

# Capstone EDA

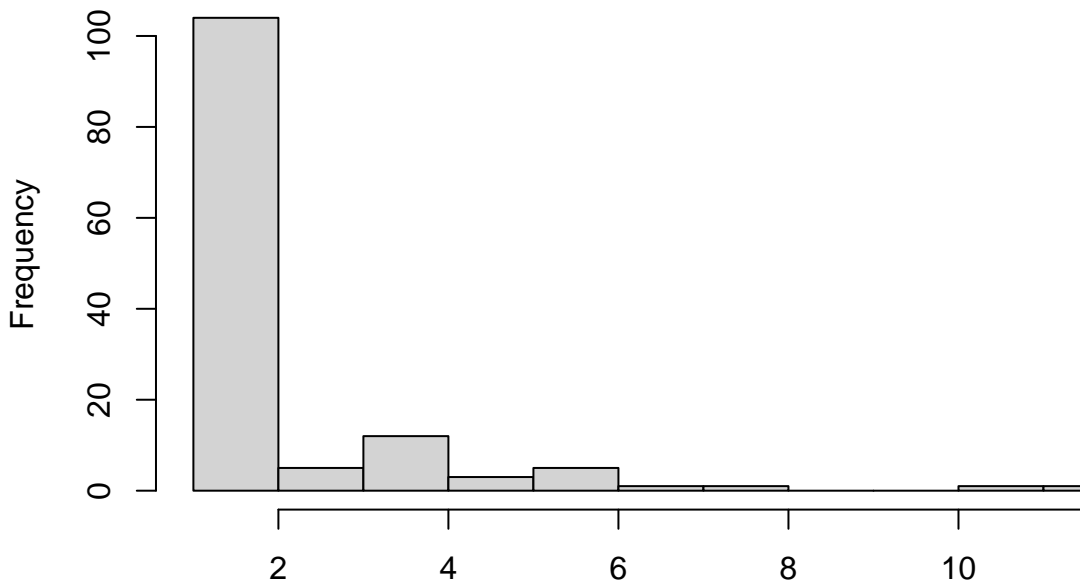
Gabby Lopez

2025-02-02

## Descriptive Statistics

Changing any cases where PreDM=1 and T2DM=1 to PreDM=0 for the descriptive statistics

**Histogram of koh.table1\$koh.counts**



distribution of KOH attendance

koh.table1\$koh.counts

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.				
##	1.000	1.000	1.000	2.045	2.000	12.000				
##	1	2	3	4	5	6	7	8	11	12
##	82	22	5	12	3	5	1	1	1	1

## Primary Objective - Hypertension pre and post analysis datasets

subsetting koh patients with hypertension

subsetting marshallese non-koh pts with HTN

## Primary objective hypertension longitudinal analysis datasets

koh patients

marshallese non-koh patients

## Primary Objective - Diabetes pre and post analysis datasets

subsetting koh patients with prediabetes or T2DM

subsetting marshallese non-koh pts with diabetes

## Primary objective Diabetes longitudinal analysis datasets

koh pts

marshallese non-koh pts

make table 1 dataset for primary objective 249 unique patients 127 from bp and 221 from alc

table 1 for primary objective population

## Adding missing grouping variables: 'UniqueIdentifier'

	KOH Participant	Non-Participant
Baseline Characteristics	N = 71	N = 178
<b>Age</b>	62 (32, 80)	56 (19, 89)
<b>Sex</b>		
Female	45 (63%)	107 (60%)
Male	26 (37%)	71 (40%)
<b>KOH Meetings Attended</b>	2.00 (1.00, 11.00)	0.00 (0.00, 0.00)
<b>Hypertension</b>	42 (59%)	98 (55%)
<b>Pre-Diabetes</b>	3 (4.2%)	8 (4.5%)
<b>Type II Diabetes</b>	68 (96%)	164 (92%)
<b>Hypertension and T2DM</b>	42 (59%)	92 (52%)
<b>Clinic Location</b>		
Maple	14 (23%)	67 (38%)
Market	46 (77%)	111 (62%)
Missing	11	0
<b>Housing Status</b>		
Housed	28 (39%)	108 (61%)
Doubling Up	43 (61%)	62 (35%)
Permanent Supportive Housing	0 (0%)	5 (2.8%)

Baseline Characteristics	KOH Participant	Non-Participant
	N = 71	N = 178
Transitional	0 (0%)	1 (0.6%)
Other	0 (0%)	2 (1.1%)
Unknown	0 (0%)	0 (0%)
<b>Income Level</b>	0 (0, 294)	43 (0, 295)
Missing	1	5
<b>Risk Score</b>	2 (0, 9)	2 (0, 12)
Missing	0	9

table 1 for all Marshallese and non-Marshallese

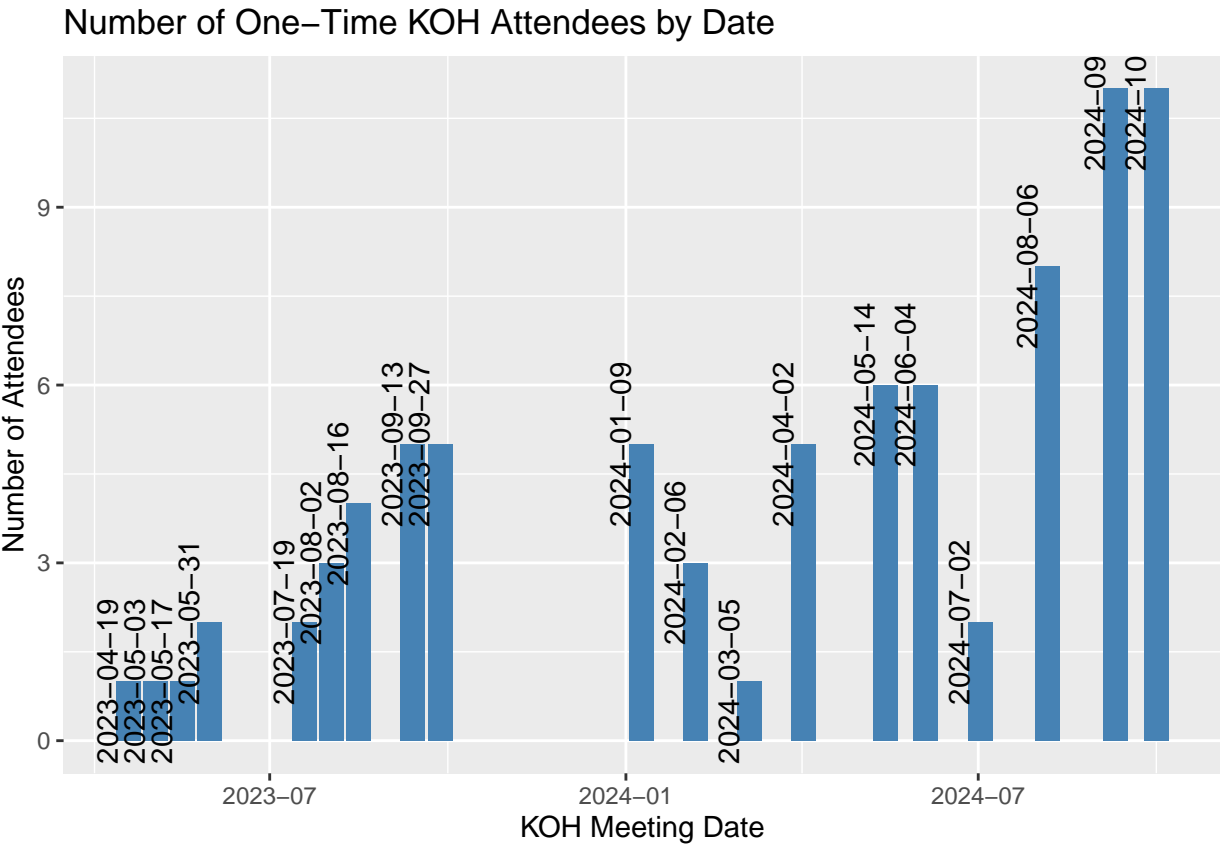
Baseline Characteristics	Overall N = 27,687 <sup>1</sup>	Marshallese N = 869	Non-Marshallese N = 26818
<b>Age</b>			
Median (Min, Max)	41 (18, 102)	46 (18, 89)	41 (18, 102)
<b>Sex</b>			
Female	16,223 (59%)	576 (66%)	15,647 (58%)
Male	11,464 (41%)	293 (34%)	11,171 (42%)
<b>Hypertension</b>	6,639 (24%)	211 (24%)	6,428 (24%)
<b>Pre-Diabetes</b>	1,306 (4.7%)	35 (4.0%)	1,271 (4.7%)
<b>Type II Diabetes</b>	3,045 (11%)	386 (44%)	2,659 (9.9%)
<b>Hypertension and T2DM</b>	2,708 (9.8%)	190 (22%)	2,518 (9.4%)
<b>Clinic Location</b>			
Maple	19,016 (69%)	429 (52%)	18,587 (69%)
Market	8,622 (31%)	391 (48%)	8,231 (31%)
Missing	49	49	0
<b>Housing Status</b>			
Housed	20,628 (75%)	547 (63%)	20,081 (75%)
Doubling Up	5,585 (20%)	304 (35%)	5,281 (20%)
Permanent Supportive Housing	350 (1.3%)	7 (0.8%)	343 (1.3%)
Transitional	626 (2.3%)	4 (0.5%)	622 (2.3%)
Other	444 (1.6%)	6 (0.7%)	438 (1.6%)
Unknown	54 (0.2%)	1 (0.1%)	53 (0.2%)
<b>Income Level</b>			
Median (Min, Max)	117 (0, 25,568)	43 (0, 461)	120 (0, 25,568)

Baseline Characteristics	Overall N = 27,687 <sup>1</sup>	Marshallese N = 869	Non-Marshallese N = 26818
Missing	4,607	68	4,539
<b>Risk Score</b>			
Median (Min, Max)	1 (0, 22)	1 (0, 12)	1 (0, 22)
Missing	3,148	73	3,075

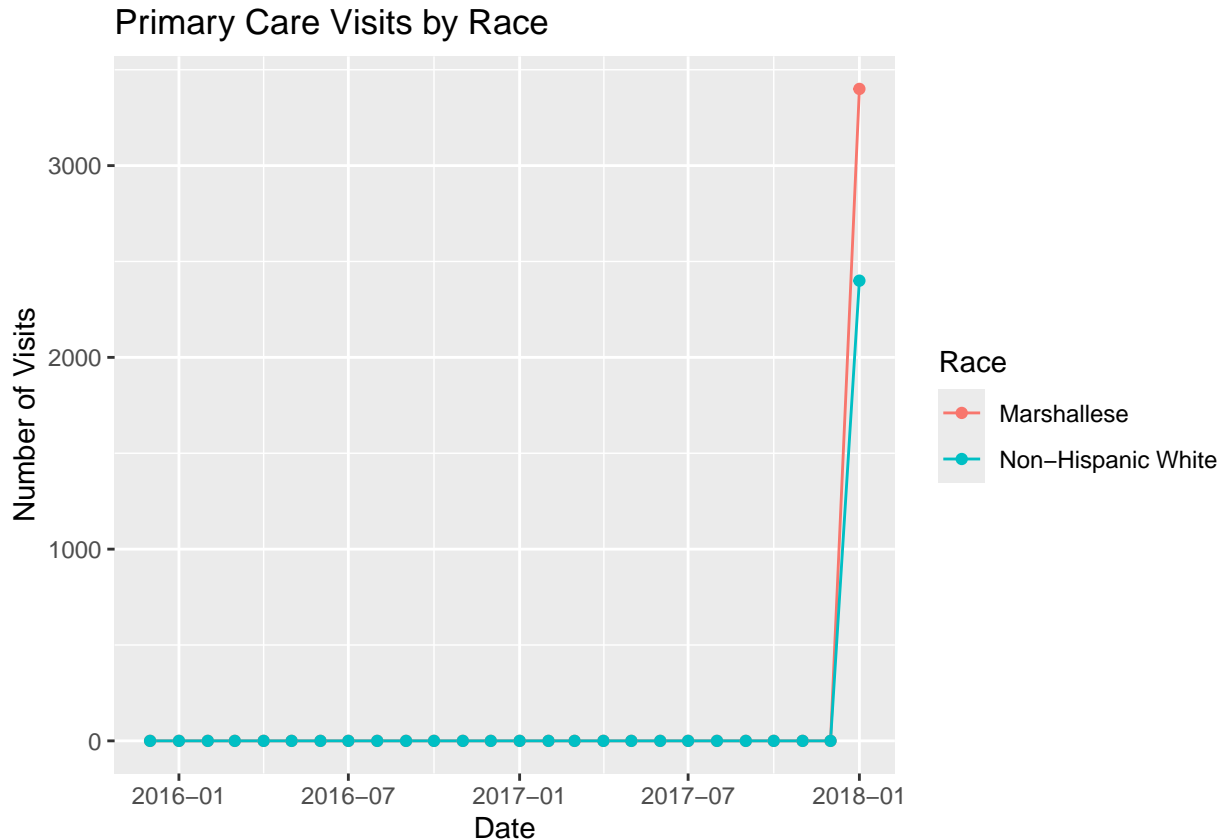
<sup>1</sup>n (%)

## One-Time KOH Attendance Distribution

## Adding missing grouping variables: 'UniqueIdentifier'



## Preliminary Parallel Trends EDA



## Encounters Table

i think i still need to adjust this table a bit but it's fine for now. I removed visit types (Health Equity and Other) that had 0 visits for multiple years as well. 2016, 2017, and 2025 didn't have much data either so I removed them. Then I removed years 2018 and 2019 just to make the table a bit smaller since it was so wide.

```
## Warning: fonts used in 'flectable' are ignored because the 'pdflatex' engine is
## used and not 'xelatex' or 'lualatex'. You can avoid this warning by using the
## 'set_flectable_defaults(fonts_ignore=TRUE)' command or use a compatible engine
## by defining 'latex_engine: xelatex' in the YAML header of the R Markdown
## document.
```

	2020		2021	
	Marshallese	Non-Marshallese	Marshallese	Non-Marshalles
Mean Visits per Patient	N = 354 <sup>1</sup>	N = 14849 <sup>1</sup>	N = 407 <sup>1</sup>	N = 16300 <sup>1</sup>
Primary Care	2.1 (0.0, 23.0)	2.6 (0.0, 81.0)	3.8 (0.0, 25.0)	5.9 (0.0, 217.0)
Emergency	0.6 (0.0, 11.0)	0.9 (0.0, 44.0)	1.1 (0.0, 16.0)	1.5 (0.0, 101.0)
Urgent Care	0.5 (0.0, 5.0)	0.6 (0.0, 14.0)	0.9 (0.0, 12.0)	1.1 (0.0, 32.0)

	2020		2021	
	Marshallese	Non-Marshallese	Marshallese	Non-Marshallese
Mean Visits per Patient	N = 354 <sup>1</sup>	N = 14849 <sup>1</sup>	N = 407 <sup>1</sup>	N = 16300 <sup>1</sup>
<b>Behavioral Health</b>	0.0 (0.0, 3.0)	0.7 (0.0, 49.0)	0.1 (0.0, 8.0)	1.3 (0.0, 96.0)
<b>Pharmacy</b>	0.6 (0.0, 14.0)	0.2 (0.0, 71.0)	1.3 (0.0, 38.0)	0.4 (0.0, 117.0)
<b>Dental</b>	0.5 (0.0, 12.0)	0.8 (0.0, 26.0)	1.3 (0.0, 16.0)	2.1 (0.0, 52.0)
<b>Multispecialty</b>	0.6 (0.0, 29.0)	0.2 (0.0, 25.0)	1.3 (0.0, 49.0)	0.5 (0.0, 56.0)

<sup>1</sup>Mean (Minimum, Maximum)