Midterm Report

YU XIA 805230474

1 Staff Demographics

1.1 Methods

We calculate the frequency (mean, median), proportion of variable from the data set, based on different classifications. The frequency is calculated by counting each unique element in the table and the proportion is calculated by finding the percentage of each unique element appearing in the total number of elements.
First, before calculating the frequency and proportion, we create two new variables, *age* and *educational level,* and use these two variables to divide the whole tables into subgroups.

Age

Criteria for *Age*	Label
$17 \leq \text{Age} \leq 30$	"17-30"
$31 \leq \text{Age} \leq 45$	"31-45"
$46 \leq \text{Age} \leq 60$	"46-60"
$61 \leq \text{Age} \leq 75$	"61-75"
$\text{Age} \geq 76$	"76 or above"

Educational Level

Criteria for *Educational Level*	Label
A	Associate Degree
C or B	Baccalaureate
D	Doctorate
U, Y, or F	Fifth year
M or V	Master's Degree
N or S	Not Reported or Special

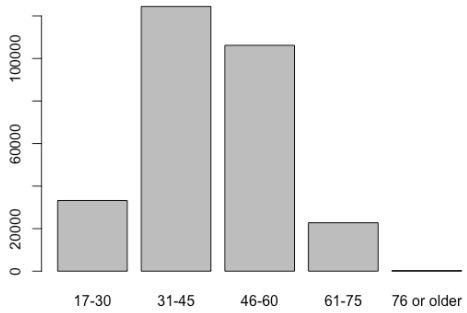
1.2 Tables/Charts

Gender

	Females	Males
Frequency	210046	76798
Percent	0.7322656	0.2677344

Age

	17-30	31-45	46-60	61-75	76 or above
Frequency	33207	124445	106192	22766	234
Percent	0.1157667582	0.4338420884	0.3702081968	0.0793671822	0.0008157744



Education Level

	Associate Degree	Baccalaureate	Doctorate	Fifth year	Master's Degree	Not Reported or Special
Frequency	1008	144013	3874	1814	130513	5622
Percent	0.003514105	0.502060353	0.013505599	0.006323995	0.454996444	0.019599504

Ethnic

	0	1	2	3	4	5	6	7	9
Frequency	10237	1397	16466	832	4762	62924	13300	174692	2234
Percent	0.0356 88388	0.0048 70243	0.0574 04024	0.0029 00531	0.0166 01358	0.2193 66624	0.0463 66666	0.6090 13959	0.0077 88205

*Code 0 = Not reported; Code 1 = American Indian or Alaska Native, Not Hispanic; Code 2 = Asian, Not Hispanic; Code 3 = Pacific Islander, Not Hispanic; Code 4 = Filipino, Not Hispanic; Code 5 = Hispanic or Latino; Code 6 = African American, not Hispanic (formerly known as Black, not Hispanic); Code 7 = White, not Hispanic; Code 9 = Two or More Races, Not Hispanic

Employment Status

	Long term	Probationary	Tenured	Other	Not reported
Frequency	14722	41812	195806	33868	636
Percent	0.051324065	0.145765643	0.682621913	0.118071147	0.002217233

Employment Statistics

	Mean	Standard Deviation
Years Teaching	14.00005	9.357039
Years in District	11.8331	8.898668
Full-Time Equivalent (FTE) Teaching Duties	82.06741	37.96761
Full-Time Equivalent (FTE) Administrative Duties	7.190041	25.74655
Full-Time Equivalent (FTE) Pupil Services Duties	8.217073	27.08383

Years Teaching:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1	6	13	14	20	93

Years in District:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	3.00	11.00	11.83	18.00	63.00

Full-Time Equivalent (FTE) Teaching Duties:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00	100.00	100.00	82.07	100.00	300.00

Full-Time Equivalent (FTE) Administrative Duties:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.00	0.00	0.00	7.19	0.00	500.00

Full-Time Equivalent (FTE) Pupil Services Duties:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.000	0.000	0.000	8.217	0.000	366.000

1.3 Results

- Gender: The number of female are much more than male staff.
- Age: The majority of the staff is between the age of 31 and 60. About 11.58% of the staff is between 17 to 30, and only 0.081% of the staff is over 76 years old.
- Education Level: Most staff have a Master's degree or Baccalaureate (45.5%, 50.2%).
- 60.9% of population are White and not Hispanic, which means staff are lack of diversity.
- Employment Status: About half of staff have tenured employment (51.32%), about 14.57% staff have probationary employment status.
- Years Teaching: The average total years among the staff is about 14.00 years with a standard deviation of around 9.36 years. (median:13 years)
- Years in District: The average total years of service in a certificated position in the district among the staff is about 11.83 years with a standard deviation of around 8.9 years.
- (FTE) Teaching Duties: The average amount of teaching duties is about 82.07% of one full-time equivalent.
- (FTE) Administrative Duties: The average amount of administrative duties is about 7.19% of one full-time equivalent, but 75% of staffs' amounts of administrative duties are 0%, which means very few staff have very high amount of administrative duties.
- (FTE) Pupil Services Duties: The average amount of pupil services is about 8.2% of one full-time equivalent, but 75% of staffs' amounts of Pupil Services Duties are 0%, which means very few staff have very high amount of Pupil Services Duties.

2 District-Level Enrolled Student Demographic Characteristics

2.1 Methods

Firstly, create a new list which contain all unique district-code in table **Dropouts**. Create a new variable named *district_code* in table **Dropouts** which contains variable *CDS_CODE's* first seven digits. Next, we use *district_code* to divide table **Dropouts** into 750 subtables. In each subgroup, the *district_code* is the same.
After data preprocessing, we count total number of enrollments and dropouts by different classifications (such as gender, ethnic and grade) in 750 subtables and achieve 750 data. Next, we calculate these 750 data's 1st Quartile, Median and 3rd Quartile.

2.2 Tables/Charts

Gender

	1st Quartile	Median	3rd Quartile
Female	53.25	345	1242.75
Male	58	373	1333

Ethnic

	1st Quartile	Median	3rd Quartile
American Indian or Alaska Native, not Hispanic	1	4.5	17
Asian, not Hispanic	1	10	107.8
Pacific Islander, not Hispanic	0	1	9
Filipino, not Hispanic	0	4	35
Hispanic or Latino	35.25	271	1262.75
African American, not Hispanic	1	9	74.75
White, not Hispanic	31	158	673.8
Two or More Races, not Hispanic	2	12	62
Not Reported	0	1	7

Grade

	1st Quartile	Median	3rd Quartile
7-8	50	249	978.5
9-10	0	56	777.8
11-12	0	79	845.5

2.3 Results

- The enrollments of male are slighter higher then female.
- Students who are Hispanic or Latino have the highest enrollments in all ethnic.
- Students whose grades are 7-8 have the highest enrollments in all grades.

3 District-Level Dropout Rates by Grade, Gender and Race/Ethnicity

3.1 Methods

We can calculate drop rates by dividing drop-out number by enrollment number in different classifications (such as gender, ethnic and grade) in 750 subtables and achieve 750 drop-out rate for each classification (such as gender, ethnic and grade).
After calculating drop rates, we calculate means and standard deviations in different classifications.
*All numbers in the tables below are dropouts per 100 enrolled students.

3.2 Tables/Charts

Gender

	Mean	Standard Deviation
Female	1.7	6.19
Male	2.14	7.34

Ethnic Group

	Mean	Standard Deviation
American Indian or Alaska Native, not Hispanic	3.06	10.91
Asian, not Hispanic	1.39	7.33
Pacific Islander, not Hispanic	2.07	9.82
Filipino, not Hispanic	0.74	3.90
Hispanic or Latino	1.97	7.22
African American, not Hispanic	3.43	11.46
White, not Hispanic	1.96	7.01
Two or More Races, not Hispanic	2.40	9.79
Not Reported	8.30	33.81

Grade

	Mean	Standard Deviation
7-8	0.467	2.63
9-10	2.0	6.02
11-12	5.20	10.81

3.3 Results

- There is a little higher dropout rate in male group then female group.
- Students whose ethnicity is not reported have the highest dropout rates and students whose ethnicity is Filipino, not Hispanic have the lowest dropout rates
- Among the students with known ethnicity background, American Indian or Alaska Native, not Hispanic and African American, not Hispanic have much higher dropout rates than others.
- Students who are in higher grades tend to have a higher dropout rate.

4 District-Level Student Dropout Rates by Staff Employment Characteristics

4.1 Methods

First, we calculate the mean of different variables (*age of staff, staff Years Teaching and staff Years in District*), *total dropout rate* (total dropout number divided by total dropout number), *Percentage of staff with a Masters or Doctorate* and *Tenured staff in different district* and put these values into a new table named **new_merge.**Next, we create a new variable *median* (when the total dropout rate is above median, we put value True. When the total dropout rate is below median dropout rate, we put value False.) to divide **new_merge** table into two subgroups. Finally, we calculate mean and standard deviation of variables in these two subgroups (above median dropout rate and under median dropout rate).

4.2 Tables/Charts

above median dropout rate:

	Mean	Standard Deviation
Percentage of staff with a Masters or Doctorate	0.2856162	0.1551788
Average age of staff	44.7864	2.880804
Percentage of Tenured staff	0.5718011	0.2340119
Average staff Years Teaching	13.18018	2.501607
Average staff Years in District	10.07437	3.098078

below median dropout rate:

	Mean	Standard Deviation
Percentage of staff with a Masters or Doctorate	0.2283484	0.1595031
Average age of staff	45.24685	4.261559
Percentage of Tenured staff	0.5627117	0.2592419
Average staff Years Teaching	13.32285	3.690048
Average staff Years in District	9.62149	3.121573

4.3 Results

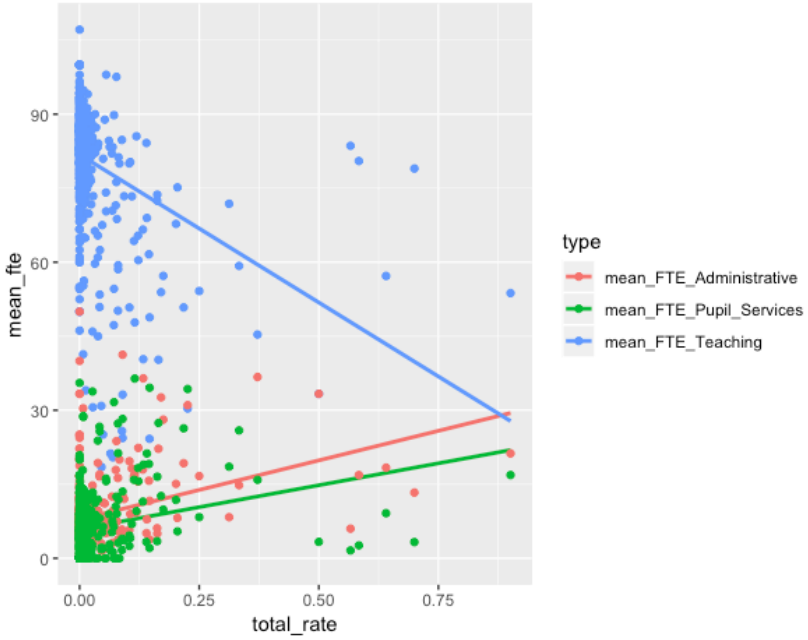
- There is a little higher Percentage of staff with a Masters or Doctorate in table1(above median dropout rate).
- Average staff Years in District in table1 are higher than it in table2 (below median dropout rate).
- Average age of staff is almost same in both tables while the standard deviation in table2 is higher.
- Average staff Years Teaching and average staff Years in District seem like the same in both tables.

5 Plot of District-Level Student Dropout Rates by FTE

5.1 Methods

First, we calculate the mean of different variables (*FTE Teaching, FTE Administrative and FTE Pupil Service*) and *total dropout rate* in different districts and put these values into a new table named **FTE**.
Next, we plot y (*Mean FTE Teaching, Mean FTE Administrative and Mean FTE Pupil Service*) and x (*total dropout rate)* in one picture.

5.2 Plot



5.3 Results

- When total dropout rate increases, the mean FTE Administrative and FTE Pupil Services increase, and the mean FTE Administrative increases faster.
- When total dropout rate increases, the mean FTE Teaching goes down.