



U.S. Department of Transportation



# Analyzing Nonresponse in National Census of Ferry Operators

2022 FCSM Research & Policy Conference

*Session E-3: Modernizing Bias Reduction Techniques*

2022-10-26

Young-Jun Kweon  
Bureau of Transportation Statistics

# Disclaimer

---

This study was performed under the sponsorship of the Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for its contents or use thereof.

# Outline

---

- INTRODUCTION  
NCFO / NRB Study
- DATA  
Questionnaire / Examples / Released Tables
- METHODS  
Data Quality Metrics / Bias Estimation / Influential Variables
- RESULTS
- CONCLUSIONS

# INTRODUCTION



U.S. Department  
of Transportation

# National Census of Ferry Operators

## INTRODUCTION

- The Safe, Accountable, Flexible Efficient, Transportation Equity Act—A Legacy for Users (SAFETEA-LU) of 2005 (P.L. 114-94) requires BTS to maintain a national ferry database.
- BTS conducts a biennial census of all ferry operators in the U.S. and its territories.
  - Who should be included?
    - Ferry operators providing itinerant, fixed route, common carrier passenger/vehicle roll-on, roll-off (RoRo) ferry service, and railroad car float operations
  - Who should NOT be included?
    - Non-itinerant operations (e.g., cruise-to-nowhere services)
    - Excursion (e.g., whale watches, casino boats, dinner cruises, etc.)
    - Passenger only water taxi services not operating on a fixed route
    - LoLo (Lift-on/Lift-off) freight/auto carrier services
    - Long distance passenger only cruise ship services

# 2020 Nonresponse Bias Study

## INTRODUCTION

- **Motivations:**

- ✓ Recommended by OMB Guidelines (2006)
- ✓ Improve data quality of NCFO
- ✓ Perform trend analysis of national total boarding counts

- **Purposes:**

- ✓ To improve data quality of NCFO
- ✓ To inform data users of potential bias in analyzing 2020 NCFO data

- **2020 NCFO**

- ✓ 247 operators invited (frame)
- ✓ 164 operators participated

# DATA



U.S. Department  
of Transportation

# 2020 NCFO Questionnaire

- 5 Sections: (1) Operator Info, (2) Vessel Info, (3) Terminal Info, (4) Segment Info, and (5) Funding Info.

1. Please provide Company	2. Are you a vessel owner or operator?	
3. Please indicate if your operation is a vessel or a terminal.		
4. Please indicate if your operation is a vessel or a terminal.		
5. Please list and provide the vessel number and name of each vessel in your fleet during calendar year 2019.		
6. For each vessel in your fleet during calendar year 2019, please indicate the percentage of time the vessel was used for passenger service.		
7. For each vessel in your fleet during calendar year 2019, please indicate the percentage of time the vessel was used for cargo service.		
8. For each vessel in your fleet during calendar year 2019, please indicate the percentage of time the vessel was used for both passenger and cargo service.		
9. For each vessel in your fleet during calendar year 2019, please indicate the percentage of time the vessel was used for other services.		
10. For each vessel in your fleet during calendar year 2019, please indicate the percentage of time the vessel was used for non-passenger services.		
11. For each vessel in your fleet during calendar year 2019, please indicate the percentage of time the vessel was used for non-cargo services.		
12. Please list the name and address of each terminal in your fleet during calendar year 2019.		
13. For each terminal in your fleet during calendar year 2019, please indicate the percentage of time the terminal was used for passenger service.		
14. For each terminal in your fleet during calendar year 2019, please indicate the percentage of time the terminal was used for cargo service.		
15. For each terminal in your fleet during calendar year 2019, please indicate the percentage of time the terminal was used for both passenger and cargo service.		
16. Please list the individual route segments defined by the direction of travel.		
17. Individual route segments are made up of multiple segments.		
18. Individual route segments may be made up of more than one segment.		
19. For each route segment, please indicate the percentage of time the route segment is regularly traveled.		
20. For each route segment, please indicate the percentage of time the route segment is irregularly traveled.		
21. Please indicate the percentage of time the route segment is not traveled.		
22. Does this Operation Receive Public Funding Sources?		
<input type="radio"/> NO <input checked="" type="radio"/> Public Funding Sources are NOT Accepted		
23. Please indicate whether your operation's boarding information or any other information you provided is business-sensitive information. (Please note: Information that you release to the public on a routine basis generally does not qualify as business-sensitive information).		
<input type="radio"/> Boarding Information is NOT business-sensitive		
Thank you for completing the 2020 NCFO!		
<b>Please return this survey in the enclosed envelope or send to:</b> NCFO Project Manager, US Department of Transportation 1200 New Jersey Avenue SE, RTS-32, Room E32-316, Washington, D.C. 20590		
If you have any questions please contact us at: <a href="mailto:FERRY@DOT.GOV">FERRY@DOT.GOV</a> or 1-800-853-1351		
23 of 23		
20 of 23		

# Ferry Operation: Example 1

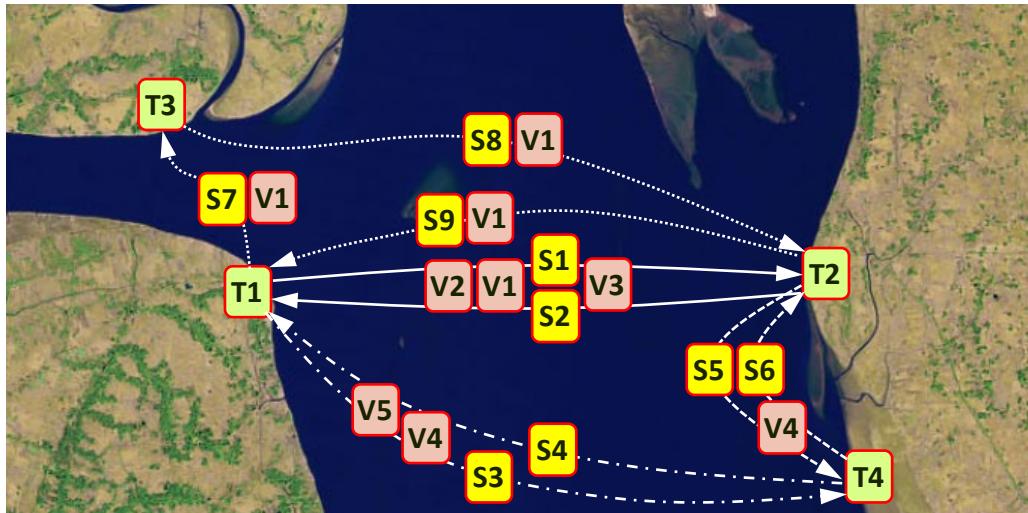
- An operator serves 2 segments connecting 2 terminals with 1 vessel



# Ferry Operation: Example 2

DATA

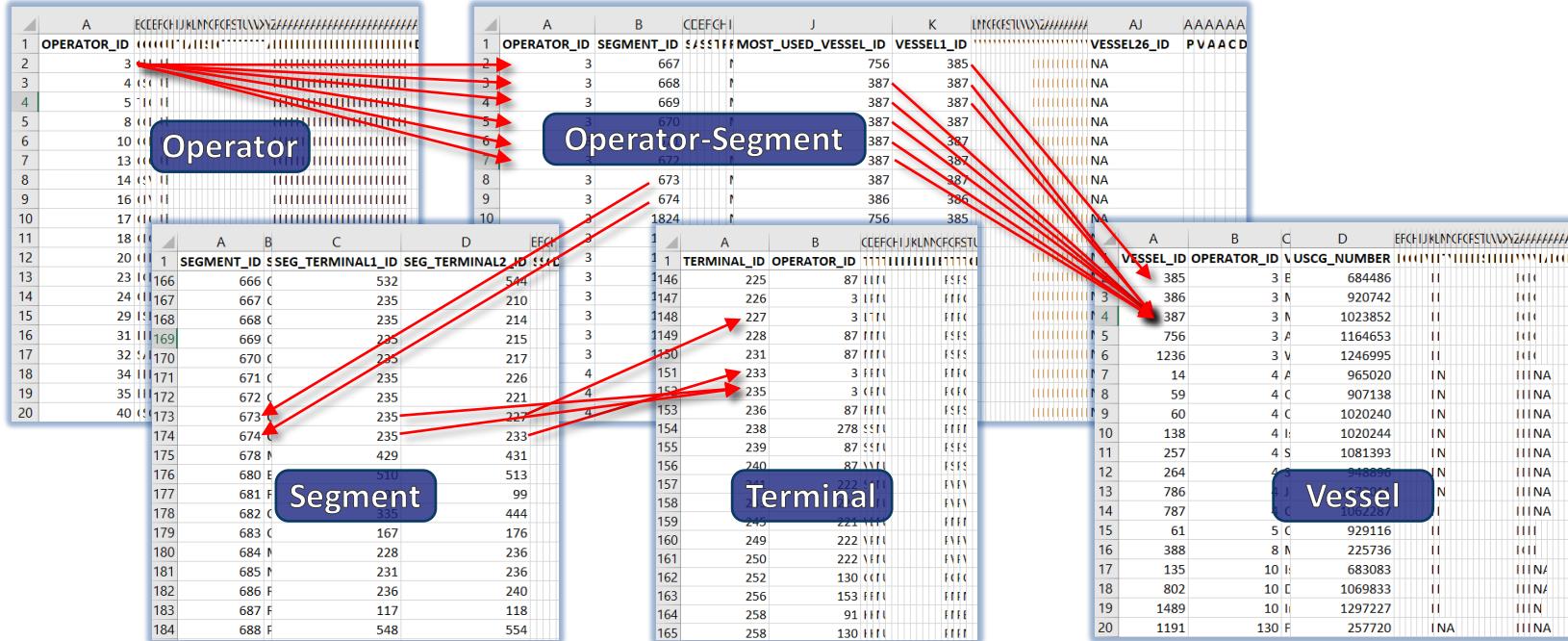
- An operator serves 9 segments connecting 4 terminals with 5 vessels



# 2020 NCFO Data Release

DATA

5 Tables	Operator	Operator-Segment	Segment	Terminal	Vessel
158 Variables	51	42	9	21	35



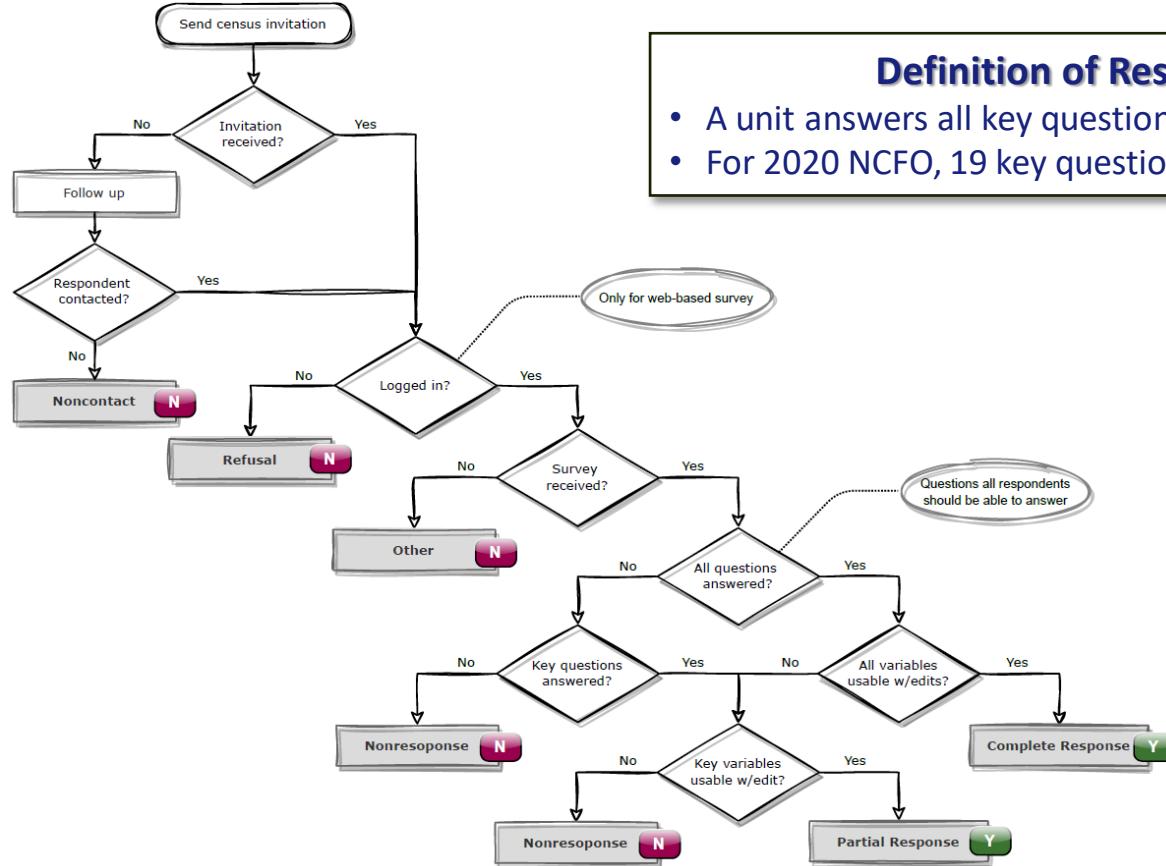
# METHODS



U.S. Department  
of Transportation

# Unit Response

## METHODS



## Definition of Response

- A unit answers all key questions of NCFO
- For 2020 NCFO, 19 key questions were designated.

# Data Quality Metrics

- **6 Data Quality Rates**

- **2 Unit-Level Rates**

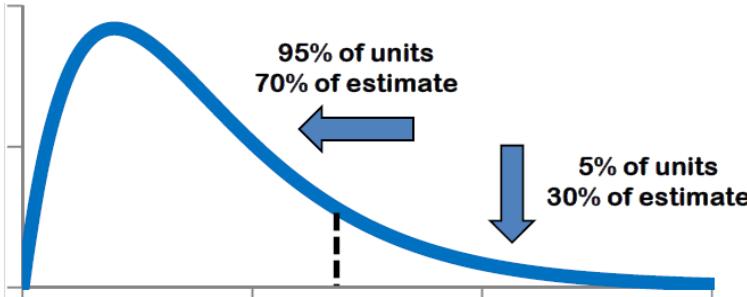
- ① Unit Participation Rate (UPR) 
- ② Unit Response Rate (URR)

- **4 Item-Level Rates**

- ③ Item Response Rate (IRR)
- ④ Total Item Response Rate (TIRR)
- ⑤ Overall Item Response Rate (OIRR) 
- ⑥ Modified Quantity Response Rate (MQRR) 

✓ Proportion of observed total value of item  $x$

Lineback and Thompson (2010). *Conducting Nonresponse Bias Analysis for Business Surveys*. JSM 2010.



# Bias Estimation ①

- **Define Nonresponse Bias**

- In sample mean of survey (OMB, 2006)

$$Bias(\bar{x}_R) = \bar{x}_R - \bar{x}_S = \left( \frac{n_{NR}}{n_S} \right) (\bar{x}_R - \bar{x}_{NR})$$

- In sample **total of census** (2020 NCFO)

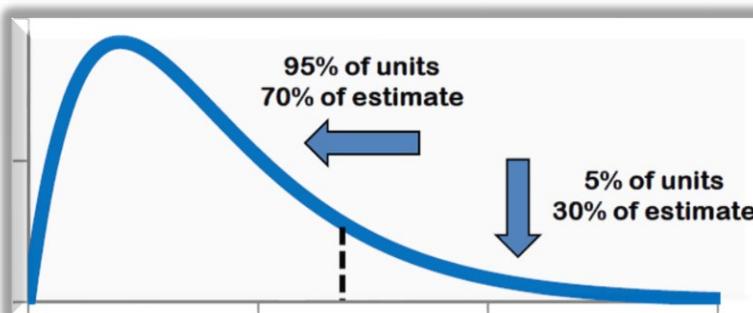
$$Bias(\hat{T}_{x,R}) = \hat{T}_{x,R} - \hat{T}_{x,S} \text{ where } \hat{T}_{x,S} = \hat{T}_{x,R} + \hat{T}_{x,NR}$$

# Bias Estimation ②

- Calculate  $\hat{T}_{x,NR}$  for Passenger Boarding

- Use growth ratio of 2018-2020:  $Growth\ Ratio(PASS) = \frac{\hat{T}_{PASS,R,2020}}{\hat{T}_{PASS,R,2018}}$
- (e.g.)  $\widehat{PASS}_{i,NR,2020} = Growth\ Ratio(PASS) \times PASS_{i,R,2018}$
- Growth ratio by Measure of Size (MOS)

Group	Description	Definition	Num. of Operators	Total Boarding	Proportion
1	Extra Large (XL)	$\geq 10,000,000$	3	60,942,777	40%
2	Large (L)			14,403,641	27%
3	Medium (M)			14,583,713	23%
4	Small (S)			4,845,905	10%





# Influential Variables

- **Method:** Conditional Tree Analysis
- **Variables considered:**
  - ① Accept Public Fund (Yes, No)
  - ② Serve National Park Service (Yes, No)
  - ③ Report on behalf of government (Yes, No)
  - ④ Segment Type (Interstate, Intrastate, International)
  - ⑤ Population (Based on American Community Survey's 2019 estimate)
  - ⑥ State (Based on operator's address)



*→ Data Collection Year*

**2020 NCFO**

**2022 NCFO**

*→ Data Year*

# RESULTS



# 2 Unit-Level Rates

- Unit Participation & Response Rates: UPR & URR

Subgroup	Category	UPR	URR
All	None	66%	50%
Report On Behalf of Government	No	67%	46%
	Yes	90%	73%
Ticket Revenue	< 50%	87%	67%
	≥ 50%	73%	56%
Number of Segments	2	77%	59%
	3-6	72%	57%
	6+	87%	61%
Number of Vessels	1-2	71%	55%
	3-6	82%	63%
	6+	82%	56%

Higher UPR & URR (Report On Behalf of Gov.)

Higher UPR & URR (Ticket Revenue < 50%)

Higher UPR (6+ Segments)

Higher UPR (3+ Vessels)

# 4 Item-Level Rates ①

- Item Response Rate (IRR) and Total Item Response Rate (TIRR)

19 Key Response Variable (x)	Table	IRR <sub>x</sub>	TIRR <sub>x</sub>
Accept Public Funding	Operator	100.0%	50.0%
Operator City		100.0%	50.0%
Operator State		100.0%	50.0%
Operator Name		100.0%	50.0%
Average Trip Time	Operator Segment	81.1%	40.5%
Most Used Vessel		81.1%	40.5%
Passenger Boarding		81.1%	40.5%
Segment Length		81.1%	40.5%
Season End Date	Segment	78.7%	39.3%
Season Start Date		78.7%	39.3%
Vehicle Boarding		81.1%	40.5%
Segment Name		81.1%	40.5%
Terminal 1 City	Terminal	98.8%	49.4%
Terminal 2 City		89.0%	44.5%
Terminal 1 State		98.8%	49.4%
Terminal 2 State		89.0%	44.5%
Terminal 1 Name	Vessel	98.8%	49.4%
Terminal 2 Name		89.0%	44.5%
Vessel Name (of Most User Vessel)	Vessel	81.1%	40.5%

Lower IRR & TIRR

# 4 Item-Level Rates ②

- Overall Item Response Rate (OIRR) & Modified Quantity Response Rate (MQRR)

Variable ( $x$ )	OIRR <sub>x</sub>	MQRR <sub>x</sub>
Passenger Boarding	54.3%	77.2%

Based on the number  
of respondents in  $x$

Based on values in  $x$   
respondents provided

# Bias Estimate (Passenger Boardings)

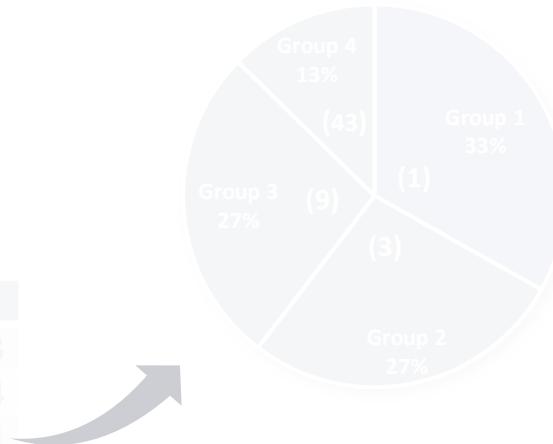
- Growth Ratios by MOS Group

Group	Growth Ratio
1 (Extra-Large Operators)	0.992
2 (Large Operators)	0.981
3 (Medium Operators)	1.105
4 (Small Operators)	1.053

- Estimated Bias in Total Boardings

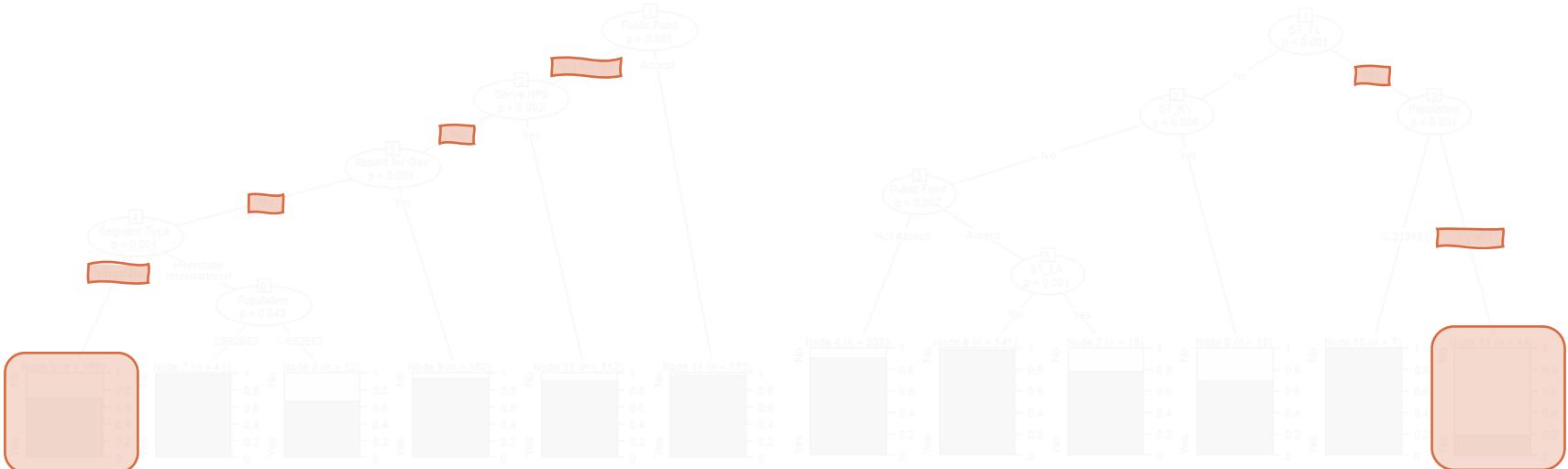
Group	Nobs	Type	2020 NCFO
All	177	Total Bias	
		Observed	113,990,893
		Estimated Number	153,852,554
		Percentage	+35,861,661
			-35%

- Contribution to Total Bias in 2020 NCFO



# Influential Variables

- Nonresponse to All 19 Key Variables
  - Without State Variables
  - With State Variables



Not accept public fund  
 Not serve National Park,  
 Not report on behalf of government  
 Serve intrastate segment

In Florida  
 Terminals in an area with population  $> \sim 300K$

# CONCLUSIONS



U.S. Department  
of Transportation

# Conclusions

- Rates vary across subgroups (e.g., reporting on behalf of government, ticket revenue, and fleet size)
  - ✓ Data users should be cautious when focusing their analysis on a specific group
- Bias in total passenger boarding is affected largely by nonresponding big operators
  - ✓ Estimating national total boarding count is challenging
- Operators in few States with certain characteristics have lower responses.

# Recommendations

- **RECOMMENDATION:** Develop a monitoring application keeping track of response on each of the key items
  - ✓ **Use:** The application identifies nonresponding ferry operators as a census rolls out.
  - ✓ **Implication:** This recommendation would increase the unit response rate (URR).
- **RECOMMENDATION:** Develop a list of ferry operators with a group indication based on boarding counts.
  - ✓ **Use:** This list helps BTS identify which ferry operators would be critical in reducing bias in the national total boarding estimate.
  - ✓ **Implication:** This recommendation would increase the unit response rate (URR) and the item-level rates on the 2 key items (passenger and vehicle boardings).

# Acknowledgement

- **Aubrey Nguyen**, IT Auditor, U.S. Government Accountability Office (GAO)  
*ORISE Fellow, USDOT (Formerly)*
- **Clara Reschovsky**, NCFO Program Manager/Survey Statistician, USDOT



**U.S. Department of Transportation**



# Contact

Young-Jun Kweon

[young-jun.kweon@dot.gov](mailto:young-jun.kweon@dot.gov)

NCFO

[ferry@dot.gov](mailto:ferry@dot.gov)



Thank You!

Questions?



U.S. Department  
of Transportation