviations in other areas, leading us to refrain from confidently asserting normality. The limited size and low diversity of our dataset, coupled with the non-dynamic nature of price information for pool villas collected for the next year, contribute to the complexity of assuming normal distribution characteristics. Therefore, it is prudent to exercise caution when making assumptions about the normality of our dataset.

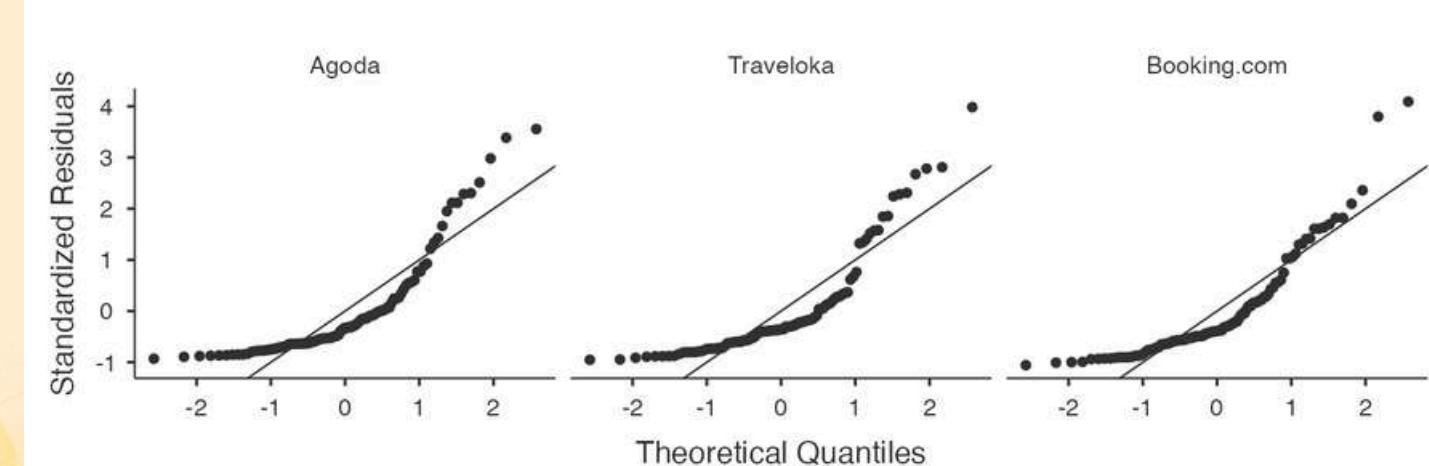
## Proportion Test (N Outcomes)

### Proportions - OTA

Level		Count	Proportion
Agoda	Observed	556855	0.260
	Expected	715041	0.333
Traveloka	Observed	714725	0.333
	Expected	715041	0.333
Booking.com	Observed	873544	0.407
	Expected	715041	0.333

### $\chi^2$ Goodness of Fit

$\chi^2$	df	p
70130	2	<.001



(Analysis Of Variance)

# Anova

ANOVA or Analysis of Variance is a statistical technique employed to investigate variations in the averages of multiple groups, determining if noteworthy differences are present among these groups. By utilizing data from various groups, ANOVA dissects the observed variance between and within groups. This analysis helps scrutinize and quantify the existing distinctions.

# (Analysis Of Variance)

## Anova

7

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	501458107	2	250729054	6.28517761	0.00212113	3.02615337
Within Groups	1.1848E+10	297	39892119.1			
Total	1.2349E+10	299				

The ANOVA test looks at differences in means across different groups. In this case:

- The variation between groups is 501,458,107.4, and we have 2 degrees of freedom for this.
- The variation within groups is 11,847,959,363, and there are 297 degrees of freedom within the groups.

The F-value, which compares the differences between and within groups, is 6.2851, and the associated p-value is 0.002121128. This suggests that there are significant differences between at least two of the groups.

## (Analysis Of Variance)

## Anova

Anova: Single Factor					
SUMMARY					
Groups	Count	Sum	Average	Variance	
Agoda	100	556855.47	5568.5547	25272229	
Traveloka	100	714724.78	7147.2478	41362393.9	
Booking.com	100	873543.52	8735.4352	53041734.3	

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	501458107	2	250729054	6.28517761	0.00212113	3.02615337
Within Groups	1.1848E+10	297	39892119.1			
Total	1.2349E+10	299				

The ANOVA results for Agoda, Traveloka, and Booking show no significant differences among their mean values. The p-value (0.002) exceeds the 0.05 significance level, leading to a failure to reject the null hypothesis.

Therefore, we cannot conclude that there are meaningful distinctions in mean values among the three groups



# Conclusion

In summary, the study on price disparities of beach pool villas among OTAs (Agoda, Traveloka, and Booking) indicates that the p-value exceeds the alpha level, leading to the inability to reject the null hypothesis. This suggests there is insufficient statistical evidence to support differences in prices of beach pool villas among Agoda, Traveloka, and Booking within a 80% confidence interval.

Moreover, the collected data does not exhibit a normal distribution, as observed in the Q-Q plot graph.

Furthermore, it's important to note that this project, while providing valuable insights, represents just one aspect of a complex landscape. Continuous refinement of data collection methods and a nuanced approach to analysis are crucial for a more comprehensive understanding of the factors influencing pricing strategies in the dynamic realm of online travel agencies.

# Reference

หลักการใช้ chi-square

<https://www.youtube.com/watch?v=Zs8IA3K1-4I>

หลักการทำ jamovi

[https://youtu.be/0mlvdN-Wh\\_A?feature=shared](https://youtu.be/0mlvdN-Wh_A?feature=shared)

Thank  
You

Bye bye!!