Group 4: Han Li, Huy Le, [Sourabh Gupta](https://seattleu.instructure.com/groups/251997/users/6199754)

**LMA 2nd Report**

1. Research questions and geographic area for context:

* In the state of Washington, how do earnings vary by educational attainment?
* Does the premium for higher education vary by race and ethnicity?
* State of Washington

1. Insights obtained from the EDA

* 78.41% of Pop is White alone
* 5.12% is two or more than two races
* 3.2% persons are under 3 years old
* 17% of population have Bachelor’ Degree
* 15% of population obtained high school diploma
* 25.30% people are high school or lower than high school educated

Earning:

* + Median: $14000
  + Mean: $33898
  + Min: $-5100
  + Max: $850000
  + 1st Quarter tile: 0
  + 3rd Quarter tile: $48000
  + NA's :11810
* There are observations that their total earnings are <= 0

1. List of sample selection criteria

We will only select:

1. one race people and remove those have two and more races
2. 2.who is older than 3 years old
3. Not NA for Earning

R Code:

prewadata <- subset(wadata, RAC1P != 9, select = c(ID, RAC1P, PERNP, SCHL))

prewadata <- subset(prewadata, !is.na(PERNP))

summary(prewadata)

1. Rationale for each sample selection criteria

* The people who is less than 3 years old have no earning, so they should not be considered in sample.
* The people who have two and more races are hard to be distinguished as one specific race, so we should take them out from the samples.
* The people whose earning is NA should be taken out from the sample.

1. descriptive statistics (after imposing sample selection criteria)

* The number of observations is 56865
* Earning:
  + - Median: $14000
    - Mean: $34127
    - Min: $-5100
    - Max: $850000
    - 1st Quarter tile: 0
    - 3rd Quarter tile: $49000

Cross-tab: Education attainment and Earnings

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Earning | Level1 | Level2 | Level3 | Level4 | Level5 | Level6 |
|  | <= 5000$ | 5001 – 20000$ | 20001-60000$ | 60001-150000$ | 150001-300000$ | >3000000$ |
| HighSchool | 15547 | 819 | 1150 | 191 | 12 | 11 |
| Associate | 13018 | 4105 | 7914 | 2573 | 154 | 80 |
| HigherEduc | 7780 | 2471 | 6481 | 6345 | 1019 | 396 |

Cross-tab Races and Education Attainment

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Races | | | | | | | | | |
| EDU | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | Sum |
| Associate | 22767 | 840 | 637 | 35 | 63 | 1417 | 176 | 744 | 1165 | 27844 |
| HigherEduc | 20352 | 510 | 179 | 12 | 20 | 2431 | 55 | 227 | 706 | 24492 |
| HighSchool | 13634 | 678 | 624 | 25 | 68 | 1588 | 139 | 1454 | 1837 | 20047 |
| Sum | 56753 | 2028 | 1440 | 72 | 151 | 5436 | 370 | 2425 | 3708 | 72383 |