



# AIRLINE PASSENGER SATISFACTION

СТАТИСТИЧНЕ ВИВЕДЕННЯ

TEAM FOUR  
ГАРМАШ ОЛЕКСІЙ  
ХОК МАКСИМ  
КУЦАЛАБА НАЗАРІЙ  
МАСЛОВ НАЗАР

# ПОМИЛКИ ЛР1

01 ПОМИЛКИ

02 ПОМИЛКИ

03 ПОМИЛКИ

04 ПОМИЛКИ

05 ПОМИЛКИ

06 ПОМИЛКИ

07 ПОМИЛКИ

08 ПОМИЛКИ

# ВИСНОВКИ ЛР1

01 ВИСНОВКИ

02 ВИСНОВКИ

03 ВИСНОВКИ

04 ВИСНОВКИ

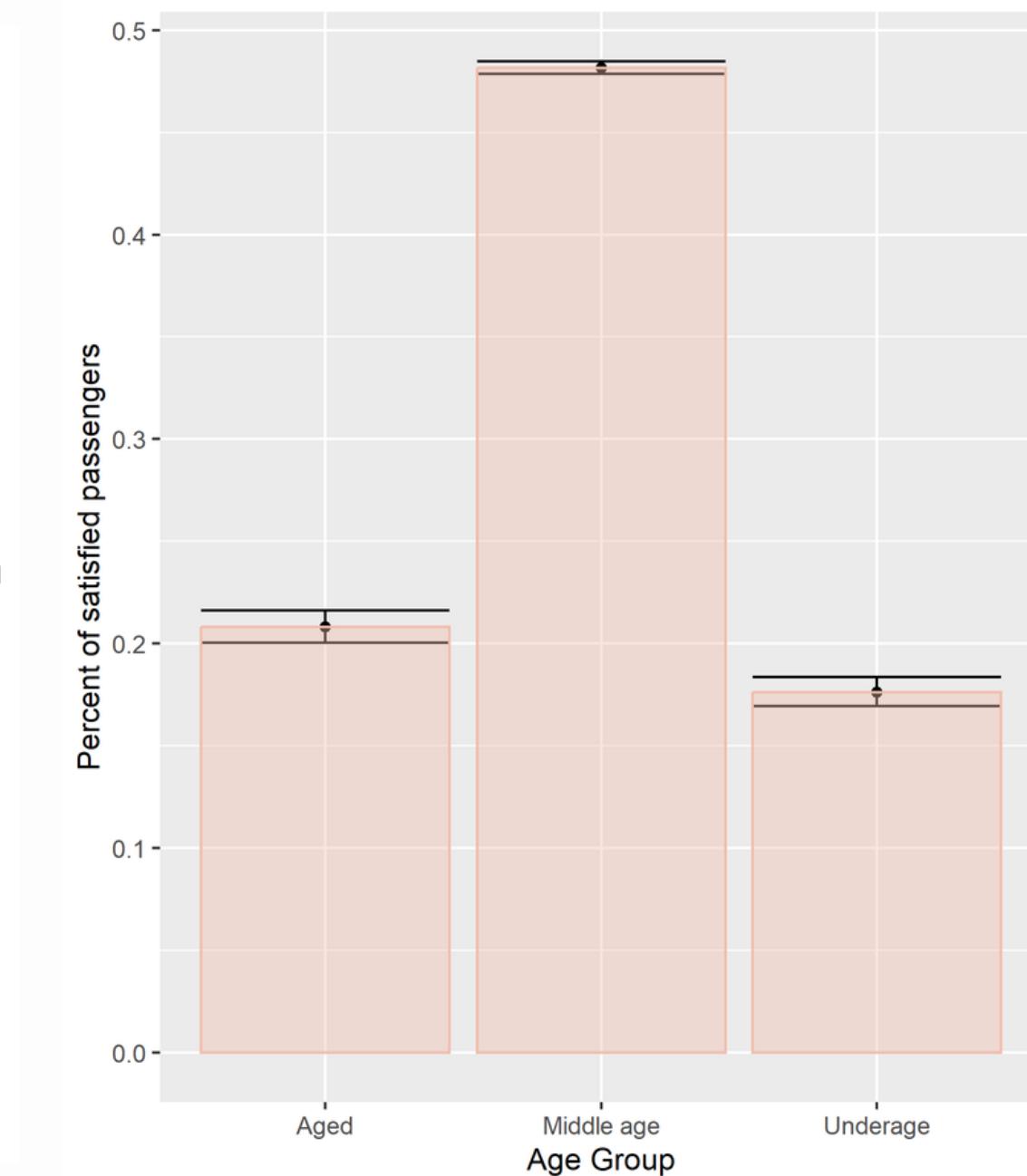
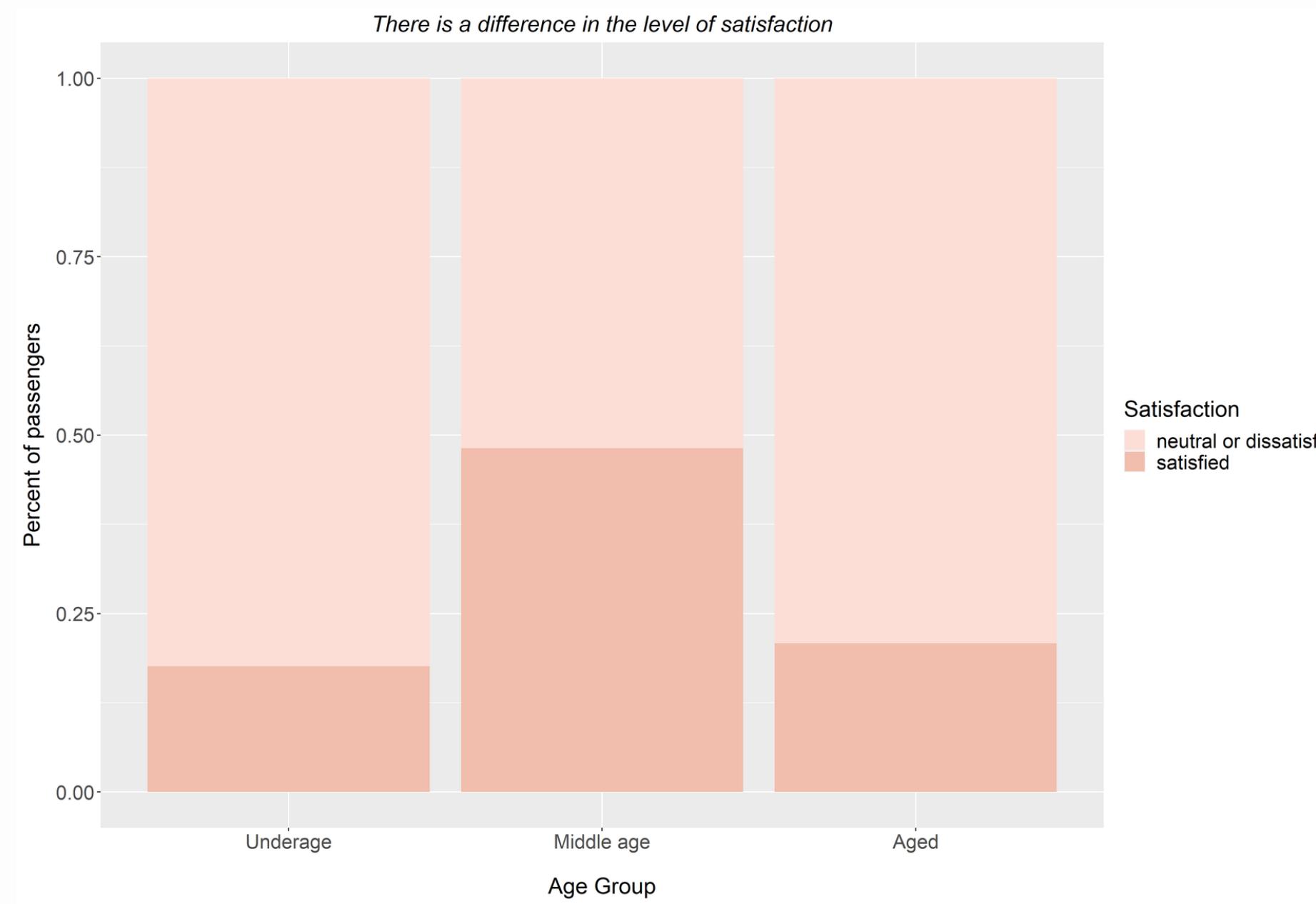
05 ВИСНОВКИ

06 ВИСНОВКИ

07 ВИСНОВКИ

08 ВИСНОВКИ

# СТАТИСТИЧНЕ ВИВЕДЕНЯ. ПЕРША ГРУПА СТАТИСТИК



H0: Статистично значущої різниці між частками спостережень немає.

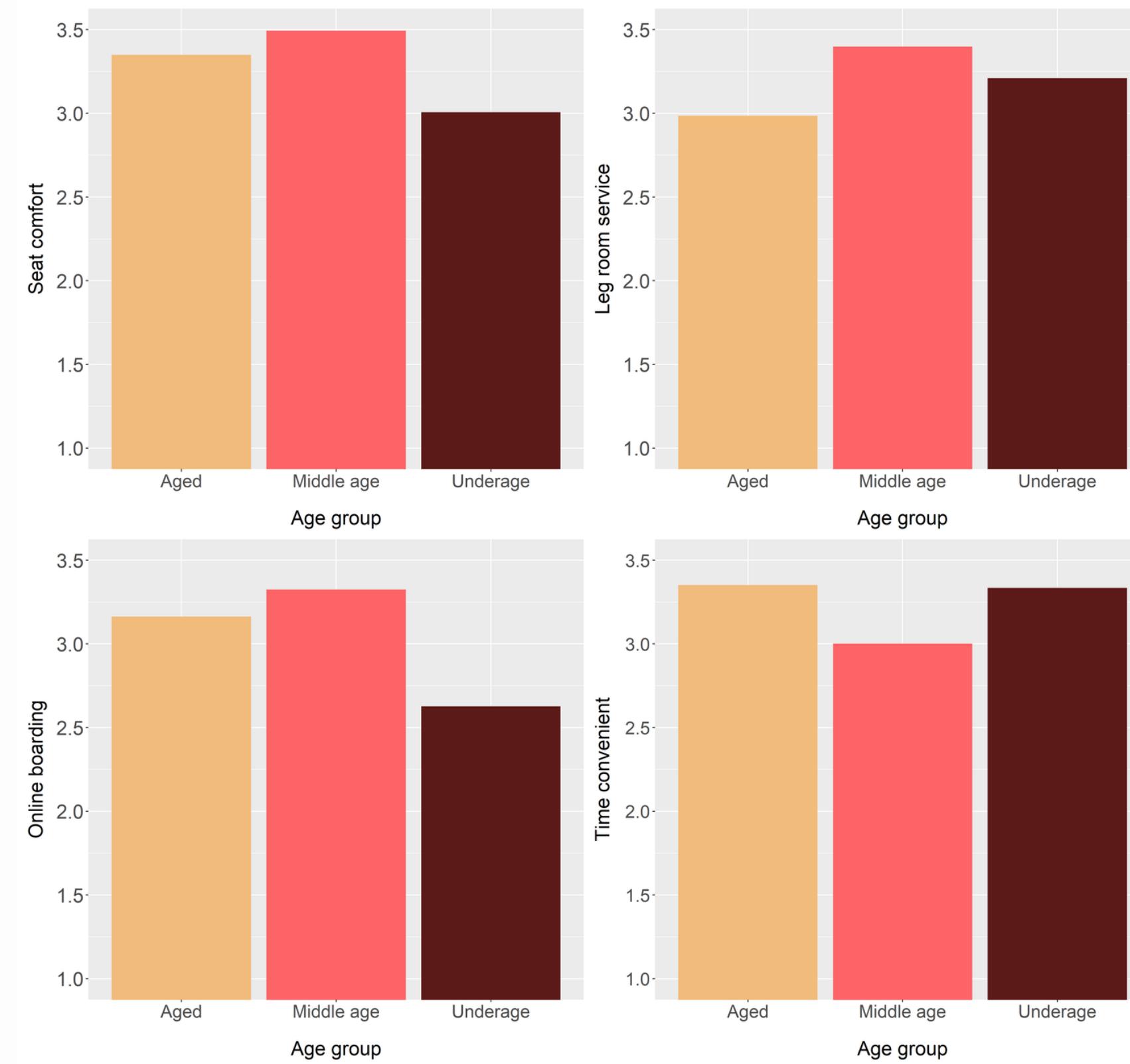
H1: Існує статистично значуща різниця між частками спостережень.

"Two sided test between Underage and Aged"  
 X-squared = 34.563, **p-value = 4.128e-09**  
 95 percent confidence interval:  
 -0.04270875 -0.02122165

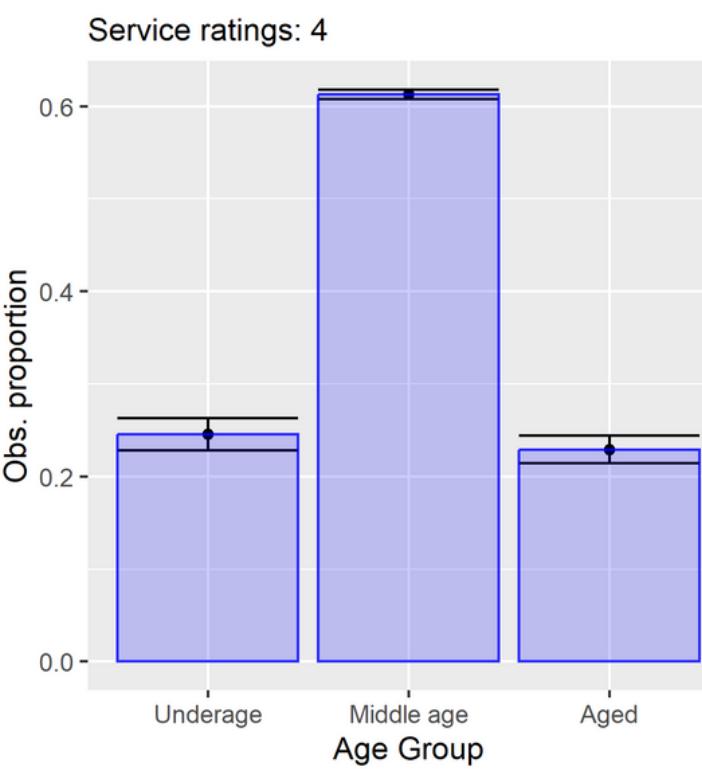
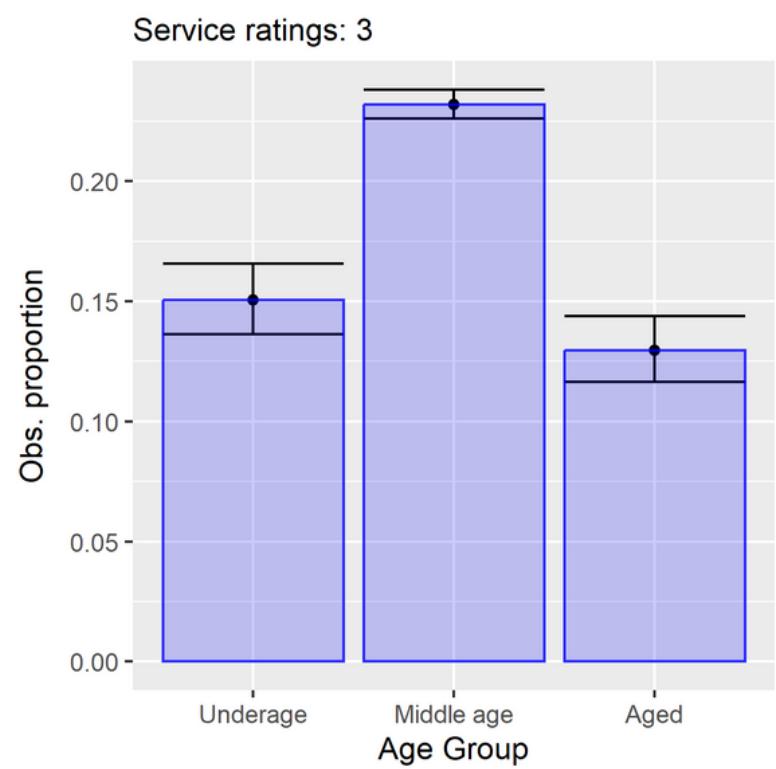
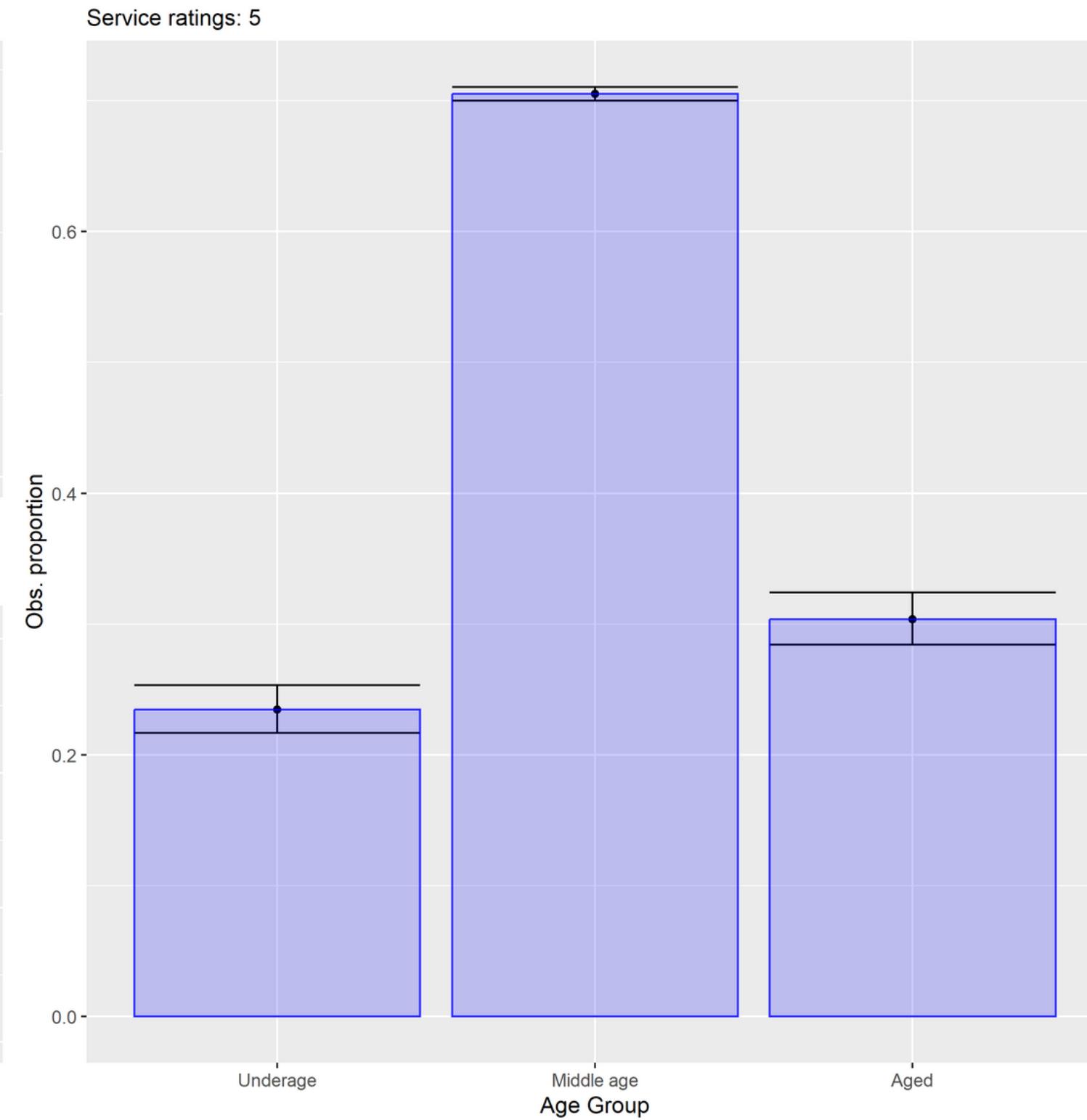
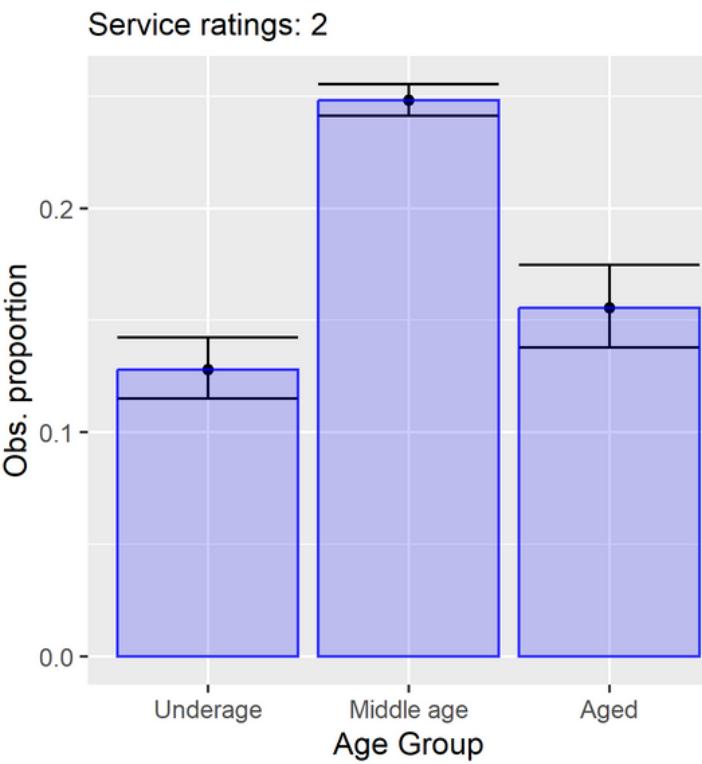
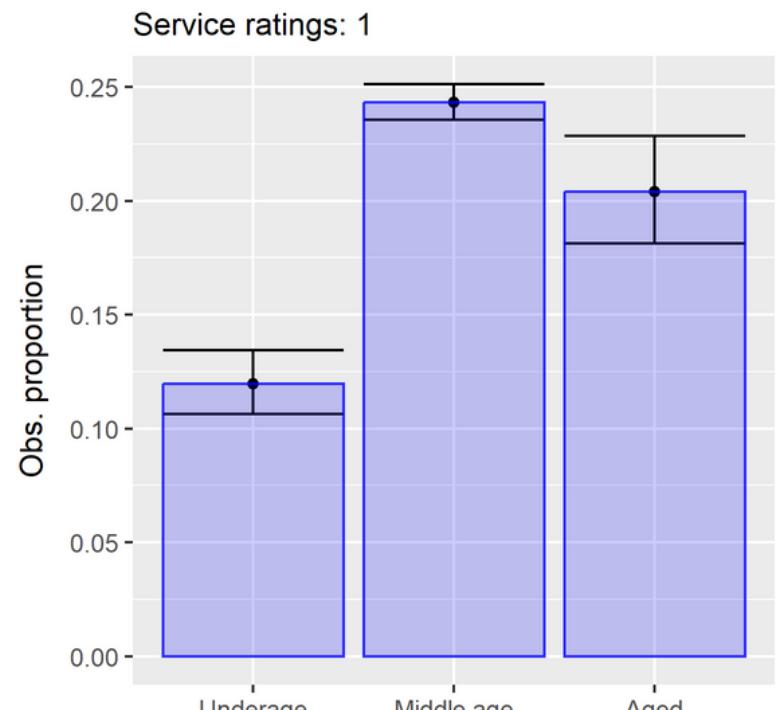
"Two sided test between Underage and Middle age"  
 X-squared = 3780 **p-value < 2.2e-16**  
 95 percent confidence interval:  
 -0.3132165 -0.2977259

"Two sided test between Middle age and Aged"  
 X-squared = 2771.4, df = 1, **p-value < 2.2e-16**  
 95 percent confidence interval:  
 0.2649774 0.2820346

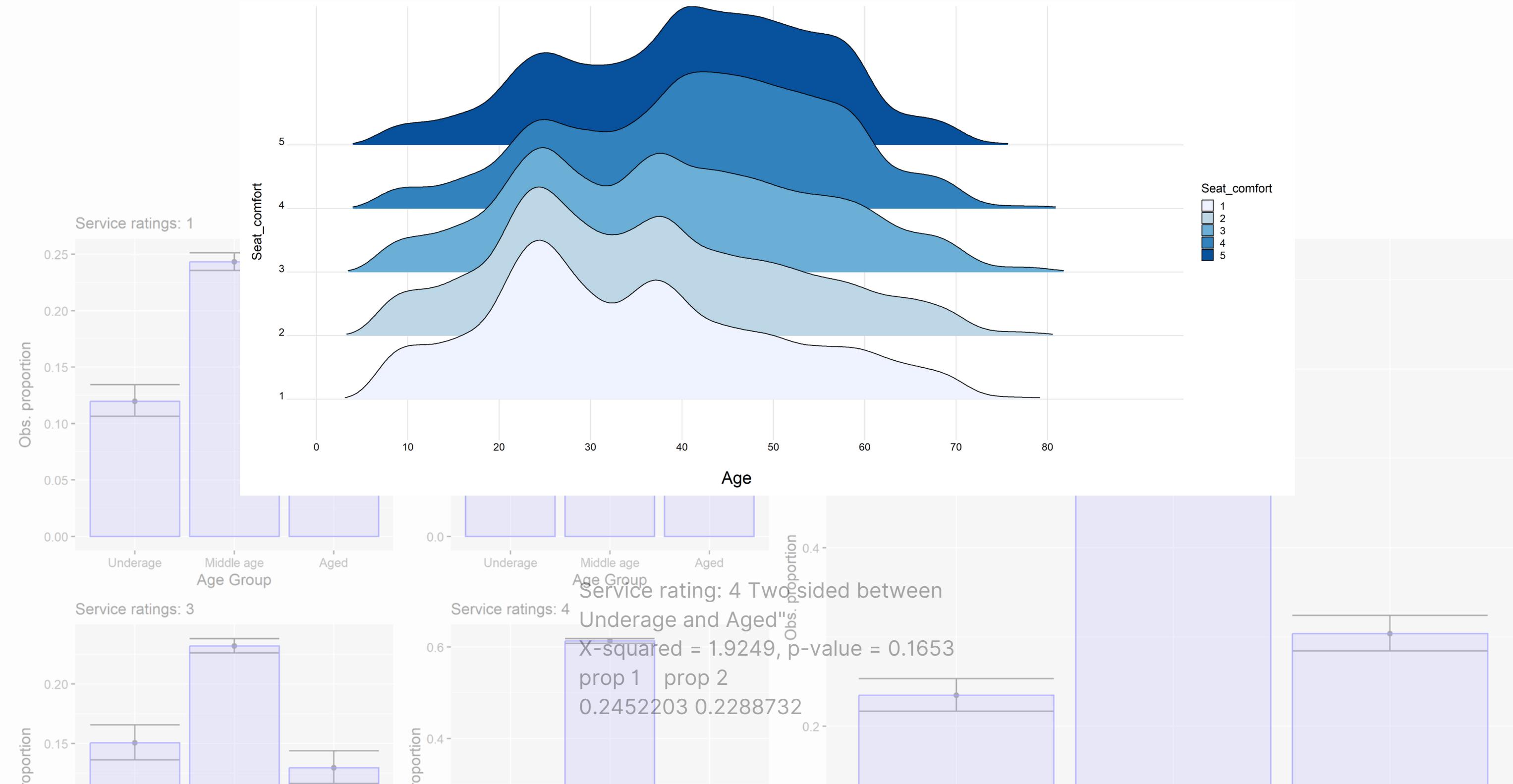
## ІСНУЮТЬ ВІКОВІ КАТЕГОРІЇ, ДЛЯ ЯКИХ ВПЛИВ ДЕЯКИХ ФАКТОРІВ ЗАДОВОЛЕНОСТІ МЕНШИЙ ЗА ІНШІ



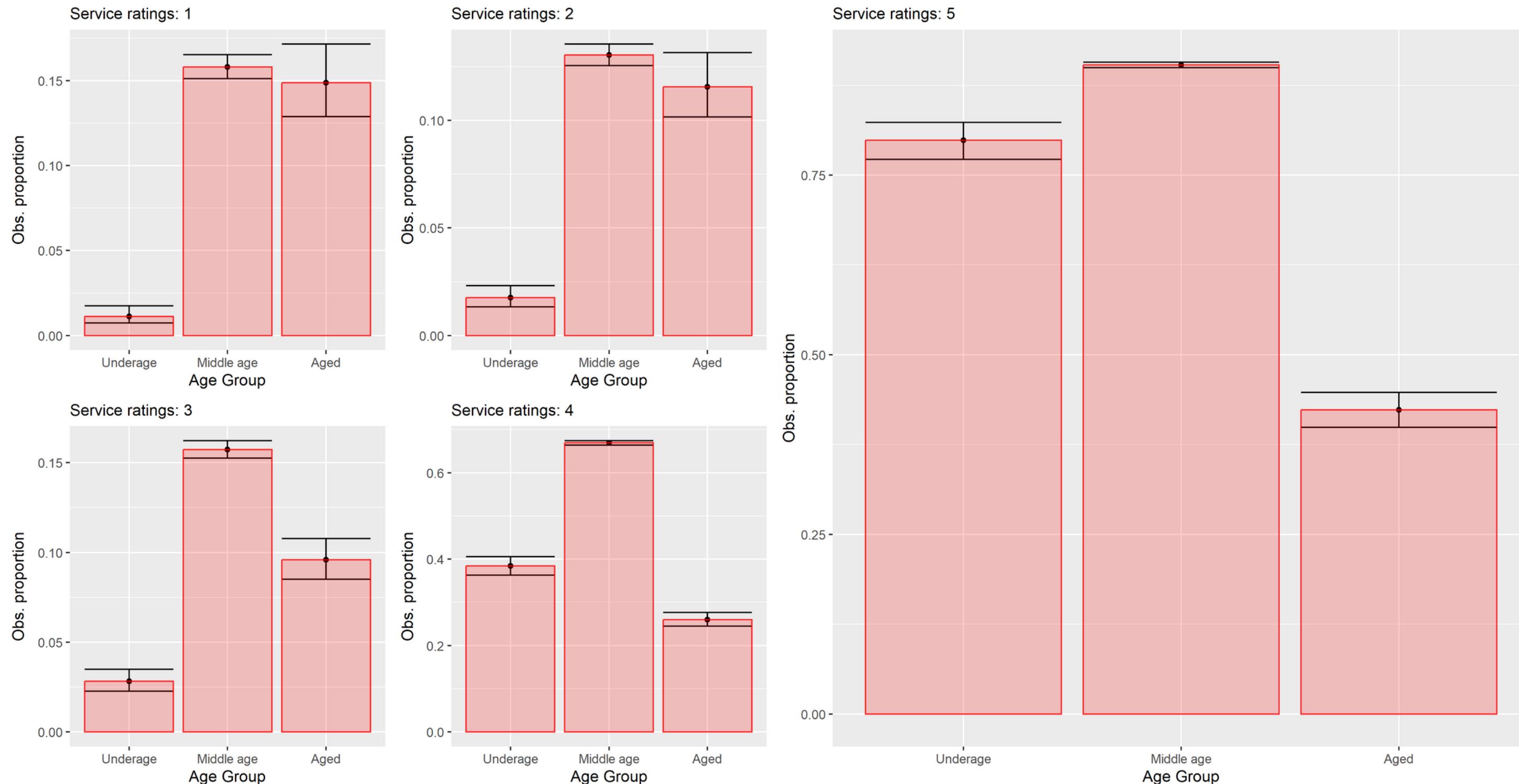
CI for observations proportion of Seat\_comfort



Service rating: 4 Two sided between  
Underage and Aged"  
X-squared = 1.9249, p-value = 0.1653  
prop 1 prop 2  
0.2452203 0.2288732



CI for observations proportion of Online\_boarding



**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 1, 2 сервісу Online Boarding між особами середнього віку та похилого віку.

**H1:** Існує значуча різниця в частках задоволеності пасажирів з рейтингом 1, 2 сервісу Online Boarding між особами середнього віку та похилого віку.

[1] "Service rating: 1 Two sided between Middle age and Aged"

X-squared = 0.55953, **p-value = 0.4545**

prop 1 prop 2

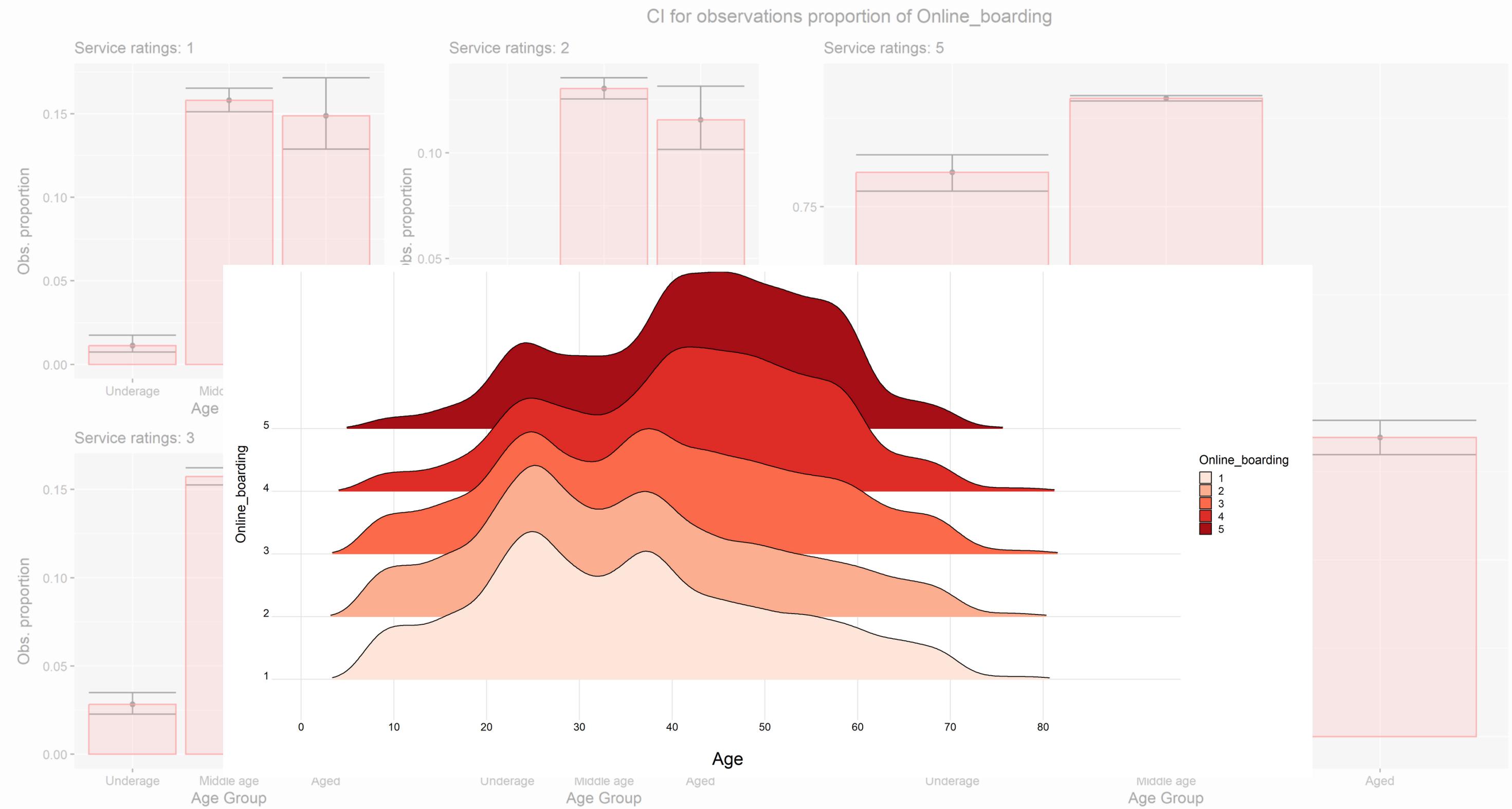
0.1581422 0.1488764

[1] "Service rating: 2 Two sided between Middle age and Aged"

X-squared = 2.9166, **p-value = 0.08767**

prop 1 prop 2

0.1303219 0.1156695



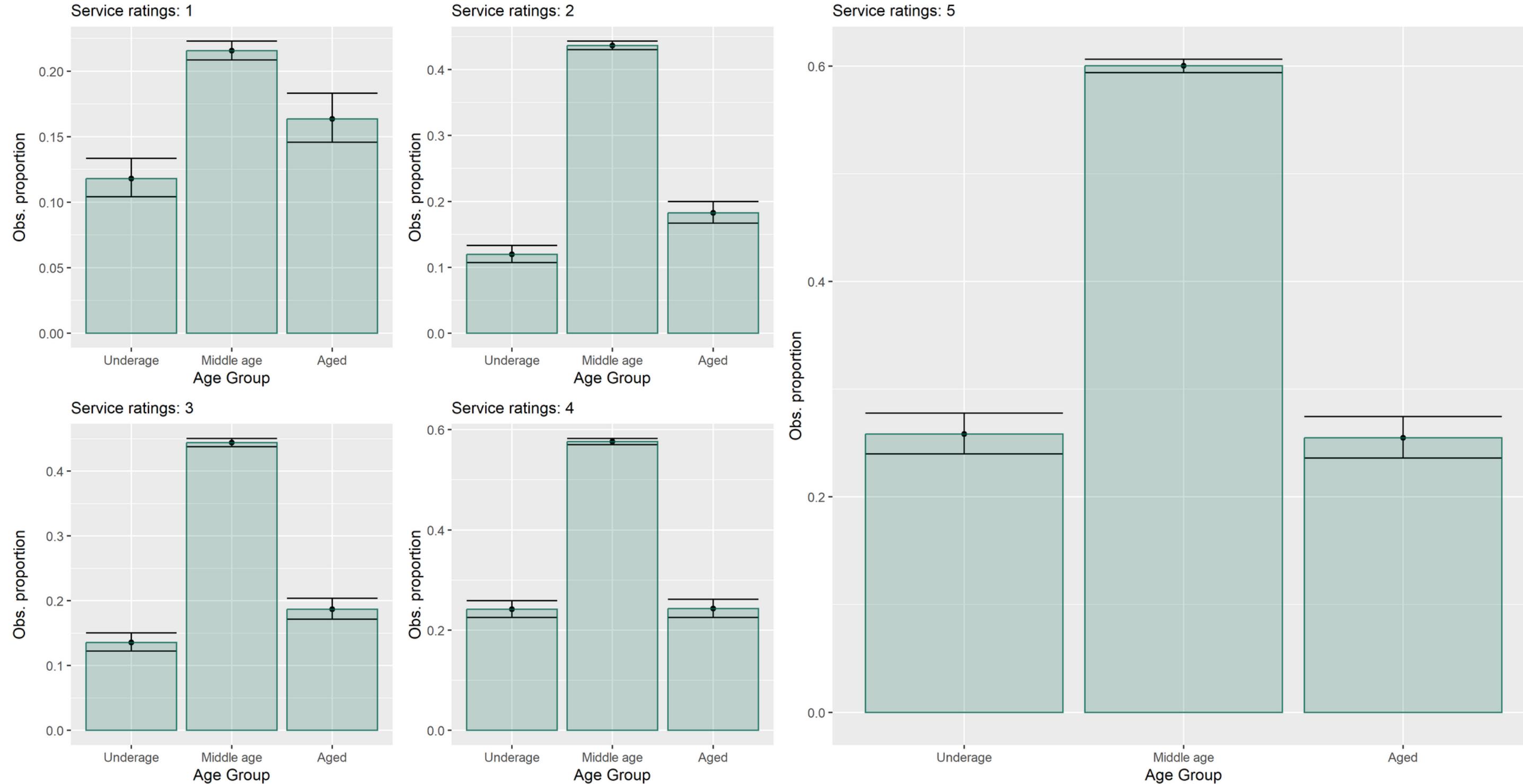
**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 1, 2 сервісу Online Boarding між особами середнього віку та похилого віку.

**H1:** Існує значуча різниця в частках задоволеності пасажирів з рейтингом 1, 2 сервісу Online Boarding між особами середнього віку та похилого віку.

[1] "Service rating: 1 Two sided between Middle age and Aged"  
 X-squared = 0.55953, **p-value = 0.4545**  
 prop 1 prop 2  
 0.1581422 0.1488764

[1] "Service rating: 2 Two sided between Middle age and Aged"  
 X-squared = 2.9166, **p-value = 0.08767**  
 prop 1 prop 2  
 0.1303219 0.1156695

CI for observations proportion of Food\_and\_drink



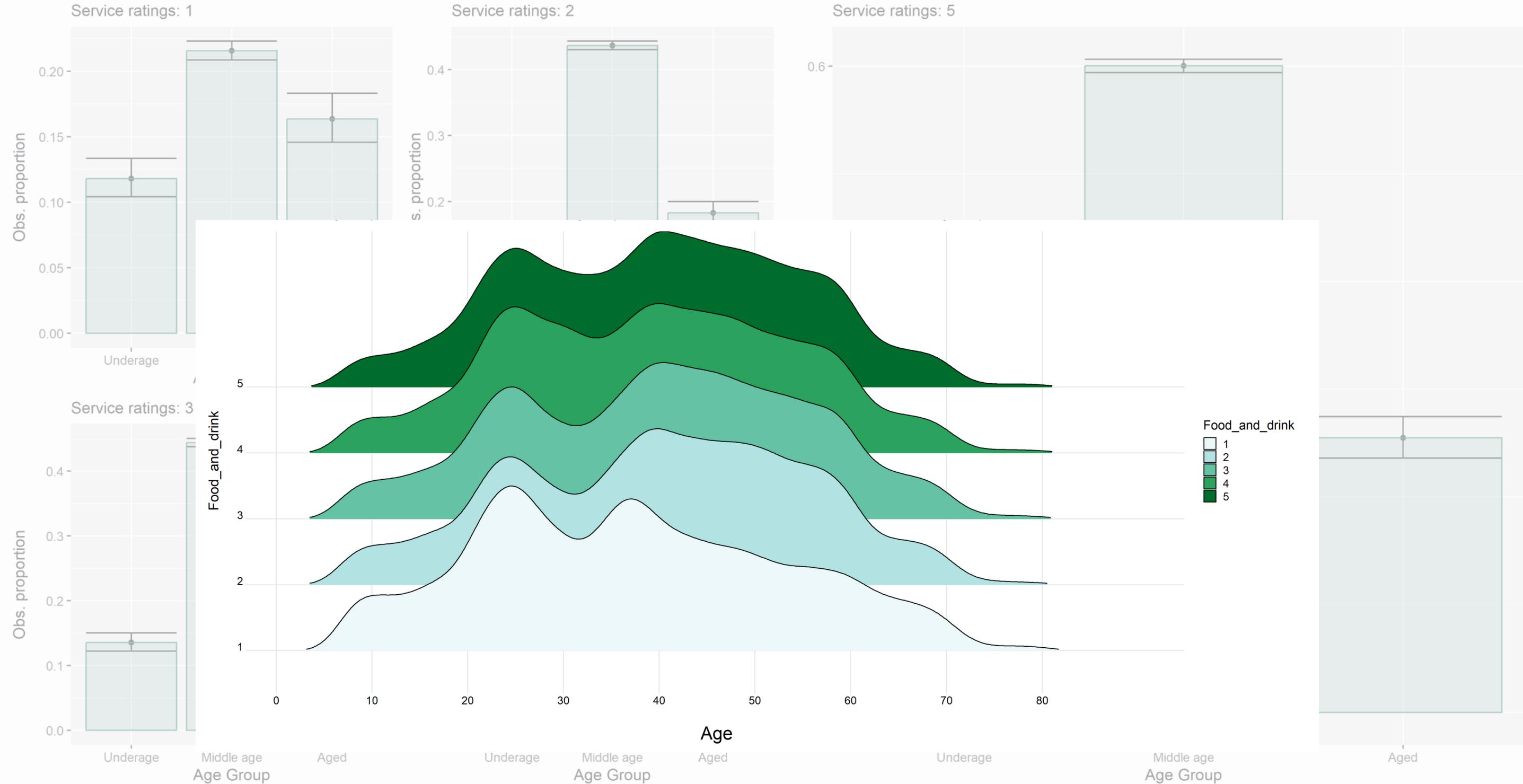
**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 4, 5 сервісу Food and drink між особами похилого віку дітьми.

**H1:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 4, 5 сервісу Food and drink між особами похилого віку та дітьми.

```
[1] "Service rating: 4 Two sided prop.test between Underage and Aged"
X-squared = 0.0039733, p-value = 0.9497
prop 1 prop 2
0.2417983 0.2430297
```

```
[1] "Service rating: 5 Two sided prop.test between Underage and Aged"
X-squared = 0.047532, p-value = 0.8274
prop 1 prop 2
0.2584600 0.2549518
```

CI for observations proportion of Food\_and\_drink



**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 4, 5 сервісу Food and drink між особами похилого віку дітьми.

**H1:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 4, 5 сервісу Food and drink між особами похилого віку та дітьми.

[1] "Service rating: 4 Two sided prop.test between Underage and Aged"

X-squared = 0.0039733, p-value = 0.9497

prop 1 prop 2

0.2417983 0.2430297

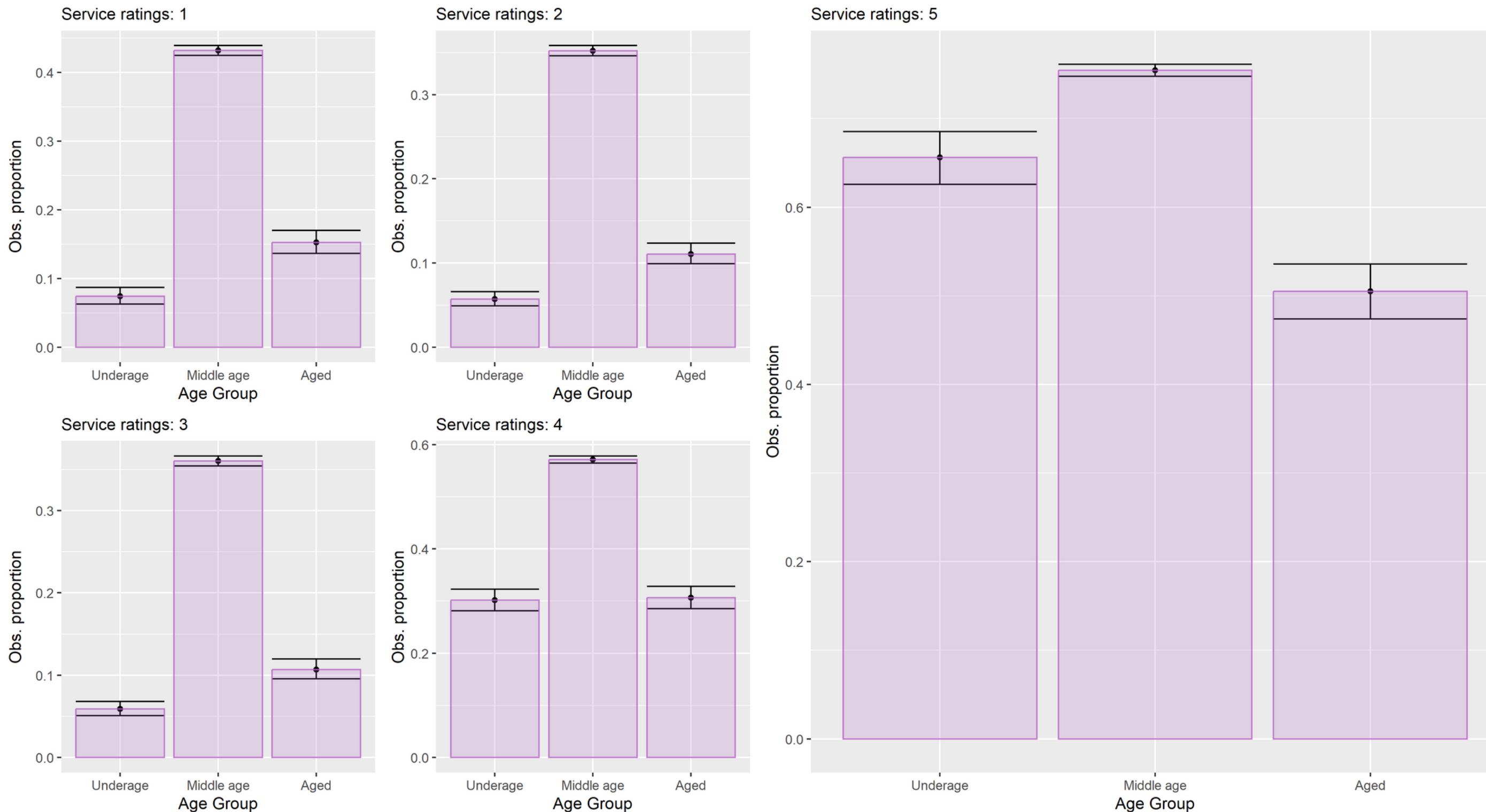
[1] "Service rating: 5 Two sided prop.test between Underage and Aged"

X-squared = 0.047532, p-value = 0.8274

prop 1 prop 2

0.2584600 0.2549518

### CI for observations proportion of Ease\_of\_Online\_booking



**H<sub>0</sub>:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 4 сервісу Online booking між особами похилого віку дітьми.

**H<sub>1</sub>:** Існує значуча різниця в частках задоволеності пасажирів з рейтингом 4 сервісу Online booking між особами похилого віку та дітьми.

```
[1] "Service rating: 4 Two sided prop.test between Underage and Aged"
X-squared = 0.07401, p-value = 0.7856
95 percent confidence interval:
-0.03517774 0.02577029
sample estimates:
prop 1 prop 2
0.3017751 0.3064789
```

### CI for observations proportion of Ease\_of\_Online\_booking

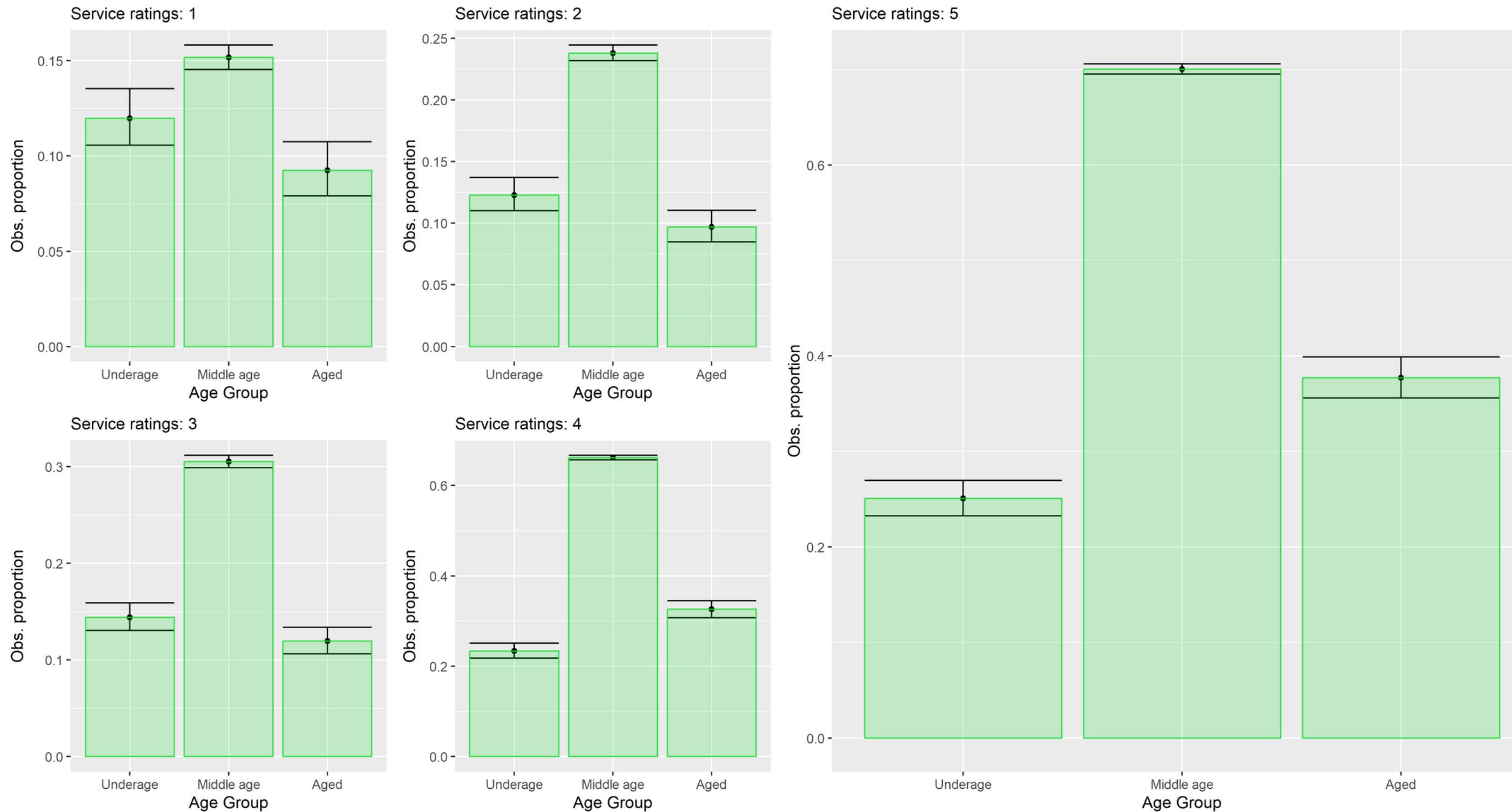


**H<sub>0</sub>:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 4 сервісу Online booking між особами похилого віку дітьми.

**H<sub>1</sub>:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 4 сервісу Online booking між особами похилого віку та дітьми.

```
[1] "Service rating: 4 Two sided prop.test between Underage and Aged"
X-squared = 0.07401, p-value = 0.7856
95 percent confidence interval:
-0.03517774 0.02577029
sample estimates:
prop 1 prop 2
0.3017751 0.3064789
```

CI for observations proportion of Inflight\_entertainment

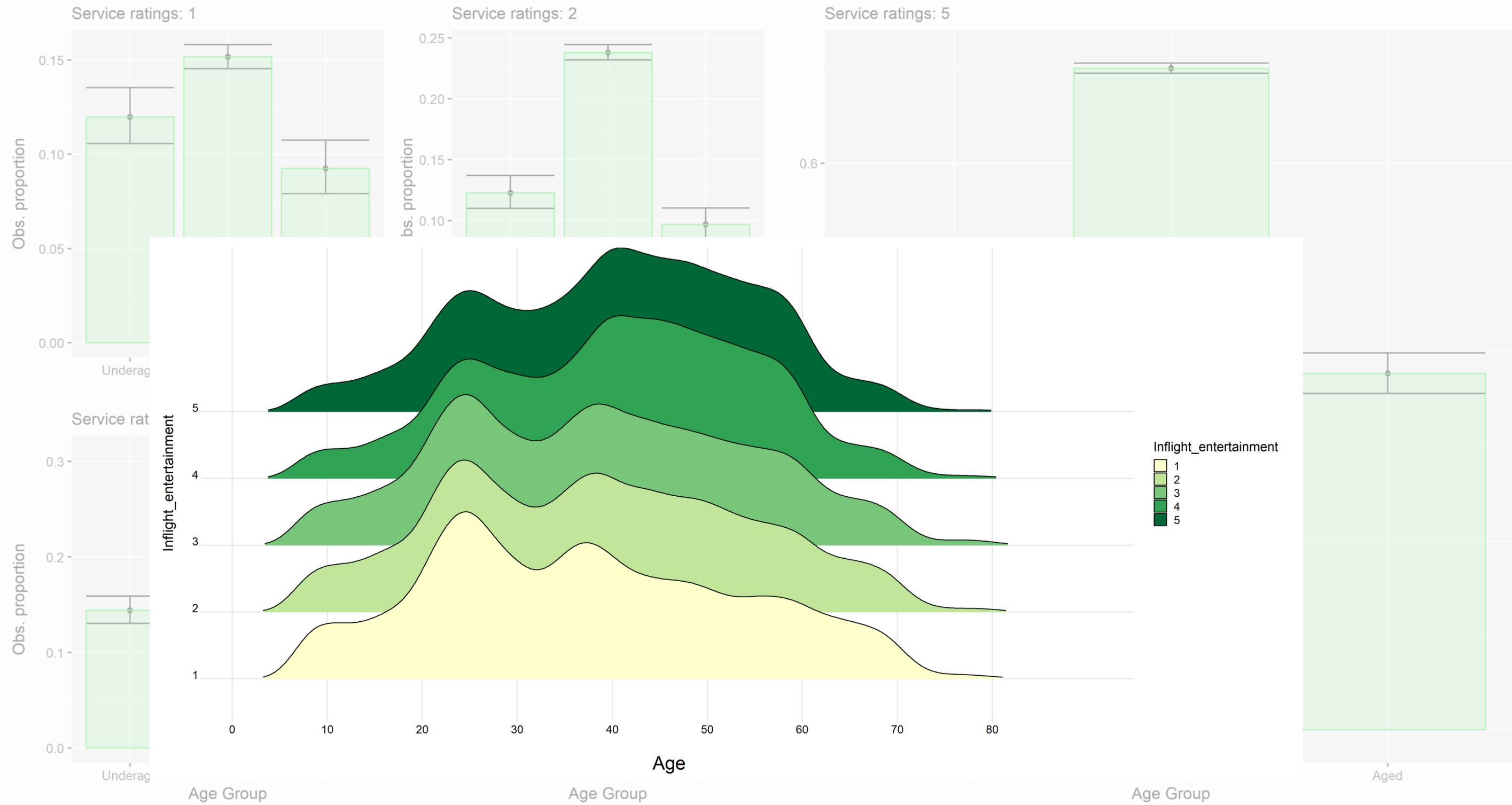


**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 3 сервісу Inflight entertainment між особами похилого віку дітьми.

**H1:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 3 сервісу Inflight entertainment між особами похилого віку та дітьми.

```
[1] "Service rating: 3 Two sided prop.test between Underage and Aged"
X-squared = 5.698, p-value = 0.01698
95 percent confidence interval:
0.004435422 0.045080097
sample estimates:
prop 1 prop 2
0.1442101 0.1194523
```

CI for observations proportion of Inflight\_entertainment

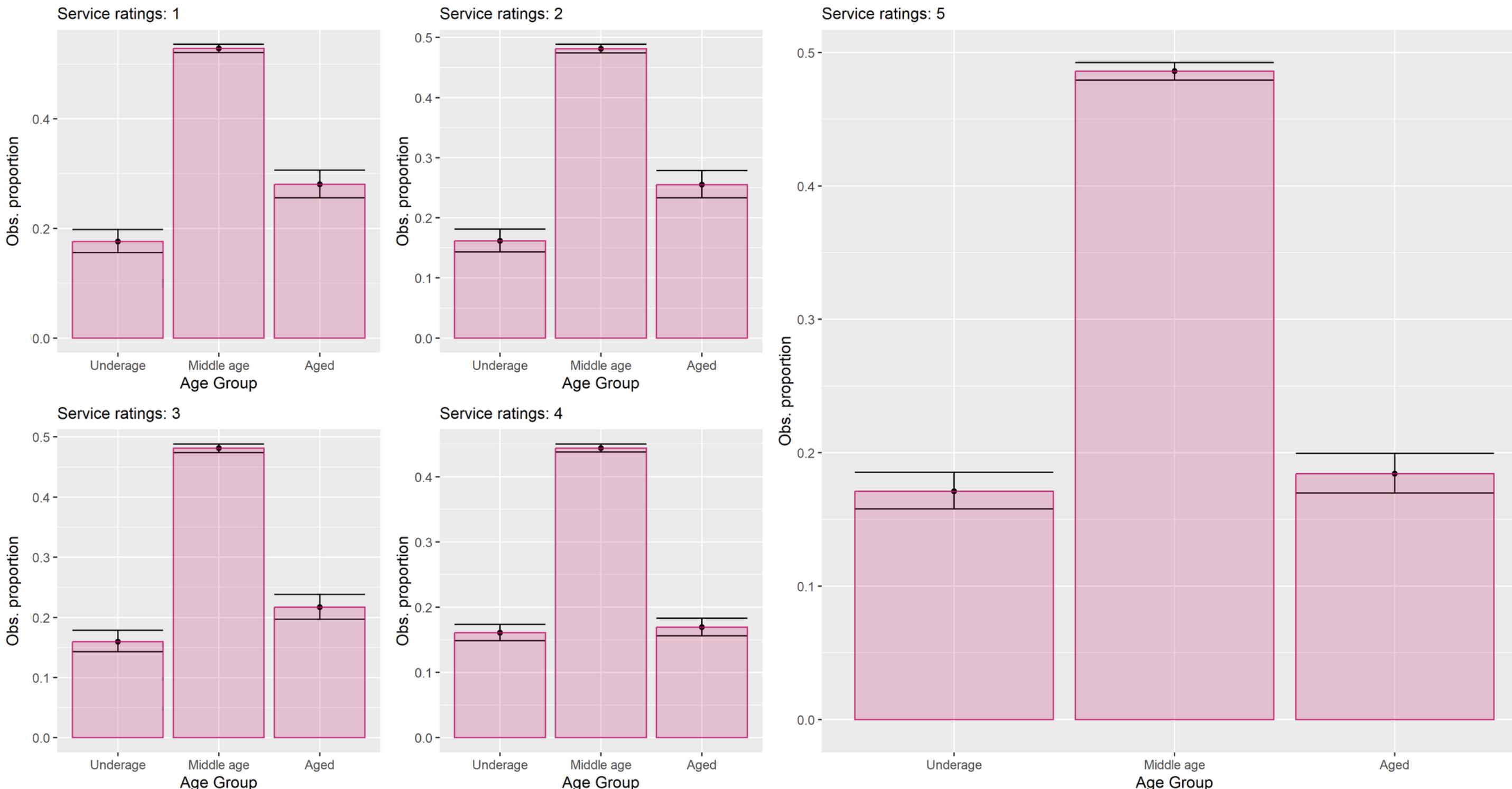


**H<sub>0</sub>:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 3 сервісів Inflight entertainment між особами похилого віку дітьми.

**H<sub>1</sub>:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 3 сервісів Inflight entertainment між особами похилого віку та дітьми.

```
[1] "Service rating: 3 Two sided prop.test between Underage and Aged"
X-squared = 5.698, p-value = 0.01698
95 percent confidence interval:
0.004435422 0.045080097
sample estimates:
prop 1 prop 2
0.1442101 0.1194523
```

### CI for observations proportion of Time\_convenient



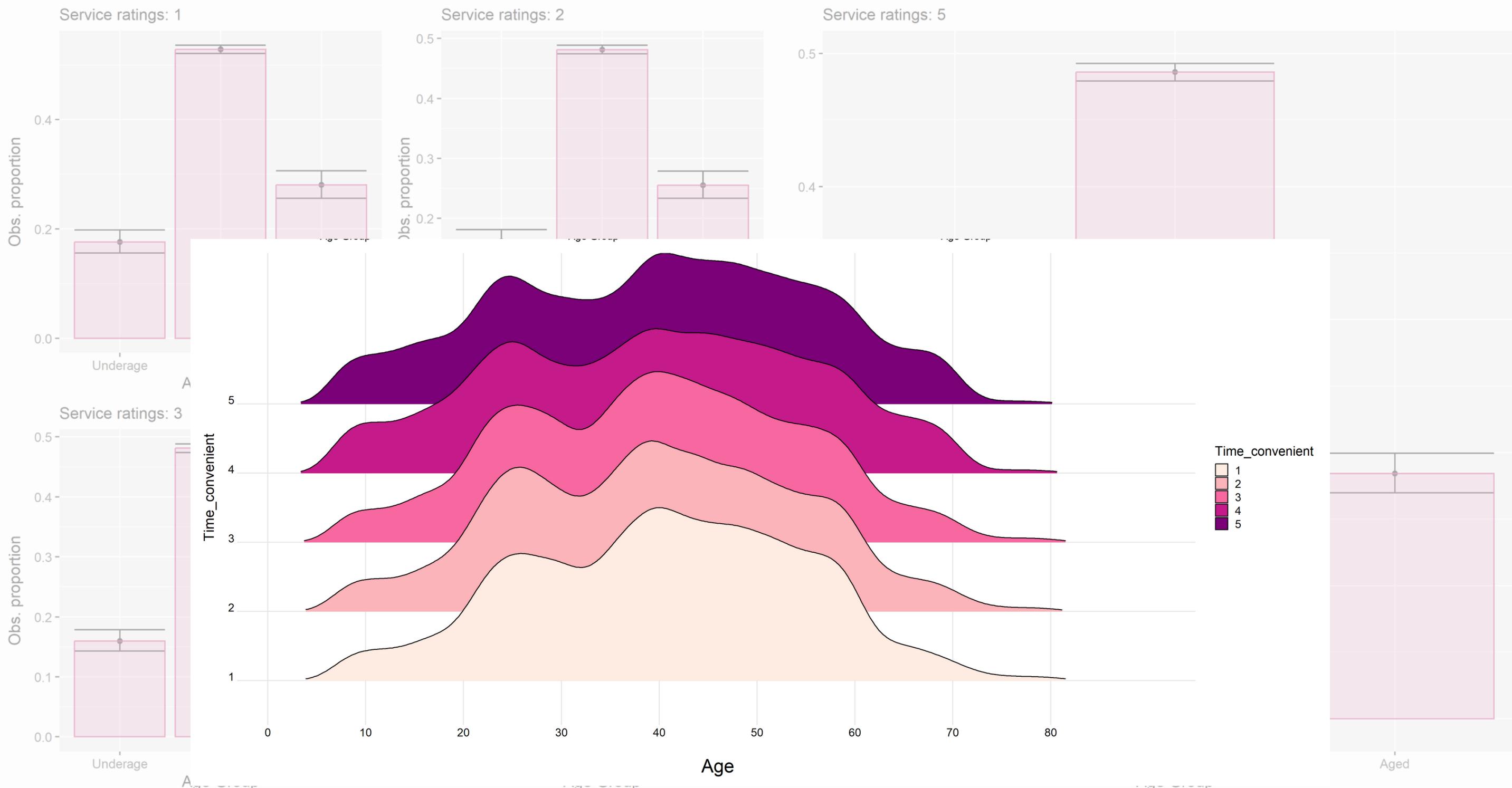
**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 4,5 сервісу Time convenient між особами похилого віку дітьми.

**H1:** Існує значуча різниця в частках задоволеності пасажирів з рейтингом 4,5 сервісу Time convenient між особами похилого віку та дітьми.

```
[1] "Service rating: 4 Two sided prop.test between Underage and Aged"
X-squared = 0.73454, p-value = 0.3914
prop 1 prop 2
0.1607251 0.1690611
```

```
[1] "Service rating: 5 Two sided prop.test between Underage and Aged"
X-squared = 1.5179, p-value = 0.2179
prop 1 prop 2
0.1711433 0.1841607
```

### CI for observations proportion of Time\_convenient

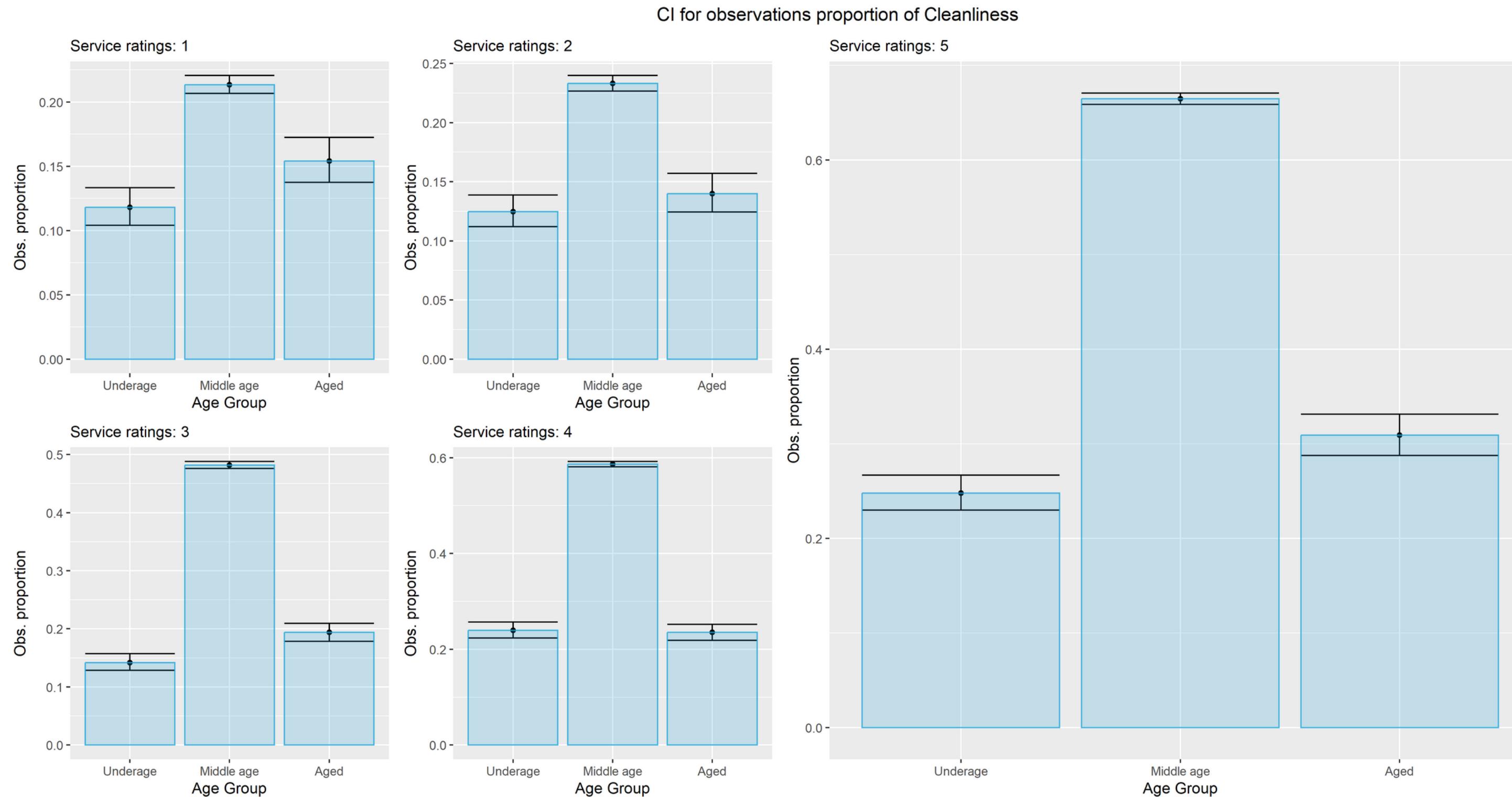


**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 4,5 сервісу Time convenient між особами похилого віку дітьми.

**H1:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 4,5 сервісу Time convenient між особами похилого віку та дітьми.

```
[1] "Service rating: 4 Two sided prop.test between Underage and Aged"
X-squared = 0.73454, p-value = 0.3914
prop 1 prop 2
0.1607251 0.1690611
```

```
[1] "Service rating: 5 Two sided prop.test between Underage and Aged"
X-squared = 1.5179, p-value = 0.2179
prop 1 prop 2
0.1711433 0.1841607
```



**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 2,4 сервісу Time convenient між особами похилого віку дітьми.

**H1:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 2,4 сервісу Time convenient між особами похилого віку та дітьми.

[1] "Service rating: 2 Two sided prop.test between Underage and Aged"

X-squared = 1.8503, **p-value = 0.1737**

prop 1 prop 2

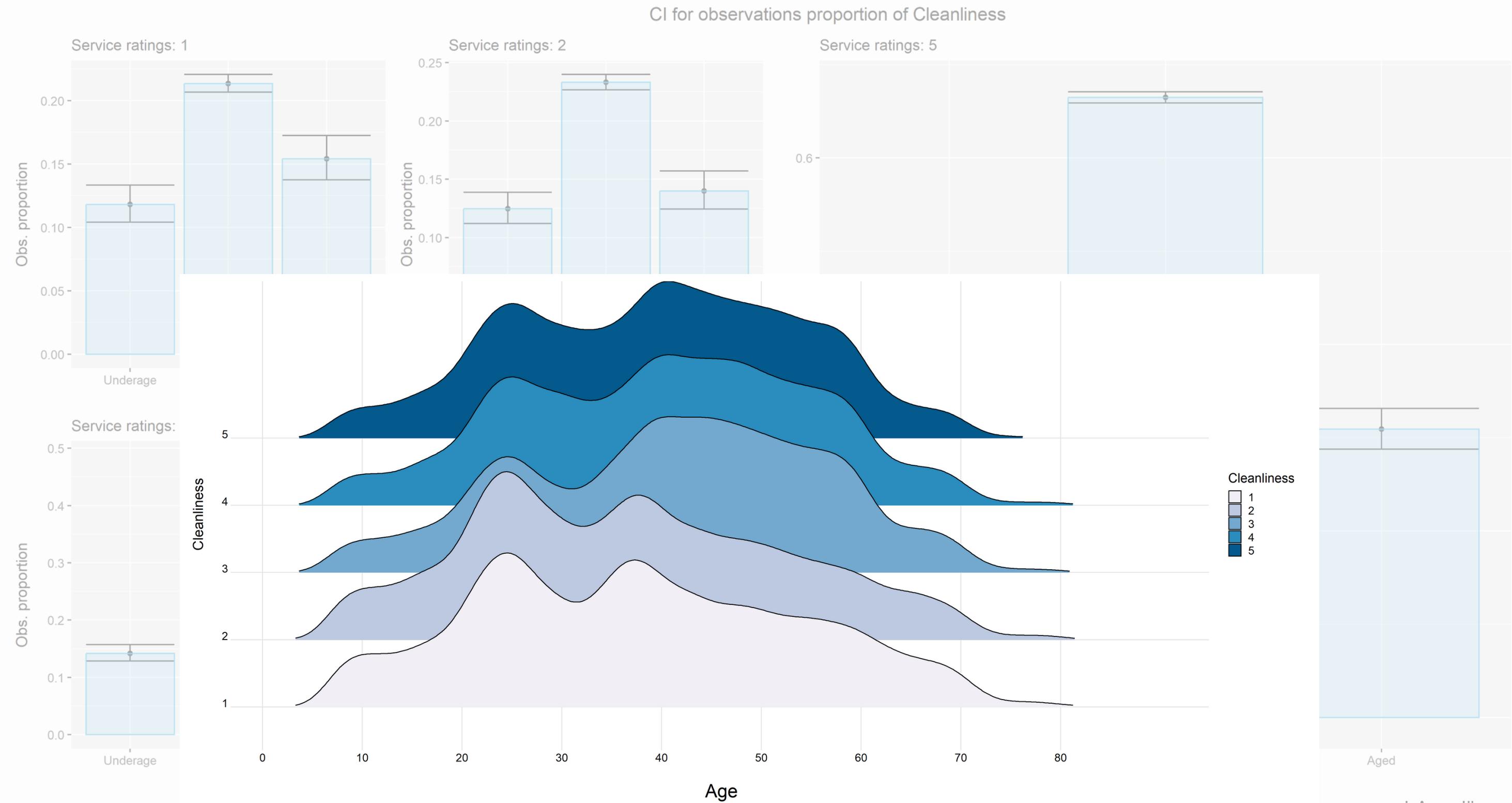
0.1247849 0.1399297

[1] "Service rating: 4 Two sided prop.test between Underage and Aged"

X-squared = 0.10759, **p-value = 0.7429**

prop 1 prop 2

0.2395330 0.2351763



**H0:** Немає значущої різниці в частках задоволеності пасажирів з рейтингом 2,4 сервісу Time convenient між особами похилого віку та дітьми.

**H1:** Існує значуща різниця в частках задоволеності пасажирів з рейтингом 2,4 сервісу Time convenient між особами похилого віку та дітьми.

[1] "Service rating: 2 Two sided prop.test between Underage and Aged"

X-squared = 1.8503, p-value = 0.1737

prop 1 prop 2

0.1247849 0.1399297

[1] "Service rating: 4 Two sided prop.test between Underage and Aged"

X-squared = 0.10759, p-value = 0.7429

prop 1 prop 2

0.2395330 0.2351763

## СТАТИСТИЧНЕ ВИВЕДЕННЯ. ДРУГА ГРУПА СТАТИСТИК

*Eco Type*

	Eco Type									
Satisfaction										
Delay_overtake	0									
Arrival_Delay_in_Minutes	-0.1	0.2								
Departure_Delay_in_Minutes	0.7	-0.1	-0.4							
Cleanliness	0	0	0.2	0						
Inflight_service	0	0	0	0	0					
Checkin_service	0.2	0.1	0	0	0	0				
Baggage_handling	0.2	0.5	0	0	0	0	0			
Leg_room_service	0.2	0.1	0.2	0	0	0	0	0.1	0	
On.board_service	0.1	0.4	0.2	0.4	0	0	0	0	0.1	0
Inflight_entertainment	0.2	0.1	0.2	0	0.2	0.8	0	0	0.2	0
Seat_comfort	0.6	0	0	0	0.1	0	0.7	0	0	0.1
Online_boarding	0.3	0.1	0	-0.1	-0.1	0.1	-0.1	0.2	0	0.3
Food_and_drink	0.2	0.7	0.8	0	0	0	0	0.8	0	0.2
Gate_location	0	0	0	0	-0.1	0	0	-0.1	0	0
Ease_of_Online_booking	0.2	0	0.6	0	0	0.1	0	0	0	0.2
Flight_Arrival_time_convenient	0.2	0.2	0	0.1	0	0	0.2	0	0.2	0
Inflight_wifi_service	0	0.7	0	0.2	0.7	0.1	0.2	0.1	0	-0.1
Flight_Distance	0	0	0	0	0	0	0	0	0	-0.1
Age	0	0	0.1	0	0	0.2	0.1	0	0	0

95% довірчий інтервал для Inflight wifi service:

Correlation for Eco Type(Satisfaction, Inflight\_wifi\_service):

0.4703533

Lower Bound: 0.4624195

Upper Bound: 0.4786391

Інтервал досить вузький, то ж приймаємо цей показник кореляції. З розглянутої матриці кореляцій видно, що фактор Inflight\_service має значну кореляцію зі задоволеністю користувачів економ-класу. Знаходячись у діапазоні 0.4624 до 0.4786, цей показник кореляції виявляється статистично значущим для даної групи людей.

## Business

	Delay_overtake											
	Satisfaction											
	Arrival_Delay_in_Minutes											
	Departure_Delay_in_Minutes											
	Cleanliness											
	Inflight_service											
	Checkin_service											
	Baggage_handling											
	Leg_room_service											
	On.board_service											
	Inflight_entertainment											
	Seat_comfort											
	Online_boarding											
	Food_and_drink											
	Gate_location											
	Ease_of_Online_booking											
ture.Arrival_time_convenient	0.7	0.7	0	0.1	0	0	0	0.1	0	0	0	0
Inflight_wifi_service	0.6	0.8	0.6	0.1	0.2	0.1	0.2	0.1	0.1	0.2	0.1	0
Flight_Distance	0	0	0	0	0	0.1	0.1	0.1	0	0	0.1	0
Age	0.1	0	0	0	0	0.2	0.2	0.1	0.1	0	0	0.1

Online Boarding та Inflight Entertainment. Знайдемо 95% довірчі інтервали для даних показників

Confidence Interval for the Correlation:  
Correlation for Business Type(Satisfaction, Inflight\_entertainment): 0.4911049

**Lower Bound: 0.4844964**

**Upper Bound: 0.4980903**

Confidence Interval for the Correlation:  
Correlation for Business Type(Satisfaction, Online\_boarding): 0.5543269

**Lower Bound: 0.5487447**

**Upper Bound: 0.5595996**

Інтервали досить вузькі, то ж приймаємо ці показники кореляції

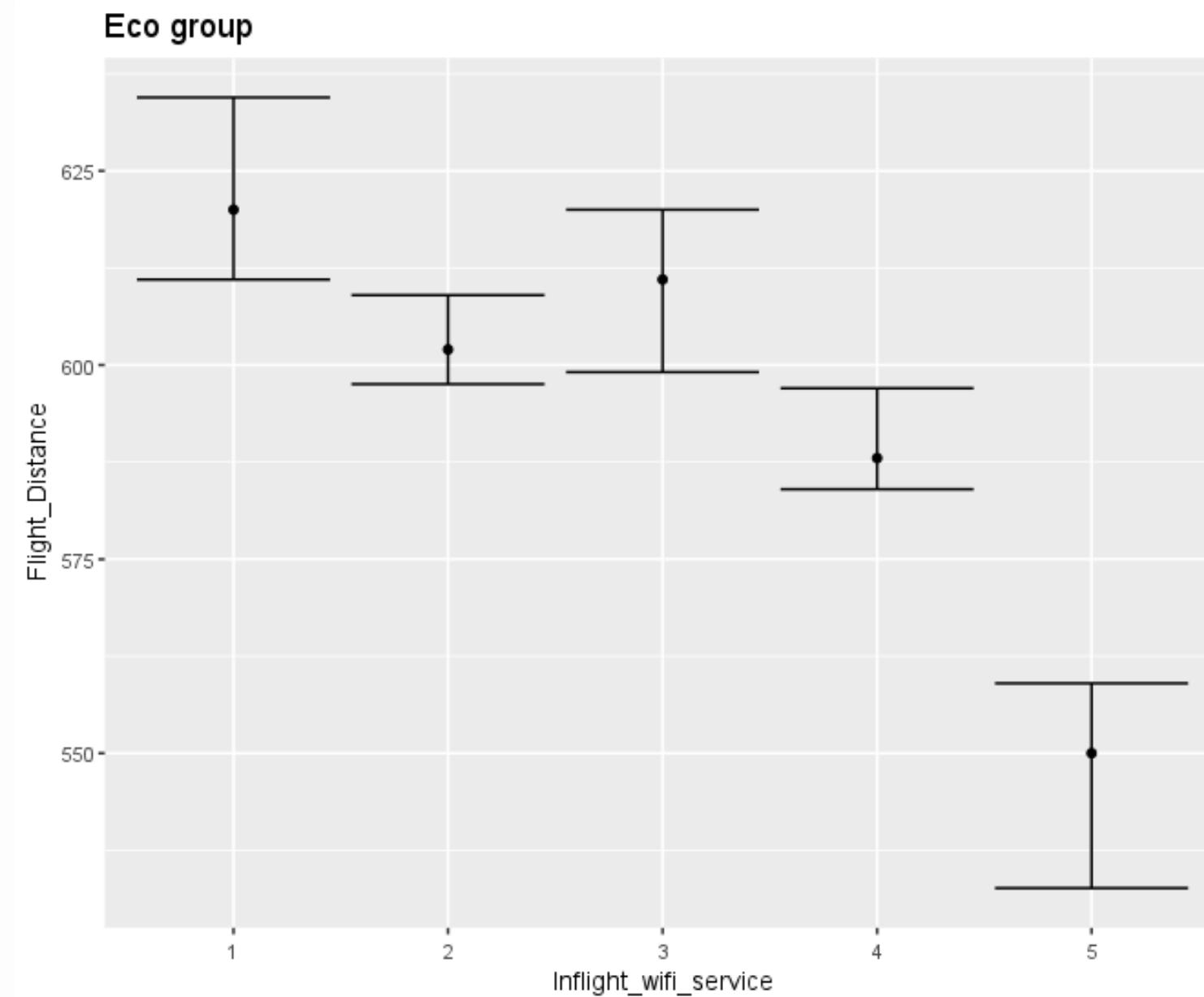
З даної візуалізації, бачимо, що чим далі подорож користувача, тим нижчу оцінку сервісу Inflight wifi service дають

**Гіпотеза:** клієнти дають оцінку “3” для Inflight wifi service коли подорожують на більшу відстань, ніж коли дають оцінку “2”.

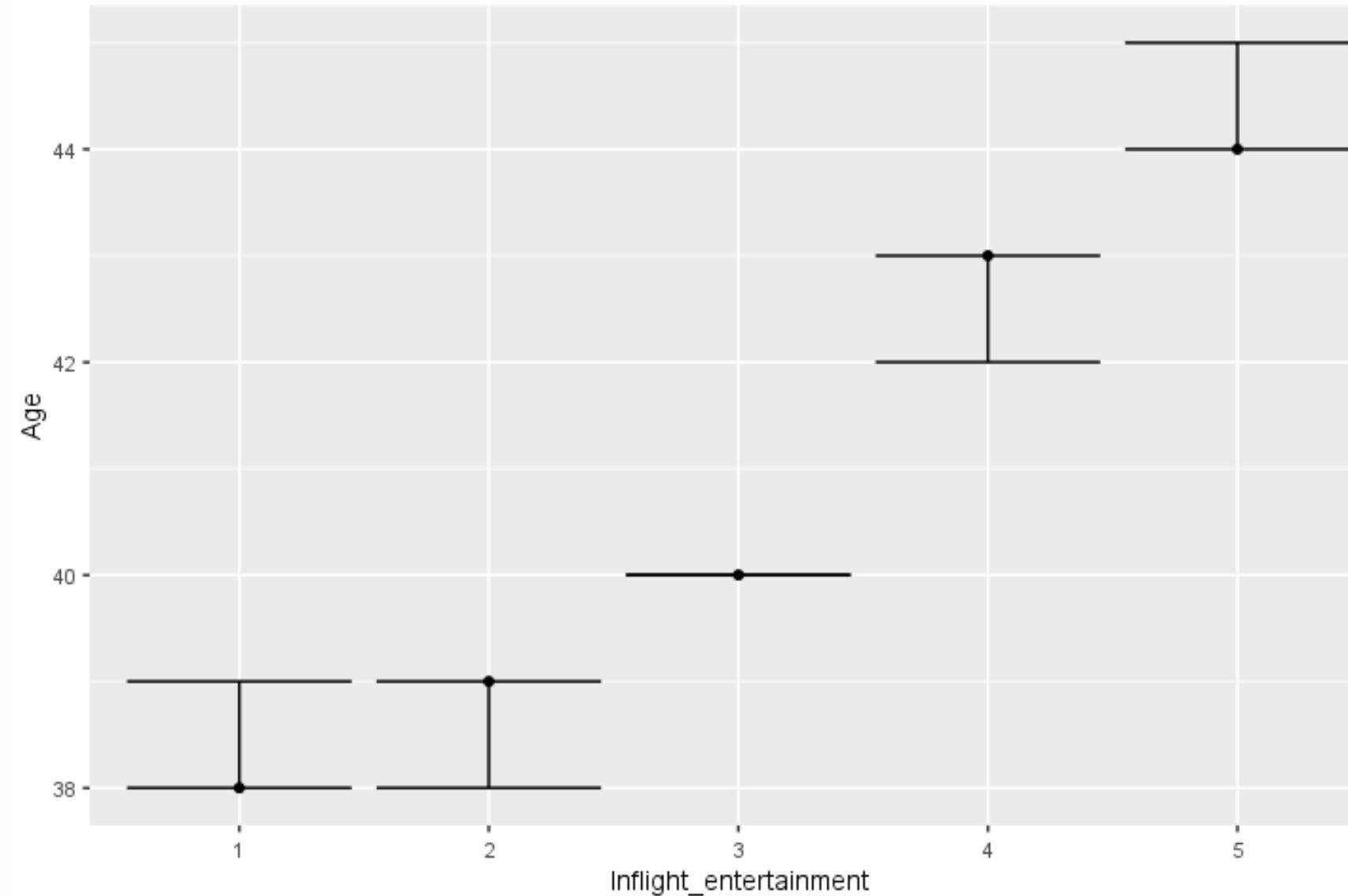
**Н0:** Медіана для значення 2 менше або дорівнює медіані для значень 3.

**p-value = 1**

Отримали результат 1, це означає, що немає достатніх доказів для того, щоб стверджувати статистично значимо більшу медіану в першій вибірці(2) на заданому рівні значущості (зазвичай 0.05).



## Business group



З даної візуалізації, бачимо, що чим старші користувачі, тим більшу оцінку сервісу Inflight\_entertainment дають.

**Гіпотеза:** Оцінки 1, 2 дають користувачі одного віку.

**H0:** Медіана значення 1 менше або дорівнює медіані значення 2.

**p-value** = 1

Отримали результат 1, це означає, що немає достатніх доказів для того, щоб стверджувати статистично значимо більшу медіану в першій вибірці(2) на заданому рівні значущості (зазвичай 0.05).

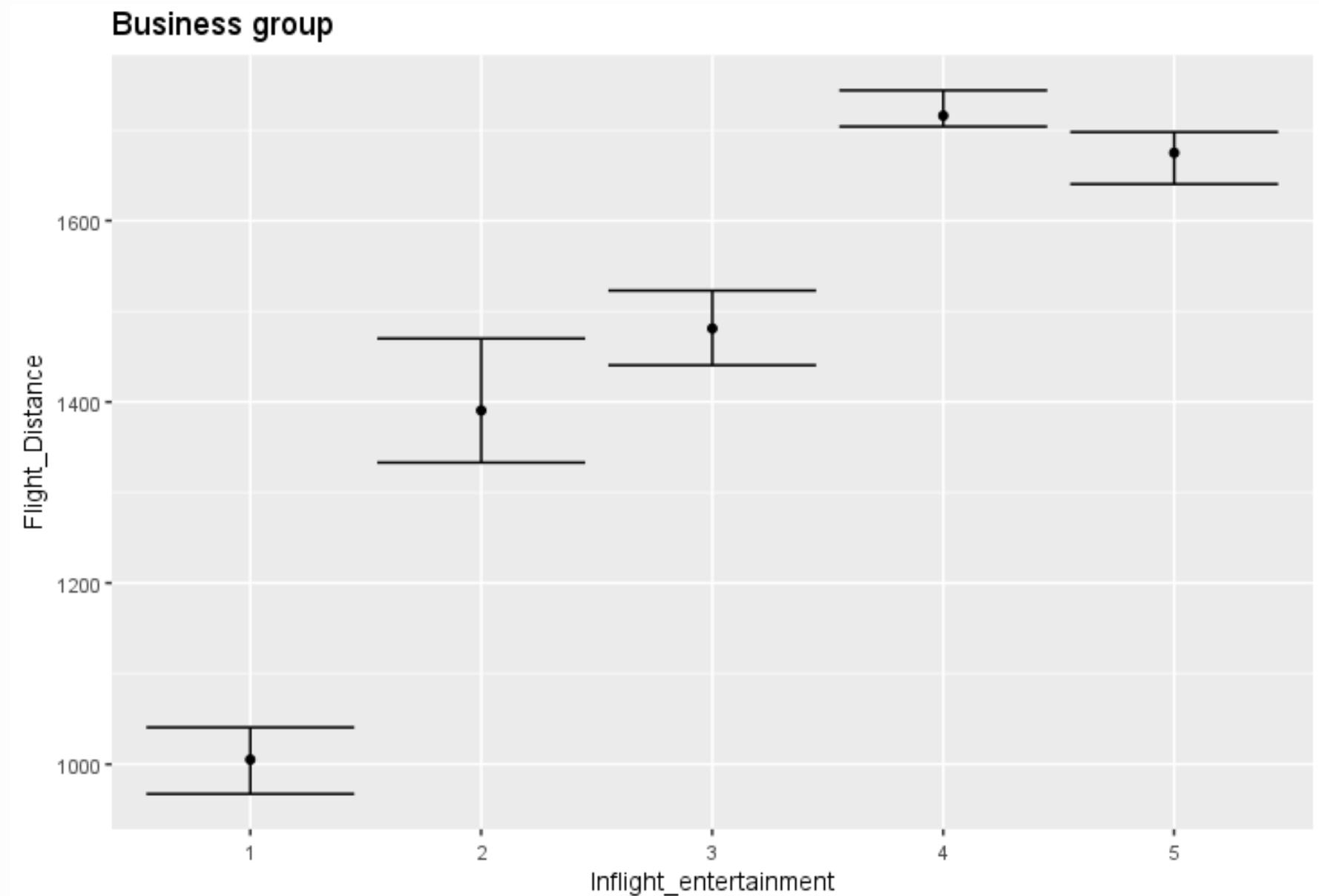
З даної візуалізації, бачимо, що далі подорож користувача, тим більшу оцінку сервісу Inflight\_entertainment дають. Це може бути показником, що даний сервіс актуальніший для на великих перельотах

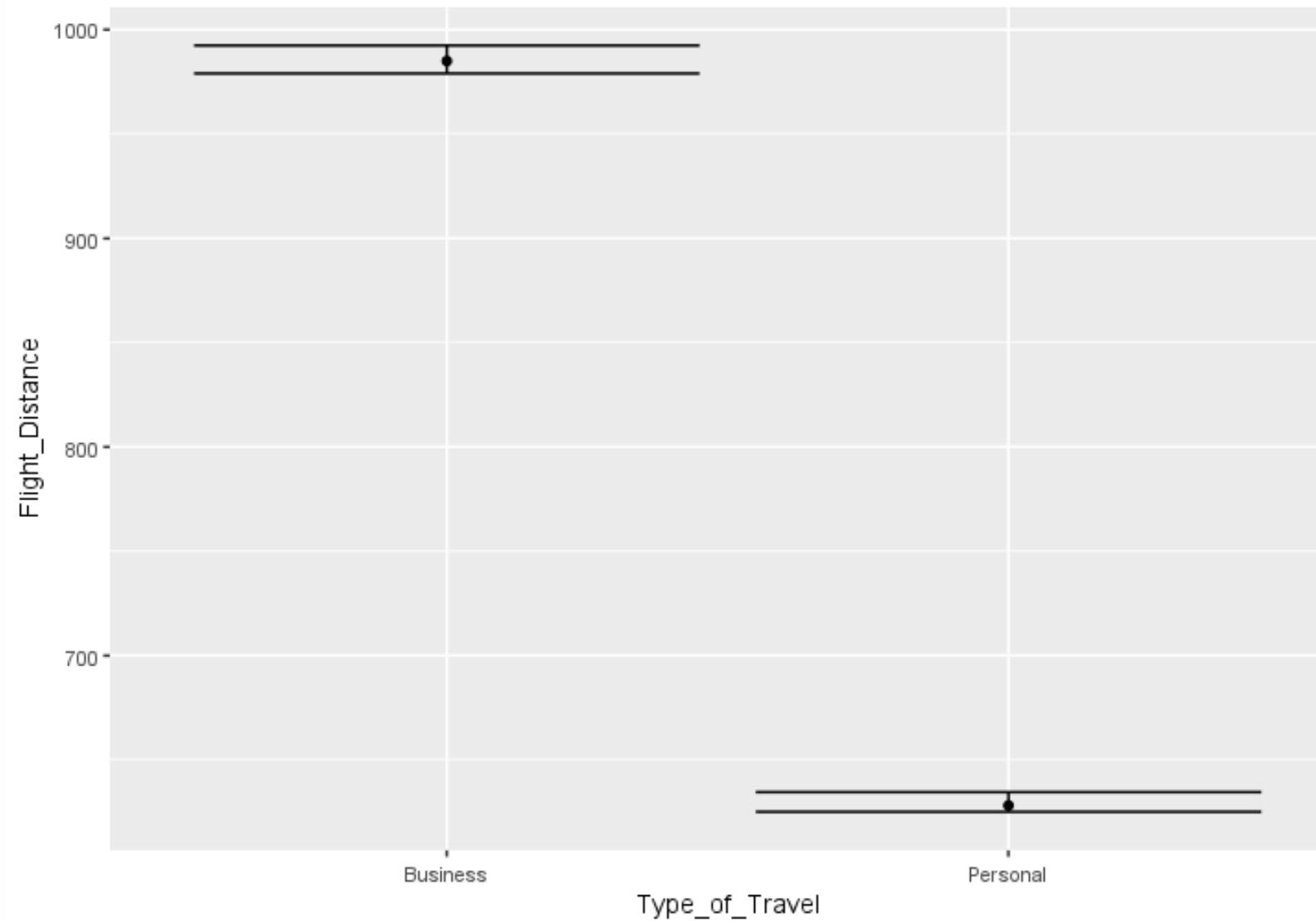
З даної візуалізації можемо сформувати **гіпотезу** Клієнти дають оцінку “4” для Inflight\_wifi\_service коли подорожують на більшу відстань, ніж коли дають оцінку “5”.

**H0:** Медіана значення “4” менше або дорівнює медіані значення “5”.

**p-value** = 2.2e-16

Отримали результат 2.2e-16, це означає, що є статистично значимі докази на підтримку того, що медіана першої вибірки є більшою за медіану другої вибірки на заданому рівні значущості (зазвичай 0.05).





**Гіпотеза:** Пасажири бізнес класу подорожують далі

**H0:** Медіана для Business менше або дорівнює медіані для Personal.

**p-value** = 2.2e-16

Отримали результат 2.2e-16, це означає, що є статистично значимі докази на підтримку того, що медіана першої вибірки є більшою за медіану другої вибірки на заданому рівні значущості (зазвичай 0.05).

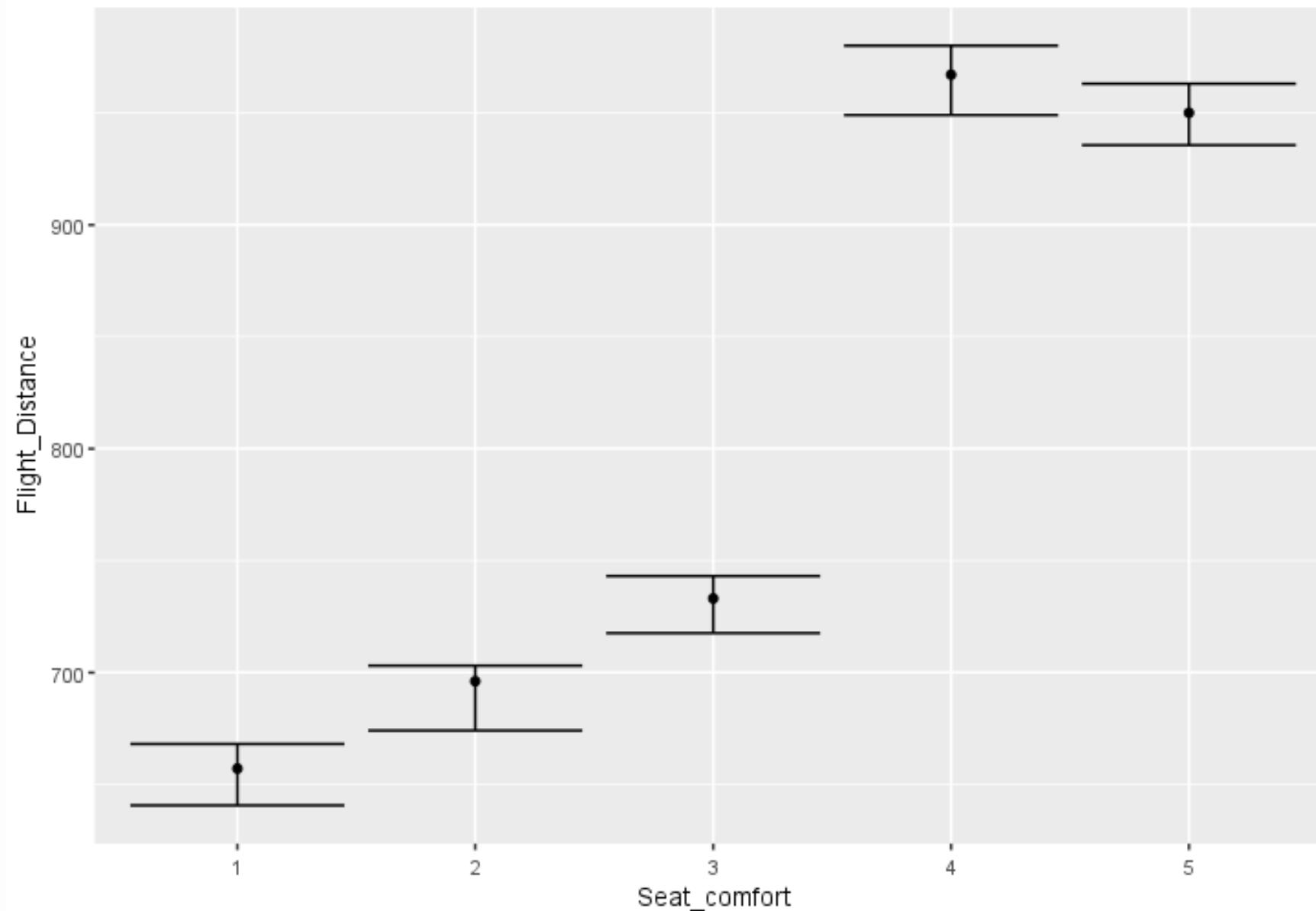
Це дозволяє нам підтвердити гіпотезу H0, що пасажири бізнес-класу подорожують далі, порівняно з пасажирами економ-класу.

Гіпотеза: Клієнти дають оцінку "4" за комфорт сидіння на більшій відстані Flight\_Distance, ніж "5".

**H0:** Медіана для значення 4 менше або дорівнює медіані для значення 5.

**p-value** = 2.2e-16

Отримали результат 2.2e-16, це означає, що є статистично значимі докази на підтримку того, що медіана першої вибірки є більшою за медіану другої вибірки на заданому рівні значущості (зазвичай 0.05).



# ВИСНОВКИ

01

Не можна стверджувати, що люди похилого віку поставили нижчу оцінку, ніж люди середнього віку.

02

Середній рейтинг сервісу Online boarding може відрізнятися між особами середнього віку та похилого віку

03

Люди похилого віку можуть бути менш вимогливі, ніж діти, щодо рейтингу сервісу щодо їжі та напоїв

04

Дітям і людям похилого віку однаково зручно оформлювати онлайн-бронювання.

05

Діти та люди похилого віку ставлять різні оцінки сервісу щодо розваг на борту



# ВИСНОВКИ

06

Людям похилого віку і дітям однаково важливі чисті умови

07

Немає достатніх доказів для підтримки заявленої залежності між віком користувачів і оцінкою сервісу Inflight\_entertainment

08

Існує статистично значима залежність між оцінкою Inflight wifi service та довжиною перельоту

09

Пасажири бізнес-класу подорожують на більші відстані порівняно з пасажирами економ-класу.

10

Існує статистично значима залежність між оцінкою комфорту сидіння та відстанню перельоту



# СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Smith, J., Johnson, A., & Brown, L. (2020). *Customer Satisfaction in the Airline Industry: A Comparative Study*. *Journal of Air Transportation*, 45(2), 78-93.
2. Davis, R., Thompson, M., & Wilson, S. (2019). *Analyzing Airline Passenger Satisfaction: A Data Mining Approach*. *International Journal of Business Analytics*, 15(3), 112-128.
3. Lee, H., Kim, S., & Park, C. (2021). *Understanding Factors Influencing Airline Passenger Satisfaction: A Case Study of XYZ Airlines*. *Journal of Travel and Tourism Research*, 28(1), 45-6