v2

Maiko Hata

A. Table of 10 exit reasons

Table 1: Table of Exit Reasons

Exit Reasons	Exit Category Codes
Program completion	Category (C) 1: A child is no longer eligible for Part C prior to reaching age three
Exit at age three	C2: A child is exiting Part C and has been determined to be eligible for Part B
Exit at age three	C3: Part B eligible, continuing in Part C
Exit at age three	C4: Not eligible for Part B, exit with referrals to other programs
Exit at age three	C5: Not eligible for Part B, exit with no referrals
Exit at age three	C6: Part B eligibility not determined
Not receiving services	C7: Deceased
Not receiving services	C8: Moved out of state
Not receiving services	C9: Withdrawal by parent (or guardian)
Not receiving services	C10: Attempts to contact the parents and/or child were unsuccessful

B. National and Oregon CHILD COUNTS

NOTE TO SELF: ADD THE CENSUS NUMBER FOR FINAL PROJECT! BIND_ROWS!! WEEK 2? 3? Labs.

B-1. Load data

where did the data go wrong? Did i combine it in below? But i think I just chose and selected the one i don't need? The one below is still correct. Where is the error?

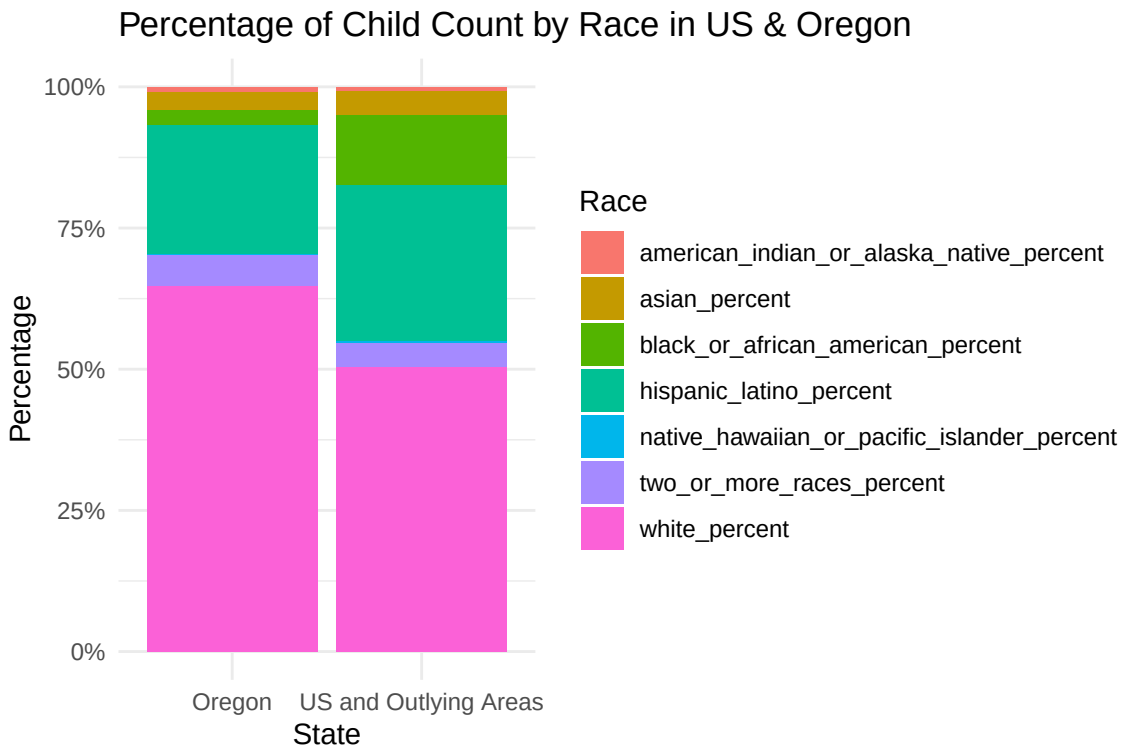
B-2: chart 1:
THIS CHUNK TO ROUND TO 2 DIGITS CONVERTED THE COLUMN TOO the distinction between the OR/US somehow. I NEED TO FIX IT

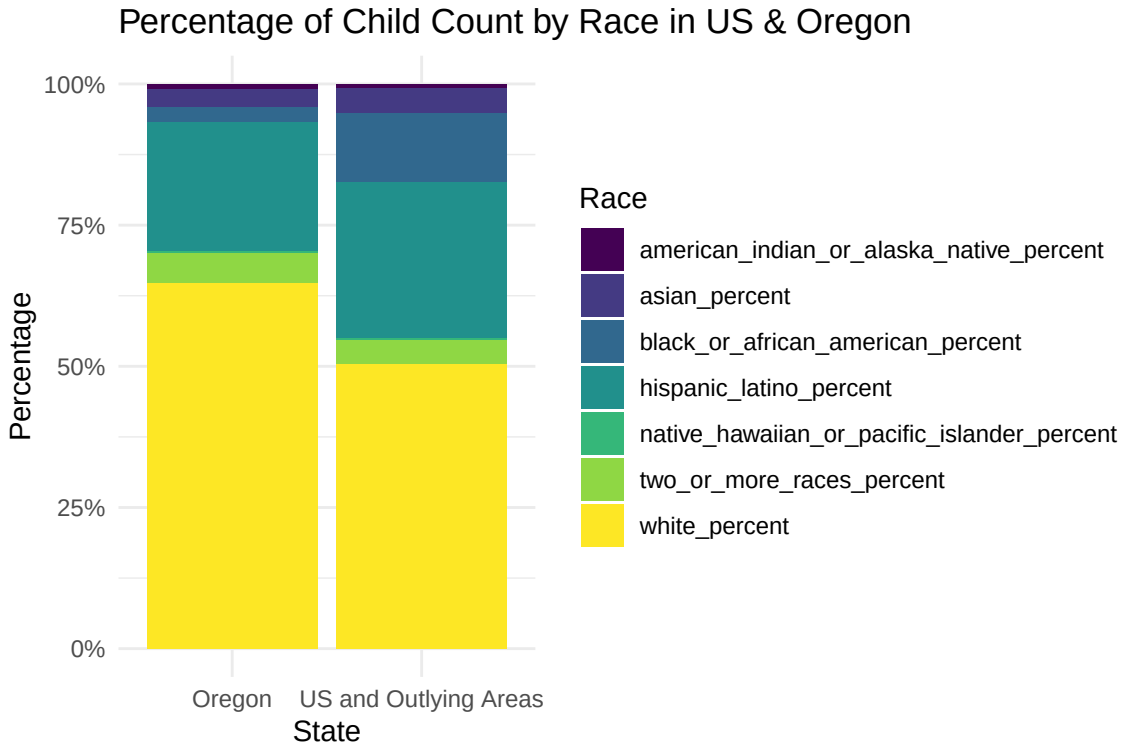
Table 2: Child Count (US & Oregon)

Category	V1	V2
state	Oregon	US and Outlying Areas
american_indian_or_alaska_native_percent	0.87	0.69
asian_percent	3.27	4.39
black_or_african_american_percent	2.69	12.35
hispanic_latino_percent	22.77	27.65
native_hawaiian_or_pacific_islander_percent	0.3	0.3
two_or_more_races_percent	5.41	4.23
white_percent	64.69	50.38

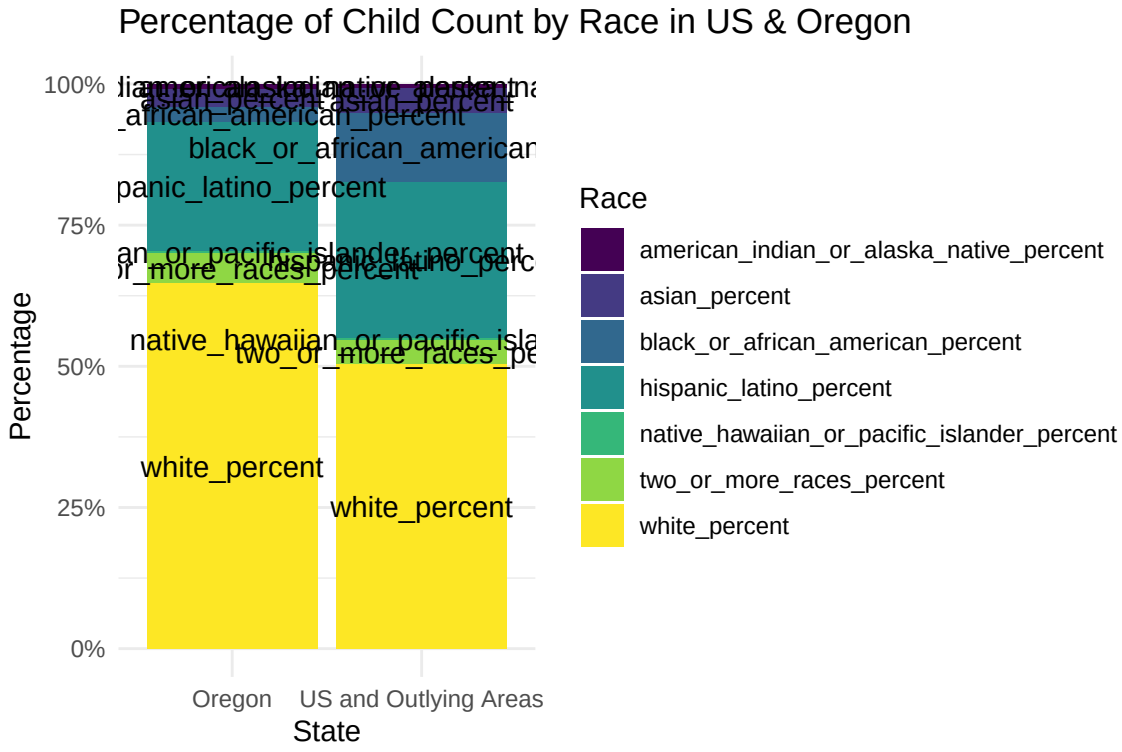
B-2: visualization 2 FIX IT!!! (OLD: SOMEWHERE ALONG THE LINE I LOST THE DATA ROWS IN DF)

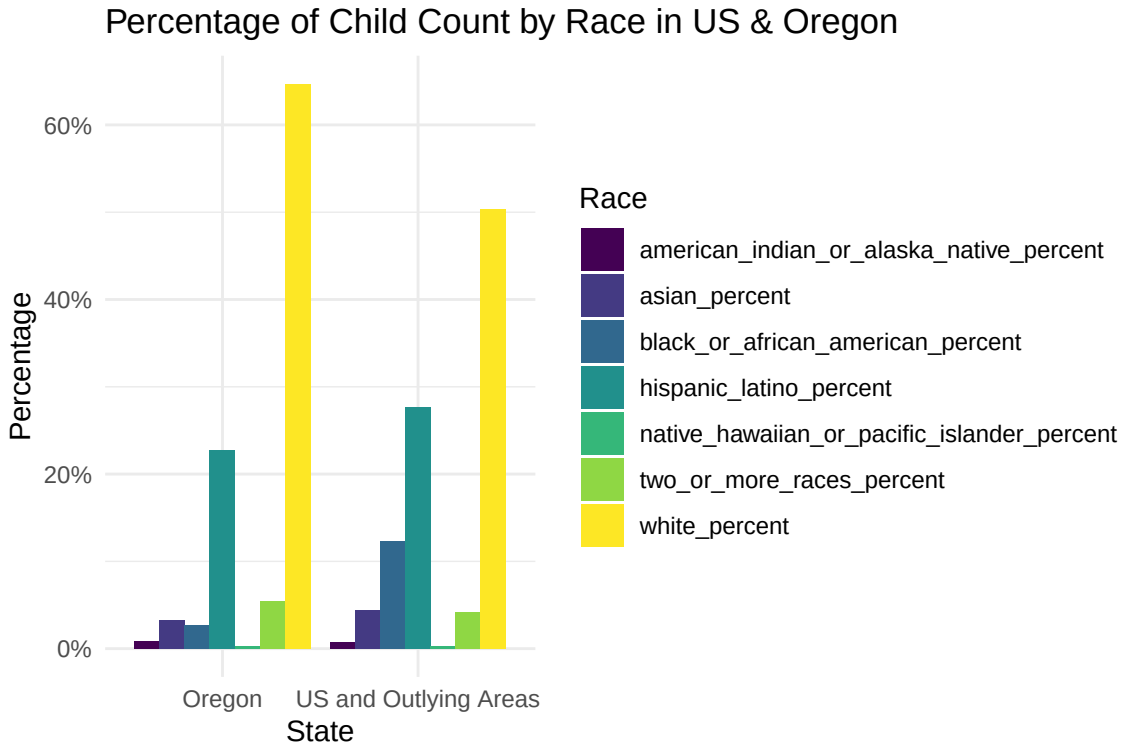
```
chr [1:14] "0.87" "0.69" "3.27" "4.39" "2.69" "12.35" "22.77" "27.65" ...
```

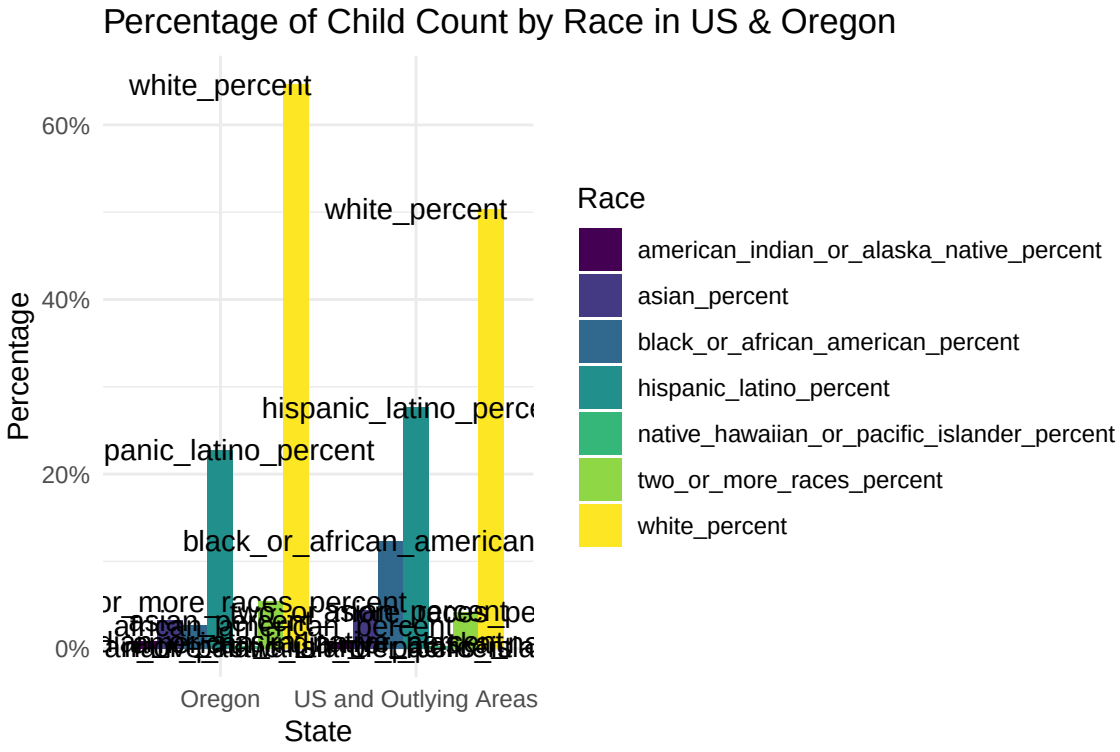


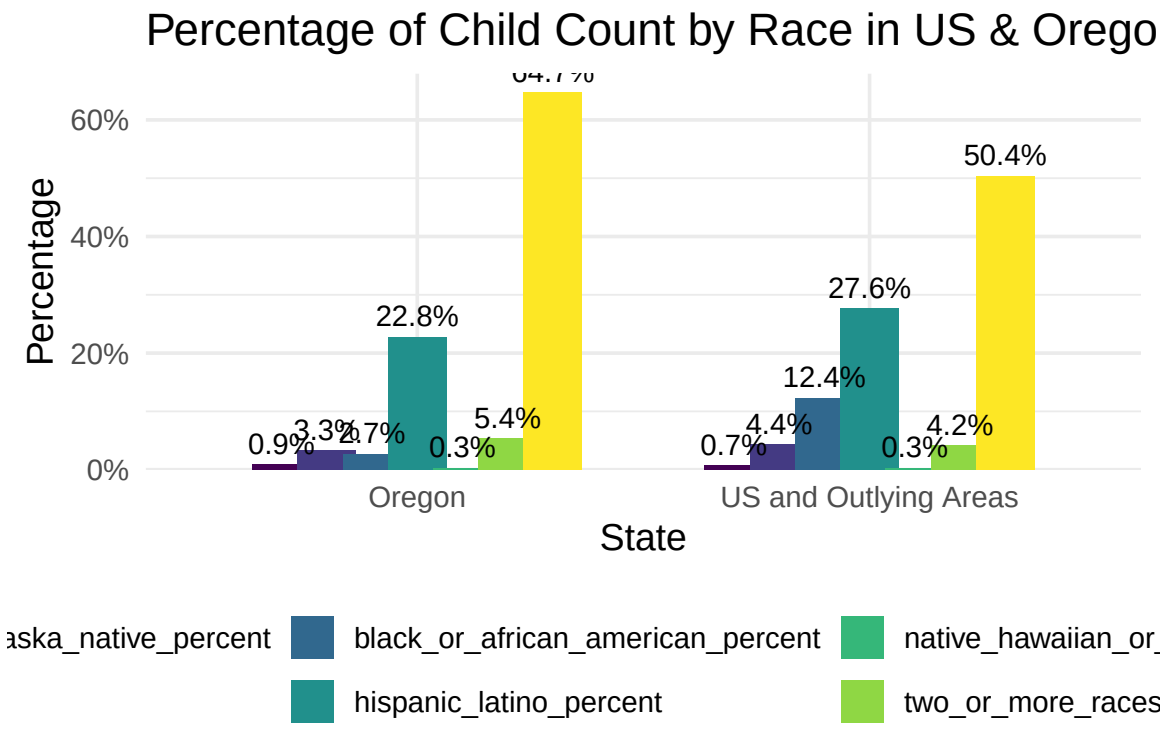


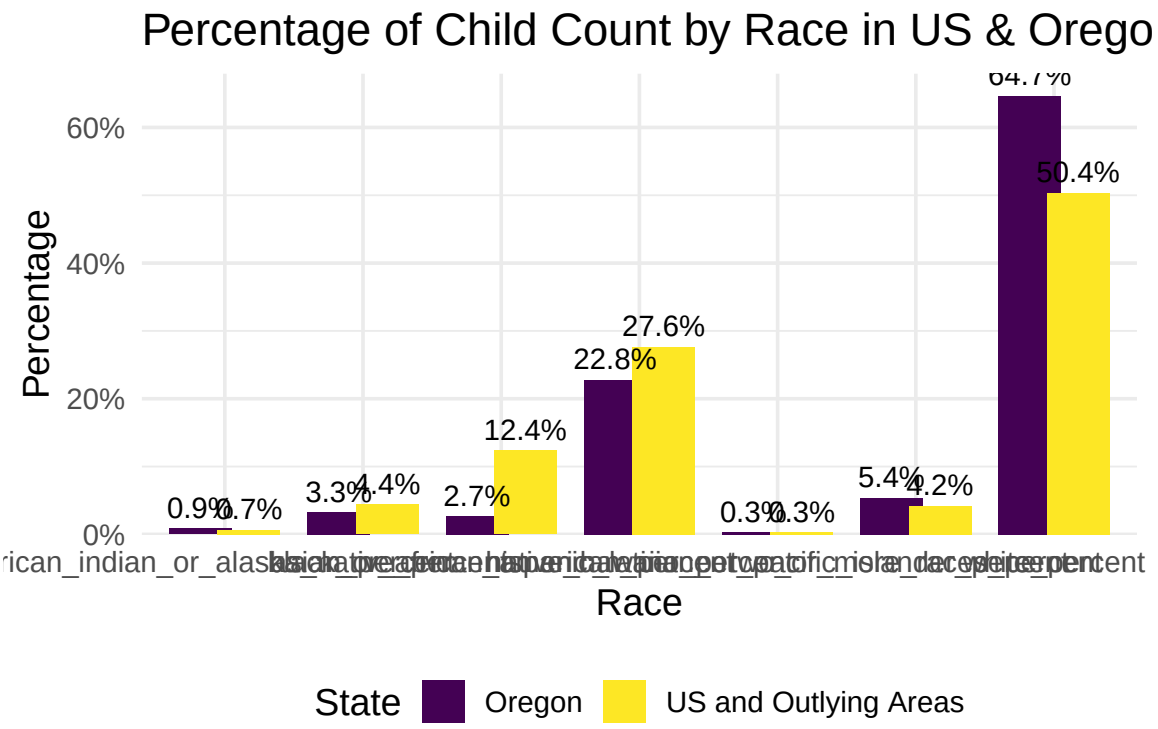
B2 Visualization 2 v.2 - it has labels on the bar but it's ugly as hell

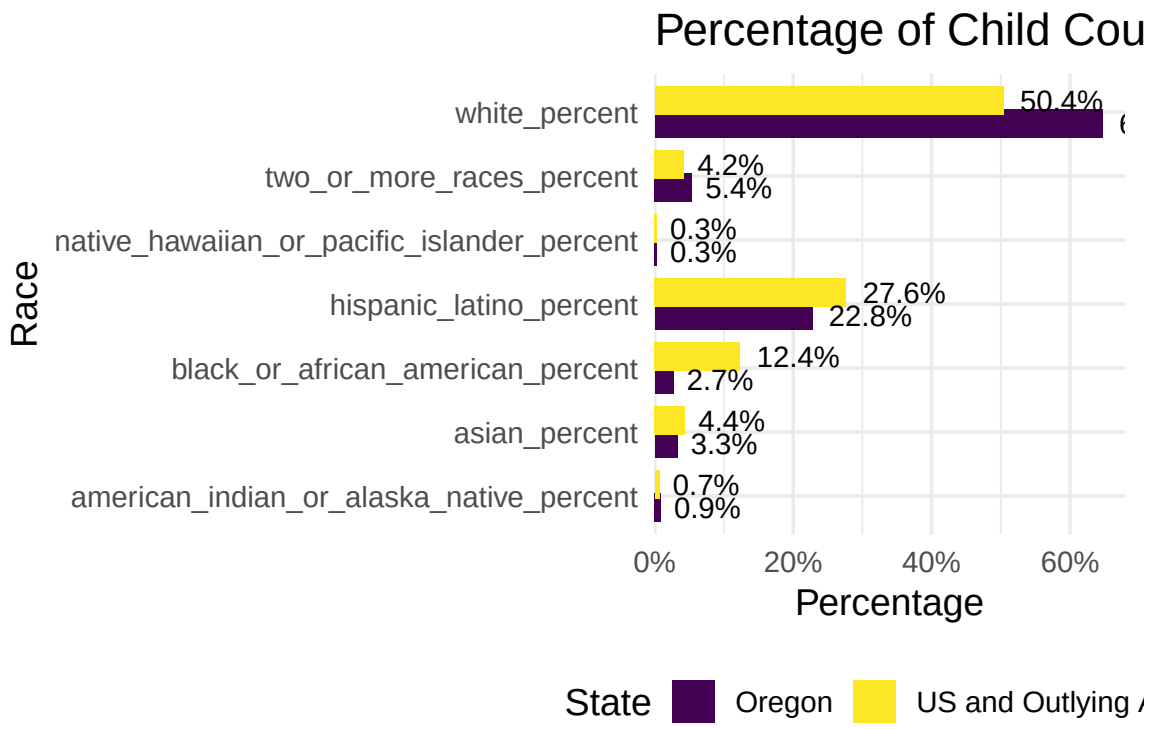

Percentage of Child Count by Race in US & Oregon
100%
75%
50%
25%
0%
Percentage
white_percent
white_percent
Oregon
US and Outlying Areas
State
Race
american_indian_or_alaska_native_percent
asian_percent
black_or_african_american_percent
hispanic_latino_percent
native_hawaiian_or_pacific_islander_percent
two_or_more_races_percent
white_percent


Percentage of Child Count by Race in US & Oregon
60%
40%
20%
0%
Percentage
Oregon
US and Outlying Areas
State
Race
american_indian_or_alaska_native_percent
asian_percent
black_or_african_american_percent
hispanic_latino_percent
native_hawaiian_or_pacific_islander_percent
two_or_more_races_percent
white_percent


Percentage of Child Count by Race in US & Oregon
white_percent
60%
white_percent
40%
hispanic_latino_perc
panic_latino_percent
20%
black_or_african_american
0%
Oregon
US and Outlying Areas
State
Percentage
Race
american_indian_or_alaska_native_percent
asian_percent
black_or_african_american_percent
hispanic_latino_percent
native_hawaiian_or_pacific_islander_percent
two_or_more_races_percent
white_percent


Percentage of Child Count by Race in US & Orego
Percentage
60%
40%
20%
0%
0.9%
3.3%
2.7%
22.8%
0.3%
5.4%
0.7%
4.4%
12.4%
27.6%
0.3%
4.2%
50.4%
Oregon
US and Outlying Areas
State
aska_native_percent
black_or_african_american_percent
native_hawaiian_or_
hispanic_latino_percent
two_or_more_races


Percentage of Child Count by Race in US & Orego
Percentage
60%
40%
20%
0%
0.9%
0.7%
3.3%
4.4%
2.7%
12.4%
22.8%
27.6%
0.3%
0.3%
5.4%
4.2%
64.7%
50.4%
Race
State
Oregon
US and Outlying Areas



C. National and Oregon EXIT data by RACE

I FIXED THE MISSING COLUMN by adding back part_b_eligibility_not_determined. I think this is what I can use for CHI-SQUARE WITH RESIDUALS?

I should be able to export df to excel this way but haven't tried it yet.

agg_by_race_and_state

OH NO where did Part B eligibility not determined go?!?!?

I'm trying out to see if I can do the chi-square with residuals (per https://chatgpt.com/share/67a1833d-9fc4-8012-8193-b6fc358a9687)

Chi-square with Residuals 1:

R doesn't like spaces or dashes / - that's why we did clean names, it could work but it can be tricky later

Chi-square with Residuals 2:

```
Pearson's Chi-squared test
```

```
data:  race_matrix
X-squared = 88130, df = 36, p-value < 2.2e-16
```

Chi-square with Residuals 3:

Cameron: Residuals are what we are measuring anyway. If nothing was happening, what would be the expected values in the cells in the matrix (so the residuals = differences between expected and what we see) so it's a raw differences

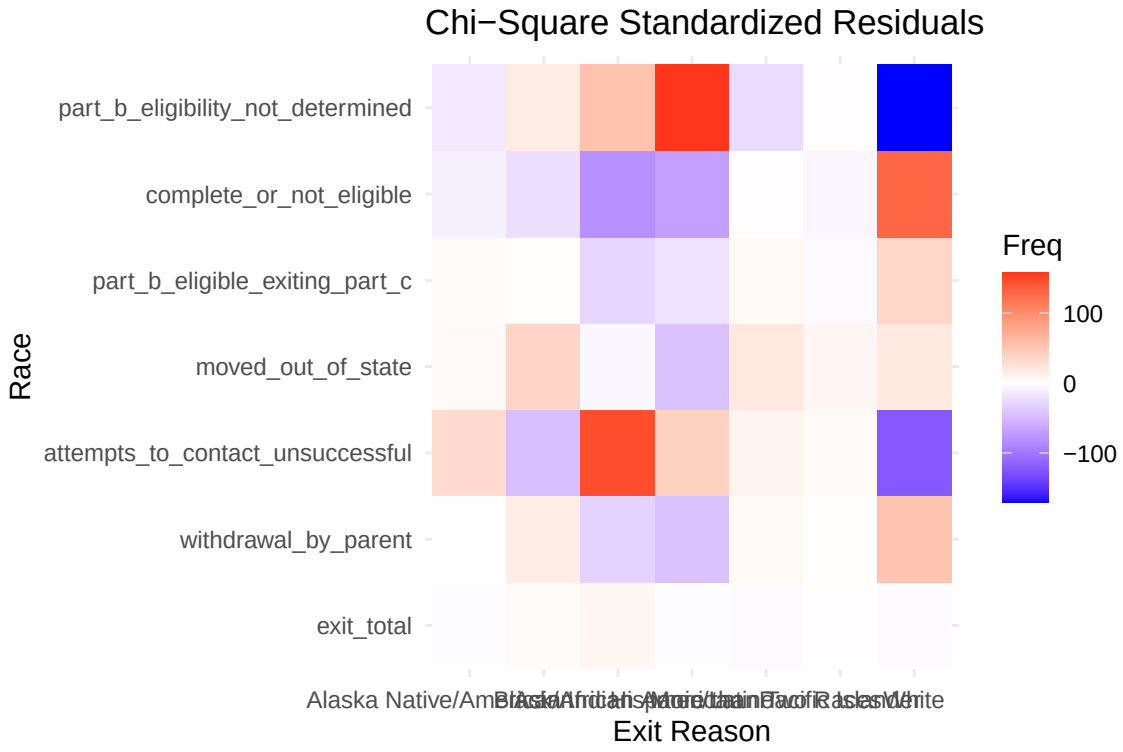
```
                            exit_total withdrawal_by_parent
Alaska Native/American Indian  -2.271421            -1.529332
Asian                           3.969401            16.392776
Black/African American          7.392204           -32.563128
Hispanic/Latino                -2.007889           -44.502779
More than Two Races            -4.250400             5.768394
Pacific Islander                1.568056             1.271574
White                          -2.817433            52.320627
                            attempts_to_contact_unsuccessful
Alaska Native/American Indian                       32.766855
Asian                                              -46.849271
Black/African American                             145.362643
Hispanic/Latino                                     41.042318
More than Two Races                                  9.926835
Pacific Islander                                     5.482329
White                                             -123.606382
                            moved_out_of_state part_b_eligible_exiting_part_c
Alaska Native/American Indian          5.200877                       4.008193
Asian                                 37.157114                       1.334661
Black/African American                -5.841365                     -30.058085
Hispanic/Latino                      -44.978596                     -21.284323
More than Two Races                   20.190835                       5.754274
Pacific Islander                       8.578719                      -4.113435
White                                 18.873265                      35.792944
                            complete_or_not_eligible
Alaska Native/American Indian                -9.771775
Asian                                       -23.542186
Black/African American                      -82.092145
Hispanic/Latino                             -70.955800
More than Two Races                           1.450087
Pacific Islander                             -6.844091
White                                       129.036997
                            part_b_eligibility_not_determined
Alaska Native/American Indian                       -16.004425
```

```
Asian                                18.345906
Black/African American               54.190397
Hispanic/Latino                     157.395198
More than Two Races                 -26.019725
Pacific Islander                      1.420595
White                              -170.810660
```

Chi-square with Residuals 4:

Cameron: How can I reverse the order of Y axis (and I should delete the exit total row too)

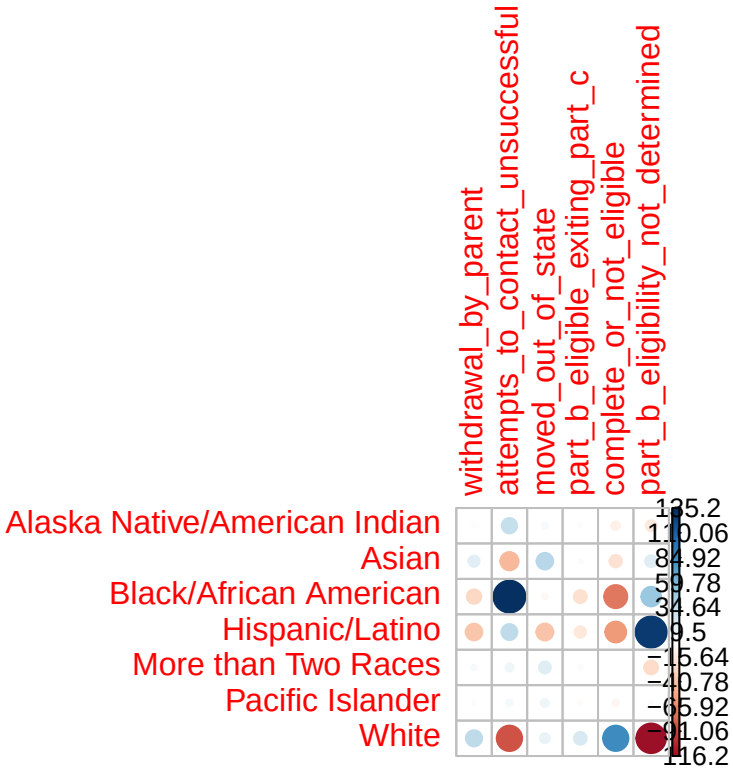
Chi-square with Residuals: Viz 1 (HEATMAP)



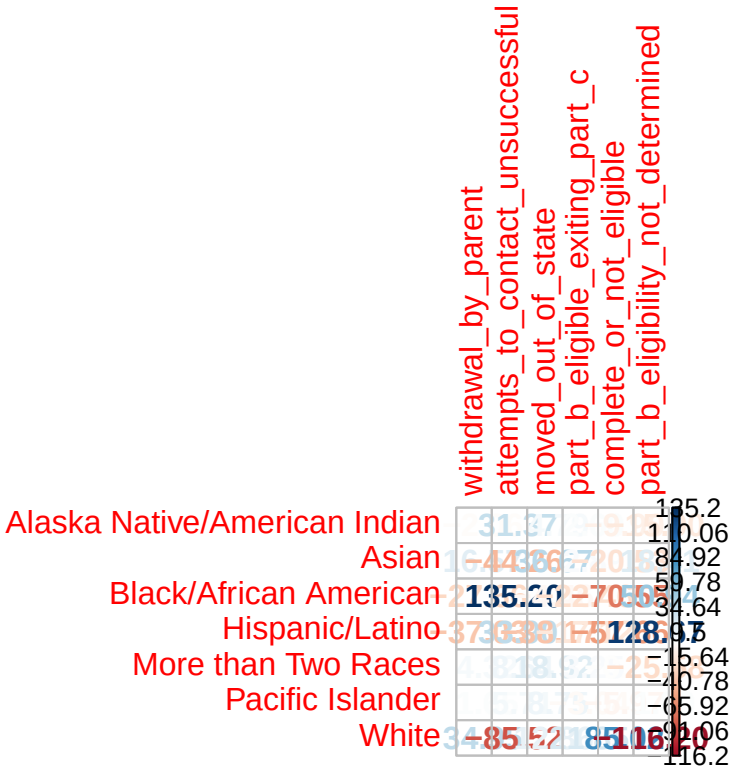
Chi-square with Residuals: Viz 2 (CORRPLOT: https://www.sthda.com/english/wiki/chi-square-test-of-independence-in-r#google_vignette)

```
[1] "exit_total"                      "withdrawal_by_parent"
[3] "attempts_to_contact_unsuccessful" "moved_out_of_state"
[5] "part_b_eligible_exiting_part_c"   "complete_or_not_eligible"
[7] "part_b_eligibility_not_determined"
```

corrplot :) Trial 1: https://cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html



corrplot trial 2:


withdrawal_by_parent
attempts_to_contact_unsuccessful
moved_out_of_state
part_b_eligible_exiting_part_c
complete_or_not_eligible
part_b_eligibility_not_determined
Alaska Native/American Indian
Asian
Black/African American
Hispanic/Latino
More than Two Races
Pacific Islander
White
31.37
135.29
−70
128.97
−85
−116.2
135.2
110.06
84.92
59.78
34.64
9.5
−15.64
−40.78
−65.92
−91.06
−116.2